ELECTRONIC SUPPLEMENTARY INFORMATION

On the Degradation Pathway of Glyphosate and Glycine

Anderson José Lopes Catão[a] and Alejandro López-Castillo[a]
[a] Departamento de Química, Universidade Federal de São Carlos (UFSCar), São Carlos, SP, Brazil, 13560-970.

The Electronic Supplementary Information (ESI) contains: all the results obtained via polarizable continuum solvation model (PCM); i.e., the transition and ground states optimizations and the intrinsic coordinate reaction calculation; the Cartesian coordinates of the transition state structures; the energies of optimized systems at DFT/aug-cc-pVDZ level of theory; and the relative orientation of dihedrals (see main text for references).

This ESI is divided into three parts. Firstly, the results concerning energy and geometry of the glycine molecule and the glyphosate molecule are presented. Secondly, the results from PCM model, which are subdivided into five sections. Thirdly, the Cartesian coordinates of the systems (ground states and transition states) are presented.

1 – Energy and Geometry

Table S1 shows the energy differences between the zwitterionic glycine and the neutral glycine optimized in liquid-phase (implicit solvent) using three DFT functionals (see also Table S2). As expected, charge separation in zwitterions is stabilized by solvent effects (Table S1 for more details). Analogous results can be found in the literature [1-4].

Table S1. Relative Energy between the zwitterionic form and the neutral form of glycine and glyphosate.

system	DFT functional	relative energy (eV)[a]
Gly(Z)	B3LYP[5, 6]	0.08
	M062X	0.17
	PBE[7]	0.06
GlyP(Z)	B3LYP	0.07
	M062X	0.18
	PBE	-0.03
Gly(Z) + $(H_2O)_1$	M062X	0.06
Gly(Z) + $(H_2O)_2$	M062X	-0.03
Gly(Z) + $(H_2O)_3$	M062X	-0.12
Gly(Z) + $(H_2O)_4$	M062X	-0.15

[a] Neutral specie in PCM (water as solvent) considered as reference.
Gly(Z) stands for zwitterionic glycine. GlyP(Z) stands for carboxylate zwitterion form of glyphosate.

Table S2 shows the ground state energies for the glycine molecule and the glyphosate molecule using different DFT functionals in gas-phase and also in liquid-phase (PCM).

Table S2 – Total and Relative Energy of the glycine and the glyphosate molecule in two different environments.

system	DFT functional	total energy (a.u.)			relative energy (eV)*	
		Neutral-Vacuum	Neutral-PCM	Zwitterionic-PCM	Neutral-PCM	Zwitterionic-PCM
Gly	B3LYP	-284.4813	-284.4961	-284.4930	-0.40	-0.32
	M062X	-284.3681	-284.3828	-284.3765	-0.40	-0.23
	PBE	-284.1646	-284.1794	-284.1771	-0.40	-0.34
GlyP	B3LYP	-891.5349	-891.5599	-891.5575	-0.68	-0.61
	M062X	-891.3023	-891.3277	-891.3212	-0.69	-0.51
	PBE	-890.7712	-890.7944	-890.7956	-0.63	-0.66

* Neutral specie in vacuum considered as reference.

The effects of explicit solvation on the systems can be seen in Table S1 and Table S3. The later presents the energies and the dipole moments of the zwitterionic glycine and the neutral glycine in different environments and number of explicit water molecules. The position of water molecules was made by taking into account that the explicit water tends to surround sites where charges or dipoles are large, for instance, NH_3^+ or NH_2 groups and carboxylic group (neutral and negatively charged). The cluster, Gly-$(H_2O)_1$, were sequentially built and optimized. The biggest cluster had seven explicit water molecules. One can see that 2 water molecules are minimally necessary to stabilize the zwitterion comparing to the neutral form, in agreement with the literature[8, 9]. The stabilization increased with the number of water molecules. It seemed to stabilize around 3.14 Kcal/mol.

Table S3 – Explicit water molecules influence in zwitterion stabilization.

	non-zwitterionic total energy (a.u.)	dipole (D)	zwitterionic total energy (a.u.)	dipole (D)	relative energy (a.u.)
0 Water	-284.3828	7.26	-284.3765	13.05	0.0064
1 Water	-360.8068	7.78	-360.8045	12.77	0.0023
2 Water	-437.2297	8.34	-437.2306	13.47	-0.0009
3 Water	-513.6522	11.07	-513.6565	15.89	-0.0043
4 Water	-590.0759	13.17	-590.0815	14.82	-0.0056
5 Water	-666.5089	7.57	-666.5097	14.17	-0.0008
6 Water	-742.9288	9.3	-742.9340	15.26	-0.0052
7 Water	-819.3556	10.58	-819.3613	17.3	-0.0057

Relaxed scans of the dihedrals Phi (O—C—C—N) and Psi (C—C—N—H) (Figures S1, S2) showed that the most stable configuration (rotamer) of the zwitterionic glycine is the one that align the carbonyl group with the hydrogen of amino group, that is, Psi and Phi angles are 0° (or 180°). This higher stability was due to the hydrogen bond formed between hydrogen from NH_3^+ group and oxygen from carbonyl group. Similar results were obtained by Valverde et al. (2018) [2] using another methodology. This configuration was start point in the transition state searching. Figure S3 highlights the dihedrals used in relaxed scan calculation. Figure S4 shows the two-dimensional relaxed scan of these dihedrals.

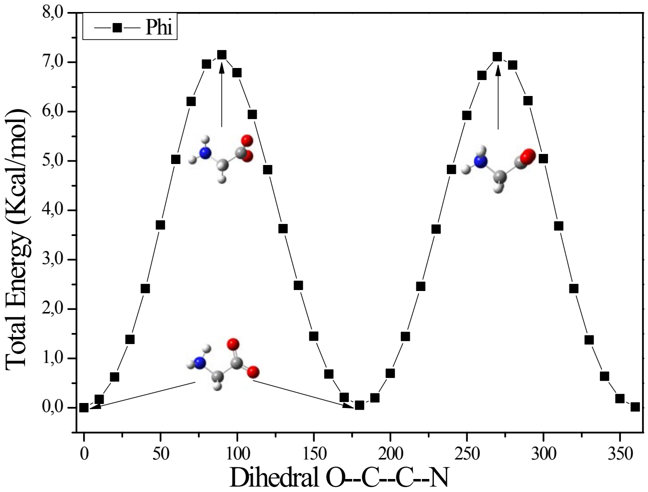


Figure S1. Relaxed Scan of Phi Dihedral Angle, (O—C—C—N), at M062X/aug-cc-pVDZ level of theory.

