Electronic Supplementary Material (ESI) for Chemical Science.

Supplementary Information

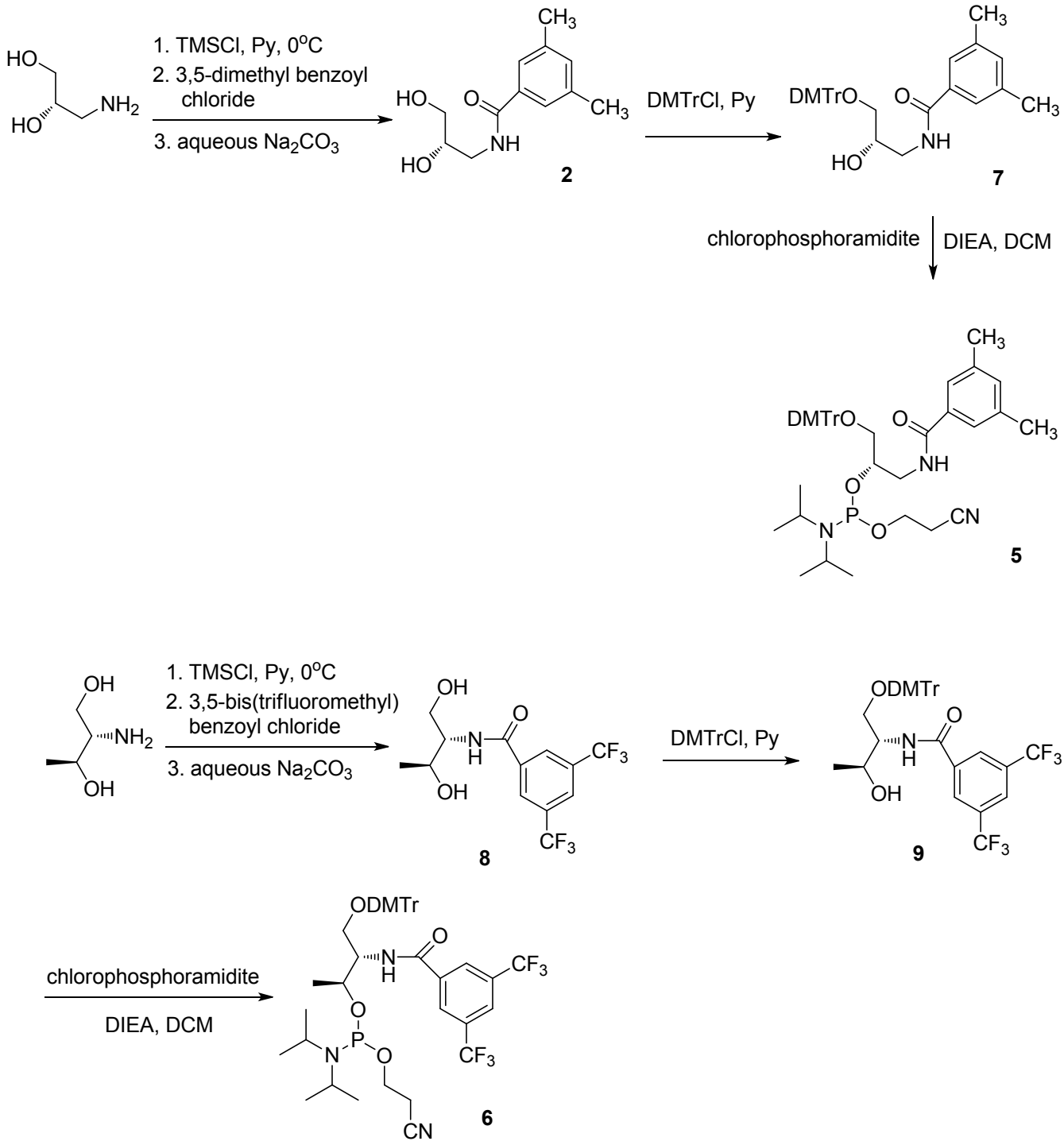
Trifluoromethylated Nucleic Acid Analogues Capable of Self-Assembly through Hydrophobic Interactions

RuoWen Wang, Chunming Wang, Yang Cao, Zhi Zhu, Chaoyong Yang, Jianzhong Chen, Feng-Ling Qing, and Weihong Tan

Table of Contents

I. Synthetic protocols

Scheme S1. The synthetic routes to phosphoramidite **5** and **6**.



Synthesis of compound 2

To a solution of (S)-3-amino-1,2-propanediol (455mg, 5 mmol) in anhydrous pyridine (25 mL) was added chlorotrimethylsilane (2.70 g, 25 mmol) dropwise at 0°C. After 30 min, 3,5-dimethylbenzoyl chloride (1.01g, 6 mmol) was added at 0°C, and the mixture was allowed to warm to RT and stirred for an additional 2 h. Then a saturated $NaHCO_3$ solution was added to terminate the reaction. The resulting reaction mixture was concentrated *in vacuo* and subjected to a flash silica gel column (CH_2Cl_2/MeOH, 10:1) to give compound **2** (950 mg, 85 % yield): ^{1}H NMR (300 MHz, CD_3OD) δ 8.32 (s, 1H), 7.40 (s, 2H), 7.09 (s, 1H), 3.82 (m, 1H), 3.55 (m, 3H), 3.40 (m, 1H), 2.28 (s, 6H); ^{13}C NMR (100 MHz, CD_3OD) δ 169.64, 137.96, 133.95, 132.75, 124.66, 70.74, 63.80, 42.59, 19.97; MS (ESI-): m/z 222.10 (Calculated M-H: 222.13).

