Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.

Supplementary Information: Probing the conformational landscape and thermochemistry of DNA dinucleotide anions via helium nanodroplet infrared action spectroscopy

Daniel A. Thomas[†], Rayoon Chang[†,‡], Eike Mucha[†], Maike Lettow[†,‡], Kim Greis[†,‡], Sandy Gewinner[†], Wieland Schöllkopf[†], Gerard Meijer[†], and Gert von Helden[†,*]

† Fritz-Haber-Institut der Max-Planck-Gesellschaft, Faradayweg 4-6, 14195 Berlin, Germany

‡ Institut für Chemie und Biochemie, Freie Universität Berlin, Takustrasse 3, 14195 Berlin, Germany

Table of Contents

Mass Spectra of DNA Dinucleotide Anions

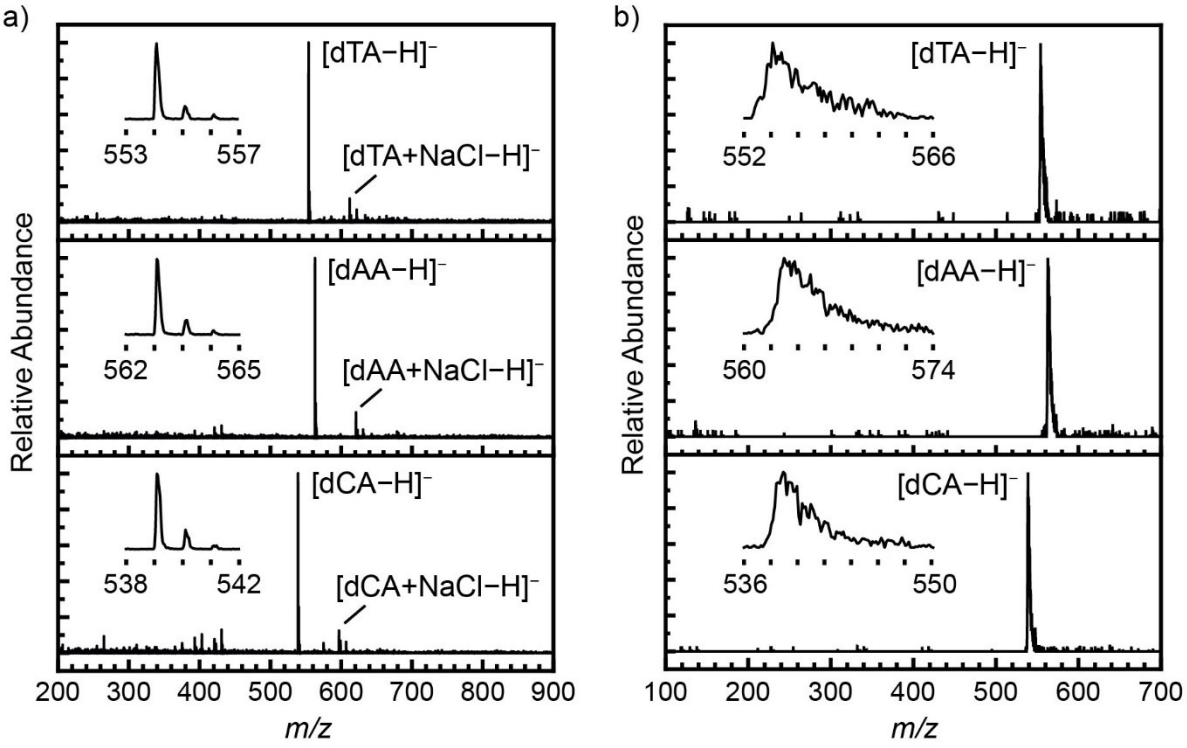

Figure S1. Mass spectra of DNA dinucleotide anions; shown in (a) are the mass spectra collected utilizing the Q-TOF without *m/z* selection, and shown in (b) are exemplary mass spectra of ions ejected from the helium nanodroplets upon IR photo excitation.

IR Spectra of [dCA−H]⁻ with Varying Pre-pickup Ion Temperature

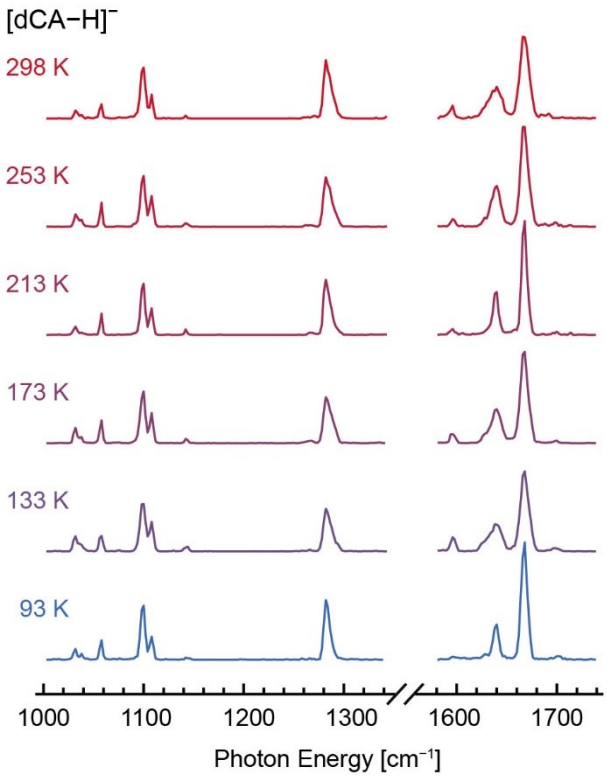

Figure S2. Helium nanodroplet infrared action spectra of the DNA dinucleotide anion [dCA−H]⁻ collected with varying ion trap temperature as indicated at left; the spectra do not show systematic changes with ion trap temperature, indicating a consistent conformer population under all measured conditions.

