

Supporting Information

Probing Local Structure and Dynamics of Nucleotides and DNA Using Vibrationally Enhanced Alkynyl Stretch

Tiantian Dong^{†‡}, Pengyun Yu^{†‡}, Juan Zhao^{†‡}, and Jianping Wang^{†‡}*

*[†] Beijing National Laboratory for Molecular Sciences, Molecular Reaction Dynamics
Laboratory, CAS Research/Education Center for Excellence in Molecular Sciences,*

Institute of Chemistry, Chinese Academy of Sciences, Beijing, 100190, P. R. China

[‡] University of Chinese Academy of Sciences, Beijing 100049, P. R. China

** Author to whom correspondences should be addressed.*

Tel: (+86)-010-62656806; Fax: (+86)-010-62563167; E-mail: jwang@iccas.ac.cn.

Table of Contents:

Section 1. Synthesis

Section 2. NMR spectroscopy characterization

Section 3. Fitting analysis of FTIR spectra and DFT calculations

Section 4. Monomer and homodimer ratio analysis

Section 5. C≡C vibrational frequency shift analysis

Section 6. Anharmonicity measurement

Section 7. Nonlinear IR spectroscopy characterization of molecule **1c**

Section 8. PED analysis of molecule **1a**

Section 1. Synthesis

5-trimethylsilylethynyl-2'-deoxycytidine, abbreviated as ^{TMS}EdC (1a). A solution of 5-I-dC (molecule **1**, 1.00 g, 2.8 mmol), [Pd(pph₃)₄] (0.487 g, 0.42 mmol), CuI (0.109 g, 0.56 mmol) in anhydrous DMF (35 ml) was bubbled with a stream of nitrogen for 30 min. After being stirred for 5 min, Et₃N (1.05 ml, 7.56 mmol) and trimethylsilylacetylene (1 ml, 7 mmol) was added to the solution. The resulting mixture was stirred at 25 °C for 24 h under nitrogen. The solvent was removed under reduced pressure, and the residue was purified by column chromatography (100 % CH₂Cl₂ to 10 % MeOH in CH₂Cl₂) to give the slightly yellow foamy solid product (0.63 g, 70 %). ¹H-NMR (DMSO-d₆, δ=2.50, 400 MHz): δ= 0.21 (m, 9H, 3×CH₃), 1.97-2.17 (m, 2H, 2'), 3.54-3.64 (m, 2H, H-5'), 3.77-3.80 (m, 1H, H-3'), 4.18-4.22 (m, 1H, H-4'), 5.05 (1H, H-1'), 5.19 (3'-OH), 6.08-6.11 (m, 1H, 5'-OH), 6.61 (brs, 1H), 7.78 (brs, 1H, NH), 8.20 (s, 1H, H-6). MS (ESI-MALDI): calculated for C₁₄H₂₁N₃O₄Si [M + Na⁺] 346.13, found 346.11946.

5-triisopropylsilylethynyl-2'-deoxycytidine, abbreviated as ^{TIPS}EdC (1b). A solution of 5-I-dC (molecule **1**, 0.50 g, 1.42 mmol), [Pd(pph₃)₄] (0.246 g, 0.21 mmol), CuI (55 mg, 0.28 mmol) in anhydrous DMF (20 ml) was bubbled with a stream of nitrogen for 30 min. After being stirred for 5 min, Et₃N (0.531 ml, 3.82 mmol) and Triisopropylsilylacetylene (0.794 ml, 5.56 mmol) was added to the solution. The resulting mixture was stirred at 25 °C for 24 h under nitrogen. The solvent was removed under reduced pressure, and the residue was purified by column chromatography (100 % CH₂Cl₂ to 5 % MeOH in CH₂Cl₂) to give the slight yellow amorphous solid product (0.35 g, 61 %). ¹H NMR (DMSO-d₆, δ=2.50, 400 MHz): δ= 1.08 (m, 21H, 6×CH₃, 3×CH), 2.00-2.21 (m, 2H, 2'), 3.52-3.64 (m, 2H, H-5'), 3.78-3.80 (m, 1H, H-3'), 4.19-4.23 (m, 1H, H-4'), 5.00-5.03 (t, *J* = 5.2 Hz, 1H, H-1'), 5.19-5.20 (d, *J* = 4.4 Hz, 3'-OH), 6.08 (m, 1H, 5'-OH), 7.88 (brs, 1H, NH), 8.28 (s, 1H, H-6). MS (ESI-MALDI): calculated for C₂₀H₃₃N₃O₄Si [M + Na⁺] 430.22, found 430.21294.

5-(4-benzaldehydeethynyl)-2'-deoxycytidine, abbreviated as ^{BA}EdC (1c). A solution of 5-I-dC (molecule **1**, 0.50 g, 1.42 mmol), [Pd(pph₃)₄] (0.246 g, 0.21 mmol), CuI (0.055 g, 0.28

mmol) in anhydrous DMF (20 ml) was bubbled with a stream of nitrogen for 30 min. After being stirred for 5 min, Et₃N (0.531 ml, 3.82 mmol) and 4-Ethynylbenzaldehyde (461 mg, 3.54 mmol) was added to the solution. The resulting mixture was stirred at 25 °C for 48 h under nitrogen. Then the mixture was filtration by water several times and yellow solid precipitated to give the product (0.45 g, 89 %). ¹H NMR (DMSO-d₆, δ=2.50, 400 MHz): δ= 2.00-2.24 (m, 2H, 2'), 3.56-3.71 (m, 2H, H-5'), 3.80-3.83 (m, 1H, H-3'), 4.21-4.26 (m, 1H, H-4'), 5.15-5.18 (t, *J* = 6.4 Hz, 1H, H-1'), 5.24-5.25 (d, *J* = 5.6 Hz, 3'-OH), 6.10-6.14 (m, 1H, 5'-OH), 7.19 (brs, 1H, NH), 7.80-7.94 (m, 5H), 8.43 (s, 1H, H-6), 10.01 (s, 1H, COH). MS (ESI): calculated for C₁₈H₁₇N₃O₅ [M + Na⁺] 378.12, found 378.1.