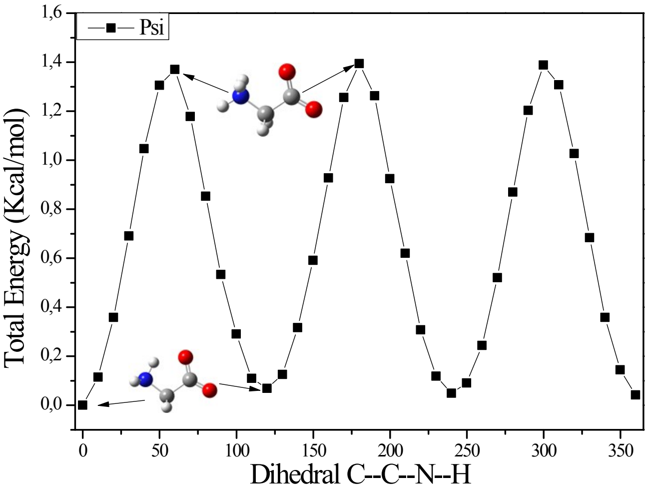
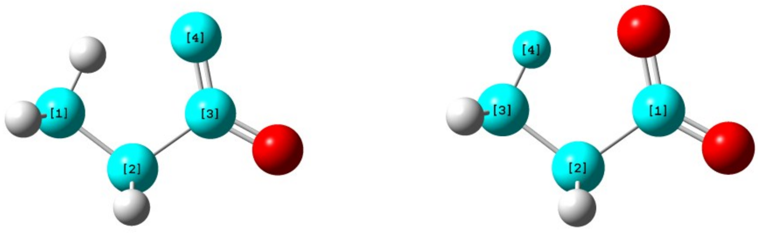


Figure S2. Relaxed Scan of Psi Dihedral Angle, (C—C—N—H), at M062X/aug-cc-pVDZ level of theory.



(a) (b)

Figure S3 – Representation of (a) Phi Dihedral Angle, (O—C—C—N) and (b) Psi Dihedral Angle, (C—C—N—H), of glycine.

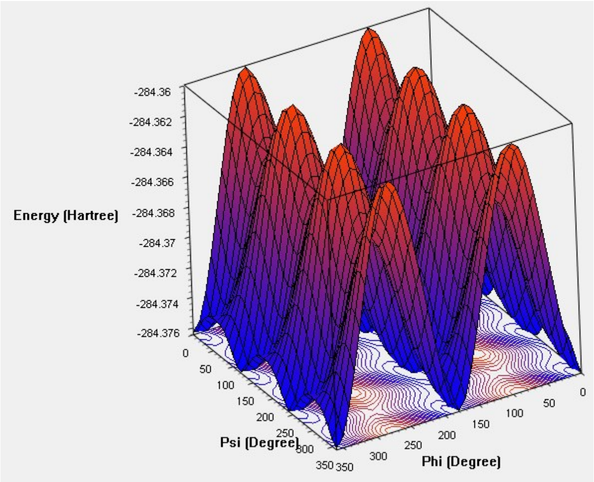
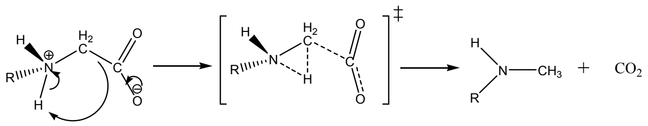


Figure S4 – Two-dimensional relaxed scan of the glycine'dihedral anles Phi (O—C—C—N) and Psi (C—C—N—H) at M062X/aug-cc-pVDZ level of theory.

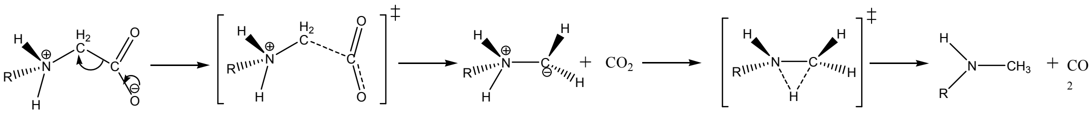
2 – PCM Results

2.1 – C—C Bond Cleavage

Two pathways are proposed: one which the proton transfer occurs simultaneously with CO_2 detachment (hereinafter called concerted mechanism) and other which consists into two sequential steps, decarboxylation and proton transfer (hereinafter called Stepwise Mechanism). The mechanism representing those pathways is schematized in Scheme S1.



Concerted Mechanism

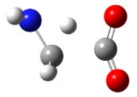
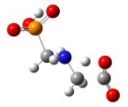
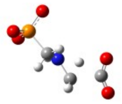
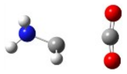
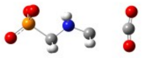
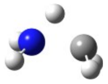
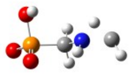
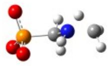


Stepwise Mechanism

Scheme S1. Glycine (R = H) and Glyphosate (R = $PO_3H_2CH_2$) degradation mechanisms: concerted (up) and stepwise (down) mechanisms.

Table S4 summarizes the transition states (TS) characteristics of the mechanism mentioned above. Figure S5 presents the Intrinsic Coordinate Reaction (IRC) path to the mechanism sketched above for the glycine (a) and the glyphosate (b) molecules.

Table S4. Transition state properties of concerted (CM) and stepwise (SM) mechanisms for glycine and glyphosate.

		glycine	**Glyphosate-pKa2**	**Glyphosate-pKa6-Conventional**
CM	**TS – proton transfer followed by decarboxylation**			
	relative energy (eV)‡	3.25	3.33	3.50
	dipole (D)	8.32	10.75	17.17
	frequency (cm^{-1}) and intensity	-1549.65 and 2878.30	-1639.15 and 3950.05	-1631.62 and 4116.02
SM	**TS1 - decarboxylation**			
	relative energy (eV) ‡	2.19	2.10	2.25
	dipole (D)	8.43	6.37	17.63
	frequency (cm^{-1}) and intensity	-98.20 and 65.33	-100.97 and 85.42	-59.75 and 78.63
	TS2 – proton transference *			
	relative energy (eV) ‡	1.22	1.26	1.36
	dipole (D)	4.06	10.41	19.28
	frequency (cm^{-1}) and intensity	-1738.65 and 2584.31	-1717.44 and 3117.94	-1734.33 and 3296.54

TS stands for Transition State; TS1 for transition state at the first step and TS2 at the second one.