Synthesis of compound 7

To a solution of compound **2** (900 mg, 4.0 mmol) in anhydrous pyridine (30 mL) was added DMTrCl (1.49 g, 4.4 mmol), and the reaction was stirred overnight. The reaction mixture was concentrated *in vacuo*, and the residue was subjected to a flash silica gel column (ethyl acetate/hexane, 1:1). Compound **7** (1.60 g, 76 % yield) was isolated as a white foam: ^{1}H NMR (300 MHz, CD_3COCD_3) δ 7.59 (s, 1H), 7.49 (d, J=5.7 Hz, 2H), 7.44 (s, 2H), 7.36 (d, J=6.6 Hz, 4H), 7.29 (t, J=5.1 Hz, 2H), 7.20 (t, J=5.1 Hz, 1H), 7.15 (s, 1H), 6.85 (d, J=6.6 Hz, 4H),4.00 (m, 1H), 3.69 (s, 6H), 3.65 (m, 1H), 3.48 (m, 1H), 3.10-3.20 (m, 2H).

Synthesis of phosphoramidite 5

To a solution of compound **7** (1.58 g, 3.0 mmol) in anhydrous DCM (30 mL) was added DIEA, followed by chlorophosphoramidite (780 mg, 3.30 mmol) at 0°C. The mixture was allowed to warm to RT and stirred for 1 h. Then the reaction mixture was diluted with 45 mL of DCM and washed with saturated $NaHCO_3$ solution and saturated saline solution. The organic phase was dried over Na_2SO_4 and then concentrated *in vacuo*. Phosphoramidite **5** (1.76 g, 81 % yield) was obtained as a white foam: ^{1}H NMR (400 MHz, acetone-d6) δ: 7.52 (d, J=8.0 Hz, 2H), 7.37 (m, J=8.8 Hz, 7H), 7.28 (m, 2H), 7.20 (m, 1H), 7.15 (s, 1H), 6.85 (m, 4H), 4.25-4.32 (m, 1H), 3.60-3.95 (m, 12H), 3.16-3.36 (m, 2H), 2.26-2.72 (m, 2H), 1.12-1.21 (m, 12H); ^{31}P NMR(acetone-d6) δ: 148.87, 148.43.

Synthesis of compound 8

To a solution of D-threoninol (1.05g, 10 mmol) in anhydrous pyridine (50 mL) was added chlorotrimethylsilane (5.43 g, 50 mmol) dropwise at 0°C. After 30 min, 3,5-bis(trifluoromethyl)benzoyl chloride (3.06 g, 11 mmol) was added at 0°C, and the mixture was allowed to warm to RT and was stirred for an additional 2 h. Then a saturated $NaHCO_3$ solution was added to terminate the reaction. The resulting reaction mixture was concentrated *in vacuo*, and the residue was subjected to a flash silica gel column (CH_2Cl_2/MeOH, 15:1) giving compound **8** (2.98 g, 86% yield) as a white solid: ^{1}H NMR (300 MHz, acetone-d6) δ 8.26 (s, 2H), 7.97 (s, 1H), 7.37 (d, *J*=7.2Hz, 1H), 4.27 (m, 1H), 4.11 (m, 1H), 3.89 (m, 2H), 3.58 (br, 2H), 1.24 (d, 3H); MS (ESI-): m/z 344.0733 (Calculated M-H: 344.0725).

Synthesis of compound 9

To a solution of compound **8** (1.73 g, 5.0 mmol) in anhydrous pyridine (35 mL) was added DMTrCl (1.86 g, 5.5 mmol), and the reaction was stirred overnight. The reaction mixture was concentrated *in vacuo*, and the residue was subjected to a flash silica gel column (ethyl acetate/hexane, 1:2). Compound **9** (2.62 g, 81 % yield) was isolated as a white foam: ^{1}H NMR (300 MHz, acetone-d6) δ 8.55 (s, 2H), 8.21 (s, 1H), 8.01 (d, *J*= 8 Hz, 1H), 7.46 (d, *J*= 7.6 Hz, 2H), 7.33 (d, *J*= 7.6 Hz, 4H), 7.25 (t, *J*= 7.6 Hz, 2H), 7.18 (t, *J*= 7.6 Hz, 1H), 6.82 (d, *J*= 8 Hz, 4H), 4.33 (m, 1H), 4.22 (m, 1H), 3.87 (m, 1H), 3.75 (s, 6H), 3.41 (t, *J*=6.8 Hz, 1H), 3.41 (t, *J*=7.2 Hz, 1H), 1.15 (d, *J*=6.4 Hz, 3H); ^{13}C NMR (100 MHz, acetone-d6) δ 164.20, 158.64, 145.39, 137.57, 136.09, 136.03, 131.25 (q, J_{C-F}= 34.0 Hz), 128.12, 128.05, 127.58, 126.57, 124.75, 124.51, 122.04, 112.90, 85.86, 65.97, 63.29, 55.99, 54.53, 19.93; ^{19}F NMR (282 MHz, acetone-d6) δ -63.70; MS (ESI+):m/z 670.1993 (Calculated M+Na: 670.1998).

Synthesis of phosphoramidite 6

To a solution of compound **9** (1.94g, 3.0 mmol) in anhydrous DCM (30 mL) was added DIEA, followed by chlorophosphoramidite (780 mg, 3.30 mmol) at 0°C. The mixture was allowed to warm to RT and stirred for 1 h. Then the reaction mixture was diluted with 50 mL of DCM and washed with saturated $NaHCO_3$ solution and saturated saline solution. The organic phase was dried over Na_2SO_4 and then concentrated *in vacuo*. The residue was subjected to a flash silica gel column (ethyl acetate/hexane, 1:8) to give **6**(2.08 g, 82% yield) as a white solid foam: ^{1}H NMR (300 MHz, acetone-d6) δ 8.48 (s, 2H), 8.25 (s, 1H), 7.88 (d, *J*=8.7Hz, 1H), 7.49 (d, J=6.4 Hz, 2H), 7.21-7.38 (m, 7H), 6.86 (d, *J*=9.0 Hz, 4H),4.45-4.58 (m, 2H), 3.77 (s, 6H), 3.57-3.70 (m, 4H), 3.45-3.50 (m, 1H), 3.29-3.35 (m, 1H), 2.63 (t,

J=6Hz, 2H), 1.10-1.25 (m, 15H); ^{19}F NMR (282 MHz, acetone-d6) δ: -63.67; ^{31}P NMR(acetone-d6) δ: 148.75.

