Supplementary Information

Autonomous Multistep Organic Synthesis in a Single Isothermal Solution Mediated by a DNA Walker

Yu He and David R. Liu*

Supplementary Methods

General Methods. DNA oligonucleotides were synthesized on a PerSeptive Biosystems Expedite 8090 DNA synthesizer using standard phosphoramidite protocols and purified either by reverse-phase HPLC using Eclipse XDB-C18 (Agilent) columns and an acetonitrile/0.1 M triethylammonium acetate (TEAA) gradient, or by 15% denaturing polyacrylamide gel electrophoresis (PAGE). Modified phosphoramidites and CPG for DNA synthesis were purchased from Glen Research. The 3'-amino modified oligonucleotides were synthesized using 3'-Amino-Modifier C7 CPG. The 5'-amino modified oligonucleotides were synthesized using the 5'-Amino-Modifier C6 phosphoramidite. The 5'-thiol modified oligonucleotides were synthesized using 1'-[(2-cyanoethyl)-(N,N-diisopropyl)]-phosphoramidite (Thiol-Modifier C6 S-S). RNA nucleotides were incorporated using the A-TOM-CE phosphoramidite and the U-TOM-CE phosphoramidite. Oligonucleotides were quantitated by UV and all modified DNAs and reagents were characterized by LC/MS on Waters Acquity UPLC equipped with a Waters Acquity UPLC BEH C18 column using an aqueous 6 mM tetraethyl ammonium bicarbonate (TEAB)/MeOH mobile phase. Electrospray mass spectrometry was carried out on a Waters Q-TOF premier instrument. All chemicals, unless otherwise noted, were purchased from Sigma-Aldrich.

Oligonucleotide Sequences

T123: 5'-ACC ATC TGT GGC ATA GCA GCG TCG ATC TTG AGC ATT GGA GCG AGT ATC TAA CGC ATG GAA GCG GCA ATC TCC TGC ATC ATC GCG

T132: 5'-ACC ATC TGT GGC ATA GCA GCG TCG ATC TTG AGC ATT GGA GCG GCA ATC TCC TGC ATC ATC GCG AGT ATC TAA CGC ATG GAA GCG

T231: 5'-ACC ATC TGT GGC ATA GCA GCG AGT ATC TAA CGC ATG GAA GCG GCA ATC TCC TGC ATC ATC GCG TCG ATC TTG AGC ATT GGA GCG

T321: 5'-ACC ATC TGT GGC ATA GCA GCG GCA ATC TCC TGC ATC ATC GCG AGT ATC TAA CGC ATG GAA GCG TCG ATC TTG AGC ATT GGA GCG

W: 5’-AGT GCT GAT TCG GAC AGG CTA GCT ACA ACG AGA GTG AC /3AmMC7/

S0: 5’-TGT CCG AAT CAG CAC TTT CGC TGC TAT GCC ACA GAT GGT

S1: 5’-/5ThioMC6-D/ GTC ACT CrArU GTC CGA ATC AGC ACT TTC GCT CCA ATG CTC AAG ATC GA

S2: 5’-/5ThioMC6-D/ GTC ACT CrArU GTC CGA ATC AGC ACT TTC GCT TCC ATG CGT TAG ATA CT

S3: 5’-/5ThioMC6-D/ GTC ACT CrArU GTC CGA ATC AGC ACT TTC GCG ATG ATG CAG GAG ATT GC

S2a: 5’-/5ThioMC6-D/ TTT TTT GTC ACT CrArU GTC CGA ATC AGC ACT TTC GCT CCA ATG CTC AAG ATC GA

S3a: 5’-/5AmMC6/ TTT TTT TTG TCA CTCr ArUG TCC GAA TCA GCA CTT TCG CGA TGA TGC AGG AGA TTG C

T in Fig. S3: 5'-ACC ATC TGT GGC ATA GCA GCG AGT ATC TAA CGC ATG GAA GCG TCG ATC TTG AGC ATT GGA GCG GCA ATC TCC TGC ATC ATC GCG

Synthesis of Substrates. Substrate oligonucleotides were synthesized as described above on CPG resin using the Thiol-Modifier C6 S-S, A-TOM-CE, and U-TOM-CE phosphoramidites. After synthesis, the oligonucleotide was purified using reverse-phase HPLC (flow rate 4mL/min, column size: 9.4x250mm, retention time ~36min). To deprotect the TOM groups on the RNA residues, the lyophilized DNA pellet was dissolved in 200 μL anhydrous DMSO and treated with 250 μL triethylamine trihydrofluoride. The reaction was incubated at 65 ºC for 2 hours. After the reaction, the mixture was subject to butanol precipitation and HPLC (flow rate 4 mL/min, column size: 9.4x250 mm, retention time ~24 min) purification to provide the deprotected oligonucleotide.

