Binding Properties of Cucurbit[7]uril to Neutral and Protonated Amino Acids: A Computational Study

Fenfen Ma,[1] Xiaoyan Zheng,[1]* Jing Xie,[1,2] and Zesheng Li[1]*

Correspondence to: **Xiaoyan Zheng** (E-mail: *xiaoyanzheng@bit.edu.cn*), **Zesheng Li** (E-mail: zeshengli@bit.edu.cn)

[1] *Key Laboratory of Cluster Science of Ministry of Education, Beijing Key laboratory of Photoelectronic/Electro-photonic Conversion Materials, School of Chemistry and Chemical Engineering, Beijing Institute of Technology, Beijing, China, 100081*

[2] *Department of Chemistry, Supercomputing Institute, Chemical Theory Center, University of Minnesota-Twin Cities, 207 Pleasant Street SE, Minneapolis, Minnesota, USA, 55455*

ABSTRACT

We systematically investigate the binding nature of CB[7] towards 20 amino acids in both neutral (AAs) and protonated (AAs^+) states by quantum chemistry methods. The result indicates molecular recognition process are enthalpy-driven. Among AAs, Arg and Asn shows the largest binding strength to CB[7], and for AAs^+, Gln^+ and Asn^+ bind to CB[7] the strongest. The binding strength of protonated CB[7]/AA^+ is much stronger than that of neutral CB[7]/AA counterpart, due to the introduction of ion-dipole interaction and the increase number and strength of hydrogen bonds. Energy decomposition analysis (EDA) indicates that electrostatic interactions play major roles in both CB[7]/AAs and CB[7]/AAs^+ complexes. Moreover, we analyzed the dependence of binding strength on single AA volume and dipole moment. This study is benefit for providing valuable information in predicting the recognition sites for sequence-based peptide or protein by CB[7] and rationally designing synthetic host molecule for specific peptide or protein recognition.

1 | Introduction

Full Proteins are key regulators in the living systems. The size, shape and polarity of proteins with different sequences are very different, but they could find corresponding target with high fidelity.[1-3] Weak noncovalent intermolecular interactions, such as electrostatic, dispersion, hydrogen bond (H-bond), play significant roles in protein functions. Designing synthetic compounds to target or recognize desired protein remains a great challenge to pharmaceuticals and medical diagnostics,[4] partially due to the lack of high resolution structures of most proteins and their surfaces. However, the recognition of amino acids (AAs) and peptide by synthetic host molecule is studied widely because AAs are easy to synthesize and to modify.[5,6] As building blocks of proteins, AAs are important components of life processes and play important roles in living systems. In addition, the binding sites between proteins and synthetic host molecule are usually structurally unique at the level of single-residue.[7-11] In particular, an increasing number of investigations are aimed at developing reliable and controllable supramolecular host molecule that target specific AAs in proteins[7, 8, 12, 13] and peptides.[14-16] Therefore, studying AAs recognition by synthetic host molecule and transferring the strategies into protein would be exceptionally valuable.

Cucurbit[7]uril (CB[7]) is a synthetic supramolecular host molecule, with macrocyclic

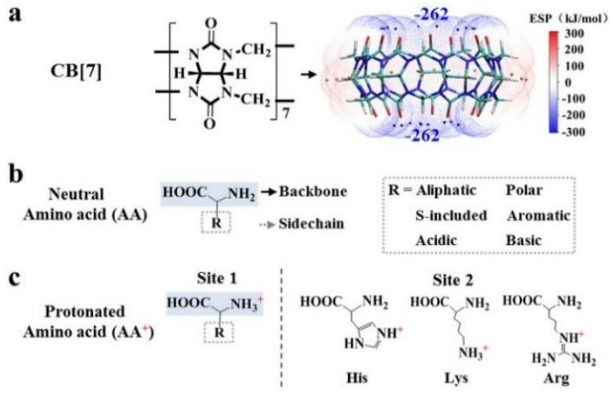


Figure 1 (a) Chemical structure and electrostatic potential (ESP, in kJ/mol) surface of CB[7] (isovalue = 0.001 a.u.). The orange balls indicate to the ESP maximum points and the blue balls refer to the ESP minimum points. C in cyan, H in white, N in blue, O in red. The color scheme in other parts are all the same in this work. (b) The structural formula and classification of twenty neutral amino acids (AAs). (c) The structural formula of two kinds of protonated amino acids (AAs^+). The protonation sites are labelled as Site 1 and Site 2, respectively.

barrel-shape connected by seven glycoluril units (Figure 1a).[17] Due to its good water solubility[18] and proper cavity size, CB[7] exhibits promising applications in biochemical systems,[18, 19] including drug delivery,[11] molecular recognition for AAs,[14] peptide,[15] and insulin,[7] as well as inhibition of amyloid fibrillation.[9] In these applications, CB[7] mainly binds to single residue of the biological systems *via* the intermolecular noncovalent interactions. Therefore, studies of molecular recognition of CB[7] on AAs could elucidate the origin of high selectivity of CB[7] on protein and provide more insights for protein recognition by synthetic host molecule.

To unravel the mechanism of molecular recognition between host CB[7] and AAs, several experimental and theoretical studies were performed in recent years. By NMR and UV-Vis measurements, Cong *et al.* reported CB[7] and AAs always form 1:1 ratio complexes.[20] Gao *et al.* reported the binding and selectivity of AAs to the inverted CB[7].[21] Lee *et al.* studied the recognition of CB[7] towards aromatic AAs and basic AAs, which revealed a higher stability of the basic AA complexes than aromatic AA complexes in the gas phase, and a reversed trend in aqueous solution.[22] Kovalenko *et al.* studied the host-guest interactions between CB[7] and a series of AAs in different ratios by electrospray ionization mass spectrometry (MS) techniques and collision induced dissociation, they emphasis the difference between the gas-phase behavior of AAs on protonation.[23] In addition, protonated AAs (AAs^+) also commonly exist and are important in biochemical systems.[22, 24] As far as we know, there is no systematic study on the binding between CB[7] and all 20 amino acids in both neutral (AAs) and protonated (AAs^+) states. The following questions remain to be answer: (i) the side chain effects of AAs on the molecular recognition by CB[7]; (ii) the specific binding pattern of the most stable configuration of each class of CB[7]/AAs or CB[7]/AAs^+ complexes; (iii) the range of binding strength and the relative contributions of different noncovalent interactions in each class of host-guest complexes; (iv) the relationship between the structural parameters (volume, V or dipole moment, μ) of single AA/AA^+ and the host-guest binding strength; (v) the last but not least, protonation effect on the host-guest binding strength after protonating each AA to AA^+. All these points will be useful in predicting the recognition sites for sequence-based peptide or protein by synthetic host molecule CB[7] and contributes to rational design of proper host molecule for specific protein recognition.

Herein we systematically studied the molecular recognition of CB[7] towards all 20 AAs in both neutral (Figure 1b) and single protonated states (Figure 1c) using density functional theory (DFT) calculations. As the smallest AA, Gly only has one hydrogen atom in the sidechain, thus we take Gly as a reference and study the sidechain effects of the other 19 AAs or AAs^+. According to the different chemical properties of sidechains, we assigned the 19 AAs into six classes, including aliphatic (Ala, Leu, Val, Ile, Pro), polar (Ser, Gln, Thr, Asn), sulfur-included (Cys and Met), aromatic (Trp, Phe, Tyr), acidic (Asp, Glu) and basic (Lys, His, Arg) AAs, as shown in Figure 1b.[25] The same classification method is used to protonated AAs^+. To be noted, because of different proton affinity, AAs have two types of protonation sites.[24] In most AAs, the protonation sites are in the $-NH_2$ group of the

backbone (Figure 1c, Site 1), while for the basic AAs, the protonation sites are in sidechains (Figure 1c, Site 2). In this work, we aim to explicitly demonstrate the driving force, binding patterns, binding strength and contributions of different noncovalent interactions to the formation of the twenty CB[7]/AA and CB[7]/AA^+ complexes. In Section II we give the details of our computational methodology, including DFT calculation method, independent gradient method (IGM)[26] and energy decomposition analysis (EDA),[27] where the latter two methods analyze the noncovalent interactions qualitatively and quantitatively. The results are presented and discussed in Section III and the paper ends with summary and conclusions in Section IV.