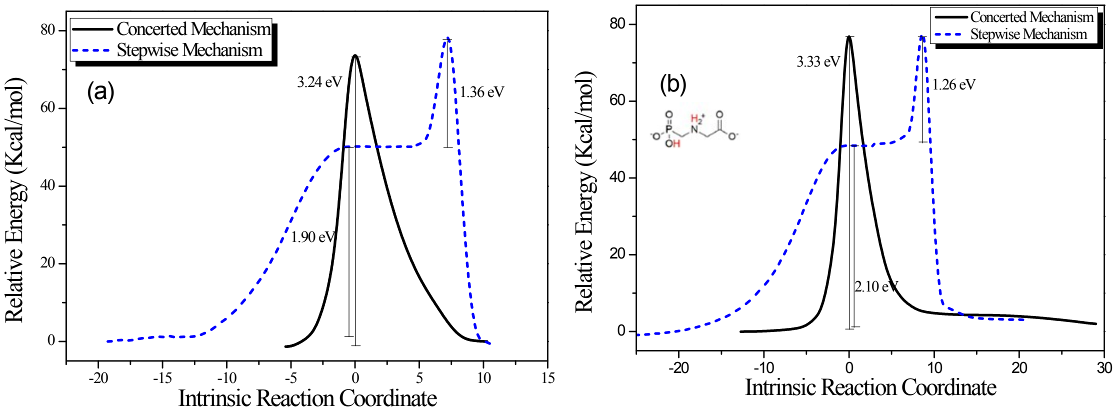


Figure S5. IRC for the (a) Glycine and (b) Glyphosate-pKa2 molecules following two different pathways.

The concerted mechanism and the stepwise mechanism are illustrated in Figure S6 for glyphosate molecule. The relative free energy of each process, as well as final product, MeAMPA and transition states are also shown.

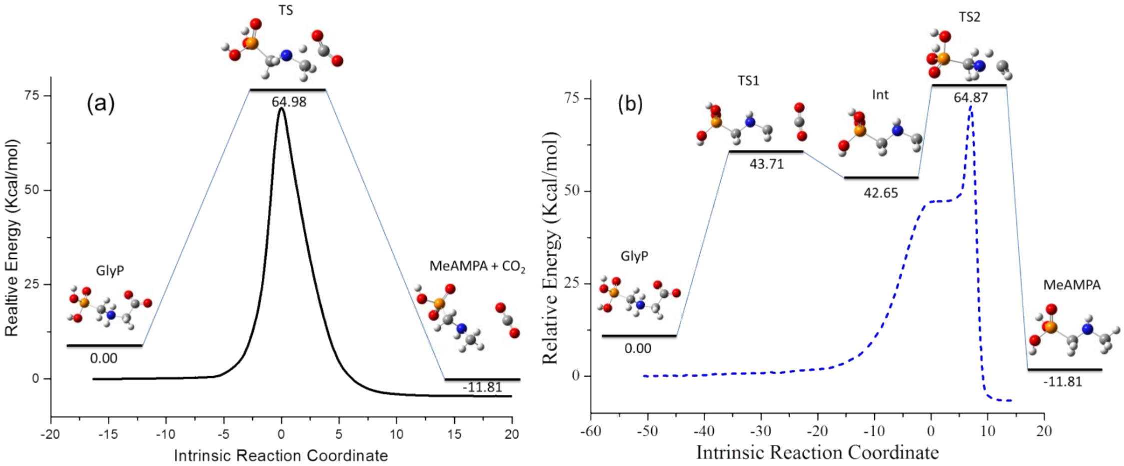
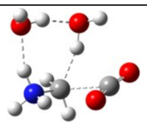
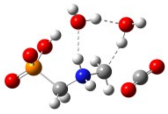
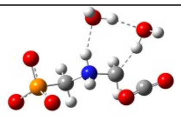
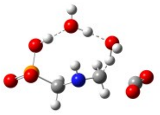


Figure S6. Glyphosate degradation pathways: (a) concerted and (b) stepwise mechanisms. After decarboxylation process CO_2 molecule is omitted in Stepwise mechanism for clarity.

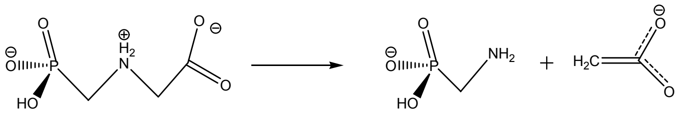
Table S5 summarizes the transition states characteristics of the concerted mechanism with two water molecules.

Table S5. Transition states characteristics of water-assisted mechanism for glycine and glyphosate.

		Gly	GlyP-pKa2	GlyP-pKa6-Conventional	GlyP-pKa6-Peixoto
CM ($2\ H_2O$)	TS – proton transfer followed by decarboxylation				
	dipole (D)	5.98	12.85	20.12	9.58
	frequency (cm^{-1}) and intensity	-382.83 and 724.34	-372.61 and 804.73	-392.68 and793.67	-1392.51 and 5834.33
	activation energy (eV)	2.03	2.00	2.15	2.50

2.2 – C—N bond cleavage

The heterolytic cleavage of C—N bond of glyphosate lead to AMPA formation, as sketched in Scheme S2. Figure S7 shows the calculated pathway for the reaction represented in Scheme S2. Energetic barrier is smaller for the pKa2 specie (see main text for references), which is 2.23 eV against 2.54 for the pKa6-Conventional specie.



Scheme S2. Glyphosate (pKa2-structure) degradation in AMPA.

