Supporting Information

Hydrogen Bond Networks in Gas-Phase Complex Anions

Zhisheng Lai[1], Minhui Shen[1], Yong Shen[1], Yu-Xin Ye[1], Fang Zhu[1], Jianqiao Xu[1]*, Gangfeng Ouyang[123]*

[1]MOE Key Laboratory of Bioinorganic and Synthetic Chemistry/KLGHEI of Environment and Energy Chemistry, School of Chemistry, Sun Yat-sen University, Guangzhou 510006, China.

[2]College of Chemistry, Center of Advanced Analysis and Gene Sequencing, Zhengzhou University, Zhengzhou 450001, China.

[3]Guangdong Provincial Key Laboratory of Emergency Testing for Dangerous Chemicals, Guangdong Institute of Analysis (China National Analytical Center Guangzhou), Guangdong Academy of Sciences, Guangzhou 510070, China

*Correspondence: xujq27@mail.sysu.edu.cn (J.X.), cesoygf@mail.sysu.edu.cn (G.O.)

Table of Contents

1. Molecular dynamics

In this paper, MD-conformers could not explain the experiment results and potential energy pathways perfectly due to the disrupted structure of $[\alpha\text{-CD–H}]^-$ in complex anions. We still discussed their possibility in the main text and showed the processes here.

For molecular dynamics (MD) simulations, the RBA molecules adopt two docking orientations (Figure S4), having their carboxyl termini interacting with the primary hydroxyl groups from the cavity of α-CD ("head" and inside) or outside α-CD ("tail" and outside). These structures were utilized as the initial docking conformers to perform the MD simulations in the Amber20 software package[1] in a vacuum environment, which was corresponding to the environment in the mass spectrometer. The GLYCAM_06j-1 force field parameters were used to construct the $[\alpha\text{-CD–H}]^-$. And the GAFF2 force field parameters were used for the RBAs. Charges were applied with the restrained electrostatic potential (RESP) method by the program Antechamber. Missing parameters were generated with the parmchk routine of Antechamber. Energy minimization was performed with the steepest descent method for the first 60000 steps, and the conjugate gradient algorithm was followed for the next 140000 steps. After minimization, the systems were brought to 400 K by ramping up the temperature over periods of 200 ps, followed by a slow annealing period with the temperature dropping to 300 K in 4 ns. Finally, the complex anions were maintained at 300 K for 2 ns to attain production simulations. 6000 geometries were output every 1 ps from the trajectories in the annealing-production process. RMSD calculations of the collecting trajectories were performed with the cpptraj module of Amber20. According to the RMSD, 100 geometries were extracted and optimized at the PM7 level[2] in the MOPAC 2016 program.[3] The attained optimized geometries were then clustered with the threshold of 0.25 kcal mol^{-1} and 1 angstrom by the Moclus program.[4] 5 representative structures with the lowest energies were selected for each complex anion. For high precision computations, the 5 structures were optimized at the B3LYP-D3(BJ)/6-311G** level of theory. Meanwhile, the counterpoise corrected energies were calculated at the M062X/def2TZVP level of theory with D3 empirical dispersion correction. CP technique was employed to correct the BSSE problem. At last, the geometries with the lowest counterpoise corrected energies were selected as the final MD-conformers for the four complex anions (Figure S5).

