Supporting Information for:

Electronic Structures of Intermolecular Hydrogen Bond Contacts on Solute in Aqueous Solution: Glycine as a Working Prototype

Lingbiao Meng,[a,b,*] Weidong Wu[a,c] and Zijing Lin,[b,*]

[a]Research Center of Laser Fusion, China Academy of Engineering Physics, Mianyang 621900, China
[b]Hefei National Laboratory for Physical Sciences at Microscale and Department of Physics, University of Science and Technology of China, Hefei 230026, China
[c]IFSA Collaborative Innovation Center, Shanghai Jiao Tong University, Shanghai 200240, China

Table of Contents

I. Computational details

The geometrical and vibrational features of all modeling clusters were investigated by the density functional theory (DFT) calculations. The hybrid functional of BHandHLY[1] is used since the functional has been shown to be accurately describing the amino acid systems that are rich in H-bonds.[2] The 6-31++G(d,p) basis set is used in most cases, but the 6-31++G(d) basis set may also be selected for better comparability with the literature studies.[3–5] The BHandHLYP vibrational frequencies are scaled by a uniform factor of 0.92 in order to account for the anharmonicity effect and be comparable with the experimental data, as suggested in the literatures.[6–8]

To test the basis set dependence of the computational results, two additional basis sets, 6-311++G* and 6-311++G** have been used. The differences in the structural parameters and relative energies obtained by different basis sets are small. For example, corresponding to Figure 4 in the main text, Figure S1 illustrates the uncertainty for the computed effective energies and potentials. The bars for the numerical uncertainty refer to the difference between the 6-31++G* results and the averages of the three basis sets of 6-31++G**, 6-311++G* and 6-311++G**. It is clear from Figure S1 that the numerical uncertainty is negligible for the energy scale of interest. The numerical uncertainties for the interested IR modes are also small, but not negligible. They are indicated in Figure 3 of the main text.

All calculations reported here were performed using the Gaussian 03/09 suite of programs.[9,10]

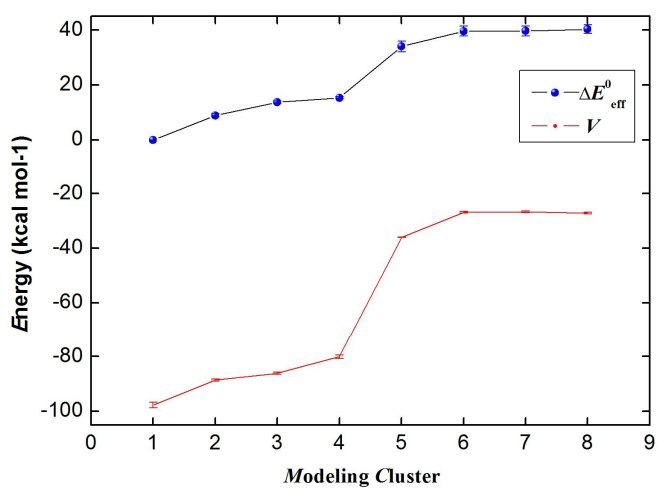


FIG. S1. The basis set dependence of the relative effective energies and potentials of modeling clusters. Fig. S1 corresponds to Figure 4 of the main text. More description may be found in the main manuscript.

II. Tables : Structures and IR frequencies of relevant modeling clusters

Table Overview:

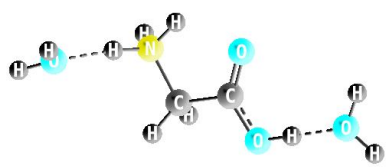
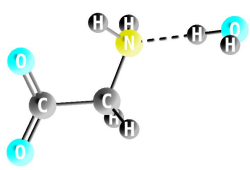
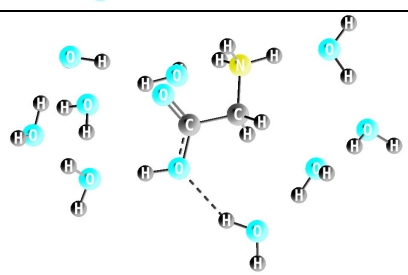
Table S1 is used to support the H-bond models of Π_1(H), Π_3(N), and Π_2(O) discussed in the main text.