5-triisopropylsilylethynyl-2'-deoxythymidine, abbreviated as TIPS^sEdT (2a). A solution of 5-I-dT (molecule **1**, 0.50 g, 1.42 mmol), [Pd(pph₃)₄] (16.21 mg, 0.14 mmol), CuI (53.40 mg, 0.27 mmol) in anhydrous DMF (20 ml) was bubbled with a stream of nitrogen for 30 min. After being stirred for 5 min, Et₃N (0.489 ml, 3.50 mmol) and Triisopropylsilylacetylene (1.107 ml, 4.92 mmol) was added to the solution. The resulting mixture was stirred at 25 °C for 48h under nitrogen. The solvent was removed under reduced presser, and the residue was purified by column chromatography (100 % CH₂Cl₂ to 7 % MeOH in CH₂Cl₂) to give the slight yellow foamy solid product (0.38 g, 65 %). ¹H NMR (DMSO-d₆, δ=2.50, 400 MHz): δ= 1.07 (m, 21H, 6×CH₃, 3×CH), 2.09-2.20 (m, 2H, 2'), 3.54-3.64 (m, 2H, H-5'), 3.78-3.80 (m, 1H, H-3'), 4.21-4.25 (m, 1H, H-4'), 5.02-5.05 (t, *J* = 5.2 Hz, 1H, H-1'), 5.22-5.23 (d, *J* = 4.4 Hz, 3'-OH), 6.08-6.11 (m, 1H, 5'-OH), 8.26 (s, 1H, H-6), 11.57 (s, 1H, NH). MS (ESI): calculated for C₂₀H₃₂N₂O₅Si [M + Na⁺] 431.21, found 431.2.

Section 2. NMR spectroscopy characterization

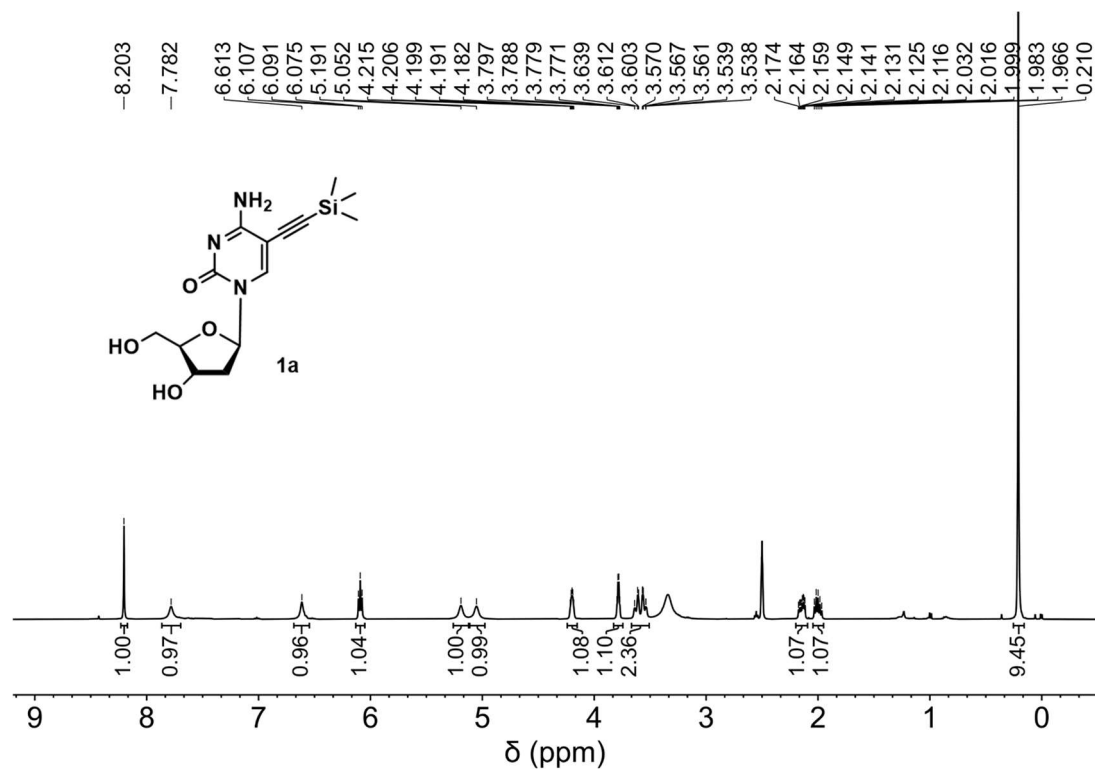


Figure S1. ¹H NMR spectrum of compound 1a.