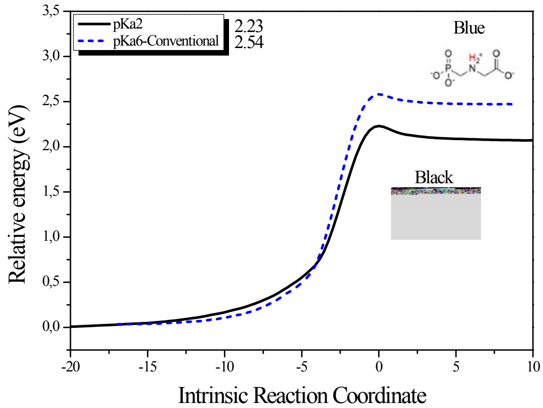
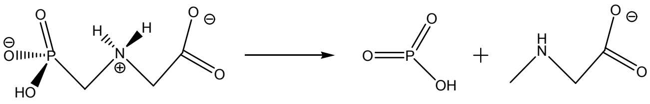


Figure S7. AMPA formation from two different protonation states of glyphosate: (black) pKa2 structure (Barrier = 2.23) pathway; (blue) pKa6 structure (Barrier = 2.54).

2.3 – C—P bond cleavage

Sarcosine is also a degradation product of glyphosate obtained via breakdown of the C—P bond. The cleavage of C—P bond and a proton transfer from amine group to vicinal methyl group lead to sarcosine formation, as sketched in Scheme S3.



Scheme S3. Glyphosate (pKa2-strucuture) degradation to sarcosine.

The proton transfer might occur from amine group or from phosphonate group in sarcosine pathway, similar to the MeAMPA case. The reaction profiles of these processes are shown in Figure S8 for two different protonation states of glyphosate. Proton transfer via phosphonate group is 0.5 eV less energetic than transfer via amine group.

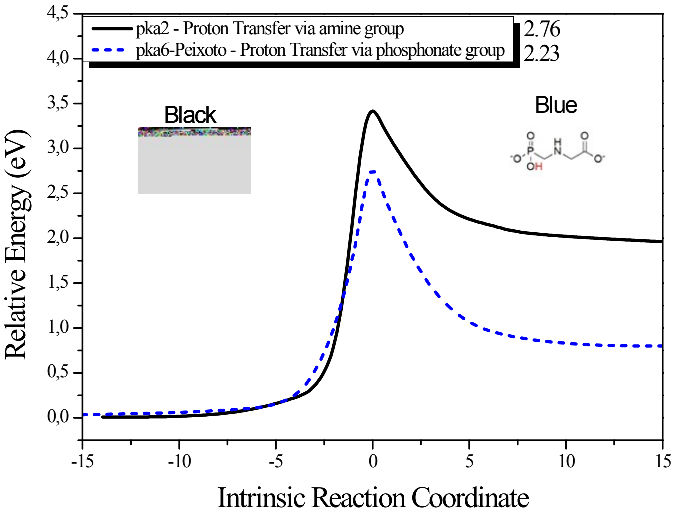
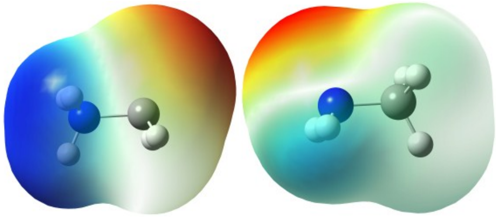


Figure S8. Sarcosine formation from two different protonation states of glyphosate: (black) pKa2 structure (Barrier = 2.76) pathway; (blue) pKa6 structure (Barrier = 2.23).

2.4 – Stepwise mechanism – Intermediate stability

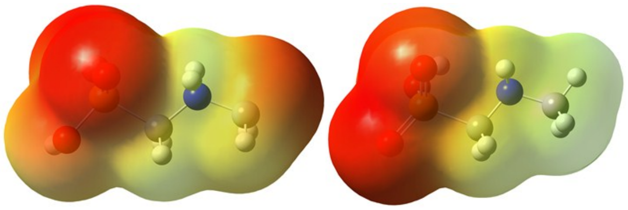
The stepwise mechanism will be plausible if the intermediate formed between TS1 and TS2 is stable enough to have a relative long half-life time. The intermediated state can be seen as formed by an ion pair, the alfa-carbanion and the protonated amino specie NH_3^+. This electrostatic interaction play an important role on the stabilization of the intermediate, and it is only possible in solvated medium. The ionic character is expressed by the electrostatic potential map (Figure S9) and by the C-N bond length, which is larger in the intermediate (1.531 Å) than in the final product (1.466 Å), methylamine. The energy difference between these two species is about 2.20 eV.



(a) (b)

Figure S9. (a) Intermediate and (b) methanamine electrostatic potential surfaces (density isovalue = 0.0004).

Glyphosate case is similar. The dipole is also high, 5.95 D for the pKa2 structure. The electrostatic potential maps are expressed in Figure S10. The energy difference between these two species is about 2.14 eV.



(a) (b)

Figure S10. (a) Intermediate and (b) MeAMPA electrostatic potential surfaces (density isovalue = 0.001).

Table S6 shows the thermochemistry data for glycine (GLY) and glyphosate (GLYP) degradation reactions following the stepwise mechanism (see Scheme S1 above), i.e., glycine decomposes into carbonic gas and methylamine passing through a carbanion, intermediate, while glyphosate decomposes into CO_2 and MeAMPA, passing through the intermediate showed in Figure S10.

Table S6. Thermochemistry data of glycine and glyphosate degradation following the proposed stepwise mechanism using DFT/PCM.

stepwise mechanism				
	glycine		glyphosate-pKa2	
	TS1 - decarboxylation	TS2 - proton transfer	TS1 - decarboxylation	TS2 - proton transfer
ε_0 (a.u.)	-284.2960	-284.2512	-890.8002	-890.7575
ε_{ZPE} (a.u.)	0.0768	0.0706	0.1185	0.1130
ε_{tot} (a.u.)	0.0832	0.0781	0.1304	0.1257
H_{corr} (a.u.)	0.0841	0.0790	0.1314	0.1267
G_{corr} (a.u.)	0.0463	0.0381	0.0778	0.0720
Δ‡G° (Kcal/mol)	46.01 (2.00 eV)	24.59 (1.07 eV)	45.58 (1.98 eV)	25.31 (1.10 eV)
k (s^{-1})	1.16E-21	5.83E-06	2.40E-21	1.73E-06

ε_0 stands to Total electronic energy; ε_{ZPE} to Zero-point correction; ε_{tot} to Thermal correction to Energy; H_{corr} to Thermal correction to Enthalpy; G_{corr} to Thermal correction to Gibbs Free Energy; Δ‡G° to Activation Free Energy (See equation 1); k to rate reaction at 298.15 K and concentration (c°) is taken to be 1 (See equation 2).