A 5’ NHS group was appended to the 5’ thiol modification as follows. In a typical process, 100 mM DTT, pH 8.5 was added to the oligonucleotide at 23 ºC for 30 min to cleave the S-S bond. The reaction was desalted by gel filtration with a NAP-5 columns (GE Healthcare Life Sciences) and added directly to 250 μL of a 40 mg/mL solution of *N*-hydroxymaleimide in 0.5 M MOPS, pH 7.5. After 30 min, the reaction was desalted by gel filtration and purified by reverse-phase HPLC (flow rate 4 mL/min, column size: 9.4x250 mm, retention time ~14 min) to generate a 5’ NHS-linked oligonucleotide.

HPLC gradient for 9.4x250 mm C18 column, 4 mL/min flow rate:

time (min)	% 0.1 M TEAA pH 7.0	% acetonitrile
0	92	8
24	80	20
34	60	40
34.2	20	80
36.8	20	80
37	92	8
41	92	8

HPLC gradient for 4.6x250 mm C18 column, 1 mL/min flow rate:

time (min)	% 0.1 M TEAA pH 7.0	% acetonitrile
0	92	8
15	60	40
20	20	80
22	20	80
22.5	92	8
25	92	8

NVOC-protected amino acids were prepared using a previously described protocol.[1-2] Briefly, an amino acid (0.3 mmol) was mixed with Na_2CO_3 (0.3 mmol) in 8 mL H_2O. An equimolar amount of 4,5-dimethoxy-2-nitrobenzyl chloroformate (0.3 mmol) dissolved in 8 mL dioxane was added slowly with stirring to the aqueous solution. After stirring at room temperature for 1 h, the reaction was diluted with 7.5 mL dichloromethane followed by acidification with 5 mL of 1 M $NaHSO_4$. The organic phase was collected and the aqueous phase was washed with dichloromethane. The combined organic extracts were dried over anhydrous Na_2SO_4 and then concentrated *in vacuo* to give a solid crude product. The crude protected amino acids were used directly for DNA conjugation.

To conjugate NVOC-protected amino acids or biotin with DNA, the desired amino acid or biotin (~2 mg) was mixed with ~0.5 mg EDC·HCl in 200 μL DMF. After 5 min, 20 μL of the mixture was added directly to an aliquot of NHS-linked oligonucleotide in 100 μL 0.1 M MES pH 6.0. After 5 min, the reaction was desalted with a NAP-5 column and purified by reverse-phase HPLC (flow rate of 1 mL/min, 4.6x250 mm C18 column). Amino acid 4-(aminomethyl) benzoic acid was conjugated to the appropriate DNA oligonucleotide to generate **S1** (with NVOC, retention time: ~13.5 min);

3'-(aminomethyl) biphenyl-4-carboxylic acid (Amatek) was conjugated to the appropriate DNA oligonucleotide to generate **S2** (with NVOC, retention time: ~16.1 min); *trans*-4-(aminomethyl) cyclohexane carboxylic acid was conjugated to the appropriate DNA oligonucleotide to generate **S3** (with NVOC, retention time: ~14.2 min); biotin was conjugated to the appropriate DNA oligonucleotide to generate **S2a** (retention time: ~11.8 min).

By adding 1% TFA to the HPLC fractions before lyophilization, the NHS ester-linked DNA could be recovered in pure form; without the addition of TFA, NHS esters would partially hydrolyze during lyophilization. The lyophilized oligonucleotide samples were stored at -80 °C before using. Exposure of the redissolved oligonucleotides (in 0.1 M NaOAc, pH 5.0) to a Spectroline® E-Series handheld UV lamp (365 nm) for 1.5 h at 4 °C resulted in the quantitative deprotection of the NVOC protection group as assessed by HPLC. The photodeprotected DNA-linked substrates, after desalting (Princeton Separations) and UV quantification were directly combined with the other components of the DNAsome.