2. Figures S1 to S7

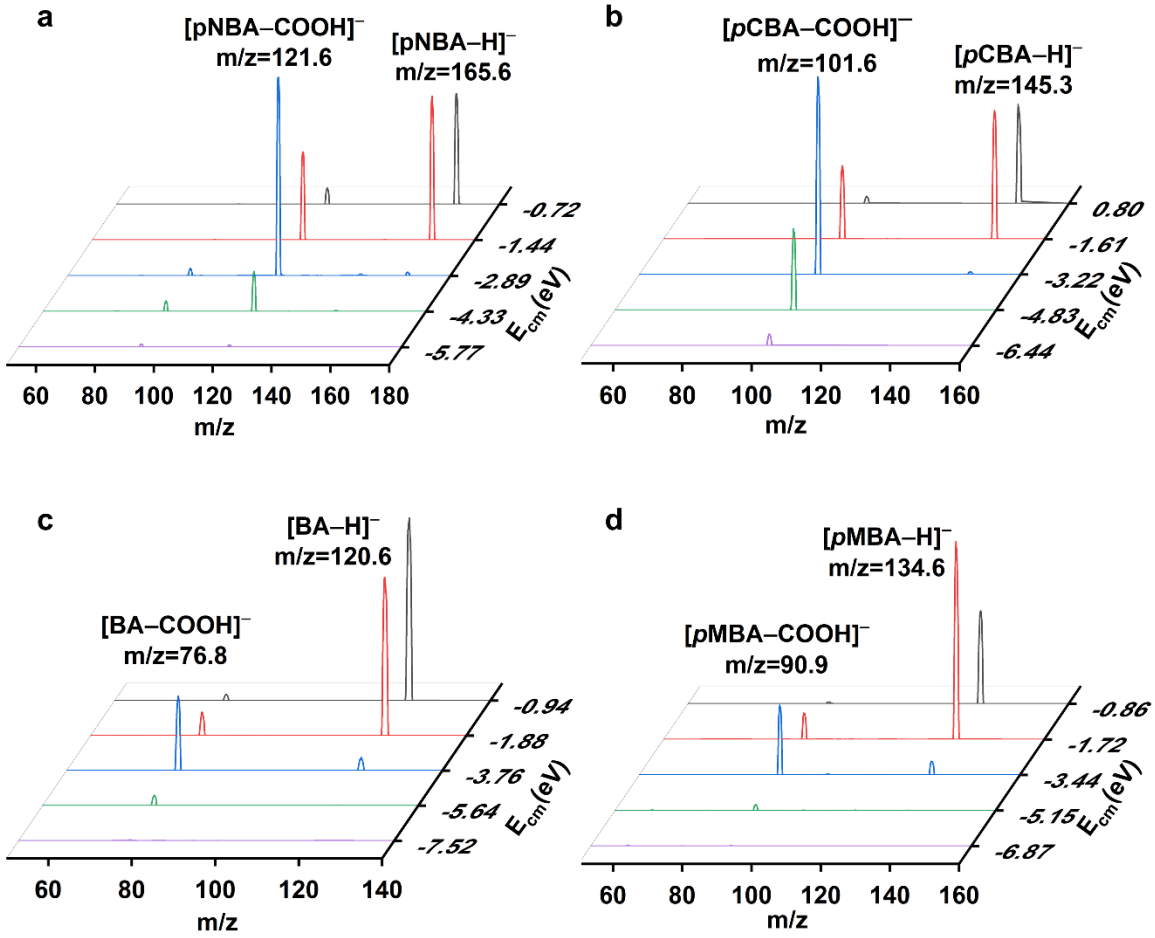

Figure S1. Tandem mass spectra of the (a) *p*NBA, (b) *p*CBA, (c) BA and (d) *p*MBA with increasing E_{cm}.

As shown in Figure S1, the parent anions $[RBA–H]^-$ and their decarboxylated products could be detected even the value of E_{cm} far exceeding 1.40 eV. By contrast, $[RBA–H]^-$ and $[RBA–COOH]^-$ could not be detected during the dissociation of $[\alpha\text{-CD+RBA–H}]^-$ complex anions in the low E_{cm} stage blow 1.40 eV. These results verified that RBAs were not fragmented into undetectable anions, but left the complex anions in their neutral forms.