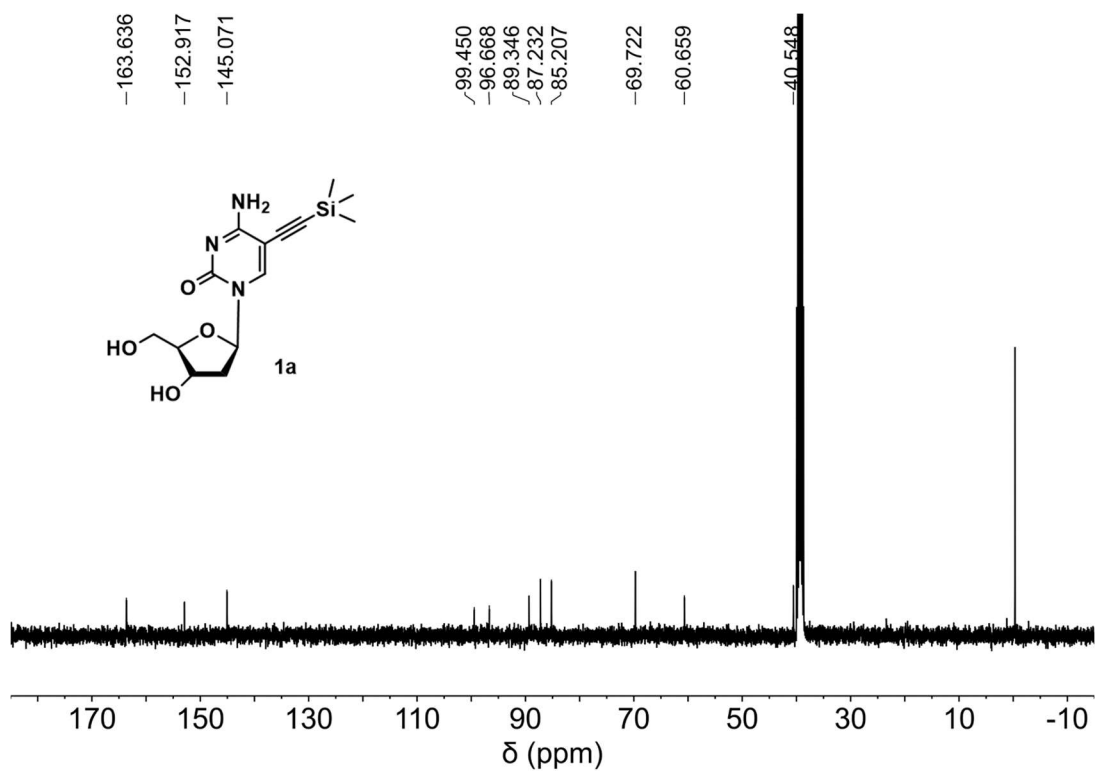
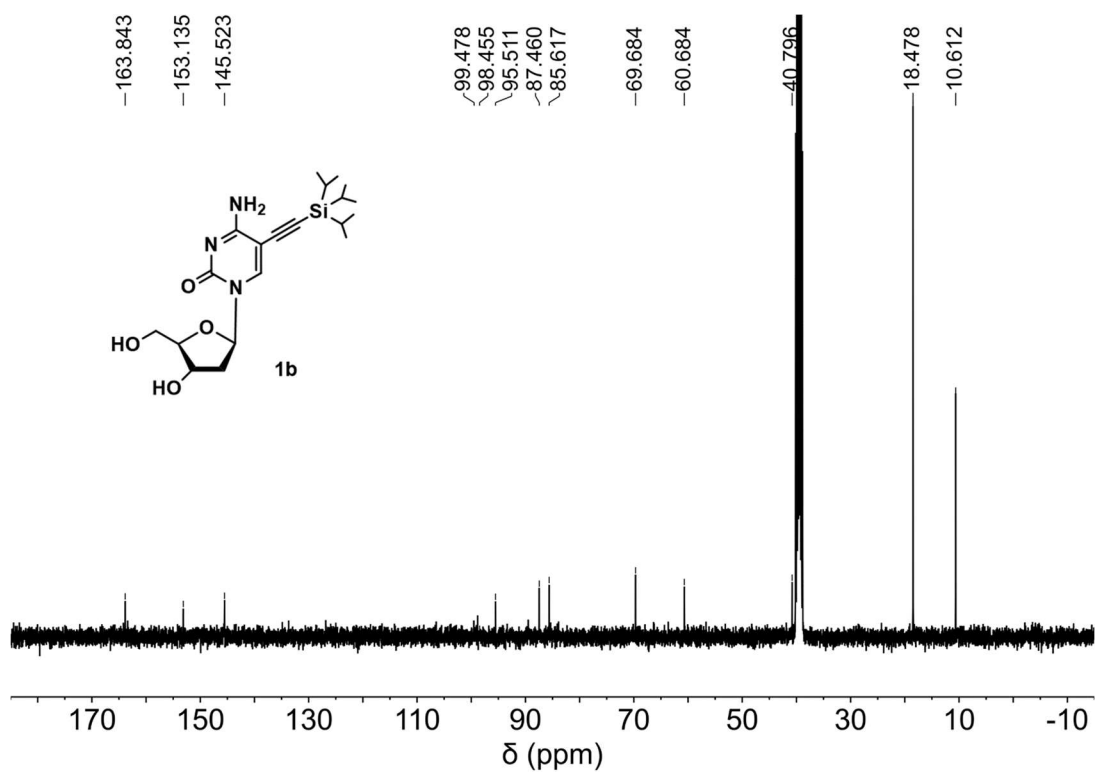
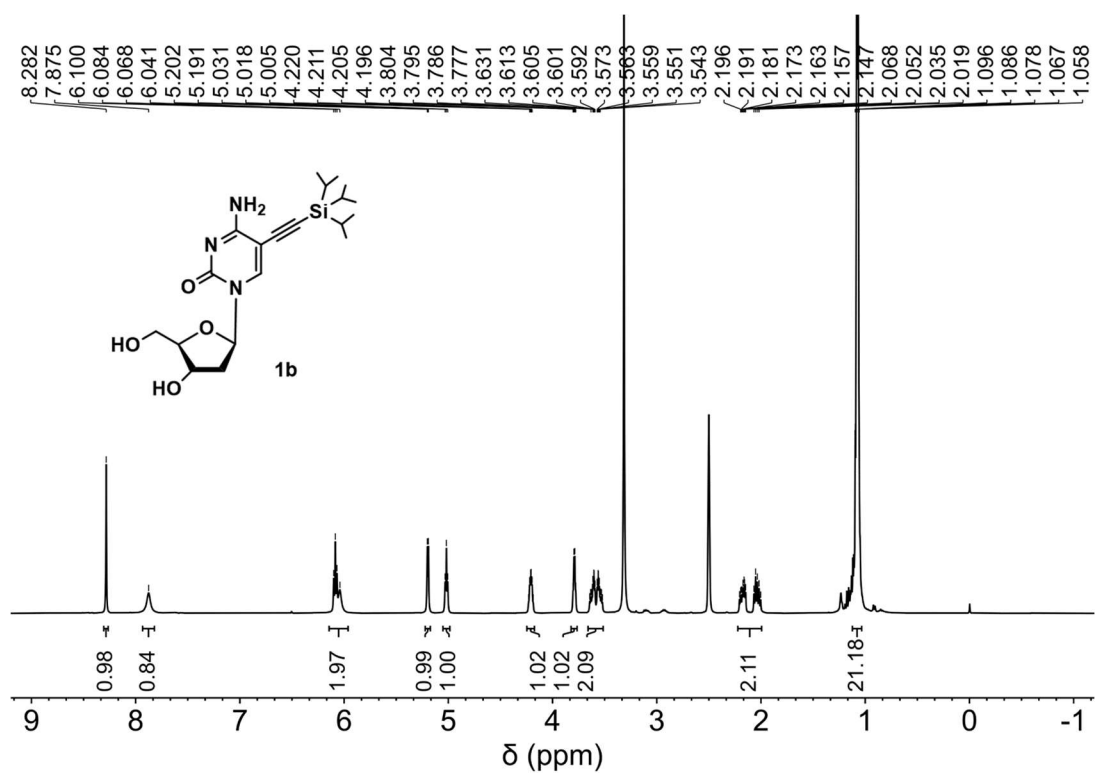
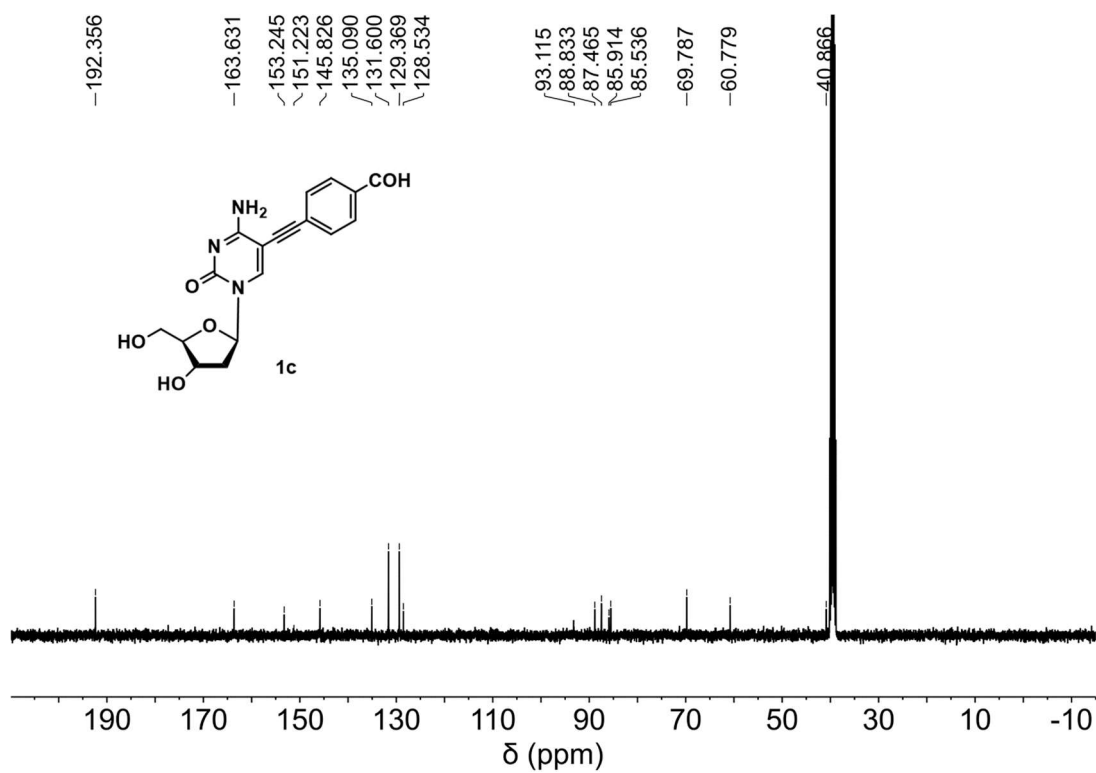