Depiction of Non-negative Matrix Factorization Process

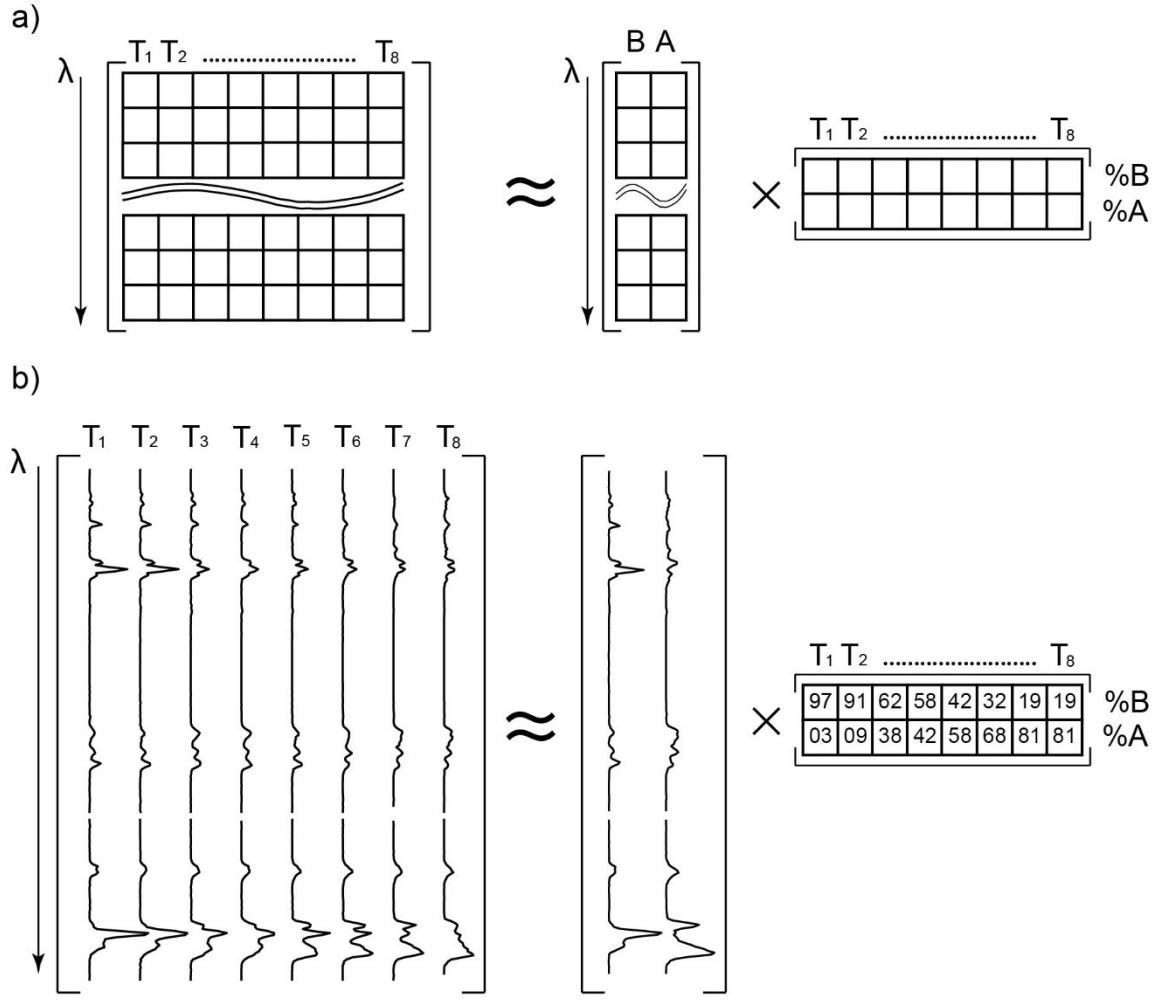

Figure S3. Illustration of non-negative matrix factorization (NMF) process, with a matrix-based visualization shown in (a) and a depiction using the data for $[dTA-H]^-$ shown in (b). In NMF, the data set is factorized into basis vectors, which in this case correspond to component spectra, and weighting factors, which here represent the relative abundance of each conformer.