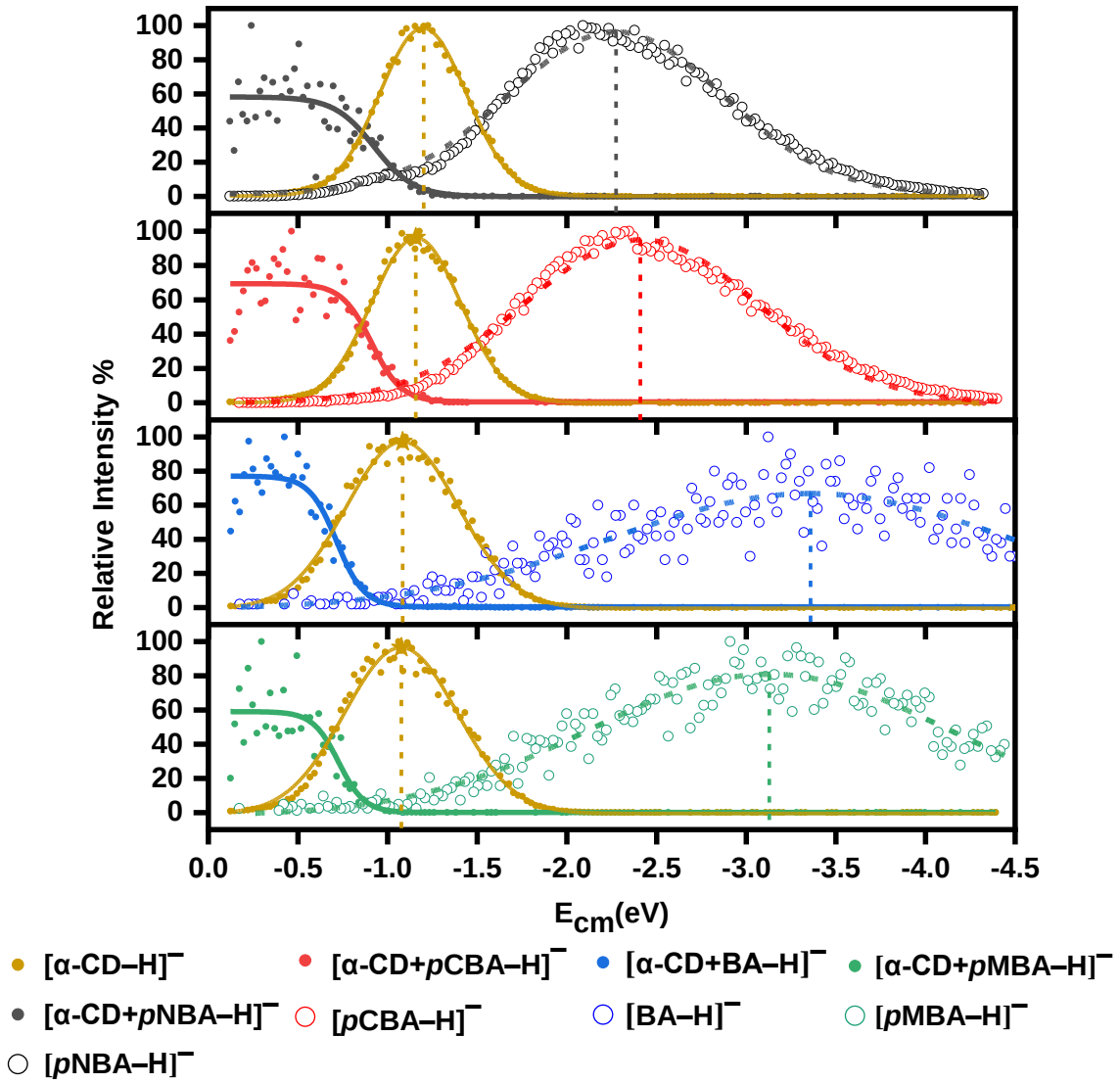

Figure S2. Abundance of parent anions and daughter anions under continuously increased collision energies. The profiles of the parent anions were fitted in the Boltzmann function, and the profiles of the daughter anions were fitted in the Gauss function. The R^2 values were 0.8532, 0.9133, 0.9506 and 0.8555 for the fitted curves of [α-CD+*p*NBA–H]$^-$, [α-CD+*p*CBA–H]$^-$, [α-CD+BA–H]$^-$ and [α-CD+*p*MBA–H]$^-$. From top to bottom, the R^2 values of the fitted curves of [α-CD–H]$^-$ are 0.9969, 0.9955, 0.9908 and 0.9879, respectively. And the R^2 values of the fitted curves of [RBA–H]$^-$ were 0.9861, 0.9838, 0.8032 and 0.9062 respectively. The fitting results were shown in Table S1.

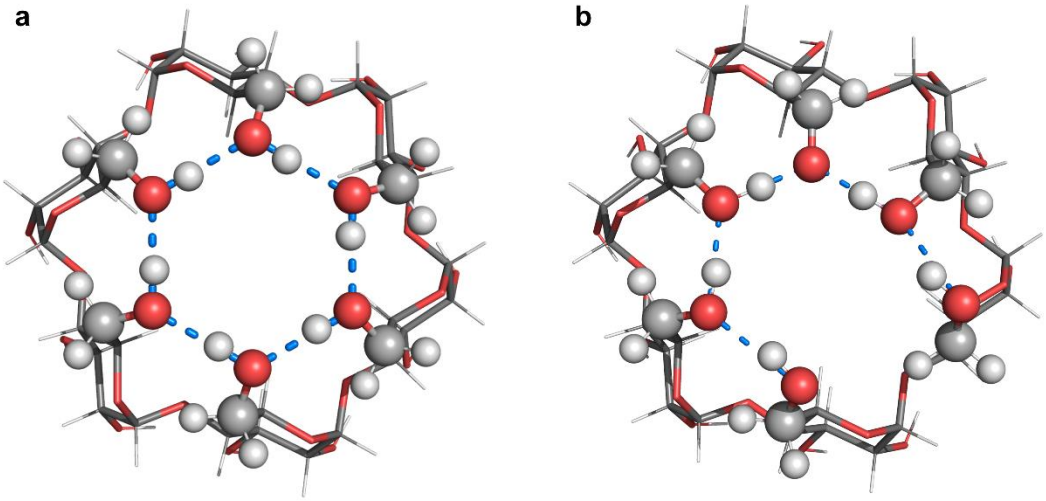

Figure S3. Calculated global minimum conformation for the (a) α-CD and (b) $[\alpha\text{-CD–H}]^-$ at the M062X/6-311G** level of theory with D3 empirical dispersion correction.