Table S2 supports the detailed discussion on Π_2(O) of Figure 3 in the main text.

Table S3 and S4 are the basis for the discussion on *z*Gly clusters in the main text.

Table S5 provides the testing results for phenylalanine and alanine in aqueous solution.

Table S1: Structures and characteristic IR frequencies of representative modeling clusters, *d*GlyW$_n$, *p*GlyW$_n$. Frequencies (cm^{-1}) are calculated at the BHandHLYP/6-31++G(d,p) level. $\nu_{ad}(NH_3^+)$ and $\nu_{sd}(NH_3^+)$ denote the NH_3^+ asymmetric and symmetric deformations, $\nu_{sc}(NH_2)$ denotes the NH_2 scissor, and ν(C=O) and ν(C-O(H)) denote the C=O and C-O(H) stretching vibrations, respectively. The experiment data are from Refs [11,12]. Note that due to the sensitivity of the N-H vibrations to the H-bond arrangement, the most satisfying agreement between the theory and experiment is found in clusters with a second hydration shell.

***G*eometries**	***P*arameters (Degree)**	***F*requencies**
	*p*GlyW$_2$ → Π_1(H): ∠(O–H···O$_W$)=179 ∠(N–H···O$_W$)=173	
	*d*GlyW$_1$ → Π_3(N): ∠(C–N···H$_W$)=116 ∠(H–N···H$_W$)=109,114	
	*p*GlyW$_9$→ Π_2(O): ∠(C–O···H$_W$)=124 ∠(H–O···H$_W$)=123 Φ[H–O–C···H$_W$]=165	ν(C-O(H))$_{calc}$=1257, $_{exp.}$ ~1258