2 | Computational Methods

2.1 | DFT calculation

All density functional theory calculations were carried out using Gaussian 09 program package.[28] Characterizing weak noncovalent interaction is challenging for correct DFTs. To correctly describe noncovalent interaction, we tested six density functionals, including B3LYP,[29] CAM-B3LYP,[30] M06-2X,[31] M05-2X,[32] PBE,[33] with Grimme's D3 dispersion energy correction,[29, 34] and ωB97X-D[35] to compared with reference MP2,[36-38] with 6-31G* basis set.[22] Since ωB97X-D/6-31G* method matches best with MP2/6-31G* result, it is selected for the geometry optimization in this work (Figure 2b). Normal mode analysis were performed at the same level of theory to ensure that the optimized structures correspond to true minima. Then, single point calculation was performed on top of these optimized structures at ωB97X-D/6-311++G** level to correct the energies. This method is denoted as ωB97X-D/6-311++G**//ωB97X-D/6-31G*.[39] The initial configurations of AAs are referred to these two works.[24, 40] And the initial structure of each CB[7]/AA and CB[7]/AA^+ complex was obtained by using ABCluster[41] to pick out the lowest energy conformation among 1000 autogenerated structures. The interaction energies for CB[7]/AA and CB[7]/AA^+ complex were corrected on basis set superposition error (BSSE)[42, 43] using the counterpoise method.

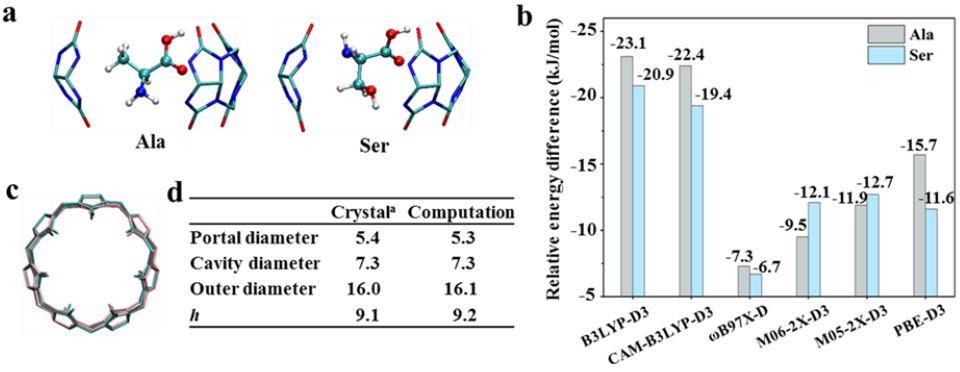
2.2 | Binding energy calculation

The binding free energy (ΔG) between CB[7] and guest molecule (AA or AA^+) is computed by equation (1):

$$\Delta G = G_{complex} - (G_{CB[7]} + G_{guest}) \quad (1)$$

where $G_{CB[7]}$, G_{guest} and $G_{complex}$ represent the Gibbs free energy of the optimized CB[7], guest molecule (AA or AA^+) and CB[7]/AA or CB[7]/AA^+ complex, respectively. Similarly, the binding enthalpy (ΔH) and binding electronic energy (ΔE) were calculated as following equations.

$$\Delta H = H_{complex} - (H_{CB[7]} + H_{guest}) \quad (2)$$



	Crystal[a]	Computation
Portal diameter	5.4	5.3
Cavity diameter	7.3	7.3
Outer diameter	16.0	16.1
h	9.1	9.2

Figure 2 (a) Structures of representative subsystems between CB[7] fragment and Ala or Ser. (b) The relative energy difference relative to MP2. (c) Alignment and (d) structural parameters comparison (in Å) of CB[7] between crystal structure (pink)[17] and optimized result (cyan).

$$\Delta E = E_{complex} - (E_{CB[7]} + E_{guest}) \quad (3)$$

The binding energy ΔE is the summation of deformation energy ΔE_{deform} and instantaneous interaction energy ΔE_{int}:

$$\Delta E = \Delta E_{deform} + \Delta E_{int} \quad (4)$$

The deformation energy is defined as:

$$\Delta E_{deform} = (E'_{CB[7]} - E_{CB[7]}) + (E'_{guest} - E_{guest}) \quad (5)$$

where $E'_{CB[7]}$ and E'_{guest} are the electronic energy of CB[7] and guest molecule with geometry that they acquire in each optimized complex, respectively. Therefore, by definition, interaction energy is calculated as

$$\Delta E_{int} = E_{complex} - (E'_{CB[7]} + E'_{guest}) \quad (6)$$

2.3 | Independent Gradient Method (IGM) analysis

To better understand the host-guest interactions, the electron density gradient (∇ρ)-based approach IGM[26] was introduced to identify and isolate the noncovalent interaction between molecules qualitively. A key descriptor in IGM is $\delta g^{inter} = |\nabla\rho^{IGM,inter}| - |\nabla\rho|$, which depends on the magnitude of the norm of the electron density gradient calculated with IGM model $\nabla\rho^{IGM,inter}$ and the density gradient ∇ρ. δg^{inter} describes intermolecular interactions between fragments and $\delta g^{inter} > 0$ indicates the presence of noncovalent interactions. Two types of IGM diagrams are displayed in Figure 3c (the scatter diagram) and Figure 3d (the isosurface diagram), respectively. In Figure 3c, the *y*-axis refers to δg^{inter}, while *x*-axis refers to $sign(\lambda_2)\rho$, that is the density multiply the sign of λ_2, where λ_2 is the second eigenvalue of electron density Hessian matrix, whose sign is able to classify the noncovalent interactions. λ_2 has a positive sign indicates the electron depletion because of steric repulsion, while λ_2 has a negative sign refers to the strong attractive interactions, such as H-bond and ion-dipole interactions. vdW interaction region has a very small ρ, thus vdW interaction are associated with $sign(\lambda_2)\rho$ approaching zero. The strong attractive interaction, vdW interaction and steric repulsion are shown in blue, green and red, respectively, in Figure 3c. The same color scheme is also used in IGM isosurface with $\delta g^{inter} = 0.01$ a.u as displayed in Figure 3d. Therefore, we could identify both the region and type of the noncovalent interactions between host and guest through IGM analysis. The IGM analysis were performed using Multiwfn[44] at ωB97X-D/6-311++G** level.

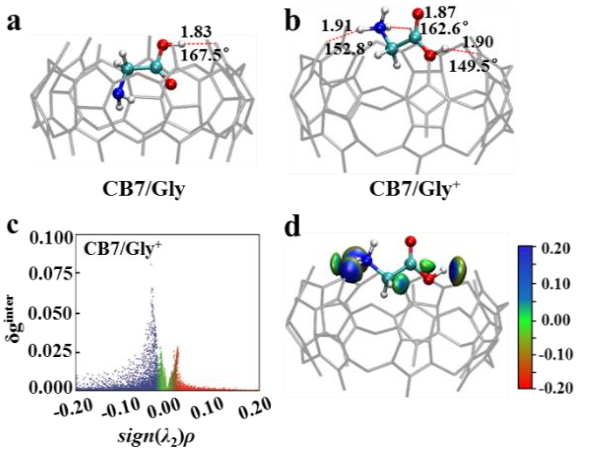


Figure 3 The optimized structures (a) CB[7]/Gly and (b) CB[7]/Gly$^+$. The H-bonds are marked by red dashed lines and distance between O of carbonyl group in CB[7] and H of -NH_2/-NH_3^+ or -COOH group of AAs/AAs$^+$ are shown nearby, in Å. IGM analysis of CB[7]/Gly$^+$: (c) 2D scatter plot with x=sign(λ_2)ρ, and y=δg^{inter}. (d) The corresponding isosurface with δg^{inter}= 0.01 a.u. The blue, green and red regions indicate strong attractive interaction, vdW and steric repulsion, respectively.

2.4 | Energy decomposition analysis (EDA)

EDA breaks the interaction energy (ΔE_{int}) between the fragments, i.e. host CB[7] and guest AA or AA$^+$, into four components, as shown in equation (7)[27]

$$\Delta E_{int}(EDA) = \Delta V_{elstat} + \Delta E_{Pauli} + \Delta E_{orb} + \Delta E_{disp} \quad (7)$$

ΔV_{elstat} corresponds to the attractive, classical electrostatic interaction between the charge distributions of the fragments. ΔE_{Pauli} is the repulsive interaction between the occupied orbitals of the fragments. ΔE_{orb} is the interaction between the occupied orbitals of one fragment and the unoccupied orbitals of the other fragment. It accounts for charge transfer, such as

donor–acceptor interactions between the H-bond fragments, and polarization such as electron density redistribution on one fragment due to the presence of another fragment. The ΔE_{disp} term corresponds to the dispersive effects between the two fragments. EDA analysis was performed at the CAM-B3LYP-D3BJ/TZP[34] level of theory using ADF (2018) program package[45, 46] on top of the geometries optimized with ωB97X-D/6-31G* method, because the interaction energies ΔE_{int} calculated by CAM-B3LYP-D3BJ/TZP is close to those by ωB97X-D/6-31++G** method.