DNAsome-Mediated Multistep Synthesis. Photodeprotected substrates **S1**-**S3** were quantified by UV absorbance at 260 nm and mixed with **T** at ~1.5 μM in a 1:1:1:1 stoichiometric ratio in 5 mM MOPS, pH 5, with 10 mM $Mg(OAc)_2$ at room temperature. Meanwhile, 1 equivalent of **W** and **D** were mixed in a separate tube in the same buffer. After 1 h incubation at 23 °C, the two solutions were combined and the buffer was adjusted to 50 mM MOPS, pH 7.5 with 10 mM $Mg(OAc)_2$. The combined solution was incubated for 6-8 hours at 23 °C. 1 M Tris pH 8.0 and 100 mM EDTA were added to quench the reaction. The solution was desalted with a Sephadex G-25 spin column (Princeton Separations). The desalted solution was directly subjected to LC/MS analysis.

LC/MS Analysis of the DNAsome Reaction Mixture. Oligonucleotide reaction mixtures were characterized by LC/MS on a Waters Acquity UPLC using a Waters BEH C18 (1.7 μm, 1.0x50 mm) column and an aqueous 6 mM tetraethyl ammonium bicarbonate (TEAB)/MeOH mobile phase. Low mobile phase flow rate (30 μL/min) was used to achieve better mass spectrometry sensitivity.

UPLC gradient for LC/MS analysis, 30 μL/min flow rate:

time (min)	% 6 mM TEAB pH 8.5	% methanol
0	100	0
2	88	12

25	60	40
26	0	100
30	0	100
31	100	0
35	100	0

After the DNAsome reaction, the solution was desalted and directly injected into the LC/MS. A typical UPLC Total Ion Count (TIC) spectrum is shown below. For each sample's mass spectrum analysis, the mass signals were obtained from the TIC spectrum from 10 min to 26 min, which encompasses all DNA-linked species injected into the mass spectrometer (the broad peak from 7-10 min contains buffer components). Under our experimental conditions, we observed that m/z values for ions of -5 charge exhibit the best signal:noise ratio. Using the -4 ions instead resulted in very similar relative ion intensities.

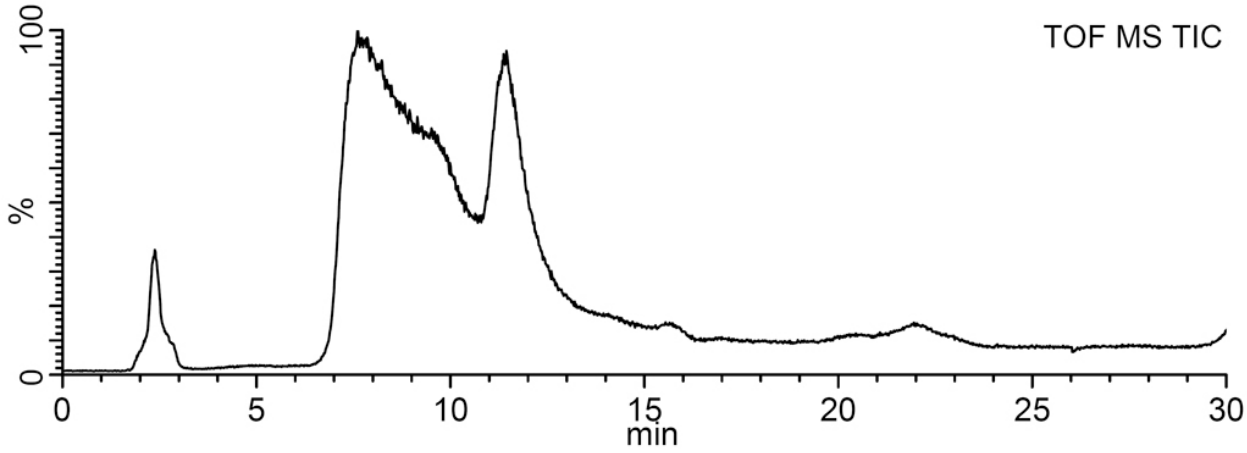


Figure S1. A typical UPLC total ion counts (TIC) trace of the DNAsome reaction mixture.