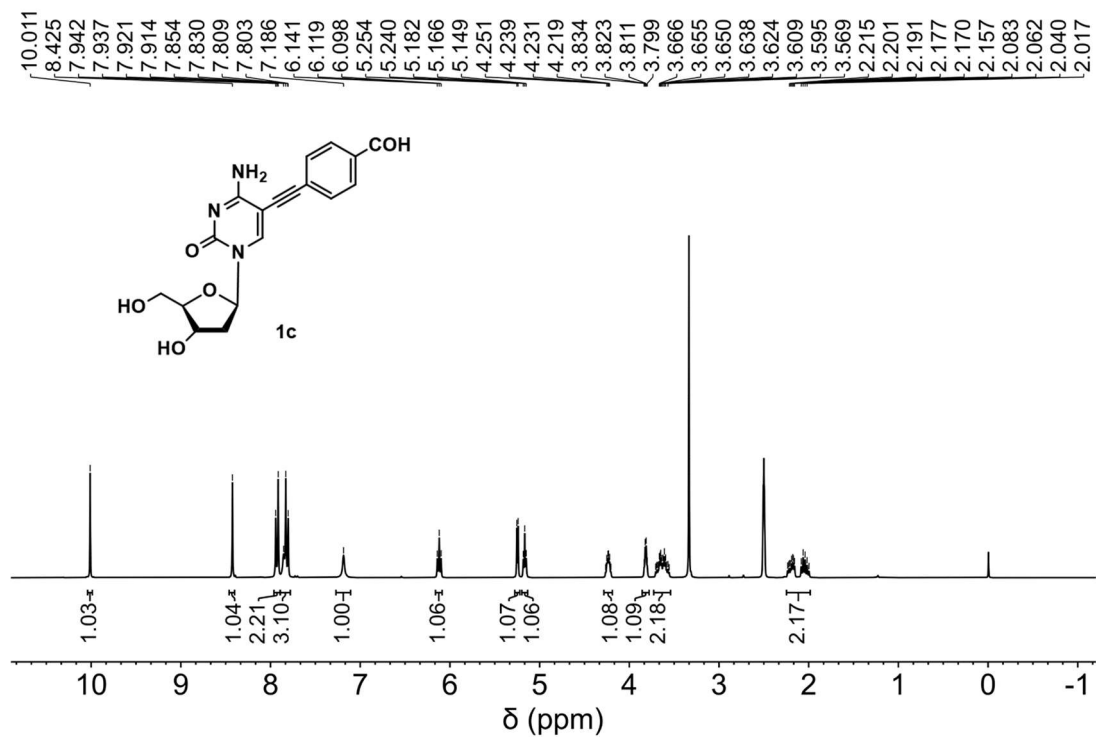
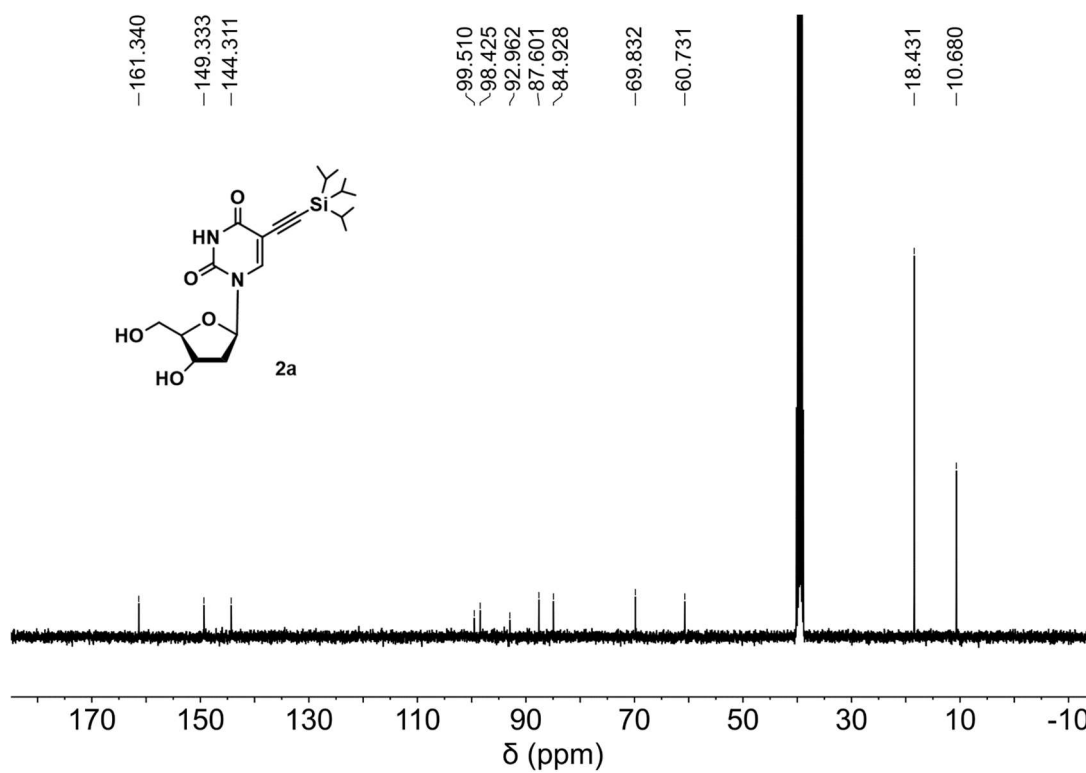
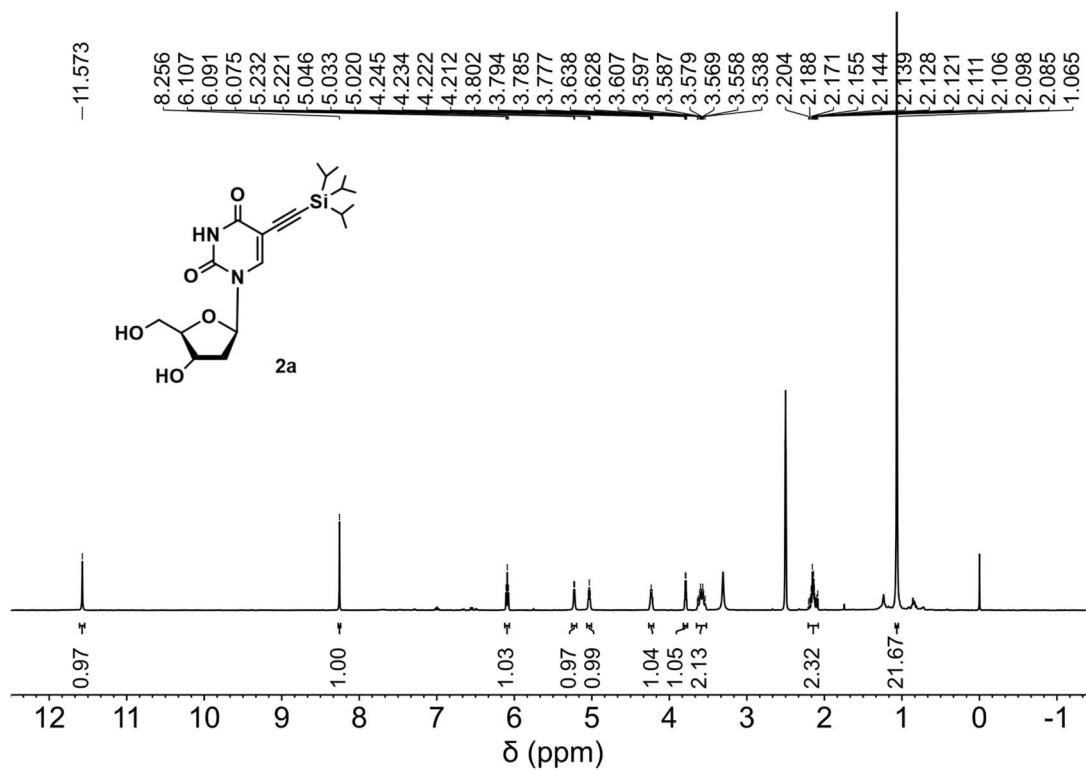