All the visualization in this work was rendered by VMD.[47]

3 | Results and Discussions

3.1 | Benchmark and geometric optimization

To select a proper DFT method, we tested the performance of six functionals, including B3LYP,[29] CAM-B3LYP,[30] M06-2X,[31] M05-2X,[32] PBE,[33] with Grimme's D3 dispersion energy correction,[29, 34] and ωB97X-D,[35] by comparing the calculated binding energies of model reactions with the values calculated by MP2[36-38] method. The model reactions are the complexation between host CB[7] and two guest molecules, i.e. Ala and Ser. To obtain the initial structures of CB[7]/Ala and CB[7]/Ser, we used ABCluster[41] to autogenerate 1000 structures and selected the conformation with lowest energy based on the CHARMM force field.[48] Then the most stable structures were optimized with ωB97X-D/6-31G* method.[39] Next, considering computation efficiency, we used a truncated model (shown in Figure 2a) and computed the single point binding energy ΔE using MP2 and abovementioned six functionals. Here, the MP2 values are used as reference values to evaluate the performance of the six functionals. As shown in Figure 2b, ωB97X-D functional gives the closest energy as compared to MP2, where relative energy differences are -7.3 kJ/mol for CB[7]/Ala and -6.7 kJ/mol for CB[7]/Ser. In addition, we further validate the performance of ωB97X-D/6-31G* by geometric optimization of host molecule CB[7] and comparing it with X-ray crystal structure.[17] It shows that the optimized structure of CB[7] aligns well with the crystal structure (Figure 2c) and the difference of the structural parameters are within 0.1 Å (Figure 2d). Therefore, ωB97X-D/6-31G* is selected to perform full structure optimization in this study. On top of these optimized structures, we further performed single point calculations at ωB97X-D/6-311++G** level to correct the energies. If not mentioned explicitly, the energies reported in this work are computed by ωB97X-D/6-311++G**//ωB97X-D/6-31G* dual-level method with BSSE correction.

3.2 | Sidechain effect

In this session, we discussed the sidechain effect of AAs towards their binding to CB[7] in seven parts, based on the classification of the AAs. To help our understanding to the binding of AA/AA$^+$ to the host CB[7], we first plotted the electrostatic potential (ESP) surface of host CB[7] at the ωB97X-D/6-311++G** level with isovalue 0.001 a.u. As shown in Figure 1a, the carbonyl-fringed portals are highly negative. There are negative ESP maximum points (blue balls) near carbonyl group of two portals with value -262 kJ/mol, and the positive ESP maximum points (orange balls) locate at the outside equator. Such dipolar nature of two portals in CB[7] make it highly attractive to positive charged groups or they can act as acceptor of H-bonds. Since the cavity of CB[7] are relative less polar, it tends to encapsulate the nonpolar functional groups.

3.2.1 | Take Gly and Gly$^+$ as references

Gly and Gly$^+$ were taken as references for neutral AAs and protonated AAs$^+$ respectively, because of only one hydrogen in sidechain. The optimized structure of CB[7]/Gly and CB[7]/Gly$^+$ are shown in Figure 3a-b. Here we define a H-bond if the distance between the donor (D) and acceptor (A) atoms is less than 3.5 Å and the angle of D-H-A is greater than 135º. The optimized structure shows that Gly forms partially inclusion complex with CB[7] by one H-bond with the portal, and

Table 1 The binding free energies (ΔG^{CP}, ΔG) with and without BSSE correction, enthalpies (ΔH^{CP}) with BSSE correction, entropies (-TΔS) and the values of BSSE correction (E_{BSSE}) (kJ/mol) for neutral CB[7]/AA, as well as volume (V, cm^3/mol) and dipole moment (μ, Debye) of each neutral AA at ωB97X-D/6-311++G** level

	Reference	Aliphatic					Polar			
	Gly	Ala	Leu	Val	Ile	Pro	Ser	Gln	Thr	Asn
ΔG^{CP}	-20.3	-21.6	-30.8	-40.2	-44.6	-46.5	-37.5	-44.0	-45.1	-75.2
ΔG	-35.6	-40.5	-50.5	-63.0	-66.9	-68.9	-58.7	-69.4	-70.8	-101.0
ΔH^{CP}	-79.9	-91.6	-105.3	-118.4	-114.1	-119.1	-102.4	-135.0	-120.3	-153.5
-TΔS	59.6	70.0	74.5	78.2	69.4	72.6	64.9	91.0	75.3	78.2
E_{BSSE}	15.3	18.9	19.7	22.8	22.3	22.4	21.2	25.4	25.7	25.8
μ	1.4	2.2	1.6	2.1	2.1	1.7	0.5	4.1	4.2	2.8
V	59.2	80.0	83.8	97.8	112.2	78.7	80.7	117.5	70.1	96.8

	S-included		Aromatic			Acidic		Basic		
	Cys	Met	Trp	Phe	Tyr	Glu	Asp	Lys	His	Arg
ΔG^{CP}	-27.0	-45.2	-36.8	-43.1	-46.5	-35.6	-37.7	-28.2	-69.0	-89.7
ΔG	-51.5	-69.2	-64.0	-67.3	-72.6	-67.5	-62.3	-49.4	-96.3	-119.3
ΔH^{CP}	-104.7	-117.0	-118.4	-122.4	-123.0	-123.7	-114.6	-110.5	-142.1	-173.9
-TΔS	77.7	71.8	81.5	79.3	76.5	88.0	76.9	82.3	73.1	84.2
E_{BSSE}	24.5	24.0	27.2	24.2	26.1	31.9	24.6	21.2	27.3	29.6
μ	0.5	3.1	2.5	1.8	1.6	7.1	5.7	2.2	4.2	8.2
V	85.4	135.5	164.1	140.1	138.6	113.4	80.4	131.7	126.0	125.8

the -NH_2 group inside the cavity of CB[7] by vdW interaction, while Gly^+ locates at one portal of CB[7] forming exclusion complexes by three H-bonds. In IGM analysis, the blue spike in scatter diagram (Figure 3c and Figure S1a) and the several blue disks in isosurface plot (Figure 3d and Figure S1b) confirm the H-bonds formation of CB[7] towards Gly and Gly^+. The protonation increases the number of H-bonds between CB[7] and Gly because of introducing the ion-dipole interaction between carbonyl group of CB[7] and -NH_3^+ group of Gly^+. Thus, the binding strength of CB[7]/Gly^+ (ΔG^{CP} = -261.2 kJ/mol, Table 2) is much larger than that of CB[7]/Gly (-20.3 kJ/mol, Table 1).

3.2.2 | Aliphatic AAs and AAs$^+$

The optimized structures of host-guest complexes of CB[7] towards aliphatic AAs (Ala, Leu, Val, Ile and Pro) and AAs^+ (Ala^+, Leu^+, Val^+, Ile^+ and Pro^+) are displayed in Figure 4a-b. We found that both aliphatic AAs and AAs^+ prefer to be completely encapsulated inside the CB[7] cavity. Comparing the orientation of the backbone and sidechain of each AA or AA^+ in CB[7] cavity, we found that the host-guest binding patterns of CB[7] with each aliphatic AA is quite similar to the corresponding aliphatic AA^+. The backbone of each aliphatic AA or AA^+ forms H-bonds with the portals of CB[7] and the sidechain is completely encapsulated inside the CB[7] cavity due to vdW interactions.