Estimation of DNAsome-Mediated Multistep Reaction Yields. Since neither PAGE nor HPLC can separate final products from starting materials and truncated products in this work, we used mass spectrometry to estimate relative product yields by integrated ion counts. Although the ionization efficiency of different molecules can differ widely, we hypothesized that the relative ionization efficiencies (RIE) of DNA-linked oligoamide products to form the observed -5 charged ions would be dominated by the DNA moiety and that oligoamide products linked to the same DNA sequence would therefore exhibit very similar negative ion mode ionization efficiencies. The determination and use of ESI-MS RIE (or relative response factors (RRF)) for LC/MS quantitation of components in a

mixture has been previously described as a way to estimate the relative abundance of each analyte in a simple mixture.[3-5] In order to characterize the RIE of DNA-linked products of interest, we synthesized authentic DNA-polyamide conjugates **W**+**S3**, **W**+**S1**+**S3**, **W**+**S2**+**S3**, and **W**+**S1**+**S2**+**S3** as standards by sequential reaction of **W** with NVOC-protected amino acid substrates, followed by UV-deprotection and HPLC purification after every coupling reaction. The authentic standards were quantitated by UV spectroscopy then combined in different stoichiometric ratios. The resulting mixtures were analyzed by LC/MS under the same conditions used to analyze DNAsome reactions (Figure S2). The results confirm that the ionization efficiencies of these DNA-linked products are comparable (Figure S2a). Next we prepared a mixture of the abovementioned four authentic DNA-linked products in a 15:25:15:45 ratio and compared the mass spectra of this mixture with that of the three-step DNAsome reaction product (Figure S2b). Finally, we prepared a DNAsome reaction using identical conditions and procedures as the DNAsome-mediated multistep synthesis described above, except **W** was replaced with a 15:25:15:45 mixture of authentic **W**+**S3**:**W**+**S1**+**S3**:**W**+**S2**+**S3**:**W**+**S1**+**S2**+**S3** (Figure S2c). The very similar relative ion counts of these three spectra further suggests that the estimated yield of ~45% of **W**+**S1**+**S2**+**S3** among detectable reaction products is reasonable (compare Figures S2b, S2c, and S2d).