Non-negative Matrix Factorization of IR Spectra of $[dAA-H]^-$

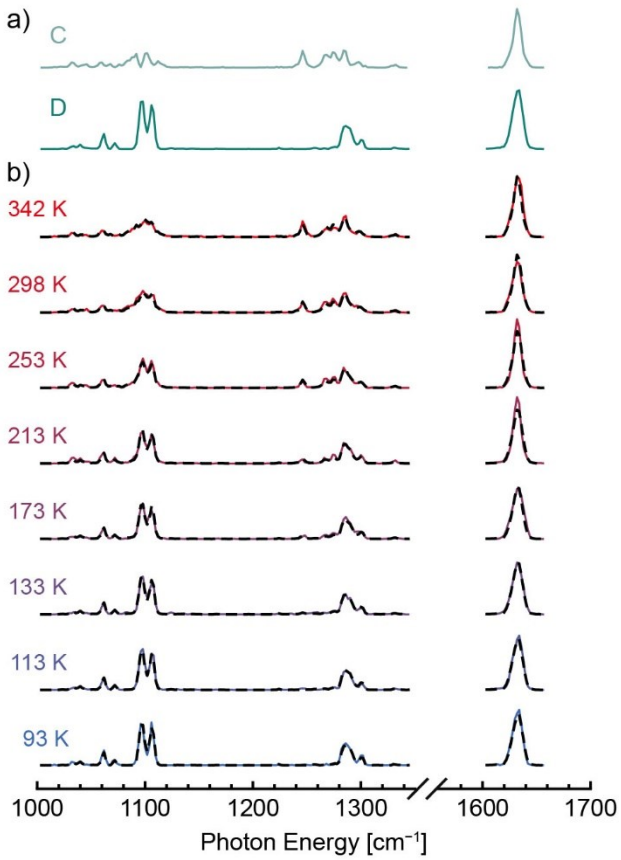

Figure S4. Non-negative matrix factorization (NMF) of experimental IR spectra of $[dAA-H]^-$; component conformer spectra were identified by NMF (a), and the experimental IR spectra (solid lines, b) were fitted as a weighted linear combination of the NMF spectra. The fitted spectra (dashed lines, b) exhibit good agreement with the experimental data.

Goodness of Fit Plots for the Determination of Equilibrium Constants

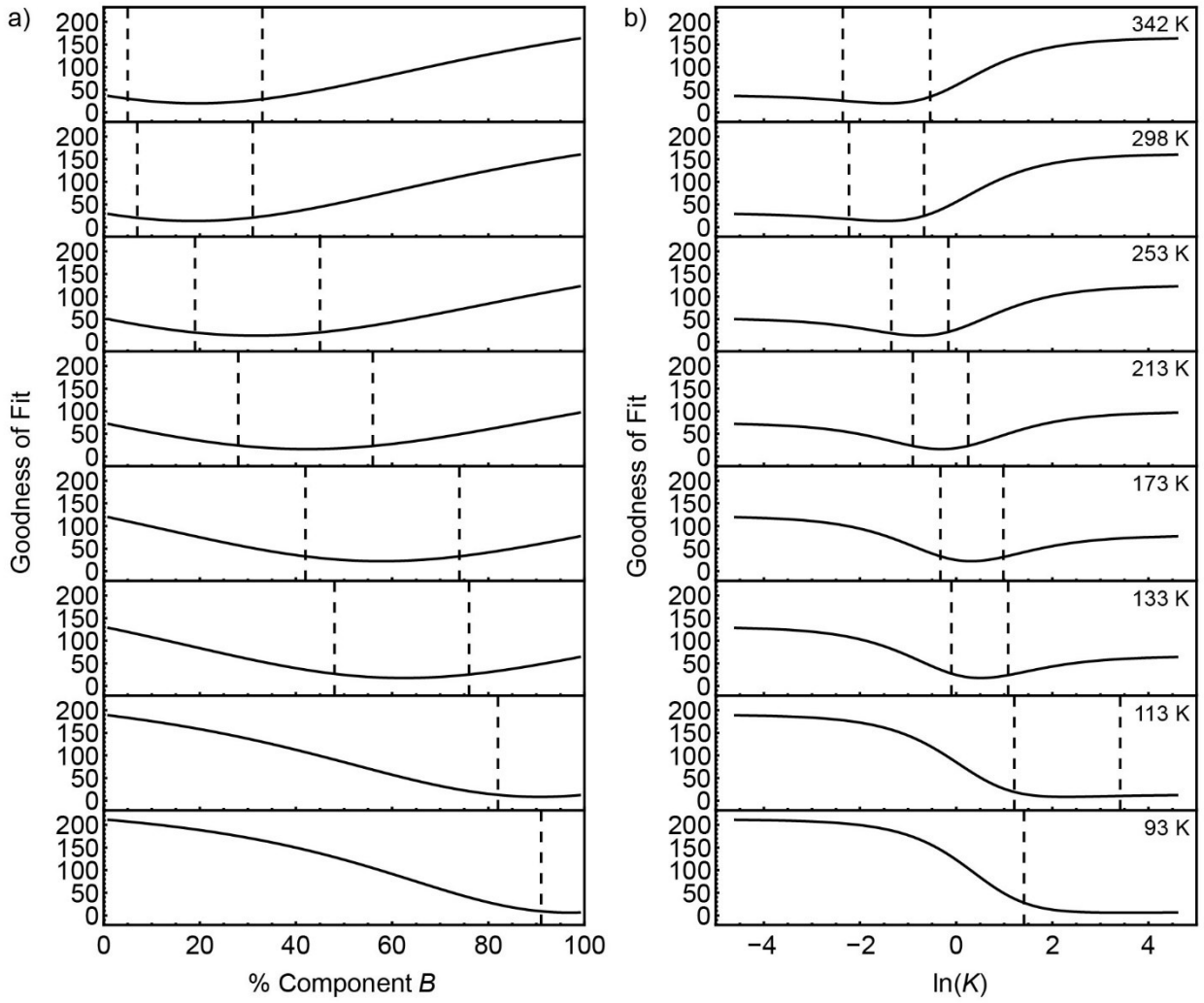

Figure S5. Goodness of fit plots used to calculate equilibrium constants of $[dTA-H]^-$; the goodness of fit parameter described in the Methods section is plotted as a function of (a) percentage or (b) ln(*K*) of the two NMF component spectra. Dashed lines mark the error estimate determined in (a) as the point at which the goodness of fit reaches 1.5 times its initial value and in (b) through propagation of uncertainty.