The binding free energy ΔG^{CP} of CB[7] with neutral aliphatic AAs ranges from -46.5 to -21.6 kJ/mol, where CB[7]/Pro shows highest binding strength (ΔG^{CP} = -46.5 kJ/mol, Table 1). After protonation, ion-dipole interaction is introduced. The binding strength between CB[7] and protonated aliphatic AAs^+ is largely increased with ΔG^{CP} in the range of -288.5 and -255.2 kJ/mol, and CB[7]/Leu^+ bears the largest binding

Table 2 The binding free energies (ΔG^{CP}, ΔG) with and without BSSE correction, enthalpies (ΔH^{CP}) with BSSE correction, entropies ($-T\Delta S$) and the values of BSSE correction (E_{BSSE}) (kJ/mol) for protonated CB[7]/AA^+, as well as volume (V, cm^3/mol) and dipole moment (μ, Debye) of each protonated AA^+, computed at ωB97X-D/6-311++G** level

	Reference	Aliphatic					Polar			
	Gly^+	Pro^+	Ala^+	Val^+	Ile^+	Leu^+	Thr^+	Ser^+	Gln_+	Asn^+
ΔG^{CP}	-261.2	-255.2	-258.3	-270.1	-283.1	-288.5	-246.6	-292.2	-319.8	-325.6
ΔG	-273.8	-278.6	-279.8	-296.6	-304.9	-310.8	-272.6	-316.3	-348.3	-350.5
ΔH^{CP}	-335.9	-332.1	-334.1	-353.1	-359.2	-367.7	-331.4	-375.5	-414.9	-405.6
$-T\Delta S$	74.7	76.9	75.8	83.0	76.2	79.2	84.8	83.3	95.1	80.0
E_{BSSE}	12.6	23.4	21.5	26.5	21.8	22.3	26.0	24.1	28.5	24.9
μ	5.8	3.5	5.0	5.2	5.8	6.5	3.8	6.2	6.2	6.4
V	43.0	80.4	71.5	115.9	120.1	81.8	79.3	78.5	106.4	93.1

	S-included		Aromatic			Acidic		Basic		
	Cys^+	Met^+	Trp^+	Phe^+	Tyr^+	Glu^+	Asp^+	His^+	Arg^+	Lys^+
ΔG^{CP}	-252.2	-292.8	-287.4	-297.1	-301.1	-258.5	-284.2	-238.0	-291.5	-294.9
ΔG	-273.8	-318.0	-318.2	-322.6	-329.3	-287.8	-310.3	-264.1	-320.3	-317.4
ΔH^{CP}	-326.9	-377.8	-372.5	-380.8	-384.6	-337.8	-366.6	-324.6	-390.2	-383.8
$-T\Delta S$	74.7	85.1	85.0	83.7	83.6	79.3	82.4	86.6	98.7	88.9
E_{BSSE}	21.6	25.2	30.8	25.5	28.2	29.3	26.1	26.1	28.8	22.5
μ	7.2	8.0	9.4	9.3	10.5	3.4	1.8	5.7	15.2	19.6
V	78.4	130.0	149.6	112.0	125.6	106.6	79.1	101.1	134.9	108.0

strength. Both the number and strength of H-bonds increases obviously after protonation. For example, CB[7]/Leu only has one H-bond with O···H distance 2.50 Å, while CB[7]/Leu^+ has three H-bonds, with O···H distances 1.81, 1.92 and 2.01 Å, respectively. This is also visualized by the large blue disk in isosurface diagram of IGM analysis in Figure S2. Dipole moment (μ) is an important factor that affect noncovalent interaction.[49] The influence of volume (V) and μ of each aliphatic AA to ΔG^{CP} is studied. For neutral aliphatic AAs, ΔG^{CP} decreases as V increases (except for Pro) and μ of five neutral aliphatic AAs are all about 2.0 Debye (see Figure 4c and Table 1); after protonation, μ is in range from 3.5 to 6.5 Debye and ΔG^{CP} decreases as μ increases monotonously (see Figure 4d and Table 2). The dipole moments of AAs^+ play a major role in host-guest interactions. Overall, the change of ΔG^{CP} are not so large, 24.9 kJ/mol and 33.3 kJ/mol for neutral and protonated aliphatic AAs, respectively.

3.2.3 | Polar AAs and AAs^+

The optimized structures of host-guest complexes of CB[7] towards polar AAs (Ser, Gln, Thr and Asn) and AAs^+ (Ser^+, Gln^+, Thr^+ and Asn^+) are shown in Figure 5a-b. Similar to aliphatic AAs and AAs^+, polar AAs and AAs^+ are also overall encapsulated into the CB[7] cavity. Different from the aliphatic ones, both backbone and sidechain of each polar AA or AA^+ form H-bonds with CB[7] portals, because both -OH group (Ser/Ser^+ or Thr/Thr^+) and $-CONH_2$ group (Gln/Gln^+ or Asn/Asn^+) could act as donors of H-bonds. As Table 1 shown, ΔG^{CP} of polar CB[7]/AA complexes are in-between -75.2 and -37.5 kJ/mol, while ΔG^{CP} of the corresponding protonated polar CB[7]/AA^+ complexes are in range from -325.6 to -246.6 kJ/mol (Table 2), therefore the binding strength of CB[7] with polar AA in both neutral and protonated states are larger than the corresponding aliphatic ones. After protonation, for CB[7]/Ser^+ and CB[7]/Thr^+, the number and strength of H-bonds increases obviously, because of introduction of ion-dipole interactions (see IGM analysis in Figure S3), and

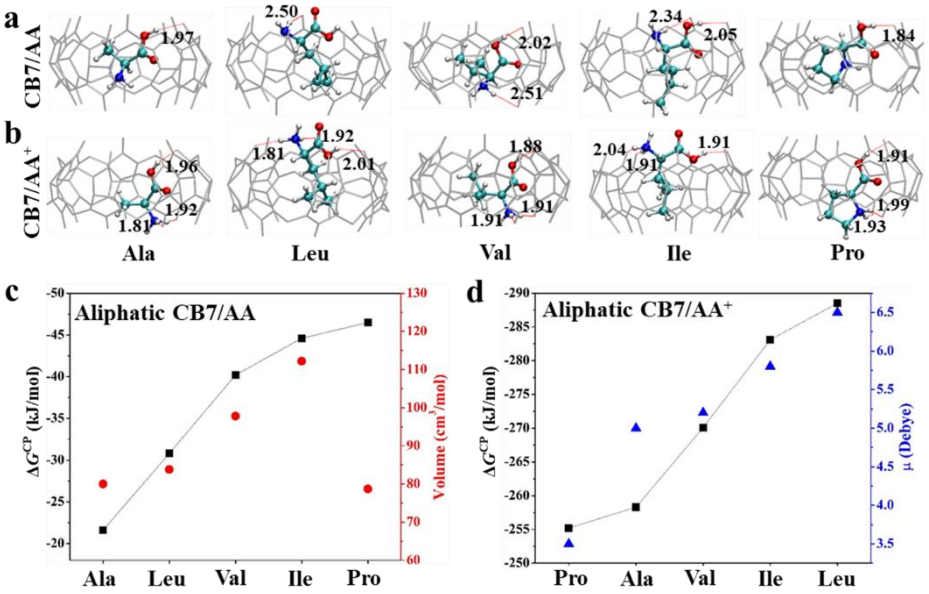


Figure 4 The optimized host-guest complex structures of CB[7] towards aliphatic (a) AAs and (b) AAs^+, respectively. (c) The correlation of ΔG^{CP} and V of each aliphatic AA. (d) The correlation of ΔG^{CP} and μ of each aliphatic AA^+.

for CB[7]/Asn^+ and CB[7]/Gln^+, compared with corresponding CB[7]/Asn and CB[7]/Gln, though the number of H-bonds is not increased, the strength of H-bonds increases obviously estimated by the decrease of O···H distances (1.83~2.05 Å vs 1.93~2.35 Å, 1.76~2.29 Å vs 1.90~2.39 Å, respectively), which significantly increase the binding strength of polar CB[7]/AAs^+.

The relationship between ΔG and V as well as μ are analyzed, which found the effect of μ to ΔG^{CP} is more obvious than V. As shown in Figure 5c-d, except for CB[7]/Asn, ΔG^{CP} decreases almost monotonously as μ increases, that is, the larger μ, the stronger binding strength between CB[7] and polar AA or polar AA^+. Overall, CB[7]/Asn and CB[7]/Asn^+ show the best binding strength

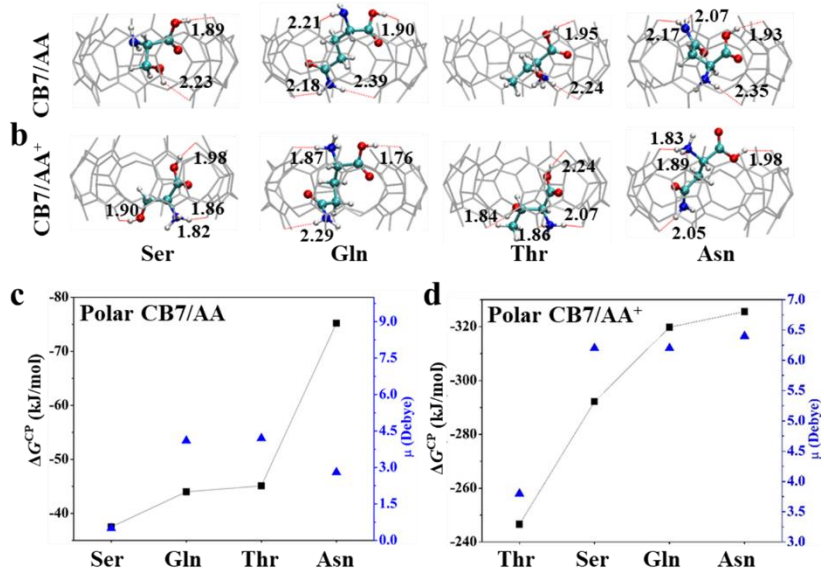


Figure 5 The optimized complexes for CB[7] binding with polar (a) AAs and (b) AAs^+, respectively. The correlation of ΔG^{CP} with (c) μ of single polar AA and (d) μ of single polar AA^+, respectively.

among polar host-guest complexes in both states, respectively. The change of ΔG^{CP} are 37.7 and 79.0 kJ/mol for neutral and protonated polar AAs, respectively.