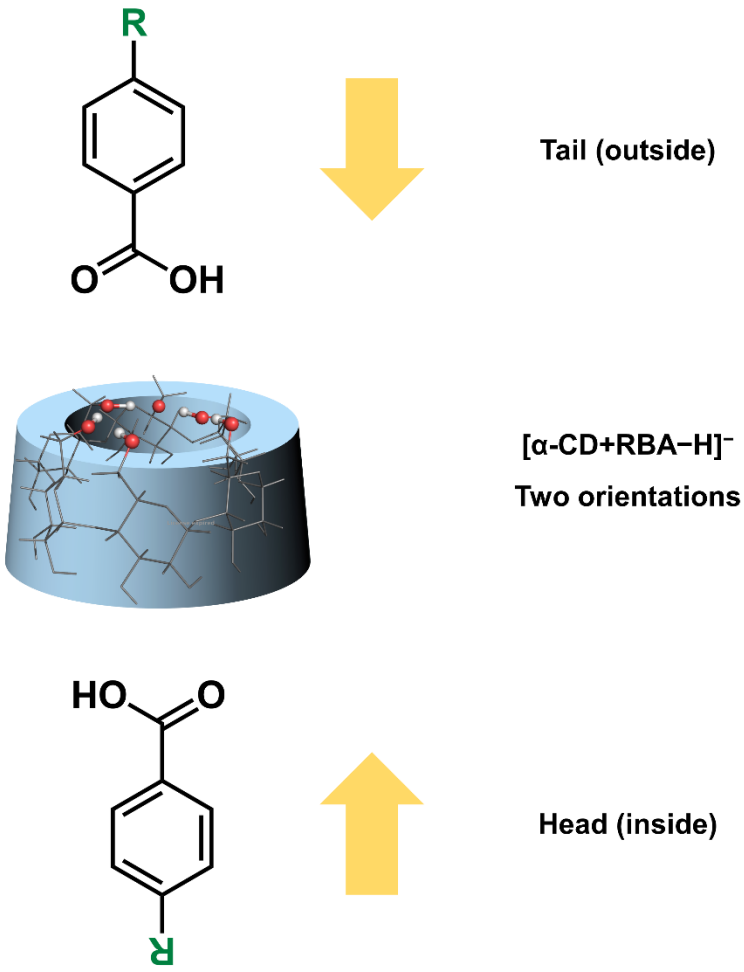

Figure S4. Two docking orientations for calculating the complex anions in the MD simulations.

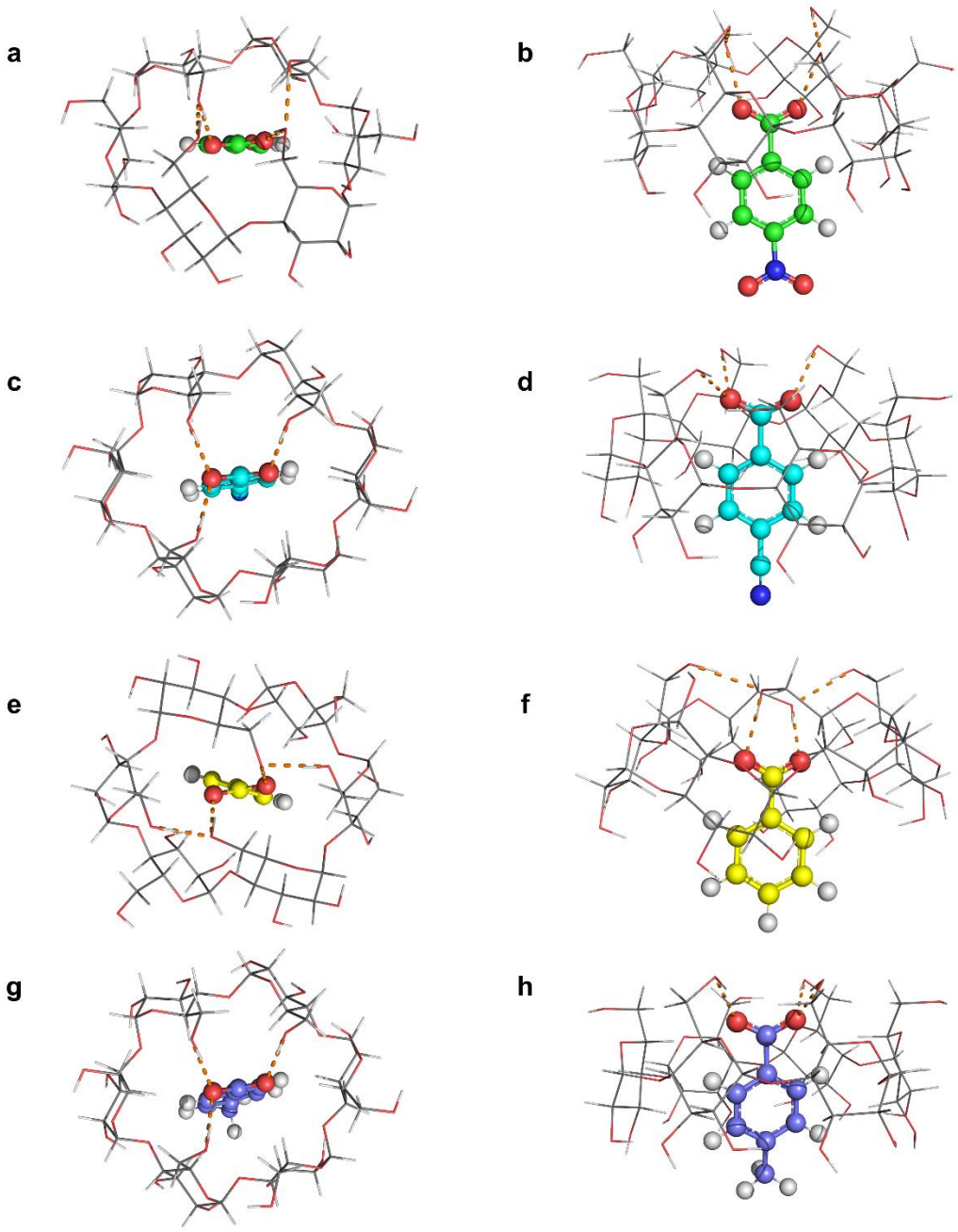

Figure S5. Geometries of MD-conformers for four complex anions, (a,b) [α-CD+*p*NBA–H]$^{-}$, (c,d) [α-CD+*p*CBA–H]$^{-}$, (e,f) [α-CD+BA–H]$^{-}$ and (g,h) [α-CD+*p*MBA–H]$^{-}$. (a,c,e,g) Top views and (b,d,f,h) side views.