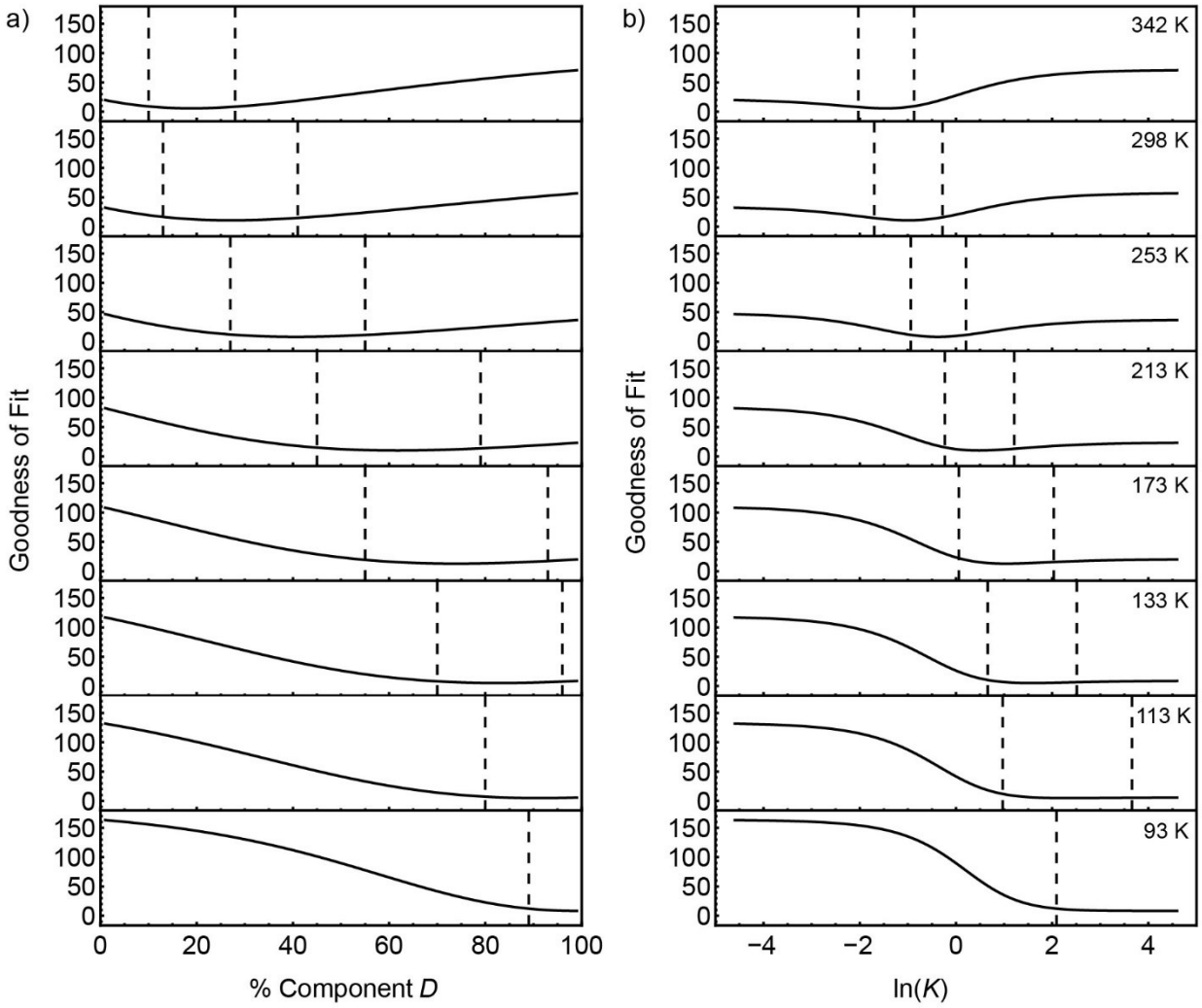

Figure S6. Goodness of fit plots used to calculate equilibrium constants of $[dAA-H]^-$; the goodness of fit parameter described in the Methods section is plotted as a function of (a) percentage or (b) ln(*K*) of the two NMF component spectra. Dashed lines mark the error estimate determined in (a) as the point at which the goodness of fit reaches 1.5 times its initial value and in (b) through propagation of uncertainty.

Equilibrium Constants Obtained by NMF and Fitting Procedure

Table S1. Population and equilibrium constants obtained from NMF analysis of data for $[dTA-H]^-$.