3.2.4 | S-included AAs and AAs$^+$

Due to the special nature, the S-included amino acid is grouped and discussed separately, such as Cys could form disulfide bonds.[6, 50] The optimized structures of CB[7] with S-included Cys and Met in the neutral and protonated states are listed in Figure 6a-b. Similar to the binding pattern of aliphatic AA, the backbone of S-included AA or AA$^+$ forms H-bonds with the portals of CB[7] and the sidechain is completely included in CB[7]. Due to the ion-dipole interaction, the number and strength of H-bonds are both increase after protonation (see Figure 6a-b and IGM analysis in Figure S4a-b). As shown in Tables 1-2 and Figure S5, Met and Met$^+$ display the better binding strength with CB[7] than that of Cys and Cys$^+$, because of its larger V and μ.

3.2.5 | Aromatic AAs and AAs$^+$

The optimized structures of host-guest complexes of CB[7] towards aromatic AAs (Trp, Phe and Tyr) and AAs$^+$ (Trp$^+$, Phe$^+$ and Tyr$^+$) are shown in Figure 6c-d. It shows that the packing pattern for CB[7] with each neutral aromatic AA is similar to that of the protonated one. The backbone of aromatic AA/AA$^+$ forms H-bonds with one of portals of CB[7] and the aromatic sidechain is included inside the cavity by vdW interactions, except for Tyr (see Figure S4c-d). Due to the hydroxyl group in sidechain, both the backbone and sidechain of Tyr forms H-bonds with portals of CB[7]. The phenyl group of Phe could completely encapsulated in CB[7] cavity, consistent with the experiment results.[14, 22] Due to the large size of indole group, the sidechain of Trp could not be completely encapsulated inside CB[7] cavity, similar to the previous result.[22] Thus, ΔG^{CP} of CB[7] towards aromatic guests obey the order: Tyr < Phe < Trp in both neutral and protonation states, that is, CB[7]/Tyr (-46.5 kJ/mol) and CB[7]/Tyr$^+$ (-301.1 kJ/mol) display the largest binding strength, respectively. In addition, the difference of ΔG^{CP} for aromatic host-guest complexes are very small (see Tables 1-2).

3.2.6 | Acidic AAs and AAs$^+$

Figure 7a-b shows the optimized structures of host-guest complexes of CB[7] towards acidic AAs (Glu and Asp) and AAs$^+$ (Glu$^+$ and Asp$^+$), respectively. We found that Asp and Glu form a six-member and seven-member ring-like structure inside the CB[7] cavity by intramolecular H-bond, due to the presence of carboxyl in sidechain (Figure 7a). Although the larger V and μ of Glu than Asp, ΔG^{CP} of CB[7]/Glu (-35.6 kJ/mol) and CB[7]/Asp (-37.7 kJ/mol)are similar (see Table 1), because the large seven-member ring-like of Glu would cause CB[7] undergo large deformation (see Figure S7a), evidenced by the large deformation energy (64.9 kJ/mol, see Table S1) and large Pauli repulsion (190.9 kJ/mol) of CB[7]/Glu, which largely

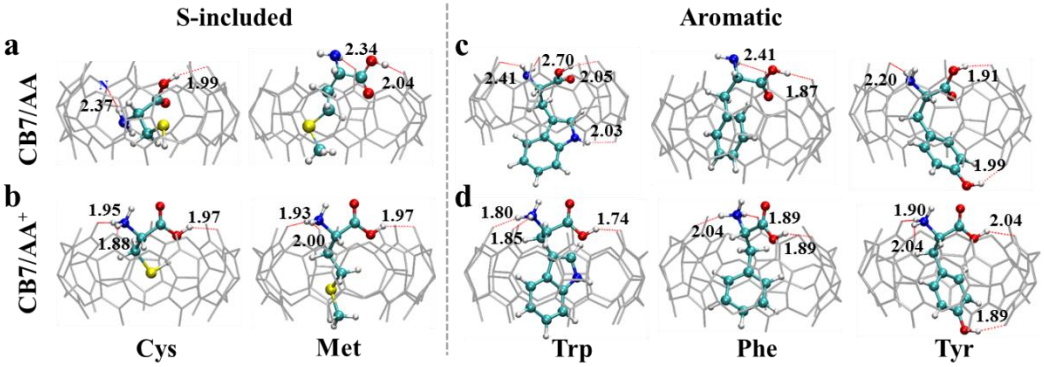


Figure 6 The most stable structures for CB[7] binding with (a,b) S-included (S is shown in yellow) and (c,d) aromatic amino acids in both neutral and protonated states.

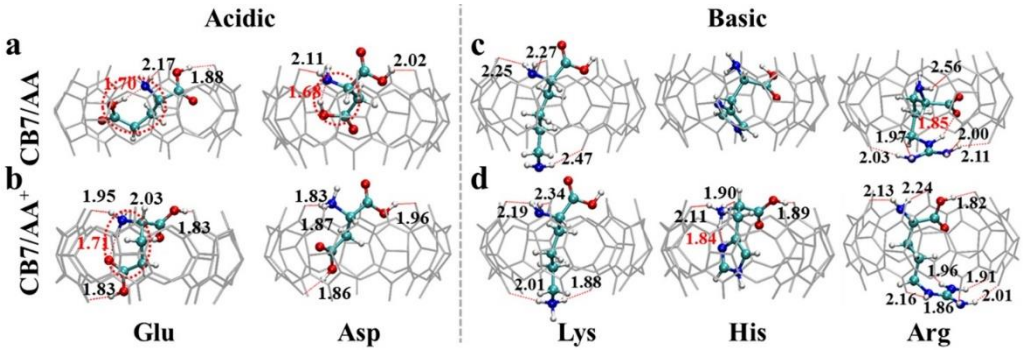


Figure 7 The most stable structures of CB[7] binding with (a,b) acidic and (c,d) basic amino acids. The red number refers to intramolecular H-bond.

counteract the attraction interaction, see Table S3. After protonation, ΔG^{CP} of CB[7]/Asp^+ is more negative than CB[7]/Glu^+. As shown in Figure 7b, the same as CB[7]/Glu, the large size ring-like structure inside CB[7] cavity reduces the strength of intermolecular H-bond and leads to stronger repulsion interaction (194.6 kJ/mol).