Figure S2. ¹³C NMR spectrum of compound 1a.







Section 3. Fitting analysis of FTIR spectra and DFT calculations

Table S1. Fitting results of the vibration intensity enhanced C≡C stretching IR spectra. All spectra are fitted with Voigt line shape function. The parameters of full width at half maximum (FWHM), Gaussian width and Lorentzian width, and calculated transition dipole moment (μ) based on the spectra are listed. Molecules are listed in the same order as in Figure 1.

Compound	$\omega_{\text{C}\equiv\text{C}}$ / cm^{-1}	FWHM / cm^{-1}	Gaussian width / cm^{-1}	Lorentzian width / cm^{-1}	μ / D
EdC (3)	2102.4	16.9	14.8	3.8	0.079
TMSE (4)	2025.4	8.5	6.3	3.7	0.109
TIPSE (5)	2022.1	7.5	6.0	2.6	0.148
^{TMS} EdC (1a)	2150.6	17.6	10.1	11.4	0.332
^{TIPS} EdC (1b)	2146.7	21.2	13.7	12.0	0.342
^{BA} EdC (1c)	2207.2	25.1	18.8	10.6	0.328
^{TIPS} EdT (2a)	2156.1	19.7	16.3	5.8	0.299

Table S2. Calculated C≡C vibrational transition frequency (ω , in cm^{-1}), transition intensity (I , in km/mol), transition dipole magnitude (μ , in Debye, D) and bond length in Å of all vibration-enhanced C≡C containing molecules at the level of B3LYP/6-31G*. The calculated frequencies were scaled by a factor of 0.960.¹ The fuzzy bond order and average value of three orbital (NBO) delocalization index (ODI) of C≡C were calculated by Multiwfn program.² The lower the value of ODI, the stronger the orbital delocalization.

No.	ω	I	μ	Bond length	Fuzzy bond order	ODI mean
6 ^[a]	2004.6	0	0	1.20497	2.909	46.83
7 ^[b]	2136.7	1.68	0.017	1.21004	2.708	—
3	2123.0	41.84	0.086	1.21240	2.685	23.26
4	2067.3	29.79	0.074	1.21541	2.719	32.22
5	2060.7	24.66	0.068	1.22444	2.712	—
1a	2152.5	292.49	0.228	1.22404	2.503	12.24
1b	2148.0	296.19	0.230	1.20120	2.496	—
1c	2196.1	465.88	0.285	1.22010	2.476	16.19
2a	2172.1	117.54	0.144	1.22116	2.505	—

[a] Compound **6** is acetylene. [b] Compound **7** is phenylacetylene.

Table S3. Calculated C≡C vibrational transition frequency (ω , in cm^{-1}), transition intensity (I , in kM/mol), transition dipole magnitude (μ , in Debye, D) at the level of B3LYP/6-31G* for all the vibration-enhanced C≡C tagged nucleosides studied in this work. The calculated frequencies were scaled by a factor of 0.960.¹

	ω	I	μ
EdA	2172.2	45.39	0.090
EdT	2142.4	6.08	0.033
EdC (3) ^[a]	2123.2	41.82	0.087
EdG	2142.7	18.53	0.057
7- ^{TMS} EdA	2155.4	317.83	0.238
5- ^{TMS} EdT	2175.4	130.05	0.151
5- ^{TMS} EdC (1a) ^[a]	2152.5	292.49	0.228
7- ^{TMS} EdG	2175.3	226.35	0.200
7- ^{TIPS} EdA	2151.0	303.66	0.233
5- ^{TIPS} EdT	2172.1	117.54	0.144
5- ^{TIPS} EdC (1b) ^[a]	2148.0	296.18	0.230
7- ^{TIPS} EdG	2169.4	215.93	0.195
7- ^{BA} EdA	2199.6	507.72	0.297
5- ^{BA} EdT	2220.2	148.19	0.160
5- ^{BA} EdC (1c) ^[a]	2196.1	465.88	0.285
7- ^{BA} EdG	2220.2	340.90	0.242

[a] Compounds have been listed in Table S2.