T **(K)**	**% *B* (NMF)[a,b]**	**% *B* (Fit)[a,c]**	**Uncertainty % *B* (Fit)[a,d]**	**ln(*K*) (NMF)[b,e]**	**ln(*K*) (Fit)[c,e]**	**Uncertainty ln(*K*) (Fit)[e,f]**
342	21	19	14	−1.3	−1.5	0.9
298	20	19	12	−1.4	−1.5	0.8
253	30	32	13	−0.8	−0.8	0.6
213	41	42	14	−0.4	−0.3	0.6
173	54	58	16	0.2	0.3	0.7
133	55	62	14	0.2	0.5	0.6
113	88	91	9	2.0	2.3	1.1
93	99	97	6	4.6	3.5	2.0

a. % *B* corresponds to the population (in percent) of the enthalpically favored conformer given by the weighting of the NMF spectrum that is dominant at low temperatures (component B, Figure 2a).
b. values obtained directly from the NMF algorithm utilized to identify component spectra.
c. values obtained using the fitting procedure described in the experimental section and shown in Figure S5.
d. uncertainty values obtained using the fitting procedure described in the experimental section and shown in Figure S5.
e. equilibrium constant *K* values derived from the relative conformer population.
f. obtained by standard propagation of uncertainty from the uncertainty in % *B* as described in the Methods section.

Table S2. Population and equilibrium constants obtained from NMF analysis of data for $[dAA-H]^-$.

T **(K)**	**%** ***D*** **(NMF)**	**%** ***D*** **(Fit)**	**Uncertainty % *D* (Fit)**	**ln(*K*) (NMF)**	**ln(*K*) (Fit)**	**Uncertainty ln(*K*) (Fit)**
342	19	19	10	−1.5	−1.5	0.6
298	29	27	11	−0.9	−1.0	0.7
253	37	41	13	−0.5	−0.4	0.6
213	50	62	19	0	0.5	0.7
173	68	74	17	0.8	1.0	1.0
133	80	83	14	1.4	1.6	0.9
113	91	91	14	2.3	2.3	1.3
93	90	99	9	2.2	4.6	2.5

a. % *D* corresponds to the population (in percent) of the enthalpically favored conformer given by the weighting of the NMF spectrum that is dominant at low temperatures (component D, Figure S3a).
b. values obtained directly from the NMF algorithm utilized to identify component spectra.
c. values obtained using the fitting procedure described in the experimental section and shown in Figure S6.
d. uncertainty values obtained using the fitting procedure described in the experimental section and shown in Figure S6.
e. equilibrium constant *K* values derived from the relative conformer population.
f. obtained by standard propagation of uncertainty from the uncertainty in % *D* as described in the Methods section.

Comparison to Predictions from Electronic Structure Methods

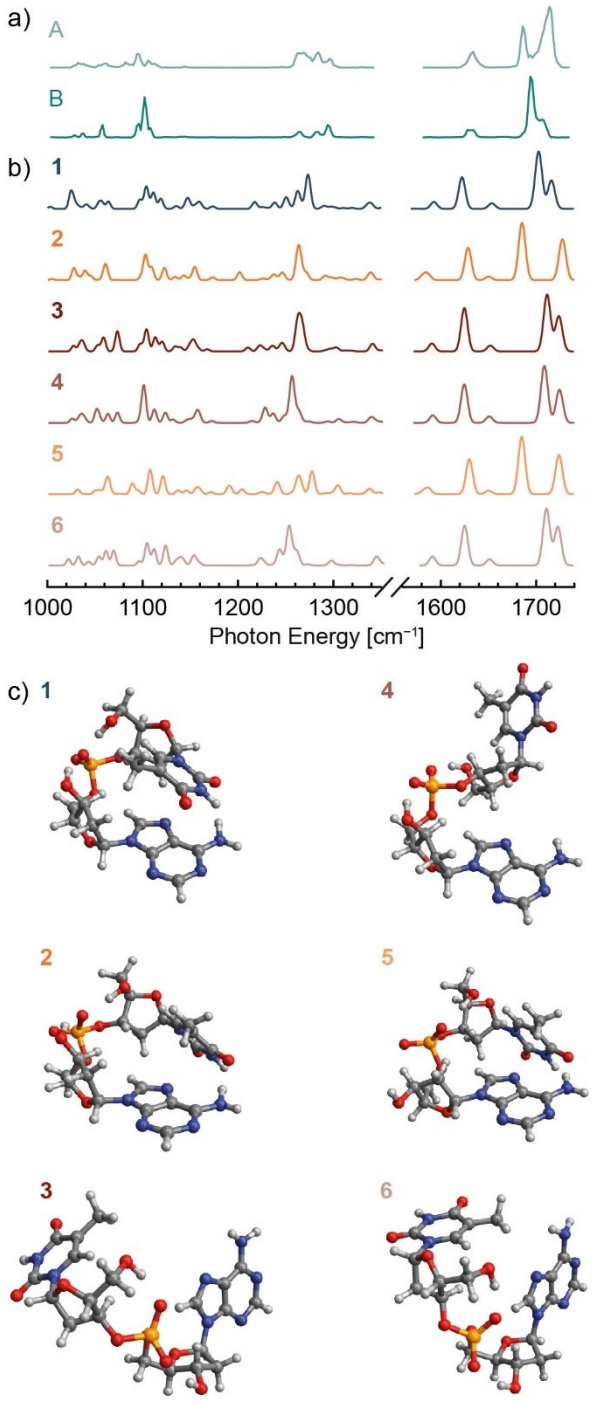

Figure S7. Low-energy conformers of $[dTA-H]^-$ identified by density functional theory (DFT) methods; the harmonic-approximation IR spectra for each conformer (PBE0-D3/def2TZVPP, 0.965 scaling factor) are shown in (b), and the spectra derived from NMF analysis of experimental data are shown in (a) for comparison. The corresponding conformer structures are shown in (c).