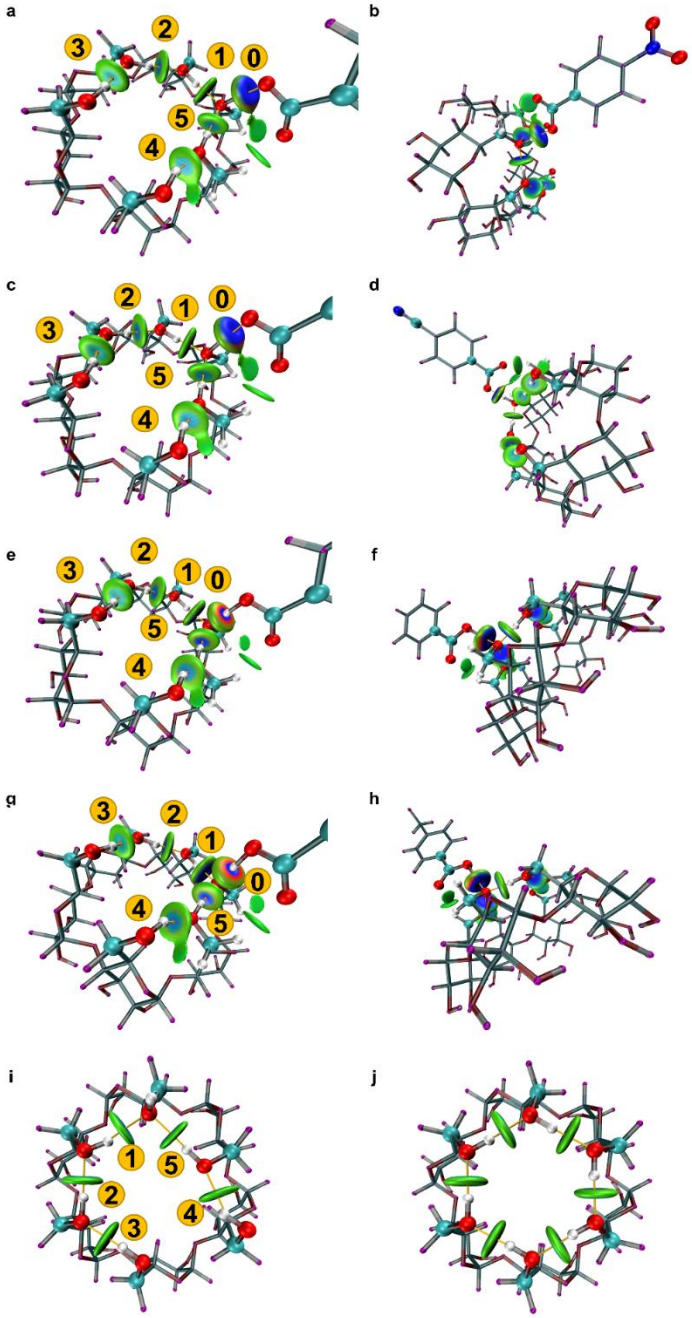

Figure S6. Visualization of HBs, BCPs and BPs in (a,b) [α-CD+*p*NBA–H]$^{-}$, (c,d) [α-CD+*p*CBA–H]$^{-}$, (e,f) [α-CD+BA–H]$^{-}$, (g,h) [α-CD+*p*MBA–H]$^{-}$, (i) α-CD_IM and (j) α-CD. (a,c,e,g) Side views and (b,d,f,h) full views.