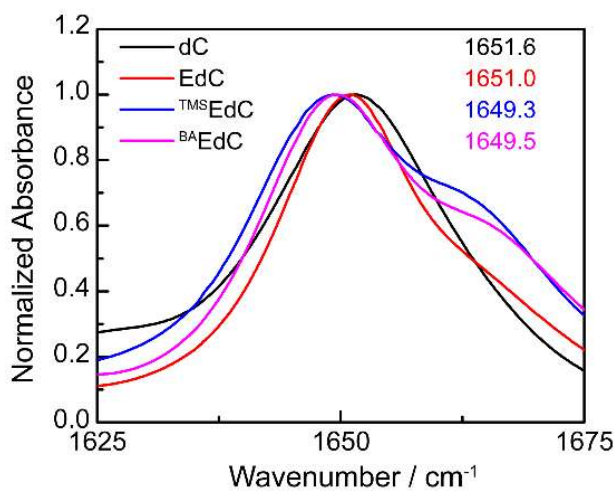


Figure S9. FTIR spectra of molecule **3** (EdC), **1a** (^{TMS}EdC) and **1c** (^{BA}EdC) and deoxycytidine (dC) in the C=O stretching vibrational region in dimethyl sulfoxide (DMSO).

Section 4. Monomer and homodimer ratio analysis

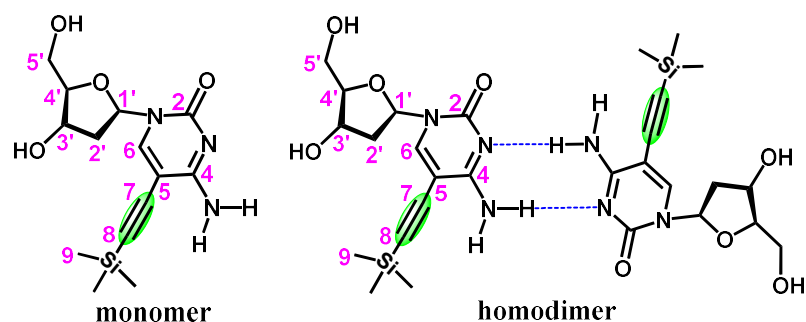


Figure S10. Monomer and homodimer structures of molecule **1a**.

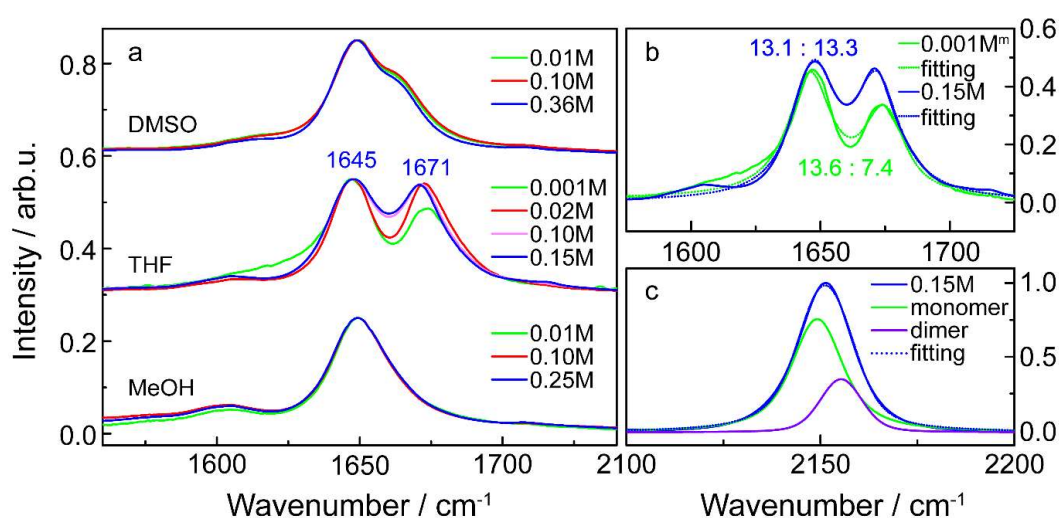


Figure S11. Normalized FTIR spectrum of molecule **1a** in three solvents at different concentration at the C=O, C=N/C=C stretching and NH₂ bending region (a), peak fitting curves (colored dotted line) with two Voigt functions in THF at selected concentration (b), and fitting curves (colored dotted line) with two components (monomer and dimer) at the C≡C stretching region in THF at 0.15 M (c). The integrated areas of two peaks are marked with values separately. The spectrum for 0.001-M case is simply scaled by optical path factor and concentration factor (labeled as 0.001 M^m) to compare with that of 0.15 M in panel (b). Data analysis is described in the text below.

To investigate whether homodimer of molecular **1a** exists in three solvents at 2D-IR experimental concentration, we carry out FTIR experiments of molecule **1a** at different concentrations. The fingerprints of the C=O, C=N/C=C stretching modes and NH₂ bending modes at higher concentration almost coincide with those at lower concentration for molecule **1a** in DMSO and MeOH but not in THF. In THF, the spectra show an intensity increase for the 1671- cm^{-1} peak as the concentration increases, which is related to the hydrogen-bonded vibration modes of dimer for molecule **1a**. The fingerprints for the 0.001-M spectrum can be regarded as being composed of monomer while partial dimer at 0.15 M. An approximate

dimer contribution can be obtained by subtracting spectra of monomer from that of higher concentration when absolute absorption coefficient is ignored. The spectrum at 0.001 M was scaled by optical path factor with value 0.15 (200 μm for 0.001 M, 30 μm for 0.15 M) and concentration factor with value of 150 (0.001 M vs. 0.15 M) to scale the spectrum at 0.001- M^m as the monomer contribution to the spectrum at 0.15 M. Therefore, the dimer ratio is approximately to the area difference between spectra at 0.15 M and 0.001 M^m as a percentage of total area. The calculated dimer population is no more than 20 %, since the absorption coefficient of the dimer mode may be larger than that of the monomer.

In addition, to further illustrate the effect of dimerization on the $\text{C}\equiv\text{C}$ stretching mode, we fit the IR spectrum of molecular **1a** in THF at 0.15M using two Voigt functions representing monomer and dimer respectively. The result is shown in panel c. The peak position and width of the fitted monomer spectrum (2149.1 cm^{-1} and 15.8 cm^{-1}) are similar to that of 0.001 M. The peak position of dimer is set to 2155.3 cm^{-1} , which is ca. 6 cm^{-1} higher than that of the monomer. DFT computation of monomer yields one $\text{C}\equiv\text{C}$ stretching mode with frequency of 2154.4 cm^{-1} , and that of dimer (Figure S10) yields one asymmetric $\text{C}\equiv\text{C}$ stretching mode with frequency of 2158.5 cm^{-1} (IR active) and one symmetric $\text{C}\equiv\text{C}$ stretching mode with frequency of 2158.8 cm^{-1} (IR inactive), so that the frequency difference between monomer and dimer agree roughly with the experimental fitting results shown in panel c. However, since in the dimer only one $\text{C}\equiv\text{C}$ stretching mode is IR active and the frequency separation between the asymmetric and symmetric modes are rather limited, it is rather difficult to resolve any cross peak between them.