Blue, orange, and red spectra denote stacked nucleobase, hydrogen-bonded, and open conformers, respectively.

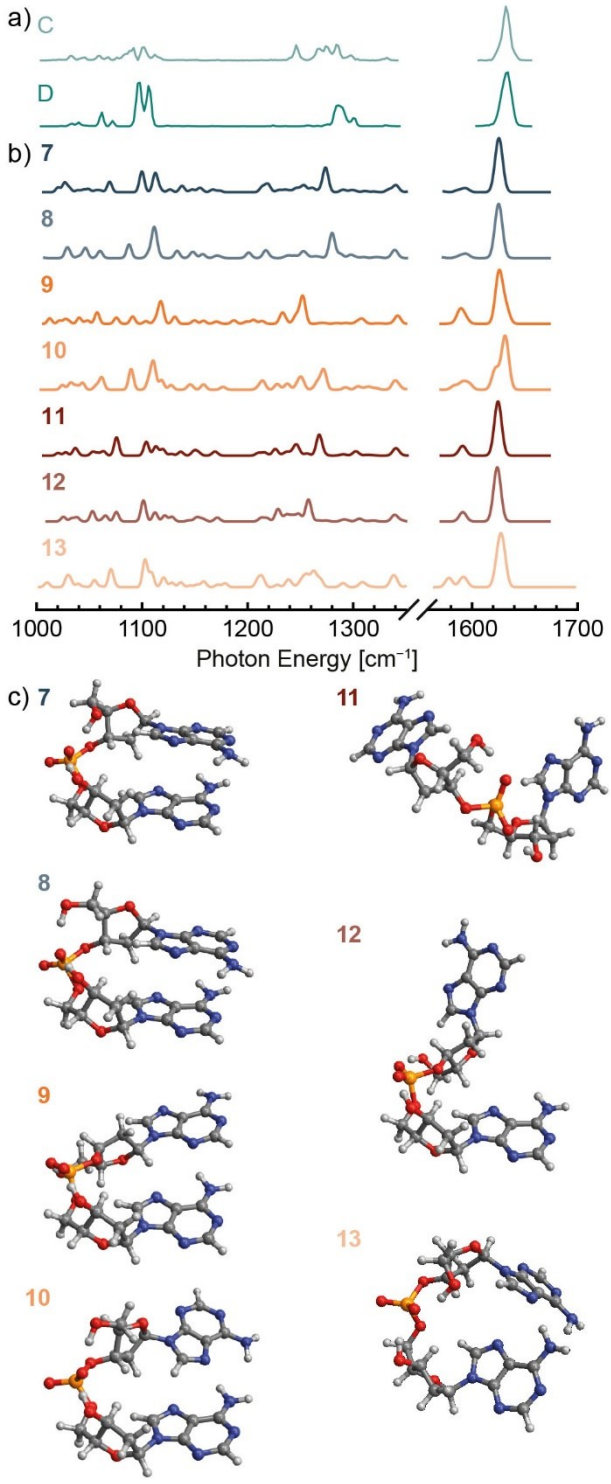

Figure S8. Low-energy conformers of $[dAA-H]^-$ identified by density functional theory (DFT) methods; the harmonic-approximation IR spectra for each conformer (PBE0-D3/def2TZVPP,

0.965 scaling factor) are shown in (b), and the spectra derived from NMF analysis of experimental data are shown in (a) for comparison. The corresponding conformer structures are shown in (c). Blue, orange, and red spectra denote stacked nucleobase, hydrogen-bonded, and open conformers, respectively.