Synthesis, labeling, and purification of oligonucleotides:

All DNA synthesis reagents were purchased from Glen Research. FNA **F6**, **F4**, **FM6** and **M6** were synthesized and purified by Sangon Biotech (Shanghai). FNA **Ft6**, **Ft8** and other oligonucleotides were synthesized on an ABI 3400 synthesizer (Applied Biosystems). Dabcyl CPG was used for all FAM-labeled FNA. The completed sequences were then deprotected in AMA (ammonium hydroxide/40% aqueous methylamine, 1:1) at 65 °C for 30 min and further purified by reversed-phase HPLC (ProStar; Varian) on a C-18 column using 0.1 M triethylamine acetate(TEAA) buffer (Glen Research) and acetonitrile (SigmaAldrich) as the eluents. The collected DNA products were dried and detritylated by dissolving and incubating DNA products in 200 μL of 80% acetic acid for 20 min. The detritylated DNA product was precipitated with NaCl (3 M, 25 μL) and ethanol (600 μL).

Table S1. Detailed sequence information of oligonucleotides[a].

FNA	Sequence [a]
Oligo1	5’ GCG TAC ACA TGC G
Oligo2	5’ CGC ATG TGT ACG C
F1-a	5’ GCG TAC **F**CA TGC G
F1-b	5’ CGC ATG **F**GT ACG C
F2-a	5’ GCG TAC **FF**C ATG CG
F2-b	5’ CGC ATG **FF**G TAC GC
F3-a	5’ GCG TAC **FFF** CAT GCG
F3-b	5’ CGC ATG **FFF** GTA CGC
F3-c	5’ GCG TA**F** C**F**C **F**AT GCG
F3-d	5’ CGC AT**F** G**F**G **F**TA CGC
F6	5’ FAM-**FFF FFF** TCT AAA TCA CTA TGG TCG C **FFF FFF**-Dabcyl 3’
F4	5’ FAM-**FFFF** TCT AAA TCA CTA TGG TCG C **FFFF**-Dabcyl 3’
FM6	5’ FAM-**FFM MFM** TCT AAA TCA CTA TGG TCG C **FMF FMM**-Dabcyl 3’
M6	5’ FAM-**MMM MMM** TCT AAA TCA CTA TGG TCG C **MMM MMM**-Dabcyl 3’
Ft6	5’ FAM-**FtFtFt FtFtFt** TCT AAA TCA CTA TGG TCG C **FtFtFt FtFtFt**-Dabcyl 3’
Ft8	5’ FAM-**FtFtFt FtFtFt FtFt** TCT AAA TCA CTA TGG TCG C **FtFtFt FtFtFt FtFt**-Dabcyl 3’

[a]Artificial bases are shown in bold font. **F** represents the unit synthesized from phosphoramidite **4**, **M** represents the unit synthesized from phosphoramidite **5**, and **Ft** represents the unit synthesized from phosphoramidite **6**.

Thermodynamic studies[1]

For the determination of melting temperatures, FNAs (1μM) and the SYBR green I dye (1X, Sigma Aldrich) were dissolved in a PIPES buffer (100 mM NaCl, 10 mM $MgCl_2$, 10 mM PIPES, pH 7.0), and the solutions were heated at 95 ^oC for 5 min. Then the samples were transferred into Real Time tubes and placed in a Real-Time PCR machine (7500, Applied Biosystems) equipped with an SYBR green filter. The samples were cooled to 15 ^oC, and then the temperature was increased from 15 to 95 ^oC at a rate of 0.5 ^oC/min. Fluorescence data from melting curves were converted into T_m by plotting the negative derivative of fluorescence *vs*. temperature (d*F*/d*T vs*.*T*). The concentrations of FNAs and SYBR green I dye were optimized with **Oligo1** and **Oligo2** prior to the study. The melting temperature (T_m) of duplex **F2-a**:**F2-b** was difficult to determine. Melting temperatures for the samples of **F3-a**:**F3-b** or **F3-c**:**F3-d**were not detected at all.

Thermodynamic study of **F6** in Tris-HCl buffer (20mM of Tris-HCl, pH7.5) was performed on the Applied Biosystems 7500 Real-Time PCR machine without the addition of SYBR green I dye. The fluorescence intensity of **F6** in buffer at temperatures ranging from 15°C to 95 °C in intervals of 0.5°C was measured and plotted against the temperature to generate the melting temperature curve. No obvious change of fluorescence signal was detected, indicating that the duplex structure of **F6** was stable at temperatures ranging from 15°C to 95 °C.

^{19}F NMR characterization of F6

^{19}F NMR experiments were run on a Varian MR400 spectrometer. A 150 nmol sample of **F6** was dissolved in 400 μL buffer (20 mM of Tris-HCl, pH7.5), and the solution was transferred to an NMR tube. The ^{19}F NMR experiment was first run at 25 ^oC until a clear ^{19}F NMR was observed (about 2 hours), which gave the ^{19}F NMR spectrum of **F6**. Then 50 nmol of cDNA in 20uL buffer (20 mM of Tris-HCl, pH7.5) was added to the tube, and the^{19}F NMR experiment was run for another 2 hours to give a ^{19}F NMR spectrum of **F6** /cDNA (3:1) mixture. The addition of 400 nmol of cDNA in 160 μL buffer to the NMR tube gave a mixture of **F6** /cDNA (1:3) in 20 mM Tris buffer, and its ^{19}F NMR spectrum was obtained after 3 hours.