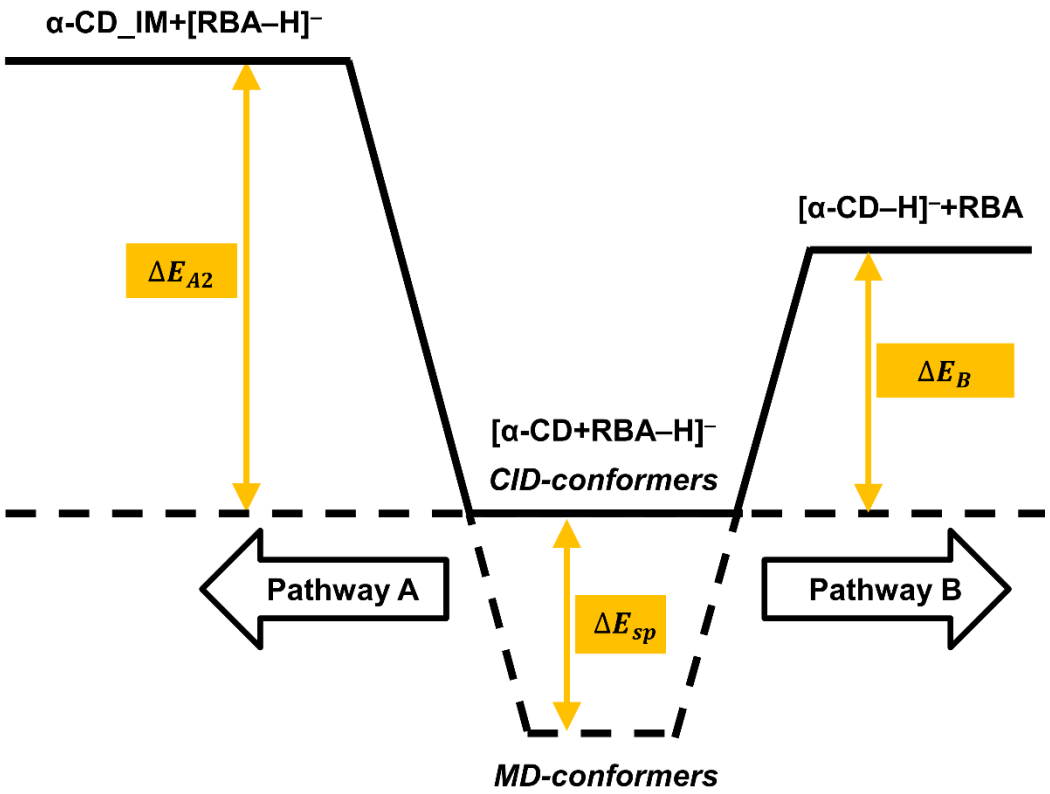

Figure S7. Potential energy diagram of two competitive dissociation pathways of complex anion [α-CD+RBA–H]⁻. The relatively stable MD-conformers, which owned lower counterpoise corrected energies, were the precursors of the CID-conformers.

3. Table S1 to S6

Table S1. Representative values of the fitted curves in **Figure S2**.

Parent anions	$E_{cm,1/2}$(eV)[a]	$E_{cm,max1}$ (eV)[b]	$E_{cm,max2}$ (eV)[c]
[α-CD+*p*NBA–H]$^-$	–0.92	–1.20	–2.27
[α-CD+*p*CBA–H]$^-$	–0.91	–1.15	–2.41
[α-CD+BA–H]$^-$	–0.72	–1.08	–3.36
[α-CD+*p*MBA–H]$^-$	–0.73	–1.08	–3.13

[a]The $E_{cm,1/2}$ values referred to the E_{cm} on the fitted curves when the dissociation ratios of the parent ions are 50%.

[b]The $E_{cm,max1}$ values indicated the E_{cm} at the peak relative abundance of [α-CD–H]$^-$ on the fitted curves.

[c]The $E_{cm,max2}$ values indicated the E_{cm} at the peak relative abundance of [RBA–H]$^-$ on the fitted curves.

The $E_{cm,1/2}$ value is commonly used to compare the stabilities of complex ions. It can be concluded from the fitted curves that gas-phase kinetic stability of the complex anions increased following as the [α-CD+pMBA–H]$^-$ ≈ [α-CD+BA–H]$^-$ < [α-CD+*p*CBA–H]$^-$ ≈ [α-CD+*p*NBA–H]$^-$. Correspondingly, |$E_{cm,max1}$| of [α-CD+*p*NBA–H]$^-$ and [α-CD+*p*CBA–H]$^-$ are larger than those of the others, matching the gas-phase kinetic stability of complex anions. Meanwhile, It was observed that stronger E_{cm} was needed to yield [RBA–H]$^-$ for [α-CD+*p*MBA–H]$^-$ and [α-CD+BA–H]$^-$.

Table S2. Calculated thermal correction factors and thermodynamic energies (Hartree) for monomers.

Species	M062X-D3/6-311G**[a]		M062X-D3/def2TZVP[b]		
	Corr.H	Corr.G	sp	H	G
α-CD	1.12	0.97	-3664.78	-3663.65	-3663.80
[α-CD–H]$^-$	1.11	0.96	-3664.24	-3663.13	-3663.28
***p*NBA**	0.13	0.08	-625.33	-625.20	-625.25
[*p*NBA–H]$^-$	0.12	0.07	-624.80	-624.68	-624.73
***p*CBA**	0.13	0.08	-513.07	-512.94	-512.99
[*p*CBA–H]$^-$	0.11	0.07	-512.53	-512.42	-512.46
BA	0.12	0.08	-420.82	-420.70	-420.74
[BA–H]$^-$	0.11	0.07	-420.27	-420.16	-420.19
***p*MBA**	0.15	0.11	-460.13	-459.98	-460.02
[*p*MBA–H]$^-$	0.14	0.10	-459.58	-459.44	-459.48

[a]Thermal corrections to Enthalpy (**Corr.H**) and Gibbs free energy (**Corr.G**) were obtained from the frequency calculations at the M062X/6-311G** level with D3 empirical dispersion correction. [b]Single point energies (sp) and thermodynamic energies (Hartree) were calculated at the M062X/def2TZVP level with D3 empirical dispersion correction. H=Corr.H+sp; G=Corr.G+sp. Gibbs free energy values of all monomers were used to calculate the △G values of deprotonation reactions in Table 1.

Table S3. Comparison of the single point energies calculated from different methods and basis sets (kcal mol^{-1}).

Method/basis set	α-CD	α-CD_IM	ΔE
M062X-D3/def2TZVP[a]	-2299683.12	-2299666.50	16.62
M062X-D3/6-31++G[b]**	-2298861.19	-2298844.89	16.31
PBE0-D3(BJ)/6-31++G[c]**	-2297343.87	-2297325.91	17.96
M062X-D3/ma-TZVPP[d]	-2299712.62	-2299695.90	16.72

[a]M062X/def2TZVP level of theory with D3 empirical dispersion correction.