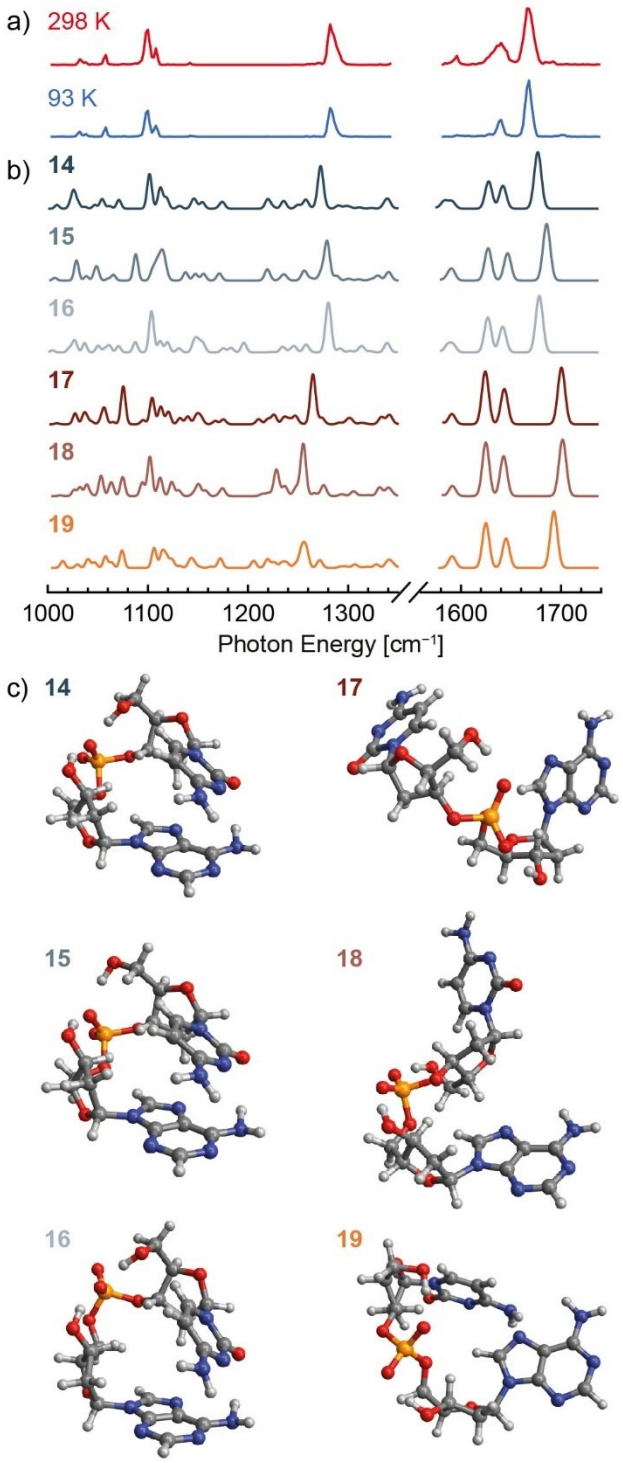

Figure S9. Low-energy conformers of $[dCA-H]^-$ identified by density functional theory (DFT) methods; the harmonic-approximation IR spectra for each conformer (PBE0-D3/def2TZVPP, 0.965 scaling factor) are shown in (b), and the experimental spectra at high and low temperature are shown in (a) for comparison. The corresponding conformer structures are shown in (c). Blue, orange, and red spectra denote stacked nucleobase, hydrogen-bonded, and open conformers, respectively.

ATDs of DNA Dinucleotide Anions

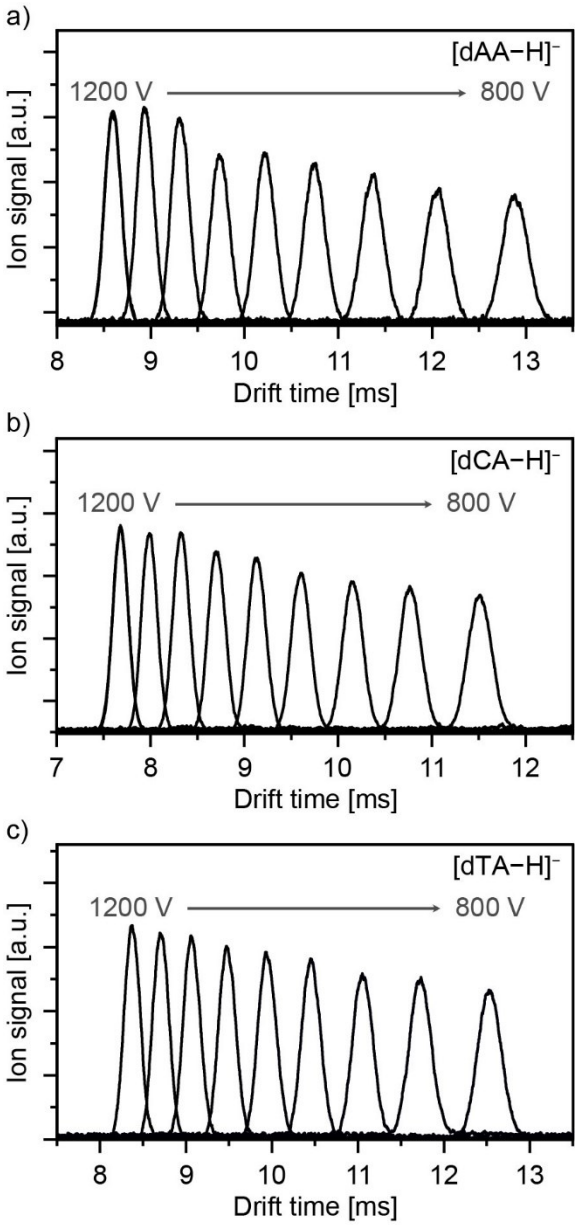

Figure S10. Arrival time distributions (ATDs) of $[dAA-H]^-$ (a), $[dCA-H]^-$ (b) and $[dTA-H]^-$ (c) obtained at drift voltages of 800–1200 V in 50 V-intervals.

Computed Thermochemistry and Collision Cross Sections of Dinucleotide Conformers

Table S3. Computed Thermochemistry and Collision Cross Sections of Conformers of $[dTA-H]^-$.

Structure	B3LYP-D3/ def2-TZVPP		ωB97X-D/ def2-TZVPP		PBE0-D3/ 6-311+G(d, p)		PBE0-D3/ def2-TZVPP			PBE0-D3/ aug-cc-pVTZ		Collision Cross Section (*CCS*)[c]		
	ΔH_{298}[a]	ΔS_{298}[b]	ΔH_{298}[a]	ΔS_{298}[b]	ΔH_{298}[a]	ΔS_{298}[b]	ΔH_{298}[a]	ΔS_{298}[b]	ΔG_{298}[a]	ΔH_{298}[a]	ΔS_{298}[b]	$^{PA}CCS_{He}$[d]	$^{TM}CCS_{He,298}$[e]	$^{TM}CCS_{He,80}$[f]
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	134.4	133.9	176.2
2	9.1	17.0	11.7	17.4	11.7	17.8	6.9	16.6	1.9	6.8	17.7	141.5	143.6	188.5
3	25.8	59.4	27.0	61.6	29.8	67.7	22.4	61.4	4.1	21.4	60.7	154.5	157.2	201.0
4	21.2	45.2	22.0	39.1	22.0	43.7	18.8	43.8	5.7	17.6	44.0	152.0	154.1	201.3
5	16.1	25.0	19.3	24.6	20.2	27.3	16.1	27.7	7.9	15.6	27.5	142.5	146.9	189.2
6	25.2	42.1	27.4	38.7	28.2	54.2	20.7	42.4	8.1	20.2	43.7	146.4	149.9	193.1
												$^{DT}CCS_{He,298}$		
												140		