Considering above results, the contribution of the dimer in 2D IR line shape of molecule **1a** in THF is not considered in 2D IR analysis.

Section 5. $\text{C}\equiv\text{C}$ vibrational frequency shift analysis

Table S4. Empirical solvent polarity, hydrogen bonding ability parameters of dimethyl sulfoxide, tetrahydrofuran, *N,N*-dimethylformamide, ethyl acetate, dichloromethane, acetone, deuterium oxide and methanol.

solvent	$E_T^N(30)$ [a]	ϵ_r [a]	α [b]	β [b]
DMSO	0.444	47.24	0	0.76
THF	0.207	7.52	0	0.55
DMF	0.386	38.25	0	0.69
EA	0.228	6.08	0	0.45
DCM	0.309	8.93	0.13	0.10
AC	0.355	21.01	0.08	0.43
D ₂ O	0.991	80.10 [c]	1.17 [c]	0.47 [c]
MeOH	0.762	33.0	0.98	0.66

[a] Empirical solvent polarity $E_T^N(30)$ and dielectric constant ϵ_r .^{3,4} [b] α hydrogen bond donating ability, β hydrogen bond accepting ability.⁵ [c] ϵ_r , α and β parameters for pure water.

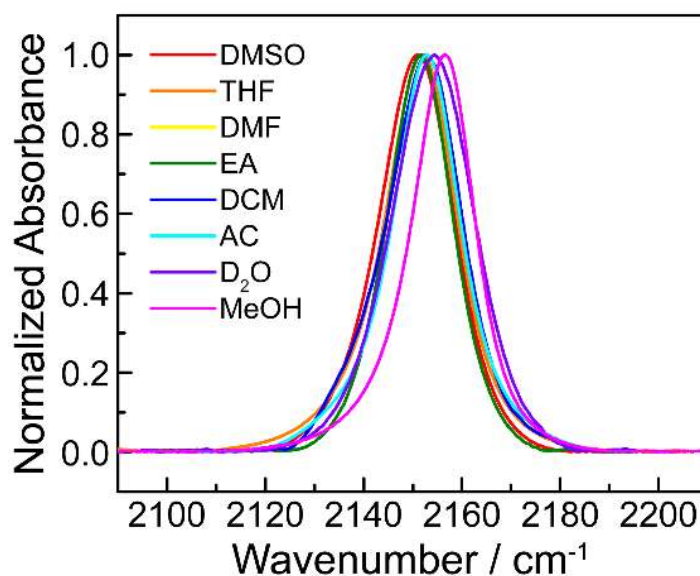


Figure S12. FTIR spectra in the C≡C stretching vibrational region of ^{TMS}EdC (**1a**) in DMSO, THF, DMF, EA, DCM, AC, D₂O and MeOH.

Molecule **1a** has a solubility in water less than 0.05 mol/L. Nonlinear IR experiments were thus carried out in three typical solvents (DMSO, THF, and MeOH) but not in water.

Table S5. Calculated bond length (in Å) of molecule **1a** of solute-solvent clusters shown in Figure 3 at level B3LYP/6-311+G**. Fuzzy bond order was calculated by Multiwfn program.²

Solvent	C≡C bond length	C≡C bond order	C5=C6 bond length	C5=C6 bond order	C2=O2 bond length	C2=O2 bond order	C4=N3 bond length	C4=N3 bond order
DMSO	1.21835	2.505	1.37108	1.462	1.23143	1.835	1.32741	1.490
THF	1.21873	2.506	1.37138	1.463	1.22319	1.879	1.32219	1.518
MeOH	1.21672	2.514	1.36974	1.477	1.24787	1.749	1.33130	1.463

Section 6. Anharmonicity measurement

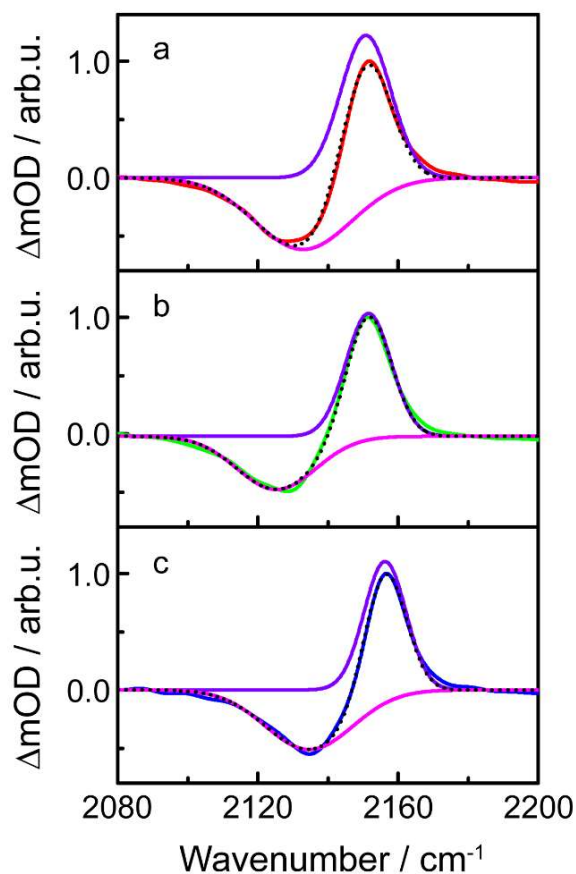


Figure S13. 2D IR spectral slices along the ω_τ axis taken at a range of ω_τ for molecule **1a** in DMSO (a), THF (b) and MeOH (c). $T_W = 0.2$ ps. $\omega_\tau = [2150 - 2155 \text{ cm}^{-1}]$ (in DMSO), $[2149 - 2154 \text{ cm}^{-1}]$ (in THF) and $[2155 - 2159 \text{ cm}^{-1}]$ (in MeOH). Spectral fittings are shown to determine the diagonal anharmonicity (Δ) in each case (Table S6).

Table S6. Fitting results of 2D IR slice of molecule **1a** in three solvents showed in Figure S13 by Gaussian function with parameters of frequency position (ω in cm^{-1}), peak width in full width at half maximum (FWHM, in cm^{-1}) and diagonal anharmonicity (Δ in cm^{-1}). Each of diagonal anharmonicity Δ constants with error bar is obtained from fitting 2D IR slice of 0.2 ps, 0.6 ps and 1.0 ps.