2.5 – Thermochemistry

Thermochemistry data for the glycine (GLY) and glyphosate (GLYP) degradation reactions following the concerted mechanism (see Scheme S1 above) are shown in Table S7:

$$\mathrm{GLY} \longrightarrow \mathrm{CO_2} + \mathrm{CH_5N}$$

$$\mathrm{GLYP} \longrightarrow \mathrm{CO_2} + \mathrm{CH_6NO_3P},$$

glycine decomposes into carbonic gas and methylamine, while glyphosate decomposes into CO_2 and methyl-aminomethylphosphonic acid (MeAMPA). This spontaneous decarboxylation is very slow due to the high activation energy for the glycine molecule, which is about 39 ± 2 kcal/mol [10]. It means that this reaction would not occur in nature without catalysts, which yield rate constants many orders of magnitude bellow the non-catalyzed reaction, e.g. 2×10^{-17} s^{-1} at 25 °C for glycine decarboxylation in neutral solution, i.e., a half-time of 1.1 billion years [10].

We calculated activation free energy, Δ‡G°, using the equation 1:

$$\Delta_r G^{\circ}(298.15K) = \sum_{Products} (\varepsilon_0 + G_{corr}) - \sum_{Reactants} (\varepsilon_0 + G_{corr}) \qquad (1),$$

and the rate reaction, $k(T)$, at 298.15 K and concentration (c°) 1, the equation 2:

$$k(T) = \frac{k_B T}{hc^{\circ}} e^{-\Delta^{\ddagger} G^{\circ}/RT} \qquad (2).$$

We used glycine decarboxylation reaction as reference for half-time estimative. The spontaneous decomposition of glycine follows a first order reaction [10], where the half-time ($t_{\frac{1}{2}}$) is given by:

$$t_{\frac{1}{2}} = \frac{\ln(2)}{k(T)} \qquad (3),$$

which is independent of the reactant concentration.

Table S7. Thermochemistry data of the glycine and glyphosate (pKa2) degradation follow the proposed concerted mechanism using DFT/PCM.

concerted mechanism				
	glycine	TS	glyphosate-pKa2	TS
ε_0 (a.u.)	-284.3765	-284.2569	-890.8811	-890.7587
ε_{ZPE} (a.u.)	0.0818	0.0717	0.1237	0.1144
ε_{tot} (a.u.)	0.0870	0.0778	0.1345	0.1259
H_{corr} (a.u.)	0.0879	0.0787	0.1355	0.1269
G_{corr} (a.u.)	0.0534	0.0421	0.0851	0.0752
Δ‡G° (Kcal/mol)	---	67.86 (2.94 eV)	---	70.67 (3.06 eV)
k (s^{-1})	---	1.10E-37	---	9.75E-40

ε_0 stands to Total electronic energy; ε_{ZPE} to Zero-point correction; ε_{tot} to Thermal correction to Energy; H_{corr} to Thermal correction to Enthalpy; G_{corr} to Thermal correction to Gibbs Free Energy; Δ‡G° to Activation Free Energy (See equation 1); k to rate reaction at 298.15 K and concentration (c°) is taken to be 1 (See equation 2).

The thermochemistry results for the systems with two water molecules combined microsolvation-continuum approach are shown in Tables S8 and S9. A naïve half-life time, first-order kinetics[10], is about 69 billion years and 11 billion years for glycine and glyphosate, respectively, in the two-water-assisted concerted mechanism in absence of tunneling effect.

Table S8. Thermochemistry data of glycine and glyphosate (pKa2) concerted mechanism degradation (DFT/PCM/microsolvation)

concerted mechanism + 2 explicit H_2O				
	glycine	TS	glyphosate-pKa2	TS
ε_0 (a.u.)	-437.2345	-437.1598	-1043.7433	-1043.6700
ε_{ZPE} (a.u.)	0.1320	0.1267	0.1755	0.1700
ε_{tot} (a.u.)	0.1426	0.1379	0.1907	0.1860
H_{corr} (a.u.)	0.1435	0.1389	0.1917	0.1870
G_{corr} (a.u.)	0.0962	0.0895	0.1328	0.1257
Δ‡G° (Kcal/mol)	---	42.68 (1.85 eV)	---	41.63 (1.80 eV)
k (s^{-1})	---	3.18E-19	---	1.89E-18

ε_0 stands to Total electronic energy; ε_{ZPE} to Zero-point correction; ε_{tot} to Thermal correction to Energy; H_{corr} to Thermal correction to Enthalpy; G_{corr} to Thermal correction to Gibbs Free Energy; Δ‡G° to Activation Free Energy (See equation 1); k to rate reaction at 298.15 K and concentration (c°) is taken to be 1 (See equation 2).

3 – Cartesian Coordinates

The Cartesian coordinates of the optimized transition states and ground states for most of the systems herein mentioned are shown in Tables from S9 to S28.

Table S9 – Cartesian coordinates (in Angstrom) of the optimized zwitterionic Gly-(H2O)$_1$ cluster.

13

Energy = -360.80446896 a.u.

C	-0.55905	1.10122	0.32343
C	-1.03273	-0.31515	-0.05899
N	0.75272	1.41328	-0.31094
H	-1.28866	1.84341	-0.00837
H	-0.43489	1.17629	1.40701
O	-2.18017	-0.60676	0.31505
O	-0.2154	-1.02658	-0.70731
H	1.47979	0.74014	0.01297
H	0.68797	1.3029	-1.32517
H	1.05646	2.36707	-0.11014
O	2.19475	-0.83208	0.47185
H	1.41481	-1.17539	-0.01768

H	2.97269	-1.14035	-0.00539

Table S10 – Cartesian coordinates (in Angstrom) of the optimized non-zwitterionic Gly-(H2O)$_1$ cluster.

13

Energy = -360.80679875 a.u.