a. values reported in kJ mol^{-1}.
b. values reported in J mol^{-1} K^{-1}.
c. values reported in square ångströms ($Å^2$).
d. projection approximation, helium drift gas.
e. trajectory method, helium drift gas at 298 K; calculated from structures and APT charges at the PBE0/aug-cc-pVTZ level of theory.
f. trajectory method, helium drift gas at 80 K; calculated from structures and APT charges at the PBE0/aug-cc-pVTZ level of theory.

Table S4. Computed Thermochemistry and Collision Cross Sections of Conformers of [dAA−H]⁻.

Structure	B3LYP-D3/ def2-TZVPP		ωB97X-D/ def2-TZVPP		PBE0-D3/ 6-311+G(d, p)		PBE0-D3/ def2-TZVPP			PBE0-D3/ aug-cc-pVTZ		Collision Cross Section (CCS)[c]		
	ΔH_{298}[a]	ΔS_{298}[b]	ΔH_{298}[a]	ΔS_{298}[b]	ΔH_{298}[a]	ΔS_{298}[b]	ΔH_{298}[a]	ΔS_{298}[b]	ΔG_{298}[a]	ΔH_{298}[a]	ΔS_{298}[b]	$^{PA}CCS_{He}$[d]	$^{TM}CCS_{He,298}$[e]	$^{TM}CCS_{He,80}$[f]
7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	137.5	138.3	182.0
8	5.0	7.4	3.9	-0.6	7.1	7.6	5.8	6.7	3.8	5.4	7.6	137.8	136.7	182.1
9	17.0	32.6	21.9	38.1	18.5	29.6	14.5	28.2	6.1	12.6	30.5	150.5	151.1	195.1
10	16.8	28.9	21.2	38.7	17.4	23.3	13.9	23.3	6.9	12.4	22.4	150.7	151.8	195.0
11	24.4	53.9	28.1	56.3	26.7	52.8	20.5	50.8	5.4	18.8	51.0	158.4	161.1	206.8
12	20.2	40.5	22.9	40.3	19.2	38.4	17.2	37.1	6.2	15.3	37.2	154.7	158.0	203.3
13	9.7	−1.7	15.3	2.7	8.7	-7.1	7.1	−7.0	9.2	7.3	−6.1	143.4	147.25	190.5
												$^{DT}CCS_{He,298}$		
												145		

a. values reported in kJ mol^{-1}.
b. values reported in J mol^{-1} K^{-1}.
c. values reported in square ångströms (Å^2).
d. projection approximation, helium drift gas.
e. trajectory method, helium drift gas at 298 K; calculated from structures and APT charges at the PBE0/aug-cc-pVTZ level of theory.
f. trajectory method, helium drift gas at 80 K; calculated from structures and APT charges at the PBE0/aug-cc-pVTZ level of theory.

Table S5. Computed Thermochemistry and Collision Cross Sections of Conformers of $[dCA-H]^-$.

Structure	B3LYP-D3/ def2-TZVPP		ωB97X-D/ def2-TZVPP		PBE0-D3/ 6-311+G(d, p)		PBE0-D3/ def2-TZVPP			Collision Cross Section (CCS)[c]		
	ΔH_{298}[a]	ΔS_{298}[b]	ΔH_{298}[a]	ΔS_{298}[b]	ΔH_{298}[a]	ΔS_{298}[b]	ΔH_{298}[a]	ΔS_{298}[b]	ΔG_{298}[a]	$^{PA}CCS_{He}$[d]	$^{TM}CCS_{He,298}$[e]	$^{TM}CCS_{He,80}$[f]
14	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	131.1	129.5	173.1
15	12.8	5.3	11.7	8.4	16.3	6.5	14.3	5.6	12.6	131.8	132.6	174.3
16	15.5	7.5	15.6	10.6	16.9	4.2	15.3	4.3	14.1	131.8	131.3	175.1
17	37.2	65.2	37.7	14.9	41.3	66.2	33.7	62.4	15.1	151.3	151.5	197.0
18	35.3	52.0	35.0	19.2	36.5	49.0	32.9	48.2	18.5	148.9	151.5	198.2
19	32.9	34.2	35.1	36.8	33.7	31.1	27.7	29.0	19.1	140.8	145.5	189.1
										$^{DT}CCS_{He,298}$		
										130		

a. values reported in kJ mol^{-1}.
b. values reported in J mol^{-1} K^{-1}.
c. values reported in square ångströms ($Å^2$).
d. projection approximation, helium drift gas.
e. trajectory method, helium drift gas at 298 K; calculated from structures and APT charges at the PBE0/aug-cc-pVTZ level of theory.
f. trajectory method, helium drift gas at 80 K; calculated from structures and APT charges at the PBE0/aug-cc-pVTZ level of theory.