3.2.7 | Basic AAs and AAs^+

The optimized structures of host-guest complexes of CB[7] towards basic AAs (Lys, His and Arg) and AAs^+ (Lys^+, His^+ and Arg^+) are shown in Figure 7c-d, respectively. We found that the binding patterns for CB[7] towards three neutral basic AAs are much different. In CB[7]/Lys, both the backbone and sidechain of Lys form H-bonds with portals of CB[7]. While in CB[7]/His, there is not H-bond interaction and His is completely encapsulated in CB[7] cavity via vdW interaction proved by IGM isosurface in Figure S6c (There is no blue disk areas in IGM isosurface). In CB[7]/Arg, due to the strong proton affinity of guanidyl in sidechain, the proton of -COOH group in neutral Arg transfers to guanidyl and the intramolecular H-bond forms simultaneously on Arg inside CB[7] cavity (see Figure S8a), which leads to large deformation of CB[7]/Arg (ΔE_{deform} = 98.4 kJ/mol, Table S1). Therefore, host-guest interaction influences the proton affinity of the functional groups in Arg. In addition, the protonated guanidyl could form four H-bonds with one portal of CB[7], which largely stabilizes CB[7]/Arg complex and makes it has the largest binding strength in neutral state (see Table 1), which is in agreement with the result of CB[7] binding to uftsin.[15] Although three more H-bonds of CB[7]/Lys than CB[7]/His, the O···H distances are all very large (2.25~2.47 Å, Figure 7c), that is, these H-bonds are weak. In addition, the vdW interactions in CB[7]/Lys are also weak. The above are proved by the lighter blue disk area and small green disk area in IGM isosurface of CB[7]/Lys in Figure S6c. As Table 1 shown, ΔG^{CP} of CB[7] with neutral basic AA are in-between -

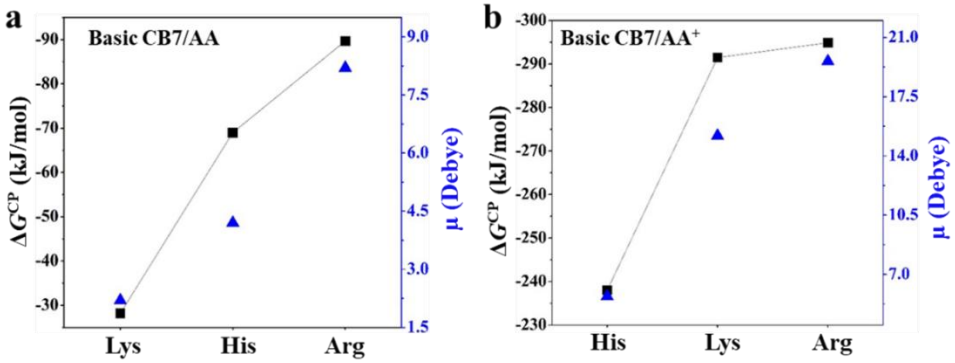


Figure 8 (a) The correlation of ΔG^{CP} and μ of each basic AA. (b) The correlation of ΔG^{CP} and μ of each basic AAs^+.

89.7 and -28.2 kJ/mol and are in the order: CB[7]/Arg < CB[7]/His < CB[7]/Lys. What's more, the binding strength between CB[7] and basic AAs enhances with μ of single basic AA (Figure 8a). In addition, the basic complexes CB[7]/Arg and CB[7]/His are more stable than aromatic complexes, because of the more negative ΔG^{CP}, consistent with the experiment results.[22]

For CB[7]/Lys$^+$, the binding pattern is almost the same as CB[7]/Lys, both backbone and sidechain form H-bonds with portals of CB[7]. For CB[7]/His$^+$, only the backbone of His$^+$ forms H-bonds with one of portals of CB[7]. It is worth to point out that the proton in the sidechain of His$^+$ transfers to the -NH_2 group of backbone and forms intramolecular H-bond after forming a complex (shown in Figure S8b). Therefore, the encapsulation of CB[7] changes the proton affinity of the functional groups in His$^+$. In CB[7]/Arg$^+$, the protonated guanidyl group could form five H-bonds with one portal of CB[7] and the backbone of Arg$^+$ forms three additional H-bonds with the other portal of CB[7]. The number and strength of H-bonds largely increase due to the introduction of ion-dipole interactions, therefore the host-guest binding strength largely increases by 266.7, 169.0 and 201.8 kJ/mol respectively between CB[7] and protonated Lys, His and Arg (see Tables 1-2). As Table 2 shown, ΔG^{CP} of CB[7] with protonated basic AAs are in the order: CB[7]/Lys$^+$ < CB[7]/Arg$^+$ < CB[7]/His$^+$. Relative to the neural case, the reverse sequence of ΔG for Lys$^+$ and His$^+$ may be related to the smaller enhancement of μ of single His (1.5 Debye) than single Lys (17.4 Debye) after protonation (see Tables 1-2, Figure 8).

3.3 | Thermodynamic properties

The contributions of entropy (-TΔS) vs enthalpy (ΔH^{CP}) of CB[7] towards AA/AA$^+$ with BSSE correction are shown in Figure 9. Six classes of AA/AA$^+$ are displayed in different colors, with aliphatic, polar, S-included, aromatic, acidic and basic, in blue, green, grey, red, yellow and purple, respectively. Here the binding enthalpy ΔH^{CP} as *x*-axis locates in range of -480 ~ -50 kJ/mol and -TΔS as *y*-axis sits in-between 55 ~ 105 kJ/mol. According to equation: $\Delta G^{CP} = \Delta H^{CP} - T\Delta S$, the host-guest complexes binding free energy ΔG^{CP} is mainly contributed by enthalpy ΔH^{CP}. It means that the molecular recognition of CB[7] towards AAs is an enthalpy-driven process (Figure S9-S10). As references, Gly and Gly$^+$ have the lowest binding strength on CB[7] in neutral and protonated states respectively among the 20 AAs and 20 AAs$^+$, except for His$^+$, Thr$^+$,Cys$^+$ and Pro$^+$ (Tables 1-2). To guide the eye, we drew two black dashed lines (with slope value = -1) that pass through CB[7]/Gly and CB[7]/Gly$^+$ as references. The host-guest complexes that siting further from the reference lines in the negative direction of ΔH^{CP} have the lower ΔG^{CP} and are more stable. As shown in Tables 1-2, the E_{BSSE} is in the range of 15.3 ~ 31.9 kJ/mol for CB[7]/AAs complexes in the neutral state, while E_{BSSE} for the CB[7]/AAs$^+$ complexes in the protonated state is in the range of 12.6 ~ 30.8 kJ/mol. The absolute values of the binding free energy and enthalpy are largely changed after BSSE correction, especially for the neutral AAs. Except for the reversed trend for CB[7]/Glu and CB[7]/Asp in the neutral state, as well as CB[7]/Arg$^+$ and CB[7]/Lys$^+$ in the protonated states, the relative trend of the binding strength for different AAs or AAs$^+$ are the same after BSSE correction.

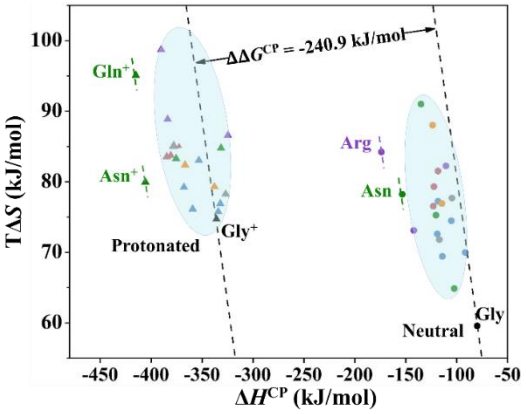


Figure 9 Plot of entropy (-TΔS) vs enthalpy (ΔH^{CP}) contributions to ΔG^{CP} for neutral CB7/AA and protonated CB7/AA$^+$ host-guest complexes. Six classes of AA are distinguished by different colors, with aliphatic, polar, S-included, aromatic, acidic and basic, in blue, green, gray, red, yellow and purple, respectively, for both states. The slope value of dashed lines is -1. Gly and Gly$^+$ are selected as references, with $\Delta\Delta G^{CP} = \Delta G_{Gly^+} - \Delta G_{Gly}$.

In neutral state, ΔG^{CP} of most CB[7]/AA are similar (marked by ellipse in cyan). Arg and Asn

show better binding strength with CB[7] (purple and green short dashed lines) than the others. CB[7] displays the best recognition ability to Arg in neutral state. After protonation, ΔG^{CP} greatly shifted to the negative direction by 240.9 kJ/mol (marked by $\Delta\Delta G^{CP}$), referred to Gly system. We found that the protonation significantly decreases ΔH^{CP} values and its influence to -TΔS is negligible. Most CB[7]/AAs$^+$ display similar ΔG^{CP} (cyan ellipse). Among all protonated host-guest complexes, the polar Gln$^+$ and Asn$^+$ shows the largest binding strength with CB[7].

3.4 | Energy decomposition analysis (EDA)

Binding energy (ΔE) is the summation of deformation energy (ΔE_{deform}) and interaction energy (ΔE_{int}). In the complexes under study, ΔE is dominated by ΔE_{int}. As shown in Tables S1-2, ΔE_{deform} of CB[7]/AA$^+$ is higher than its neutral counterpart. This suggests that for protonated cases, CB[7] and each AA$^+$ need to tune their configurations more significantly to form stable host-guest complex. ΔE_{int} of each protonated CB[7]/AA$^+$ is significantly more negative than ΔE_{int} of neutral CB[7]/AA (Table S3-4), suggesting a stronger interaction for protonated complex than neutral one. This trend of ΔE_{int} is in line with ΔE and ΔG, so we can use EDA of ΔE_{int} to understand the intrinsic reason of the trend of ΔE and ΔG.