C	0.46572	1.08749	-0.39429
C	1.07683	-0.25872	-0.02801
N	-0.75633	1.40154	0.35204
H	1.26249	1.82942	-0.26917
H	0.2147	1.03525	-1.46264
O	2.27965	-0.40671	0.02738
O	0.26761	-1.288	0.19767
H	-0.69588	-1.09098	0.05693
H	-0.53512	1.56011	1.3327
H	-1.13282	2.2809	0.00645
O	-2.34091	-0.77381	-0.06406
H	-2.70554	-0.80396	-0.95584
H	-2.01965	0.15407	0.07316

Table S11 – Cartesian coordinates (in Angstrom) of the optimized zwitterionic Gly-(H2O)$_2$ cluster.

16

Energy = -437.23060664 a.u.

C	0.19303	0.87579	0.13867
C	0.2058	-0.65071	0.33491
N	-1.13551	1.43849	0.50431
H	0.95414	1.34448	0.7674
H	0.39288	1.12419	-0.90706
O	1.3091	-1.2059	0.13887
O	-0.87563	-1.19078	0.67079
H	-1.88126	1.0281	-0.09852
H	-1.37109	1.18695	1.46705
H	-1.15193	2.456	0.42378
O	-2.79338	-0.21341	-0.99733
H	-2.31081	-0.83932	-0.42075
H	-3.72883	-0.3288	-0.79766
O	3.37966	0.38037	-0.64177
H	2.65624	-0.20805	-0.32943
H	3.83823	0.65419	0.15902

Table S12 – Cartesian coordinates (in Angstrom) of the optimized non-zwitterionic Gly-(H2O)$_2$ cluster.

16

Energy = -437.22966356 a.u.

C	-0.07531	0.83283	0.30155
C	-0.21741	-0.64569	-0.03143
N	1.0309	1.48783	-0.40256
H	-1.05376	1.28984	0.11189
H	0.12011	0.88771	1.38151
O	-1.31734	-1.16059	-0.14182
O	0.86733	-1.38513	-0.16772
H	1.72503	-0.90416	0.01029
H	0.82584	1.54567	-1.39777
H	1.10031	2.44871	-0.07659
O	3.15038	-0.11033	0.17852
H	3.46859	-0.00894	1.08289
H	2.58834	0.6845	-0.01148
O	-3.63889	0.42761	0.023
H	-3.87235	0.49513	0.95469
H	-2.8539	-0.14856	0.00595

Table S13 – Cartesian coordinates (in Angstrom) of the optimized zwitterionic Gly-(H2O)$_3$ cluster.

19

Energy = -513.65649437 a.u.

C	-0.16143	-1.37295	-0.22115
C	0.17511	0.12462	-0.20854
N	-1.55929	-1.58828	-0.68084
H	0.51616	-1.90814	-0.89106
H	-0.06464	-1.78807	0.78575
O	1.36443	0.40335	0.0846
O	-0.74568	0.92696	-0.4861
H	-2.2284	-1.09465	-0.04912
H	-1.69168	-1.19337	-1.6146
H	-1.79482	-2.58108	-0.71563
O	-2.90959	0.1453	1.01073
H	-2.31052	0.76585	0.55488
H	-3.80407	0.46176	0.84244
O	3.01351	-1.7354	0.57863
H	2.47148	-0.94224	0.38328
H	3.57865	-1.84458	-0.19291
O	1.09887	3.18948	-0.11524
H	1.41075	2.27414	0.01283
H	0.17777	3.02077	-0.35279

Table S14 – Cartesian coordinates (in Angstrom) of the optimized non-zwitterionic Gly-(H2O)$_3$ cluster.

19

Energy = -513.65223710

C	0.43357	-1.35271	-0.2749
C	-0.11508	0.05107	-0.08119
N	1.67545	-1.60962	0.45668
H	-0.38266	-2.0404	-0.02306
H	0.63387	-1.45115	-1.35084
O	-1.32383	0.24751	-0.06404
O	0.70397	1.07287	0.03209
H	1.68112	0.85496	-0.0675
H	1.48554	-1.67436	1.45433
H	2.03222	-2.5211	0.18111
O	3.24772	0.52437	-0.11999
H	3.63903	0.51256	-1.00119
H	2.95028	-0.4024	0.07103
O	-3.08895	-1.97378	0.0612
H	-3.41653	-2.16239	-0.82444
H	-2.53422	-1.18225	-0.04349
O	-1.70261	3.08599	0.12759
H	-0.78376	3.36958	0.18594
H	-1.63447	2.11847	0.06309

Table S15 – Cartesian coordinates (in Angstrom) of the optimized zwitterionic Gly-(H2O)$_4$ cluster.

22

Energy = -590.08145704

C	-0.14475	-1.24991	-0.52027
C	-0.20538	0.2768	-0.39107
N	-1.48546	-1.79185	-0.86251
H	0.56514	-1.53645	-1.30076
H	0.17694	-1.68939	0.42892
O	0.90076	0.83496	-0.14533
O	-1.31461	0.83044	-0.5307
H	-2.16995	-1.55047	-0.11174
H	-1.83198	-1.3631	-1.72386
H	-1.46209	-2.80426	-0.992
O	-2.97665	-0.61317	1.14637
H	-2.6804	0.19781	0.69707
H	-3.93878	-0.56952	1.17688
O	3.02185	-0.82116	-0.80151
H	2.36167	-0.1107	-0.66915
H	3.82694	-0.37696	-1.08742
O	-0.08417	3.47097	-0.05618
H	0.4595	2.66214	-0.03835
H	-0.94946	3.08981	-0.25345
O	2.17141	-0.75809	1.96378
H	1.67026	-0.04685	1.53679
H	2.72244	-1.07193	1.23117

Table S16 – Cartesian coordinates (in Angstrom) of the optimized non-zwitterionic Gly-(H2O)$_4$ cluster.

22

Energy = -590.07587069 a.u.

C	1.63172	0.94222	0.25345
C	0.26633	0.29316	0.09313
N	2.6892	0.28914	-0.52204
H	1.49845	2.00438	0.01759
H	1.88149	0.86259	1.32075
O	-0.74712	0.9738	0.06155
O	0.17565	-1.02101	0.02655
H	1.04907	-1.52576	0.13736
H	2.54603	0.45359	-1.51614
H	3.57757	0.7274	-0.29188
O	2.40029	-2.30067	0.18141
H	2.73211	-2.51589	1.06094
H	2.81411	-1.4349	-0.07796
O	-0.48592	3.80856	-0.05234
H	-0.72293	4.17161	0.8075
H	-0.63833	2.85261	0.0335
O	-3.48164	-0.00851	-0.03068
H	-3.28776	-0.95832	-0.08238
H	-2.59677	0.38631	0.00342
O	-2.27516	-2.59162	-0.21848
H	-1.4195	-2.14761	-0.11127
H	-2.33502	-3.19663	0.52915

Table S17 – Cartesian coordinates (in Angstrom) of the optimized zwitterionic Gly-(H2O)$_5$ cluster.