Additional Supplementary Figures

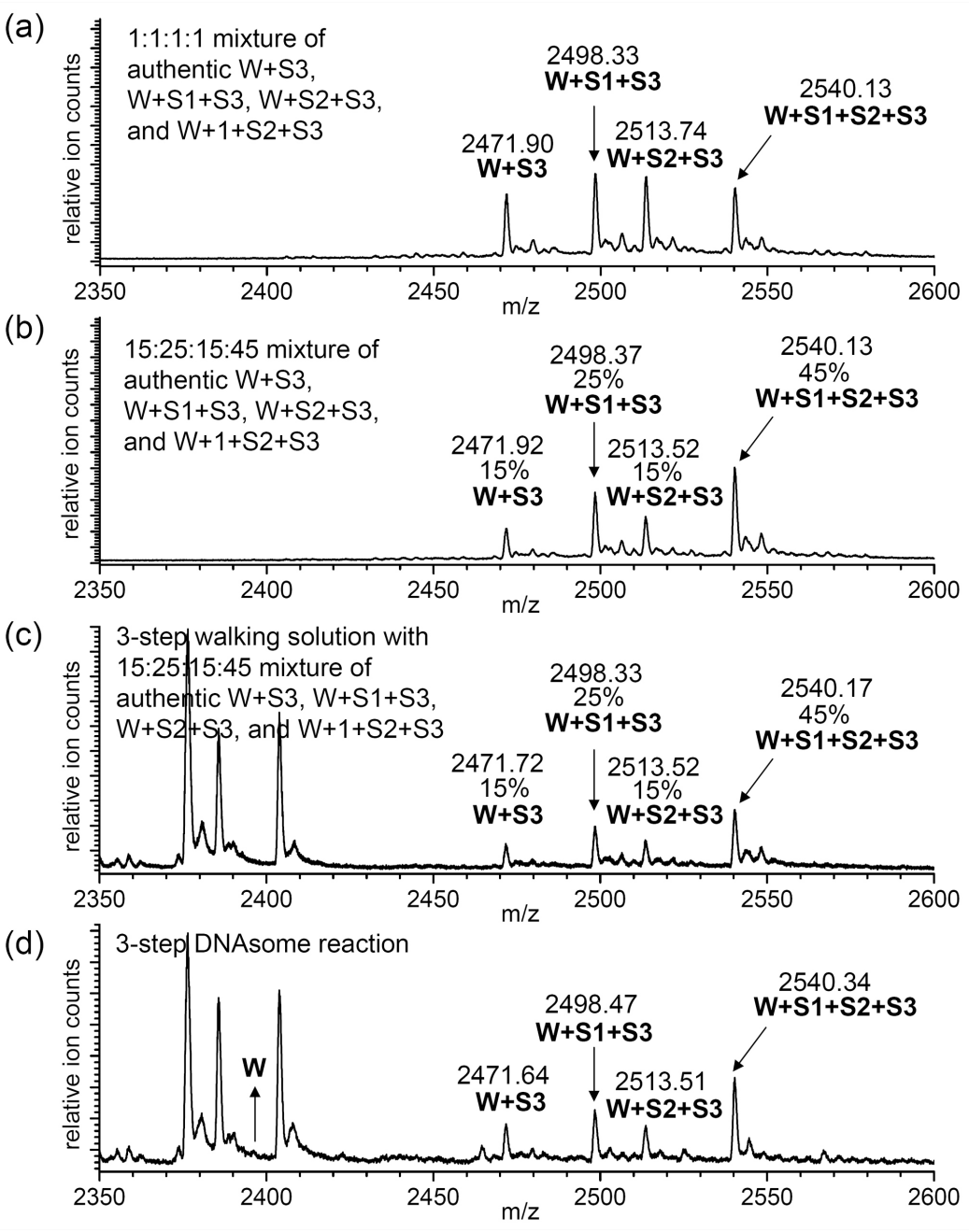


Figure S2. (a) MS analysis of an equimolar mixture of authentic **W**+**S3**, **W**+**S1**+**S3**, **W**+**S2**+**S3**, and **W**+**S1**+**S2**+**S3** (20 pmol DNA in total). (b) MS analysis of a 15:25:15:45 mixture of authentic **W**+**S3**, **W**+**S1**+**S3**, **W**+**S2**+**S3**, and **W**+**S1**+**S2**+**S3** (20 pmol DNA in total). (c) MS analysis of a DNAsome reaction mixture in which a 15:25:15:45 mixture of authentic **W**+**S3**:**W**+**S1**+**S3**:**W**+**S2**+**S3**:**W**+**S1**+**S2**+**S3** replaced strand **W**. (d) MS analysis of the three-step DNAsome reaction mixture. Observed m/z values of ions of -5 charge are shown above each peak.