[b]M062X/6-311G** level of theory with D3 empirical dispersion correction.

[c]Using PBE0[6] functional with Grimme's DFT-D3(BJ) empirical dispersion correction.

[d]The ma-TZVPP basis set was adapted, which is the "minimally augmented" version of the def2-TZVPP basis set[7,8] for which s and p type diffuse basis functions are added to the non-hydrogen atoms.

For confirming the accuracy of the calculated energies, the current method/basis set assembly for energy calculation was compared with two other ones. It was observed the variation of the energies was smaller than 1.5 kcal mol^{-1} no matter the method or the basis set was changed (Table S3).

Table S4. Properties of bond critical points (BCPs) under AIM theory analysis in Multiwfn.

Species	BCP[a]	ρ(r) (a.u.)[b]	$\nabla^2\rho(r)$[c]	$\lvert\nabla\rho(r)\rvert$[d]	Sign(λ2)*ρ[e]	δg[f]	ESP (a.u.)[g]
[α-CD+*p*NBA–H]⁻	0	0.0917	0.1020	1.03E-16	-0.0917	0.1770	0.3981
	1	0.0426	0.1217	1.10E-17	-0.0426	0.0774	0.1723
	2	0.0327	0.1140	5.05E-17	-0.0327	0.0585	0.1310
	3	0.0251	0.0969	1.61E-17	-0.0251	0.0439	0.0755
	4	0.0294	0.1123	2.26E-17	-0.0294	0.0540	0.0760
	5	0.0295	0.1009	1.19E-17	-0.0295	0.0507	0.0925
[α-CD+*p*CBA–H]⁻	0	0.0936	0.0995	2.50E-16	-0.0936	0.1809	0.4080
	1	0.0432	0.1221	1.92E-17	-0.0432	0.0786	0.1742
	2	0.0331	0.1148	1.71E-17	-0.0331	0.0592	0.1320
	3	0.0252	0.0974	2.50E-17	-0.0252	0.0442	0.0756
	4	0.0294	0.1122	8.93E-17	-0.0294	0.0540	0.0753
	5	0.0292	0.1001	2.82E-17	-0.0292	0.0501	0.0898
[α-CD+BA–H]⁻	0	0.1189	0.0322	1.77E-16	-0.1189	0.2319	0.5643
	1	0.0598	0.1197	9.03E-17	-0.0598	0.1113	0.2205
	2	0.0363	0.1184	7.30E-17	-0.0363	0.0653	0.1323
	3	0.0256	0.0984	2.30E-17	-0.0256	0.0449	0.0679
	4	0.0317	0.1172	3.72E-17	-0.0317	0.0585	0.0758
	5	0.0450	0.1175	1.48E-17	-0.0450	0.0809	0.1377
[α-CD+*p*MBA–H]⁻	0	0.1148	0.0465	2.78E-17	-0.1148	0.2236	0.5369
	1	0.0608	0.1189	3.50E-17	-0.0608	0.1131	0.2235
	2	0.0367	0.1191	2.38E-17	-0.0367	0.0662	0.1334
	3	0.0259	0.0991	1.44E-17	-0.0259	0.0454	0.0685
	4	0.0315	0.1167	6.28E-17	-0.0315	0.0581	0.0741
	5	0.0447	0.1171	7.97E-17	-0.0447	0.0802	0.1351
α-CD_IM	1	0.0249	0.0963	2.33E-17	-0.0249	0.0435	0.1966
	2	0.0287	0.1053	7.06E-17	-0.0287	0.0503	0.1923
	3	0.0256	0.0984	1.98E-17	-0.0256	0.0448	0.1495
	4	0.0207	0.0896	1.98E-17	-0.0207	0.0390	0.1034
	5	0.0210	0.0805	1.65E-17	-0.0210	0.0348	0.1712
[α-CD–H]⁻	1	0.1053	0.0783	1.63E-15	-0.1053	0.2054	0.4440
	2	0.0428	0.1248	4.33E-17	-0.0428	0.0786	0.1380
	3	0.0242	0.0947	3.04E-17	-0.0242	0.0424	0.0457

	4	0.0343	0.1228	3.00E-17	-0.0343	0.0637	0.0778
	5	0.0754	0.1223	8.92E-17	-0.0754	0.1447	0.2700
α-CD	avg.[h]	0.0330	0.1151	3.25E-17	-0.0330	0.0592	0.2101

[a]Serial numbers of BCPs were in conformity with Figures 3a to 3b and S6.

[b]ρ(r) refers to the electron density.

[c]$\nabla^2\rho(r)$ refers to the Laplacian of electron density.

[d]$|\nabla\rho(r)|$ refers to the gradient norm of electron density.

[e]Sign(λ_2) refers to the sign of the second largest eigenvalue of electron density Hessian matrix.

[f]δg function is defined in the reported IGMH method. The δg function at bond critical point in weak interaction region is shown to be closely related to interaction strength. This function can also be plotted as plane map or isosurface map to reveal all bonding regions.

[g]ESP refers to the total electrostatic potential contributed by electrons and nucleus.

[h]Average values of six intramolecular BCPs on the primary rim for α-CD.

Table S5. Properties of hydrogen bonds in the CID-conformers.