25

Energy = -666.50968641 a.u.

C	3.54E-4	-1.23924	-0.35162
C	0.01868	0.28527	-0.18667
N	-1.34937	-1.79872	-0.08889
H	0.29641	-1.50074	-1.37192
H	0.71577	-1.68489	0.34608
O	1.14985	0.81957	-0.32681
O	-1.05828	0.87382	0.07202
H	-1.67125	-1.53081	0.86269
H	-2.03398	-1.40531	-0.7634
H	-1.34923	-2.81573	-0.16778
O	-1.96492	-0.50551	2.32246
H	-1.86741	0.28476	1.76447
H	-2.84734	-0.45515	2.7063
O	3.0125	-0.9548	-1.32638
H	2.4092	-0.21493	-1.10888
H	3.72641	-0.56559	-1.84219
O	0.33244	3.48981	0.08674
H	0.80297	2.65407	-0.08746
H	-0.5637	3.16125	0.2294
O	2.90924	-0.70642	1.5308
H	2.33493	0.02457	1.26235
H	3.2197	-1.03653	0.67257
O	-2.82995	-0.1537	-1.81211
H	-2.46299	0.52587	-1.22026
H	-3.7851	-0.02816	-1.80375

Table S18 – Cartesian coordinates (in Angstrom) of the optimized non-zwitterionic Gly-(H2O)$_5$ cluster.

25

Energy = -666.50889535 a.u.

C	-0.19477	1.81325	-0.58232
C	-0.14215	0.32998	-0.90724
N	-0.8568	2.09462	0.6936
H	0.83804	2.17726	-0.62894
H	-0.76084	2.29058	-1.3946
O	0.84379	-0.17218	-1.4276
O	-1.1999	-0.41971	-0.65724
H	-2.00704	0.07944	-0.29972
H	-0.28678	1.68791	1.43692
H	-0.87237	3.1005	0.83899
O	-3.19273	0.83917	0.37936
H	-3.79719	1.2898	-0.22182
H	-2.48816	1.50318	0.63497
O	3.10081	1.09962	-0.32725
H	3.94883	0.65299	-0.42397
H	2.4914	0.63651	-0.92833
O	1.40162	-2.24017	0.53894
H	0.56479	-2.73812	0.50806
H	1.40409	-1.77197	-0.30938
O	-1.19677	-3.20594	0.11518
H	-1.3565	-2.37177	-0.352
H	-1.36306	-3.8986	-0.53372
O	1.29563	0.32369	1.75763
H	2.0259	0.66319	1.21245
H	1.25841	-0.61845	1.51123

Table S19 – Cartesian coordinates (in Angstrom) of the optimized transition state of glycine in concerted mechanism.

10

Energy = -284.25693940 a.u.

C	-0.83701	-0.73967	2.51E-4
C	0.91863	0.10064	-2.5E-5
N	-1.97489	0.23111	-1.55E-4
H	-0.77371	-1.31969	-0.91726
H	-0.77341	-1.31843	0.91854
O	1.67838	-0.83076	-1.54E-4
O	0.89065	1.3236	8.8E-5
H	-0.61047	0.47692	-4.2E-4
H	-2.52985	0.21729	-0.84685
H	-2.53033	0.21764	0.84623

Table S20 – Cartesian coordinates (in Angstrom) of the optimized transition state of glycine in stepwise mechanism - Decarboxylation.

10

Energy = -284.29600767 a.u.

C	-1.10199	-0.34683	-7.68E-4
C	1.50518	0.06061	1.93E-4
N	-2.51921	0.22929	5.32E-4
H	-1.08846	-1.00916	-0.88209
H	-1.0893	-1.01506	0.87615
O	1.6029	-1.10259	3.85E-4
O	1.64551	1.21916	-6.1E-5
H	-2.65506	0.82429	0.81682
H	-2.65412	0.83071	-0.81115
H	-3.28505	-0.45104	-0.0026

Table S21 – Cartesian coordinates (in Angstrom) of the optimized transition state of glycine in stepwise mechanism – Proton Transfer

10

Energy = -284.25116932 a.u.

C	-1.32102	-0.98769	0.06845
N	-2.05331	0.41644	-0.10291
H	-2.09406	-1.73814	-0.11751
H	-0.94257	-1.06575	1.09309
H	-3.06715	0.40876	-0.13199
H	-1.66619	1.14219	0.48714
H	-1.30917	0.00342	-0.88237
C	1.4073	0.18147	-0.01026
O	2.02209	-0.80027	-0.07952
O	0.84474	1.19674	0.06988

Table S22 – Cartesian coordinates (in Angstrom) of the optimized transition state of glyphosate (pKa2) in concerted mechanism.

17

Energy = -890.75866107 a.u.

C	-1.87667	-1.14247	-0.06571
C	-3.14436	0.37523	0.01368
N	-0.47626	-0.78027	0.35003
H	-2.29953	-1.93104	0.55258
H	-1.97308	-1.29324	-1.14143
O	-4.20747	-0.14758	-0.18374
O	-2.59597	1.44751	0.17479
H	-1.63411	-0.10521	0.54285
H	-0.07914	-1.43693	1.01471
C	0.41653	-0.39521	-0.73219
H	-0.05262	0.41719	-1.30107
H	0.60107	-1.23719	-1.41674
P	2.0249	0.22429	-0.08897
O	1.8111	1.43187	0.8209
O	2.99865	0.26931	-1.26044
O	2.44433	-1.07078	0.88941
H	2.41953	-0.78404	1.81132

Table S23 – Cartesian coordinates (in Angstrom) of the optimized transition state of glyphosate (pKa2) in stepwise mechanism - Decarboxylation.

17

Energy = -890.80021608 a.u.