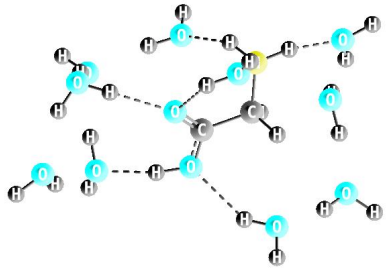
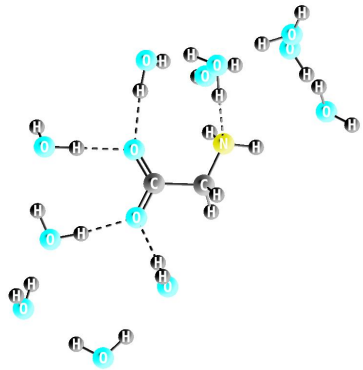
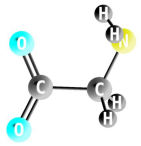
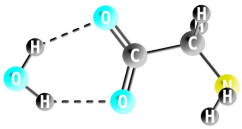
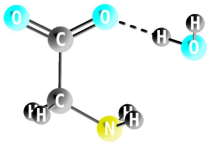
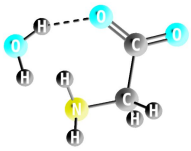
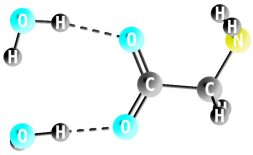
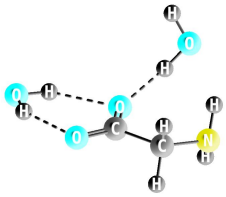
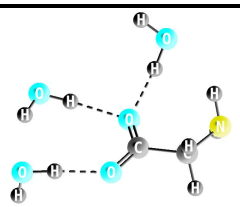
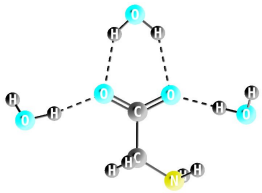
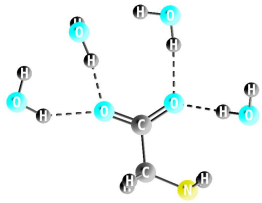
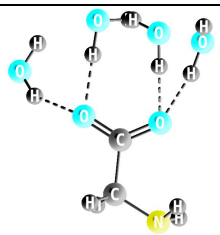
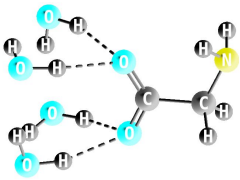
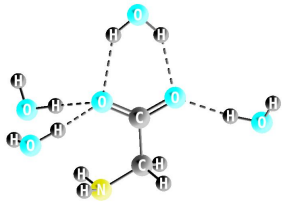
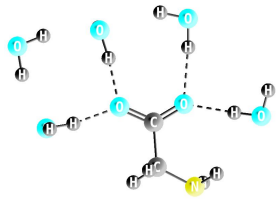
	*p*GlyW$_{10}$ → Π_1(H)Π_2(O) ∠(N–H$_3$···O$_W$)=167,163,151 ∠(O–H···O$_W$)=162 ∠(C=O···H$_W$)=120,154 Φ[H$_W$···O═C···H$_W$]=175 ∠(C–(H)O···H$_W$)=114 Φ[H–O–C···H$_W$]=145	$\nu_{ad}(NH_3^+)_{calc}$=1617/1631, $_{exp.}$ ~ 1615 $\nu_{sd}(NH_3^+)_{calc}$=1511, $_{exp.}$ ~ 1512/1521 ν(C=O)$_{calc}$=1745, $_{exp.}$ ~1740 ν(C-O(H))$_{calc}$=1245, $_{exp.}$ ~1258
	*d*GlyW$_{11}$ → Π_1(H)Π_3(N)Π_2(O) ∠(N–H$_2$···O$_W$)=149,146 ∠(C–N···H$_W$)=121 ∠(C=O···H$_W$)=119,124,132,144 Φ[H$_W$···O═C···H$_W$]= 172,158	$\nu_{sc}(NH_2)_{calc}$=~1588, $_{exp.}$ ~1590-1600 $\nu_{asym}(CO_2^-)_{calc}$=1567, $_{exp.}$ ~1561 $\nu_{sym}(CO_2^-)_{calc}$=1409, $_{exp.}$ ~1404

Table S2: Structures and characteristic IR frequencies of representative modeling clusters, *d*GlyW$_n$, for deprotonated (anionic) Gly. Frequencies (cm^{-1}) are calculated at the BHandHLYP/6-31++G(d,p) level. $\nu_{as}(CO_2^-)$ and $\nu_{ss}(CO_2^-)$ denote the CO_2^- asymmetric and symmetric stretching vibrations, respectively. Band gaps in the Table refers the differences in $\nu_{as}(CO_2^-)$ and $\nu_{ss}(CO_2^-)$; The structure of *d*GlyW$_{4d}$ is from Ref [13]; The experiment data are taken from Ref [11].

***C*lusters**	***G*eometries**	***F*requencies**		
		$\nu_{as}(CO_2^-)$	$\nu_{ss}(CO_2^-)$	Band gap
*d*GlyW$_0$		1619	1351	268
*d*GlyW$_{1a}$		1600	1369	231
*d*GlyW$_{1b}$		1616	1355	260
*d*GlyW$_{1c}$		1605	1349	257
*d*GlyW$_{2a}$		1599	1383	216
*d*GlyW$_{2b}$		1583	1370	213

*d*GlyW$_{3a}$		1385	1386	200
*d*GlyW$_{3b}$		1576	1386	190
*d*GlyW$_{4a}$		1578	1392	186
*d*GlyW$_{4b}$		1574	1393	181
*d*GlyW$_{4c}$		1574	1387	187
*d*GlyW$_{4d}$		1575	1381	194
*d*GlyW$_{5}$		1579	1409	170

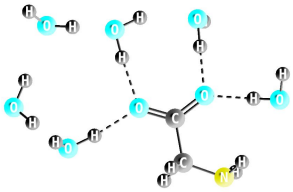
*d*GlyW$_6$		1573	1410	163
Exp.		1561	1404	157

Table S3: IR frequencies (freq) and intensities (inten) for *z*GlyW$_8$ of Fig. 2 of the main text. Frequencies (cm^{-1}) are calculated at the BHandHLYP/6-31++G(d) level. Only frequencies in the range of 1300–1700 cm^{-1} that are relevant to the CO_2^- vibrations are shown here. The experiment data are from refs [3,11].