Species	HB[b]	Bond length (Å)	Bond angle (°)	Bond energy (kcal mol^{-1})
[α-CD+*p*NBA–H]$^-$	0	2.49	176.38	-31.55
	1	2.70	175.07	-15.23
	2	2.78	172.02	-11.93
	3	2.88	170.43	-9.40
	4	2.77	155.77	-10.82
	5	2.85	172.73	-10.88
[α-CD+*p*CBA–H]$^-$	0	2.48	176.96	-32.17
	1	2.70	175.25	-15.44
	2	2.77	171.74	-12.05
	3	2.87	170.37	-9.45
	4	2.77	155.73	-10.82
	5	2.85	172.77	-10.78
[α-CD+BA–H]$^-$	0	2.44	178.40	-40.58
	1	2.61	177.40	-20.95
	2	2.75	172.10	-13.11
	3	2.87	169.56	-9.58
	4	2.75	156.92	-11.61
	5	2.70	175.66	-16.04
[α-CD+*p*MBA–H]$^-$	0	2.45	178.23	-39.23
	1	2.60	177.64	-21.26
	2	2.74	171.87	-13.26
	3	2.86	169.71	-9.66
	4	2.75	156.62	-11.53
	5	2.71	175.59	-15.92
α-CD_IM	1	2.87	165.48	-9.33
	2	2.83	175.70	-10.60
	3	2.87	173.83	-9.59
	4	2.82	141.46	-7.93
	5	2.98	177.39	-8.04
[α-CD–H]$^-$	1	2.46	178.32	-36.05
	2	2.69	170.32	-15.30
	3	2.87	163.84	-9.09
	4	2.72	157.04	-12.46
	5	2.53	174.01	-26.11

α-CD	avr.[c]	2.80	171.97	-12.04

[a]The hydrogen bond is defined by the O-O distance smaller than 0.35 nm and the O-H-O angle no smaller than 150°.[9]

[b]Serial numbers were in conformity with Figures 3a to 3b and S6.

[c]Average values of six intramolecular HBs on the primary rim for α-CD.

Table S6. Calculated energies of the conformers of the complex anions (kcal mol^{-1})[a].

Species	MD-conformers	CID-conformers	$\|\Delta E_{sp}\|$
[α-CD+*p*NBA–H]$^-$	-2691787.86	-2691773.48	14.38
[α-CD+*p*CBA–H]$^-$	-2621346.70	-2621325.34	21.36
[α-CD+BA–H]$^-$	-2563451.39	-2563431.54	19.85
[α-CD+*p*MBA–H]$^-$	-2588123.23	-2588099.02	24.21

[a]Counterpoise corrected energies (E_{sp}) were calculated at M062X/def2TZVP level with D3 empirical dispersion correction.

References

(1) Case, D. A.; Aktulga, H. M.; Belfon, K.; Ben-Shalom, I. Y.; Brozell, S. R.; Cerutti, D. S.; III, T. E. C.; Cisneros, G. A.; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Giambasu, G.; M.K. Gilson, H.; Gohlke, A. W. G.; R. Harris, S. I.; S.A. Izmailov, C. J.; Kasavajhala, K.; Kaymak, M. C.; E. King, A. K.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Machado, M.; Man, V.; Manathunga, M.; Merz, K. M.; Miao, Y.; Mikhailovskii, O.; Monard, G.; Nguyen, H.; O'Hearn, K. A.; Onufriev, A.; Pan, F.; Pantano, S.; Qi, R.; Rahnamoun, A.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shen, J.; Simmerling, C. L.; Skrynnikov, N. R.; Smith, J.; Swails, J.; Walker, R. C.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xue, Y.; York, D. M.; Zhao, S.; Kollman, P. A. *Amber20*, University of California, San Francisco., **2021**.

(2) Stewart, J. J. P. Optimization of parameters for semiempirical methods VI: more modifications to the NDDO approximations and re-optimization of parameters. *J Mol Model* **2013**, *19*, 1-32.

(3) Stewart, J. J. P. *MOPAC2016*, Stewart Computational Chemistry, Colorado Springs, CO, USA, **2016**.

(4) Lu, T. *Molclus program, Version 1.9.9.6*, http://www.keinsci.com/research/molclus.html. (accessed 2021-09-29).

(5) Ma, X.; Wei, Z.; Xiong, X.; Jiang, Y.; He, J.; Zhang, S.; Fang, X.; Zhang, X. Gas-phase fragmentation of host–guest complexes between β-cyclodextrin and small molecules. *Talanta* **2012**, *93*, 252-256.

(6) Perdew, J. P., Ernzerhof, M. & Burke, K. Rationale for mixing exact exchange with density functional approximations. *J. Phys. Chem*. **1996**, *105*, 9982–9985.

(7) F. Weigend, R. Ahlrichs, *Phys. Chem. Chem. Phys*. **2005**, *7*, 3297.

(8) J. Zheng, X. Xu, D. G. Truhlar, *Theor. Chem. Acc*. **2011**, *128*, 295.

(9) Khuntawee, W.; Wolschann, P.; Rungrotmongkol, T.; Wong-Ekkabut, J.; Hannongbua, S. Molecular dynamics simulations of the interaction of beta cyclodextrin with a lipid bilayer. *J. Chem. Inf. Model.* **2015**, *55*, 1894-902.