We performed EDA to quantify the contributions of different components of ΔE_{int}.[49, 51] Values of the decomposition terms, i.e. ΔV_{elstat}, ΔE_{Pauli}, ΔE_{orb} and ΔE_{disp}, are plotted in Figure 10a,b for twenty CB[7]/AA and CB[7]/AA$^+$ complexes, respectively (details are in Table S3-4). Three of the decomposition terms, ΔV_{elstat}, ΔE_{orb} and ΔE_{disp} are attractive interactions, and ΔE_{Pauli} is the repulsive interaction. For neutral CB[7]/AA complex, the average contributions of ΔV_{elstat}, ΔE_{orb}, and ΔE_{disp} are 50%, 20%, and 30% for the attractive interactions, respectively. For protonated CB[7]/AA$^+$ complexes, the respective percentages of ΔV_{elstat}, ΔE_{orb} and ΔE_{disp} are 60%, 25%, and 15%. According to EDA results, we found that the stronger interactions of protonated complex mainly arise from much stronger electrostatic interactions, reflecting by more number and stronger H-bonds in protonated CB[7]/AA$^+$ than its neutral counterpart. On the one hand, the average ΔV_{elstat} values of protonated complexes are more negative by about 250 kJ/mol than the values of neutral complexes (Table S3-4). On the other hand, the contribution of ΔV_{elstat} in attractive interactions is greater for protonated complex (60%) than for neutral complex (50%). Thirdly, if we sum up ΔV_{elstat} and ΔE_{Pauli}, the sum vanishes for neutral complexes and remains negative for protonated complex. In brief, we conclude that the electrostatic interactions of protonated

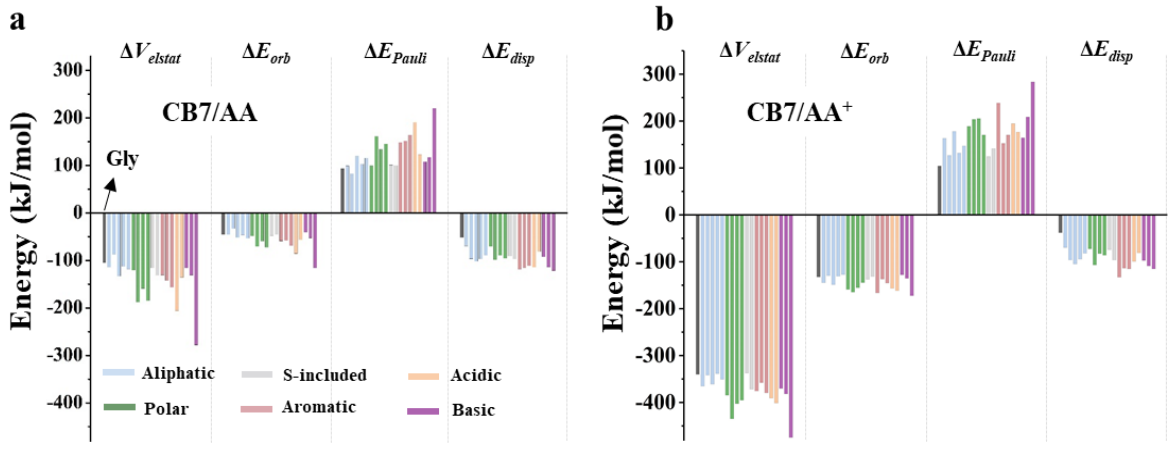


Figure 10 EDA analyses of (a) CB[7]/AAs and (b) CB[7]/AAs$^+$ at CAM-B3LYP/TZP level.

complexes lead to their stronger binding strength.

4 | Conclusions

To conclude, we have systematically investigated the binding geometries and binding free energies of host CB[7] to 20 amino acids (AAs) in both neutral and protonated states. To unravel the binding nature of CB[7] to AA with different kinds of sidechains, AAs are divided into six types, including aliphatic, polar, S-included, aromatic, acidic and basic AAs. The calculated geometric structures and trends of host-guest binding strength are in good agreement with the experimental results.[22] In addition, the effect of volume and dipole moment to binding strength of CB7/AA and CB7/AA^+ is discussed. IGM and EDA have been employed to interpret the nature of noncovalent interactions and their relative contributions to the total interaction energy between host and guest system in qualitative and quantitative.

The optimized structures show that most of stable structures of CB[7]/AA and CB[7]/AA^+ are inclusion complexes with AA or AA^+ encapsulated in the cavity of CB[7]. The molecular recognition of CB[7] to AA or AA^+ is mainly enthalpy-driven process. For the neutral CB[7]/AAs, the H-bond and vdW interaction are concurrent and both are play important roles in stabilizing the host-guest complexes. For aliphatic, S-included, aromatic and acidic AAs, the H-bonds are mainly contributed by interactions between the portal of CB[7] and the backbone of AAs, while for polar and basic AAs, both backbone and sidechains of AAs take part in the H-bonds formation. After protonation, because of the introduction of ion-dipole interaction, the number and strength of H-bonds increase, which largely increases the binding strength of CB[7]/AAs^+ complexes. Energy decomposition analysis indicated that the electrostatic interaction ΔV_{elstat} is dominated among the attractive interactions, and greatly increases the host-guest interaction energy, and also causes the more favorable binding of protonated complexes than neutral complexes.

We also evaluate the correlation between V and μ with ΔG^{CP} of host-guest complexes. For neutral aliphatic AAs, ΔG^{CP} decreases as V of single AA increases; while for aliphatic AAs^+, polar, S-included and basic AAs/AAs^+, ΔG^{CP} decreases as μ of single AA increases.

In neutral state, the basic-type CB[7]/Arg shows the strongest binding strength, and the polar-type CB[7]/Asn is the secondary stable host-guest complex. In protonated state, both the polar-type CB[7]/Gln^+ and CB[7]/Asn^+ shows the largest binding strength with CB[7]. Asn in both neutral and protonated states show strong binding strength with CB[7].

Finally, theoretical quantification of the amino acids recognition properties of CB[7] has helped us to systematically unravel the binding patterns of different kinds of amino acids and the respective contributions of concurrent noncovalent interactions in each host-guest complex. This research deepens our knowledge of molecular recognition of synthetic host molecule CB[7] to the basic units of protein and it provides useful clues of predicting the recognition sites for sequence-based peptide or protein by synthetic host molecule. The molecular recognition of CB[7] towards dipeptide, tripeptide and small proteins are in progress and will be reported in our future work.

Acknowledgments

We thank the financially supporting of the National Natural Science Foundation of China (Grants 21803007, 21473010 , 91544223), Beijing Institute of Technology (BIT) Research Fund Program for Young Scholars, and BIT Teli Young Fellow Program (No. 3100011181905). J. Xie thanks the Minnesota Supercomputing Institute for computing resources of ADF calculations.

Keywords: amino acids, cucurbit[7]uril, host-guest, noncovalent interaction, Density Functional Theory

Additional Supporting Information may be found in the online version of this article.