Fluorescence variation experiments

Fluorescence measurements were carried out on a RF-5301-PC Fluorescence Spectrophotometer (Shimadzu, Japan). The excitation and emission wavelengths were set at 490 nm and 520 nm, respectively, with a 5 nm bandwidth. The background fluorescence of 100nM FNAs (in 200ul of 20mM of Tris-HCl, pH7.5) was monitored. After a stable fluorescent signal was obtained, an excess of

complementary DNA of the loop (10X, 1mM) was added. Fluorescence intensity was recorded as signal fluorescence.

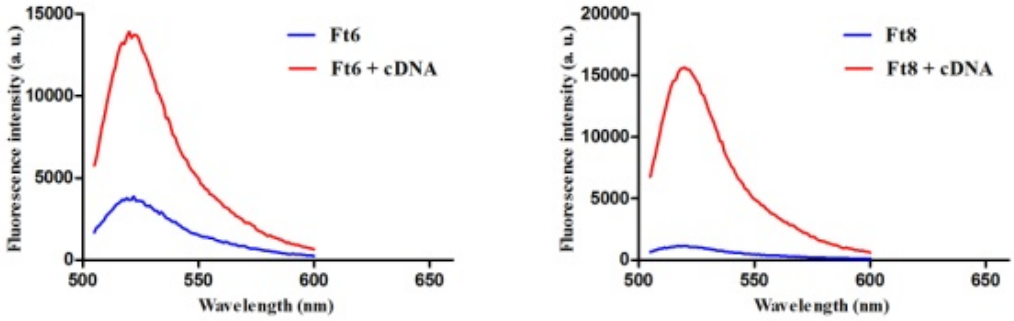


Figure S1. S/B fluorescence of **Ft6** (a) and **Ft8** (b) in 20 mM Tris buffer. Final concentration ratio of FNA:cDNA=1:10.

Agarose gel electrophoresis

FNA sample (**F4** and **F6**) was mixed with 4 μL of glycerol and analyzed by 4% agarose gel at 90V for about 45 min in 1x TBE buffer (89 mM tris(hydroxymethyl)aminomethane, 2 mM ethylenediamine tetraacetic acid and 89 mM boric acid, pH 8.0). The bands were analyzed and visualized using the ChemiDoc molecular imager (Bio-Rad) by fluorescence illumination (488 nm) and photographed by a digital camera.

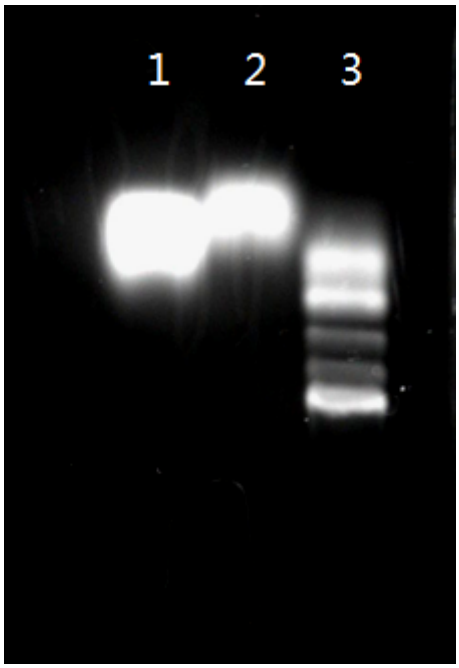


Figure S2. Agarose gel electrophoresis of FNA **F4** (lane 1), **F6** (lane 2) and DNA marker (lane 3). The gel was irradiated with 488 nm laser light in the fluorescence mode.

Computational method

The initial structures of 4-mer, 6-mer and 8-mer FNAs were constructed using GaussView 5.0[2] and VMD 1.9 (Visual Molecular Dynamics)[3] based on the stable snapshots obtained from a preliminary molecular dynamics (MD) simulation.

Prior to the final round of MD simulations, partial charges of the residue unit in FNA were determined by the restrained electrostatic potential (RESP) protocol implemented in the ANTECHAMBER module of the AMBER 11 program[4] through fitting the electrostatic potentials calculated by GAUSSIAN 09[5] at HF/6-31G* level of theory. Other parameters were taken from the ff99SB force field[6] and general amber force field (GAFF)[7]. Additional Na^+ counterions were included to balance the negative charge of the phosphate backbone, and a pre-equilibrated box of TIP3P waters extending 9Å beyond the solute in each dimension was also introduced to mimic the solvent environment in each case. After careful energy refinement conducted by the SANDER module in AMBER 11, every system was gradually heated in the NVT ensemble from 0 to 300 K by adding a 10 kcal/mol·$Å^3$ restraint on the solute. The restraint was then slowly released, and a final 200 ns MD production run with a periodic boundary condition in the NPT ensemble was carried out at 300 K and 1 atm with Particle-mesh Ewald (PME)[8] employed to control the long-range electrostatic interactions. More details will be described in our next paper focusing on the simulation work.

One thousand snapshots generated in the last stable 5 ns MD at 5 ps intervals were extracted to estimate the binding free energy in each case of 6-mer FNA and 8-mer FNA. To estimate the binding free energy of the 4-mer FNA, 1000 snapshots of the 51 to 55 ns time period at 5 ps intervals were chosen because the duplex suffered a distinct structural transformation after 55 ns during the simulation (data not shown), which may have resulted from the instability of 4-mer FNA. The binding free energy of the two strands in each FNA was calculated using the molecule mechanics/Poisson-Boltzmann surface area (MM/PBSA)[9] and molecular mechanics/generalized Born surface area (MM/GBSA)[10] methods in AMBER 11. The pair-wise GB model of Hawkins and coworkers[11,12] was implemented in the MM/GBSA calculation.