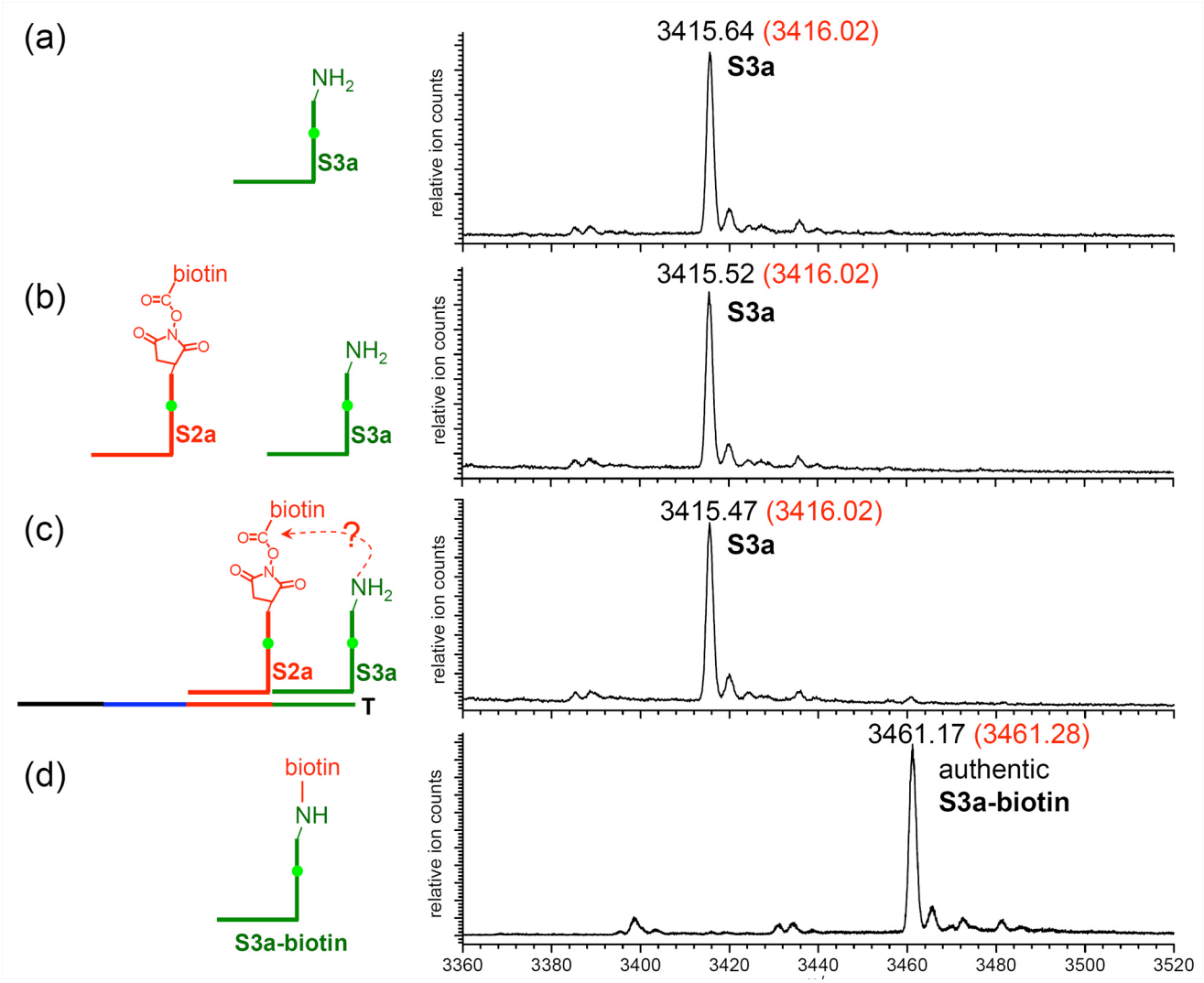


Figure S3. Mass spectroscopy analysis of the possibility of undesired intermolecular reaction and cross-reaction between adjacently hybridized substrates. Mass spectra in (a) and (d) show **S3a** and authentic **S3a-biotin**. (b) 2 μM each of S2a and S3a without any DNA track were combined in 50 mM MOPS buffer (pH 7.5) with 10 mM $Mg(OAc)_2$ and incubated at room temperature for 16 hours. No significant formation of an intermolecular reaction product is observed. (c) 2 μM each of **S2a**, **S3a**, and **T** were combined in 50 mM MOPS buffer (pH 7.5) with 10 mM $Mg(OAc)_2$ and incubated at room temperature for 16 hours. No significant cross-reaction between adjacently hybridized substrates was observed. All samples were analyzed by LC/MS analysis using an aqueous 6 mM tetraethyl ammonium bicarbonate (TEAB)/MeOH mobile phase and on a Waters Q-TOF Premier mass spectrometer. Each spectrum shows total ion counts for the entire sample. Observed *m/z* values reflect ions of -5 charge and are shown in black; expected *m/z* values are shown in red in parentheses.

Supplementary Data

The observed masses and expected masses of all significant species in Figures 2 and 3 are shown below. Observed *m/z* values reflect ions of -5 charge and are shown in black; expected *m/z* values are shown in red in parentheses.

Figure 2 with Observed and Expected *m/z* Values:

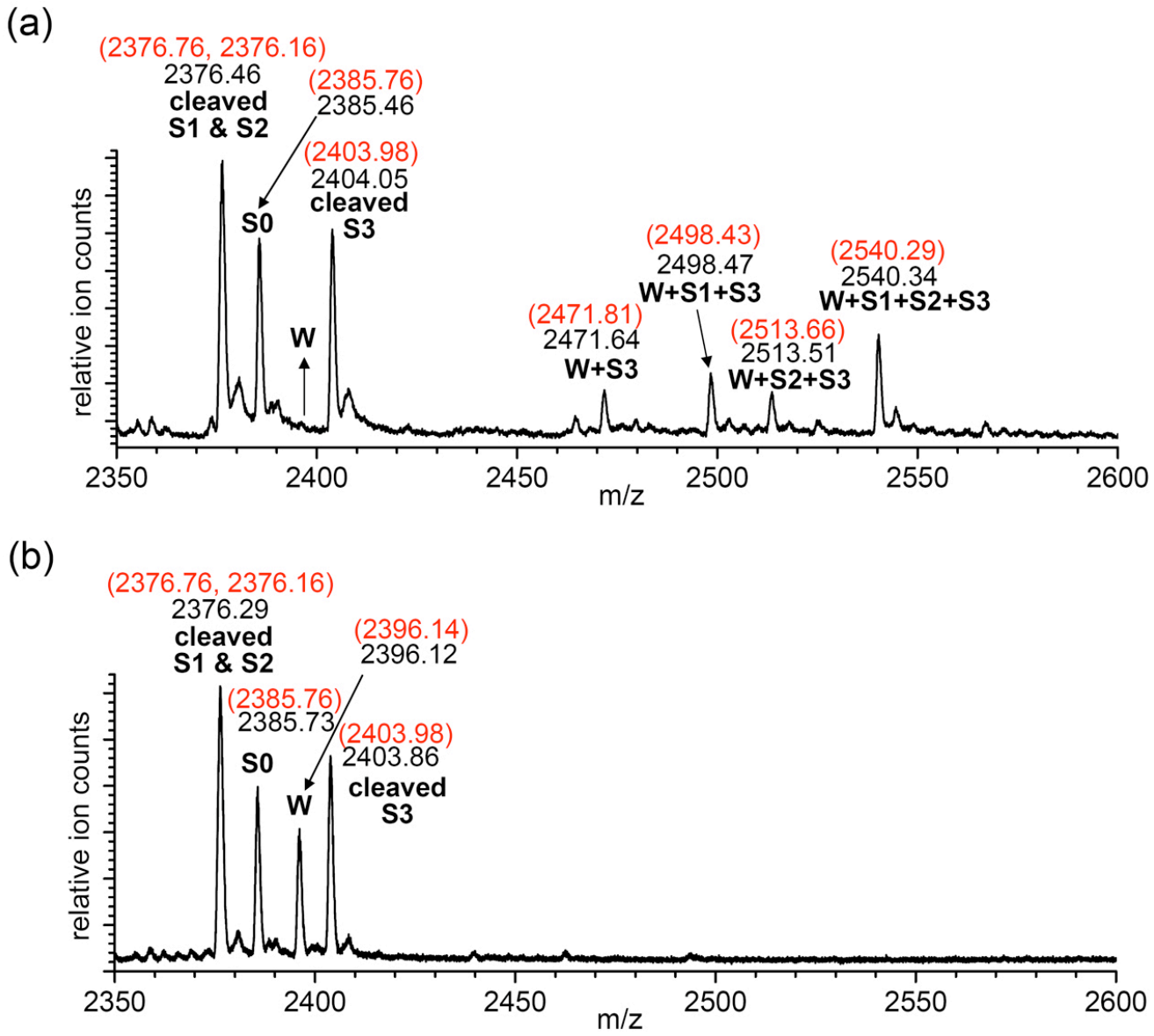
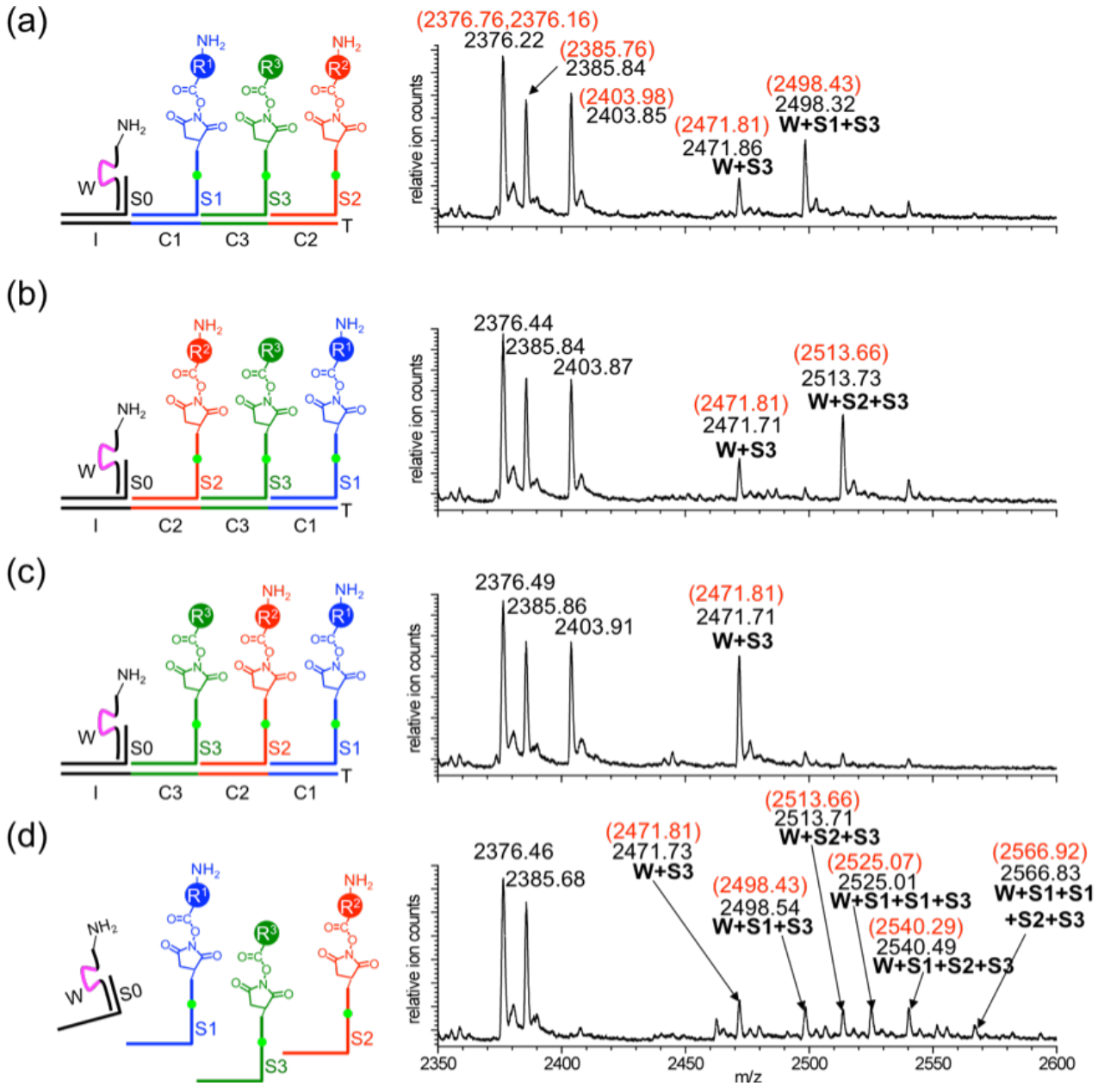