Theory			***Exp***.
freq	inten	Assignment	
1667	56	NH_3 ad. + H_2O bend + CO_2 as.	
1660	43	NH_3 ad. + H_2O bend + CO_2 as.	
1639	258	NH_3 ad. + CO_2 as. + H_2O bend	1637
1623	61	CO_2 as. + NH_3 ad. + H_2O bend	
1583	528	CO_2 as. + H_2O bend	1598/1599
1545	135	NH_3 sd + H_2O bend	1510
1433	13	CO_2 ss. + CH_2 sc.	
1412	118	CO_2 ss. + CH_2 sc.	1412 /1413
1329	108	CH_2 r. + CO_2 ss.	1331

Table S4: Relative cluster total energies (E_{tot}), solute energies (E_{eff}) and solute-solvent interaction energies (V_{GW}) of model clusters of *z*Gly and *n*Gly. All energies are in kcal/mol and calculated at the BHandHLYP/6-31++G(d,p) level with the BSSE corrections. The structure of *z*GlyW8 (*z*1) is obtained through our proposed H-bond model. The structure of reference modeling clusters, 8*Z*-*c*, 8*Z*-*a*, 8*Z*-*e*, 8*N*8-*a*, 8*N*1-*a*, 8*N*6-*a*, and 8*N*1-*d*, are from Ref [14].

No.	***Clusters***	E_{tot}	E_{eff}	V_{GW}
*z*1	*z*GlyW$_8$(z1)	0.0	0.0	-97.7
*z*2	8Z-c	-4.7	8.9	-88.6
*z*3	8Z-a	-6.0	13.7	-86.2
*z*4	8Z-e	-5.3	15.3	-80.1
*c*1	8N8-a	-7.4	34.3	-36.5
*c*2	8N1-a	-6.0	39.6	-26.9
*c*3	8N6-a	-6.5	40.3	-27.2
*c*4	8N1-d	-5.5	39.7	-26.7

Table S5: Testing results for the proposed Π_2(O) state of solvated phenylalanine and alanine. The experiment data are from Refs [15,16].

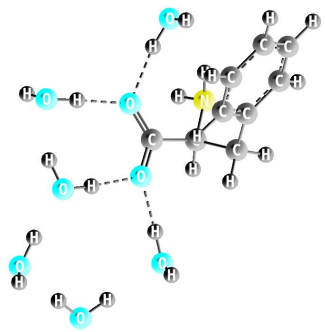
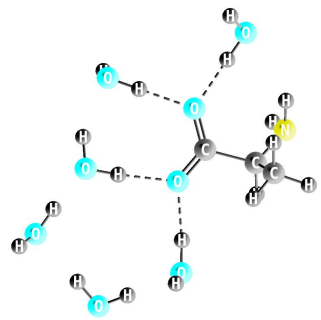
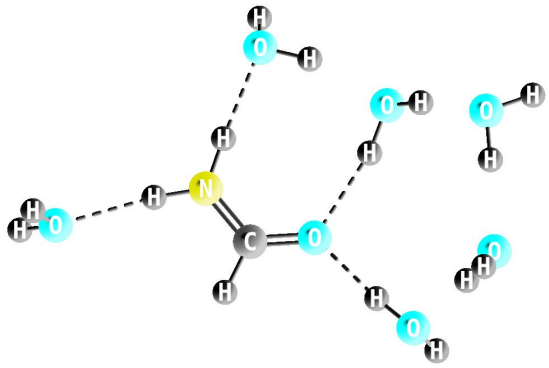
Phenylalanine			***Alanine***		
					
∠(C=O···H_W)=123,126,127,130 Φ[H_W···O═C···H_W]=161,157			∠(C=O···H_W)=124,125,131,132 Φ[H_W···O═C···H_W]=171,168		
	v_{as}(C-O)	v_{ss}(C-O)		v_{as}(C-O)	v_{ss}(C-O)
Theo.	1563	1398	*Theo.*	1571	1402
Exp.	1560	1410	*Exp.*	~1560	~1410

Table S6: Testing results for formamide in aqueous solution and in solid/liquid phase. The experiment data are from Refs [17-19].