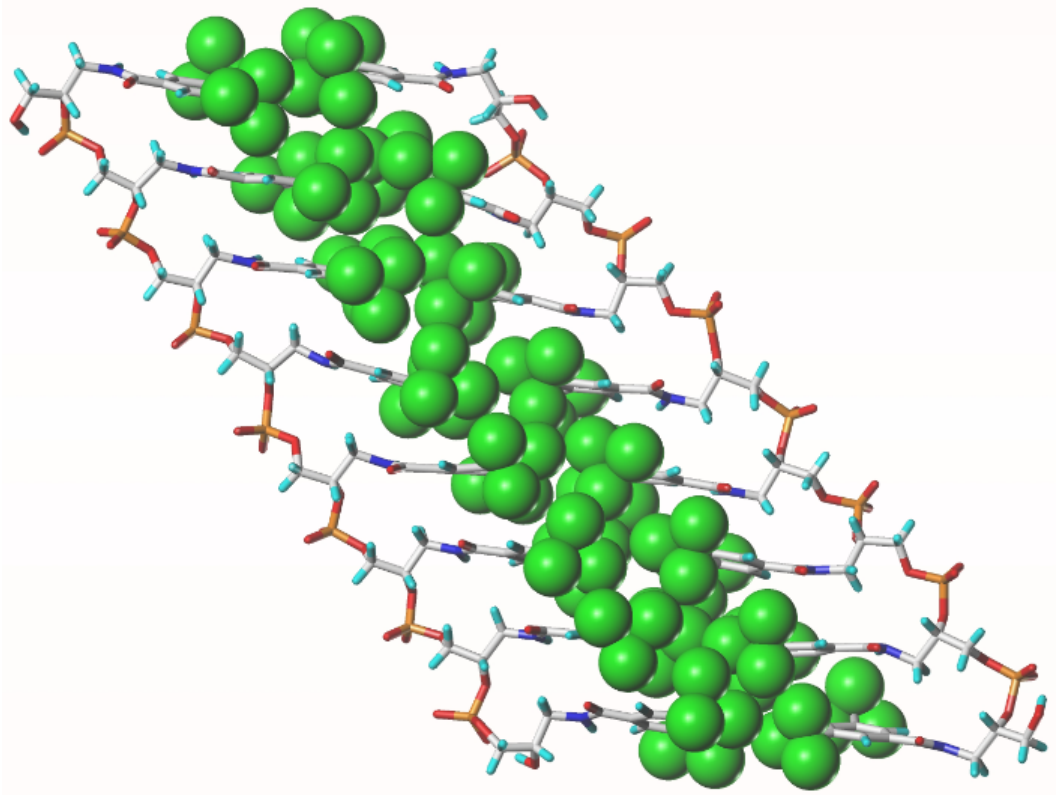


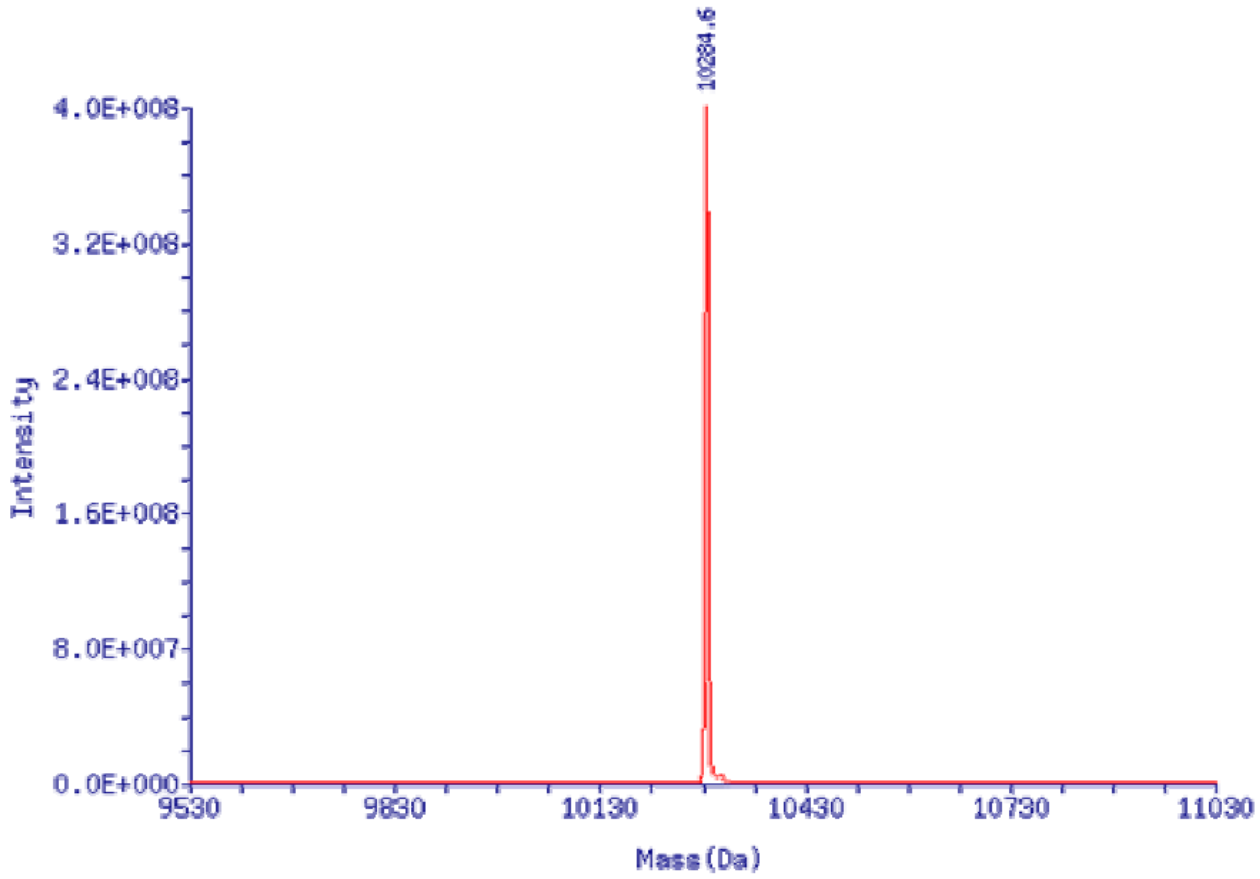
Figure S3. The Ball-and-Stick model of FNA. Green sphere represents fluorine atom.