Figure 3 with Observed and Expected *m/z* Values:



The expected masses and *m/z* values of all species in Figures 2 and 3 are as follows:

	expected mass (Da)	*m/z* at -5 charge
cleaved **S1**	11888.8	2376.76
cleaved **S2**	11885.8	2376.16
cleaved **S3**	12024.9	2403.98
S0	11933.8	2385.76
W	11985.7	2396.14
W+3	12364.0	2471.81
W+1+3	12497.2	2498.43
W+2+3	12573.3	2513.66
W+1+2+3	12707.5	2540.29
W+1+1+3	12630.4	2525.07
W+1+2+1+3	12839.6	2566.92
S3a	17085.1	3416.02
S3a+biotin	17311.4	3461.28

Supplementary References Cited

1 Snyder, T. M. & Liu, D. R. Ordered multistep synthesis in a single solution directed by DNA templates. *Angew. Chem. Int. Ed.* **44**, 7379-7382 (2005).

2 Robertson, S. A., Ellman, J. A. & Schultz, P. G. A General and Efficient Route for Chemical Aminoacylation of Transfer-Rnas. *J. Am. Chem. Soc.* **113**, 2722-2729 (1991).

3 Chalcraft, K. R., Lee, R., Mills, C. & Britz-McKibbin, P. Virtual Quantification of Metabolites by Capillary Electrophoresis-Electrospray Ionization-Mass Spectrometry: Predicting Ionization Efficiency Without Chemical Standards. *Anal. Chem.* **81**, 2506-2515 (2009).

4 Oss, M., Kruve, A., Herodes, K. & Leito, I. Electrospray Ionization Efficiency Scale of Organic Compounds. *Anal. Chem.* **82**, 2865-2872 (2010).

5 Leito, I. *et al*. Towards the electrospray ionization mass spectrometry ionization efficiency scale of organic compounds. *Rapid Commun. Mass Sp.* **22**, 379-384 (2008).