Solvent	Bleach ($\nu = 0 \rightarrow 1$)		Absorption ($\nu = 1 \rightarrow 2$)		anharmonicity Δ
	ω_{01}	FWHM	ω_{12}	FWHM	
DMSO	2150.8	13.5	2132.8	27.0	18.1 ± 0.8
THF	2151.6	12.7	2126.4	27.9	25.3 ± 1.0
MeOH	2156.2	12.0	2134.7	25.0	21.6 ± 0.3

Section 7. Nonlinear IR spectroscopy characterization of molecule 1c

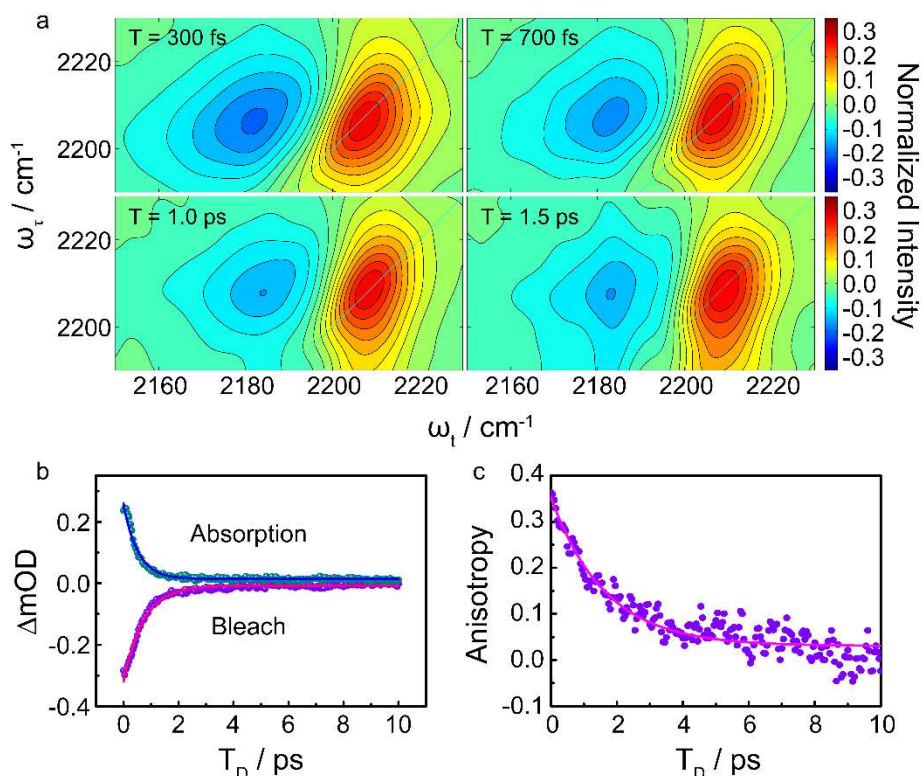


Figure S14. (a) Waiting time dependent, purely absorptive 2D IR spectra of molecule ^{BA}EdC (**1c**) in the C \equiv C stretching region in DMSO. A thin line across the red peak indicating diagonal trace in each panel. All spectra are normalized by the intensity of the positive (red) peak. (b) Population relaxation dynamics of the C \equiv C stretching mode of the molecule **1c** as a function of the delay time (T_D). Experimental data (solid circles) and single exponential fitting (solid lines) are given. Both the absorption relaxation and bleach recovery are shown. (c) Anisotropy dynamics probed at the absorption signal of the C \equiv C stretching mode of the molecule **1c** in DMSO. Experimental data (solid dots) and single exponential fitting (solid lines) are shown.

Population relaxation dynamics were measured as bleach recovery (negative) and excited-state decay (positive) signals, with probing frequency at 2209 cm^{-1} and 2181 cm^{-1} respectively. Both negative and positive signals curves can be fitted by single exponential function with a constant time of 0.72 ps and 0.52 ps separately, with the former representing the vibrational excited-state lifetime. The reorientation relaxation dynamics of the C \equiv C stretching mode in DMSO was found to be 1.70 ps, which is longer than that of molecule **1a** in DMSO (1.35 ps).

Section 8. PED analysis of molecule 1a

Table S7. Computed vibrational transition frequency (ω , in cm^{-1}), and potential energy distributions (PEDs) of selected vibrational modes in molecule **1a**, at the level B3LYP/6-311+G**. The calculated frequencies were scaled by a factor of 0.967.¹ Brackets indicate the participation ratio of component contribution to each vibration mode. Participation ratios of vibrational component that account for more than five percent are mainly listed in table.

Mode	PED (s = stretching vibration; b = bending vibration; R = ring)
2154.4	C29C30 _s (35 %), C14C29 _s (22 %), C30Si31 _s (17 %)
1683.2	C20O2 _s (16 %), N19C20 _s (9 %), C12N11C20 _b (5 %), N11C20N19 _b (7 %), N11C20O21 _b (5 %)
1621.8	H17N16H18 _b (7 %), C ^R 14C ^R 12 _s (8 %), N ^R 11C ^R 12 _s (5 %), C ^R 15N ^R 19 _s (6 %), N11C12H13 _b (7 %), H13C12C14 _b (7 %), C ^R 15N ^R 16sH17 (6 %)
1569.2	C ^R 15N16 _s (6 %), H17N16H18 _b (24 %), C15 ^R N16C18 _b (14 %), C15 ^R N16C17 _b (11 %)
1476.1	C ^R 15N ^R 19 _s (9 %), C ^R 12C ^R 14 _s (7 %), N ^R 11C ^R 12H13 _b (6 %), C15 ^R N16C18 _b (6 %), C ^R 15C ^R 14 _s (6 %)
708.6	N11C ^R 12C ^R 14C ^R 15 (5 %), C ^R 12C ^R 14C ^R 15 N ^R 19 (8 %), C29C ^R 14C ^R 15 N16 (8 %), C ^R 14C ^R 15 N16H18 (5 %), N ^R 19C ^R 15 N16H17 (7 %), C ^R 14C ^R 15 N ^R 19C ^R 20 (7 %)
563.5	C ^R 14C ^R 15N16H18 (6 %), C ^R 14C ^R 15N16H17 (8 %), N ^R 19C ^R 15N16H17 (8 %), N ^R 19C15N16H18 (6 %)
562.5	C ^R 14C ^R 15N16H18 (7%), C ^R 14C ^R 15N16H17 (9 %), N ^R 19C ^R 15N16H17 (7%)
470.2	C ^R 14C29C30C ^R 15 (11 %), C29C30Si31C36 (5 %), N11C ^R 12C ^R 14C29 (4 %)

References

1. R. Johnson, NIST Computational Chemistry Comparison and Benchmark Database, 2013, <http://cccbdb.nist.gov/>.
2. L. Tian and F. Chen, *J. Comput. Chem.*, 2012, **33**, 580-592.
3. C. Reichardt, *Chem. Rev.*, 1994, **94**, 2319-2358.
4. G. Baysinger, L. I. Berger, R. Goldberg, H. Kehiaian, K. Kuchitsu, G. Rosenblatt, D. Roth and D. Zwillinger, *CRC handbook of chemistry and physics*, CRC Press, Boca Raton, FL, 95 edn., 2015.
5. S. Spange, N. Weiß, C. H. Schmidt and K. Schreiter, *Chemistry-Methods*, 2021, **1**, 42-60.