Formamide				
				
∠(N–H···O_W)=174°, 173°, ∠(C–O···H_W)=126°, 128°, Φ[H_W···O=C···H_W]=157°				
IR mode	Theo.	Exp.		
		solid	*liquid*	*Aqueous solution*
Asym NH_2 stretch	3410	3372		
Sym NH_2 stretch	3176	3179		
CH stretch	2894	2895		
CO stretch	1685	1698	1682	1691
NH_2 scissoring	1603	1628		
CH bending	1360	1386		
CN stretch	1312	1334		

III. References

(1) See the manual of ref 9 for the half-and-half functional implemented in Gaussian03.

(2) Yu, W. B.; Liang, L.; Lin, Z. J.; Ling, S. L.; Haranczyk, M.; Gutowski. M. *J. Comput. Chem.* **2009**, *30*, 589.

(3) Derbel, N.; Hernández, B.; Pflüger, F.; Liquier, J.; Geinguenaud, F.; Jaïdane, N.; Lakhdar, Z. B.; Ghomi, M. *J. Phys. Chem. B* **2007**, *111*, 1470.

(4) Hernández, B.; Pflüger, F.; Nsangou, M.; Ghomi, M. *J. Phys. Chem. B* **2009**, *113*, 3169.

(5) Hernández, B.; Pflüger, F.; Derbel, N.; Coninck, J. D.; Ghomi, M. *J. Phys. Chem. B* **2010**, *114*, 1077.

(6) Scott, A.; Radom, L. *J. Phys. Chem.* **1996**, *100*, 16502.

(7) Merrick, J.; Moran, D.; Radom, L. *J. Phys. Chem. A* **2007**, *111*, 11683.

(8) Kuppens, T.; Vandyckc, K.; Eycken, J.; Herrebout, W.; Veken, B.; Bultinck, P. *Spectrochimica Acta Part A* **2007**, *67*, 402.

(9) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A. Gaussian 03, Revision A.1, Gaussian, Inc., Pittsburgh PA (2003).

(10) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev,

O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. Gaussian 09, Revision D.01, Gaussian, Inc., Wallingford CT, 2009.

(11) Max, J. J.; Trudel, M.; Chapados, C. *Appl. Spectrosc.* **1998**, *52*, 226.

(12) Williams, R. W.; Kalasinsky, V. F.; Lowrey, A. H. *J. Mol. Struct.* **1993**, *281*, 157.

(13) Lowrey, A. H.; Kalasinsky, V. F.; Williams, R. W. *Struct. Chem.* **1993**, *4*, 289.

(14) Aikens, C. M.; Gordon, M. S. *J. Am. Chem. Soc.* **2006**, *128*, 12835.

(15) Olsztynska, S.; Komorowska, M.; Vrielynck, L.; Dupuy, N. *Appl. Spect.* **2001**, *52*, 901.

(16) Wolpert, M.; Hellwig, P. *Spectrochimica Acta Part A* **2006**, *64*, 987.

(17) Torrie, B. H.; Brown, B. A. *J. Raman. Spect.* **1994**, *25*, 183.

(18) Sivaraman, B.; Raja Sekhar, B. N.; Nair, B. G.; Hatode, V.; Mason, N. J. *Spectrochimica Acta Part A: Mol. Bio. Spect*, **2013**, *105*, 238.

(19) Eaton, G.; Symons, M. C. R.; Rastogi, P. P.; *J. Chem. Soc., Faraday Trans. I*, **1989**, *85*, 3257.