ESI-MS spectra of FNAs

FNA **F4**

Expected mass: 10280

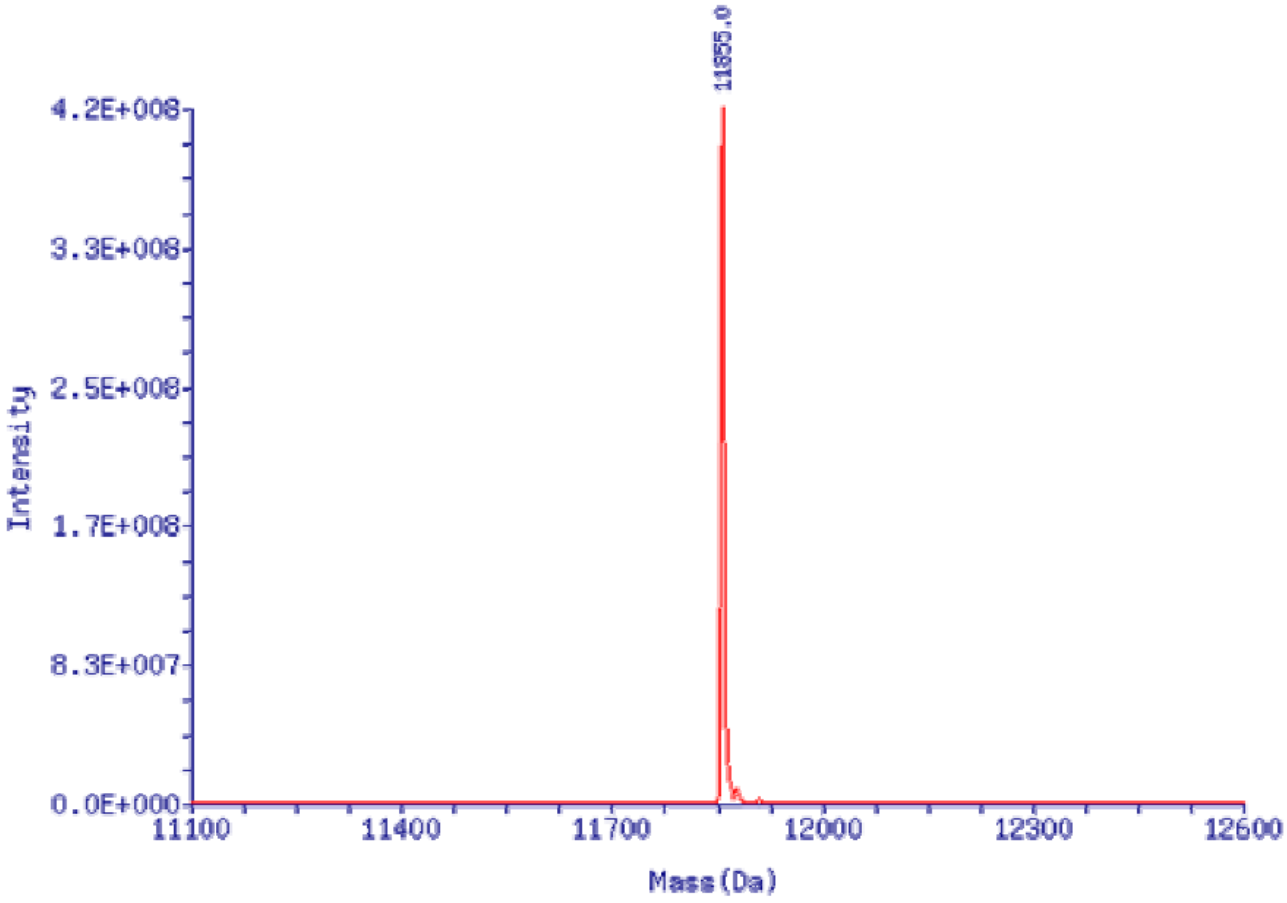
ESI spectrum:



FNA **F6**

Expected mass: 11852

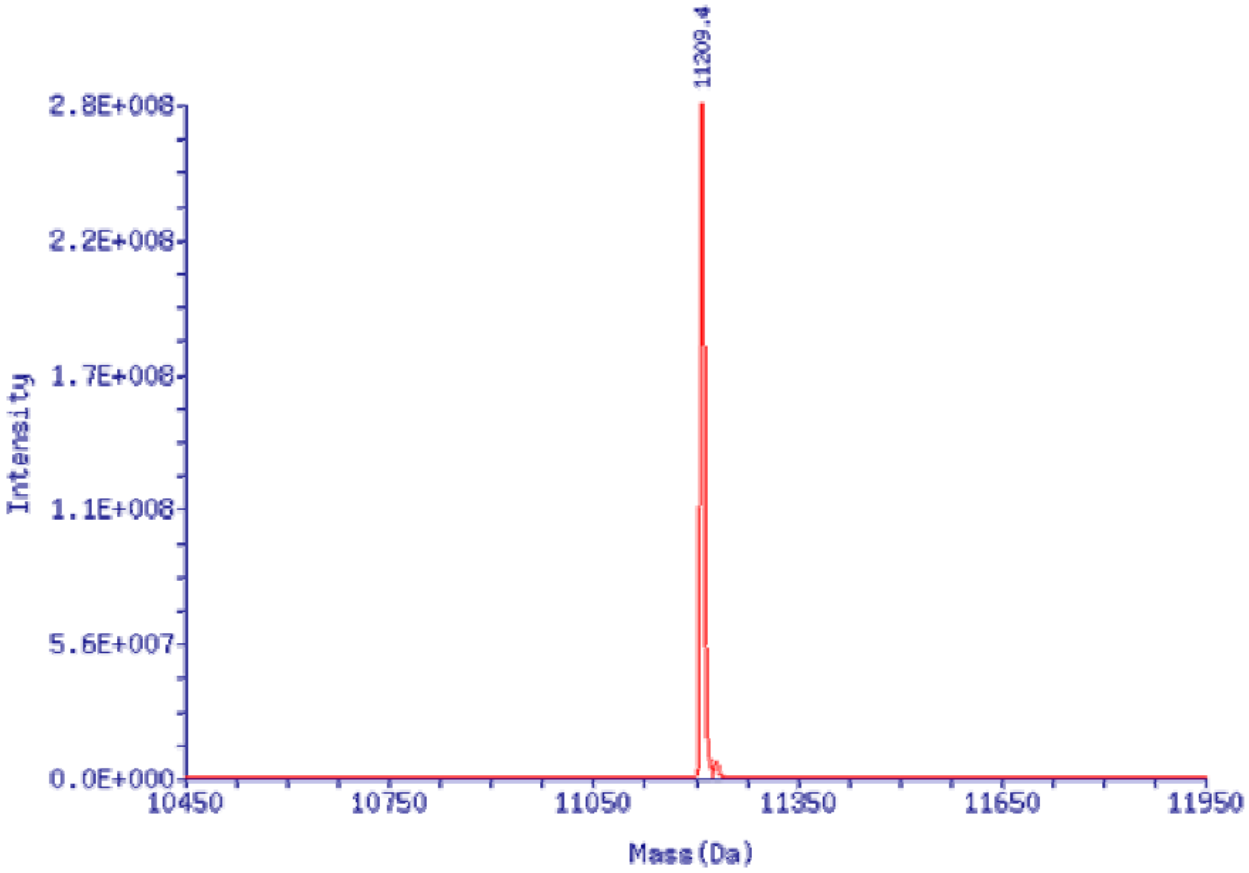
ESI spectrum:



FNA **FM6**

Expected mass: 11204

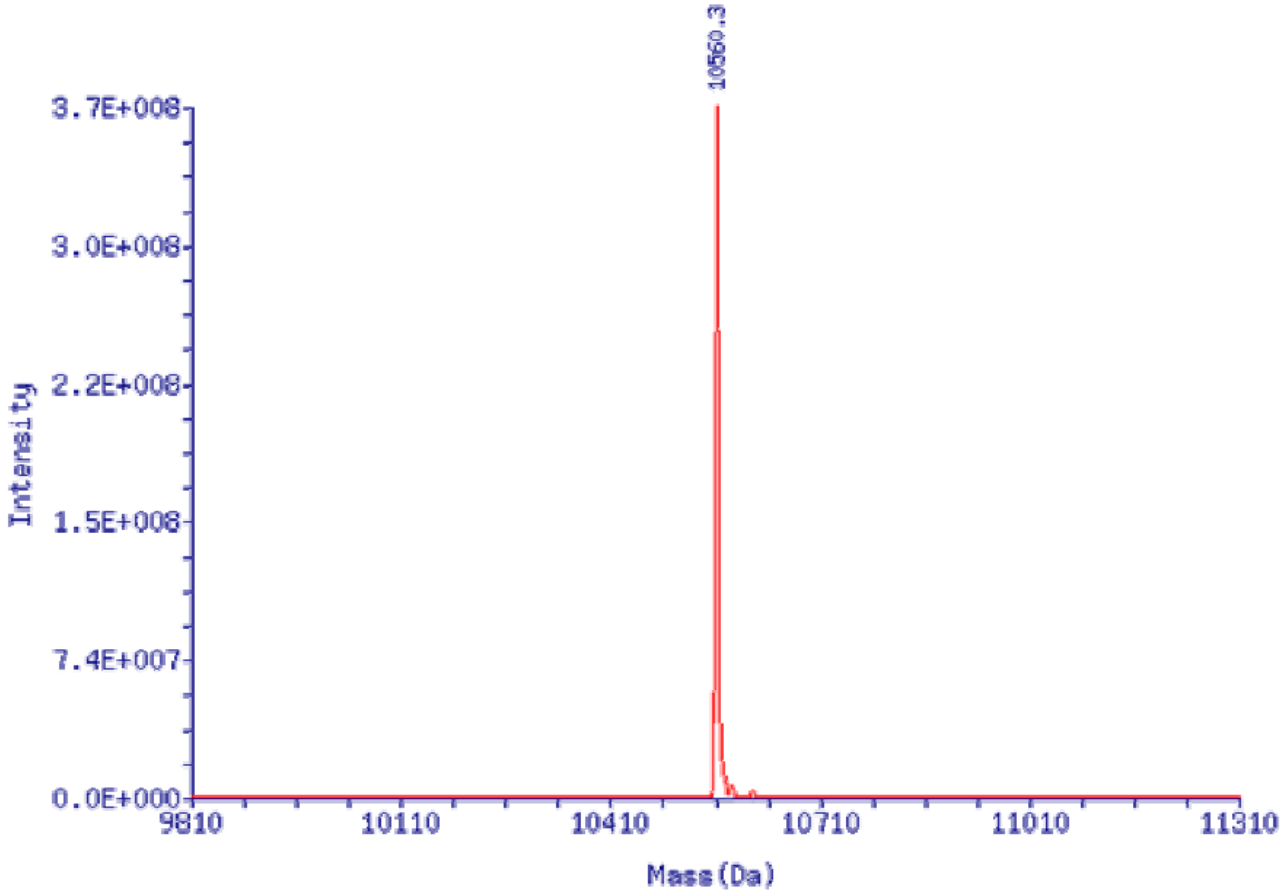
ESI spectrum:



FNA **M6**

Expected mass: 10557

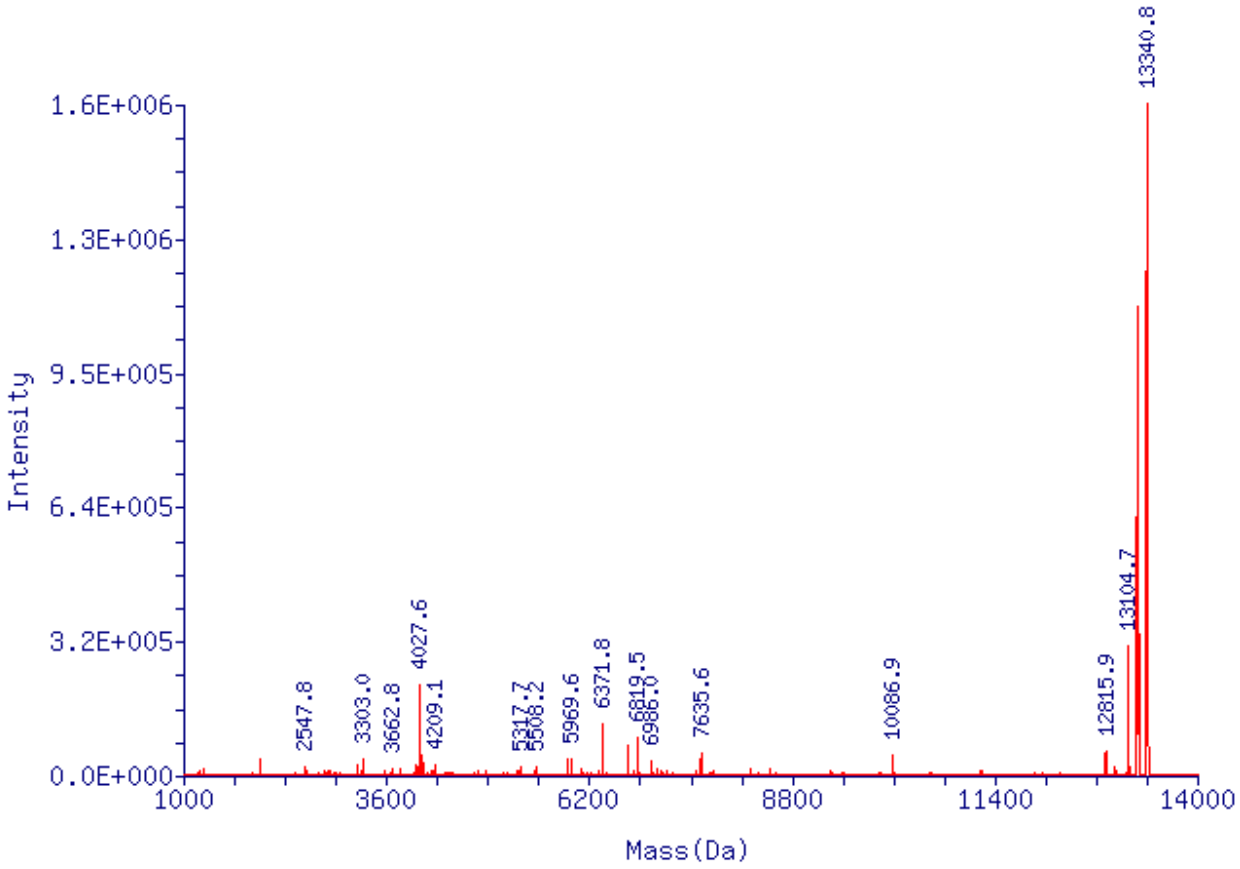
ESI spectrum:



FNA **Ft8**

Expected mass: 13313

ESI spectrum:



Supplementary References:

1. (a) Lipsky, R. H.; Mazzanti, C. M.; Rudolph, J. G.; Xu, K.; Vyas, G.; Bozak, D.; Radel, M. Q.; Goldman, D. *Clin. Chem.* **2001**, *47,* 635-644. (b) Schallon, A.; Synatschke, C. V.; Pergushov, D. V.; Jerome, V.; Muller, A. H. E.; Freitag, R. *Langmuir***2011**, *27*, 12042-12051.

2. Dennington, Roy D., Todd A. Keith, and John M. Millam. "GaussView 5.0. 8." *Gaussian Inc* (2008).

3. Humphrey, W., Dalke, A., Schulten, K.VMD: Visual Molecular Dynamics. *J. Mol. Graph.***14**, 33-38 (1996).

4. Case, D. A., T. A. Darden, T. E. Cheatham III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo et al. "AMBER 11." *University of California, San Francisco* 142 (2010).

5. Frisch, M. J., G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani et al. "J. Gaussian 09, Revision C. 01." *Gaussian Inc* (2010).

6. Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A.& Simmerling, C. Comparison of multiple Amber force fields and development of improved. *Proteins*, **65**, 712-725 (2006).

7. Wang, J., Wolf, R.M., Caldwell, J.W., Kollamn, P.A.& Case, D.A. Development and testing of a general Amber force field. *J. Comput. Chem.*, **25**, 1157-1174 (2004).

8. Essmann, U., Perera, L., Berkowitz, M. L.& Darden, T. A smooth particle mesh Ewald method. *J. Chem. Phys.***103**, 8577-9593 (1995).

9. Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., Lee, M., Lee, T., Duan, Y., Wang, W., Donini, O., Cieplak, P., Srinivasan, J., Case, D. A.& Cheatham, T. E. Calculating structures and free energies of complex molecules: combining molecular mechanicsand continuum models. *Acc. Chem. Res.***33**, 889-897 (2000).

10. Gohlke, H., Kiel, C., Case, D. A. Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes. *J. Mol. Biol.***330**, 891-913 (2003).

11. Hawkins, G. D., Cramer, C. J.& Truhlar, D. G. Pairwise SoluteDescreening of Solute Charges from a Dielectric Medium.*Chem. Phys.Lett.***246**, 122-129 (1995).

12. Hawkins, G. D., Cramer, C. J.& Truhlar, D. G. Parametrized modelsof aqueous free energies of solvation based on pairwise descreeningof solute atomic charges from a dielectric medium. *J. Phys. Chem.***100**, 19824–19839 (1996).