References and Notes

1. M. Jewginski, T. Granier, B. Langlois d'Estaintot, L. Fischer, C. D. Mackereth and I. Huc, *J. Am. Chem. Soc.* **2017**, 139, 2928-2931.
2. W. Liu, S. K. Samanta, B. D. Smith and L. Isaacs, *Chem. Soc. Rev.* **2017**, 46, 2391-2403.
3. Q. Luo, C. Hou, Y. Bai, R. Wang and J. Liu, *Chem. Rev.* **2016**, 116, 13571-13632.
4. J. Zhou, G. Yu and F. Huang, *Chem. Soc. Rev.* **2017**, 46, 7021-7053.
5. M. A. Gamal-Eldin and D. H. Macartney, *Org. Biomol. Chem.* **2013**, 11, 488-495.
6. Z. Hirani, H. F. Taylor, E. F. Babcock, A. T. Bockus, C. D. Varnado, C. W. Bielawski and A. R. Urbach, *J. Am. Chem. Soc.* **2018**, 140, 12263-12269.
7. J. M. Chinai, A. B. Taylor, L. M. Ryno, N. D. Hargreaves, C. A. Morris, P. J. Hart and A. R. Urbach, *J. Am. Chem. Soc.* **2011**, 133, 8810-8813.
8. S. van Dun, C. Ottmann, L.-G. Milroy and L. Brunsveld, *J. Am. Chem. Soc.* **2017**, 139, 13960-13968.
9. H. Shang, A. Zhou, J. Jiang, Y. Liu, J. Xie, S. Li, Y. Chen, X. Zhu, H. Tan and J. Li, *Acta Biomater.* **2018**, 78, 178-188.
10. S. Sonzini, A. Marcozzi, R. J. Gubeli, C. F. van der Walle, P. Ravn, A. Herrmann and O. A. Scherman, *Angew. Chem. Int. Ed.* **2016**, 55, 14000-14004.
11. W. Li, A. T. Bockus, B. Vinciguerra, L. Isaacs and A. R. Urbach, *Chem. Commun.* **2016**, 52, 8537-8540.
12. F. Guagnini, P. M. Antonik, M. L. Rennie, P. O'Byrne, A. R. Khan, R. Pinalli, E. Dalcanale and P. B. Crowley, *Angew. Chem. Int. Ed*. **2018**, 57, 7126-7130.
13. X. Zhou, X. Su, P. Pathak, R. Vik, B. Vinciguerra, L. Isaacs and J. Jayawickramarajah, *J. Am. Chem. Soc.* **2017**, 139, 13916-13921.
14. L. A. Logsdon, C. L. Schardon, V. Ramalingam, S. K. Kwee and A. R. Urbach, *J. Am. Chem. Soc.* **2011**, 133, 17087-17092.
15. E. A. Kovalenko, E. A. Pashkina, L. Y. Kanazhevskaya, A. N. Masliy and V. A. Kozlov, *Int. Immunopharmacol.* **2017**, 47, 199-205.
16. E. Cavatorta, P. Jonkheijm and J. Huskens, *Chem. Eur. J.* **2017**, 23, 4046-4050.
17. J. Kim, I.-S. Jung, S.-Y. Kim, E. Lee, J.-K. Kang, S. Sakamoto, K. Yamaguchi and K. Kim, *J. Am. Chem. Soc.* **2000**, 122, 540-541.
18. S. J. Barrow, S. Kasera, M. J. Rowland, J. del Barrio and O. A. Scherman, *Chem. Rev.* **2015**, 115, 12320-12406.
19. J. K. K. Dinesh Shetty, Kyeng Min Parkad and Kimoon Kim, *Chem. Soc. Rev.* **2015**, 44, 8747-8761.
20. H. Cong, L.-L. Tao, Y.-H. Yu, S.-F. Xue and Z. Tao, *Acta Chim. sin.*, **2006**, 64, 989-996.
21. Z.-Z. Gao, J.-L. Kan, L.-X. Chen, D. Bai, H.-Y. Wang, Z. Tao and X. Xiao, *ACS Omega* **2017**, 2, 5633-5640.
22. J. W. Lee, H. H. L. Lee, Y. H. Ko, K. Kim and H. I. Kim, *J. Phys. Chem. B* **2015**, 119, 4628-4636.
23. E. Kovalenko, M. Vilaseca, M. Díaz-Lobo, A. N. Masliy, C. Vicent and V. P. Fedin, *J. Am. Soc. Mass Spectrom.* **2016**, 27, 265-276.
24. S. Gronert, D. C. Simpson and K. M. Conner, *J. Am. Soc. Mass Spectr.* **2009**, 20, 2116-2123.
25. J. L. MacCallum, W. F. Bennett and D. P. Tieleman, *Biophys. J.* **2008**, 94, 3393-3404.
26. C. Lefebvre, G. Rubez, H. Khartabil, J. C. Boisson, J. Contreras-Garcia and E. Henon, *Phys. Chem. Chem. Phys.* **2017**, 19, 17928-17936.

27. L. Zhao, M. von Hopffgarten, D. M. Andrada and G. Frenking, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2018**, 8, e1345.
28. G. W. T. M.J. Frisch, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, G., V. B. Scalmani, B. Mennucci, G.A. Petersson, H. Nakatsuji, M. Caricato, X. Li,, A. F. I. H.P. Hratchian, J. Bloino, G. Zheng, J.L. Sonnenberg, M. Hada, M., K. T. Ehara, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O., H. N. Kitao, T. Vreven, J.A. Montgomery Jr., J.E. Peralta, F. Ogliaro, M. Bearpark,, E. B. J.J. Heyd, K.N. Kudin, V.N. Star-overov, T. Keith, R. Kobayashi, J., K. R. Normand, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi,, J. M. M. N. Rega, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J., R. G. Jaramillo, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C., J. W. O. Pomelli, R.L. Martin, K. Morokuma, V.G. Zakrzewski, G.A. Voth, P., J. J. D. Salvador, S. Dapprich, A.D. Daniels, O. Farkas, J.B. Foresman, J.V. and J. C. Ortiz, D.J. Fox, Gaussian G09, Revision D.01, Gaussian, Inc., Wallingford, CT, 2013.
29. S. Grimme, J. Antony, S. Ehrlich and H. Krieg, J. Chem. Phys., 2010, 132, 154104.
30. T. Yanai, D. P. Tew and N. C. Handy, *Chem. Phys. Lett*. **2004**, 393, 51-57.
31. Y. Z. D. G. Truhlar, *Theor. Chem. Acc.* **2008**, 120, 215-241.
32. N. E. S. Yan Zhao, and Donald G. Truhlar, *J. Chem. Theory Comput*. **2006**, 2, 364-382.
33. M. Ernzerhof and G. E. Scuseria, *J. Chem. Phys.* **1999**, 110, 5029-5036.
34. S. Grimme, S. Ehrlich and L. Goerigk, *J. Comput. Chem.* **2011**, 32, 1456-1465.
35. J. D. Chai and M. Head-Gordon, *Phys. Chem. Chem. Phys*. **2008**, 10, 6615-6620.
36. J. A. P. Martin Head-Gordon, *Chem. Phys. Lett.* **1988**, 153, 503-506.
37. Y. P. Yurenko, J. Novotny, V. Sklenar and R. Marek, *Phys. Chem. Chem. Phys.* **2014**, 16, 2072-2084.
38. P. Jurecka, J. Sponer, J. Cerny and P. Hobza, *Phys. Chem. Chem. Phys.* **2006**, 8, 1985-1993.
39. H. Lambert, N. Mohan and T.-C. Lee, *Phys. Chem. Chem. Phys.* **2019**, 21, 14521-14529.
40. S. Ling, W. Yu, Z. Huang, Z. Lin, M. Harañczyk and M. Gutowski, *J. Phys. Chem. A* **2006**, 110, 12282-12291.
41. J. Zhang and M. Dolg, *Phys. Chem. Chem. Phys.* **2015**, 17, 24173-24181.
42. S.F. Boys and F. Bernardi, *Mol. Phys.* **1970**, 19, 553-566.
43. S. S. Xantheas, *J. Chem. Phys.* **1996**, 104, 8821-8824.
44. T. Lu and F. Chen, *J. Comput. Chem.* **2012**, 33, 580-592.
45. G. te Velde, F.M. Bickelhaupt, E.J. Baerends, C. Fonseca Guerra, S.J.A. van Gisbergen, J.G. Snijders, T. Ziegler, *J. Comput. Chem.* **2001**, 22, 931-967.
46. E. J. Baerends et al., *Theoretical Chemistry.* (Vrije Universiteit, Amsterdam, The Netherlands, 2018).
47. W. Humphrey, A. Dalke; K. Schulten, *J. Mol. Graph.* **1996**, 14, 33-38.
48. K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov and A. D. Mackerell Jr., *J. Comput. Chem.* **2010**, 31, 671-690.
49. S. C. C. van der Lubbe, F. Zaccaria, X. Sun and C. F. Guerra, *J. Am. Chem. Soc.* **2019**, 141, 4878-4885.
50. C. B. Anfinsen, *Science* **1973**, 181, 223-230.
51. N. S. Venkataramanan, A. Suvitha and Y. Kawazoe, *J. Mol. Liq.* **2018**, 260, 18-29.

GRAPHICAL ABSTRACT

Fenfen Ma,[1] Xiaoyan Zheng,[1]* Jing Xie,[1,2] and Zesheng Li[1]*

Binding Properties of Cucurbit[7]uril to Neutral and Protonated Amino Acids: A Computational Study

The systematic study of the binding nature of CB[7] towards 20 amino acids in both neutral (AAs) and protonated (AAs^+) states provides a complete spectra of the relative binding strength of all AAs and AAs^+, which provides valuable clues in predicting the recognition sites for sequence-based peptide or protein by CB[7].