C	-1.74345	-0.34927	0.00898
C	-4.32396	0.07332	-0.0012
N	-0.33062	0.20064	0.03374
H	-1.78803	-1.01286	0.88953
H	-1.7559	-1.01389	-0.87165
O	-4.43875	-1.08857	-0.03691
O	-4.46498	1.23198	0.03425
H	-0.24285	0.8297	0.83426
H	-0.19575	0.8015	-0.78308
C	0.80466	-0.76543	0.07362
H	0.69847	-1.42834	-0.79035
H	0.7195	-1.34356	0.99845
P	2.39572	0.15992	-0.01517
O	2.38532	0.91309	-1.34189
O	3.50026	-1.07397	-0.1541
H	3.80674	-1.33775	0.72326
O	2.60725	0.88625	1.31148

Table S24 – Cartesian coordinates (in Angstrom) of the optimized transition state of glyphosate (pKa2) in stepwise mechanism – Proton Transfer

17

Energy = -890.75746605 a.u.

C	3.273	-0.75954	-0.01739
O	4.29819	-0.31135	0.29056
O	2.26631	-1.2603	-0.31661
C	1.71046	1.83568	-0.00288
N	0.29904	1.20337	-0.34489
H	1.9126	1.57502	1.0418
H	1.64242	2.91982	-0.13341
H	1.28193	0.86347	-0.8624
H	-0.29744	1.82948	-0.87842
C	-0.40194	0.51461	0.72562
H	-0.60035	1.21124	1.55566
H	0.22483	-0.30363	1.09709
P	-1.99738	-0.15261	0.10518
O	-2.62781	0.9335	-0.77076
O	-2.75071	-0.8064	1.25694
O	-1.49823	-1.39685	-0.89147
H	-1.2878	-1.04324	-1.76516

Table S25 – Cartesian coordinates (in Angstrom) of the optimized transition state of Gly-(H2O)$_2$ cluster in concerted mechanism.

16

Energy = -437.15980190 a.u.

C	0.04923	0.71539	-1.21571
C	-1.74815	-0.05791	0.34126
N	0.75874	1.67926	-0.29283
H	0.79316	0.46474	-1.9891
H	-0.73278	1.30323	-1.71706
O	-2.09589	-1.08844	-0.09222
O	-1.65952	0.91609	0.99203
H	0.32299	-1.01018	-0.68328
H	1.11898	2.52826	-0.73286
H	0.11589	1.96644	0.44758
O	0.7681	-1.88563	-0.43121
H	0.09014	-2.39222	0.02865
H	1.54251	1.18542	0.16478
O	2.59836	-0.24372	0.73524
H	2.73637	-0.35207	1.68208
H	2.00673	-0.97974	0.465

Table S26 – Cartesian coordinates (in Angstrom) of the optimized transition state of GlyP-pKa2-(H2O)$_1$ cluster in concerted mechanism.

20

Energy = -967.24171488 a.u.

C	-1.12755	-0.7736	-1.15737
C	-2.95629	-0.03801	0.35488
N	-0.27908	-1.55087	-0.16346
H	-1.84384	-1.49897	-1.56556
H	-0.45132	-0.50742	-1.98312
O	-3.73879	0.40657	-0.39385
O	-2.43463	-0.43357	1.33151
H	-1.13608	1.03428	-0.61274
H	-0.57074	-2.52792	-0.12944
C	1.1941	-1.50424	-0.36259
H	1.41779	-1.72014	-1.41045
H	1.65078	-2.26848	0.27534
P	1.89104	0.11681	0.18964
O	1.65859	1.13917	-1.08383
O	3.39163	-0.06868	0.35455
O	0.99278	0.56258	1.35584
H	0.77768	1.56291	-0.99803
O	-0.8755	1.94252	-0.26607
H	-0.41118	1.70026	0.556
H	-0.45926	-1.16416	0.77295

Table S27 – Cartesian coordinates (in Angstrom) of the optimized transition state of GlyP-pKa2-(H2O)$_2$ cluster in concerted mechanism (proton transfer via amine group).

23

Energy = -1043.66640300 a.u.

C	1.17928	-0.33152	-0.92599
C	3.33192	-0.9671	0.17966
N	0.21785	-0.79018	0.13759
H	0.60565	0.39517	-1.52478
H	1.33015	-1.20991	-1.57224
O	4.14264	-0.28417	-0.31769
O	2.80188	-1.78972	0.83035
H	2.18518	1.16369	-0.57647
H	0.71594	-1.46926	0.71558
O	2.44273	2.13211	-0.43784
H	3.29673	2.1058	0.00636
H	0.03202	0.01502	0.75204
O	0.03032	2.08864	0.7516
H	-0.13122	2.60511	1.54712
H	0.93888	2.30334	0.44416
C	-1.0907	-1.37535	-0.30878
H	-1.28809	-2.29428	0.25336
H	-1.008	-1.60939	-1.37312
P	-2.48127	-0.20816	0.01603
O	-2.02706	1.12874	-0.84776
O	-2.4211	0.12699	1.5074
O	-3.73739	-0.77005	-0.63729
H	-1.36217	1.63193	-0.33513

Table S28 – Cartesian coordinates (in Angstrom) of the optimized transition state of GlyP-pKa2-(H2O)$_2$ cluster in concerted mechanism (proton transfer via phosphonate group).

23

Energy = -1043.66485244 a.u.

C	1.55197	-0.80139	-0.54571
C	3.82898	-0.65402	0.16733
N	0.153	-0.33584	-0.22108
H	1.58025	-0.93036	-1.63704
H	1.61864	-1.8022	-0.09625
O	4.17941	-1.40087	-0.66944
O	3.86413	0.05081	1.10954
H	-0.01067	0.56904	-0.67463
C	-0.99975	-1.2249	-0.5394
H	-0.79841	-2.20198	-0.08969
H	-1.07838	-1.32456	-1.62544
P	-2.51277	-0.48883	0.22524
O	-2.76272	0.84035	-0.73014
O	-2.05065	-0.02296	1.61014
O	-3.70091	-1.41641	0.02813
H	-2.11369	1.53863	-0.51763
H	2.05846	0.96405	-0.12072
O	1.94415	1.94535	0.07022
H	2.43574	2.09097	0.88565
O	-0.61709	2.64819	-0.41173
H	-0.61297	3.28819	-1.13178
H	0.32562	2.54075	-0.15143
H	0.0783	-0.12316	0.78078

References

1. R. Perez de Tudela and D. Marx, Water-Induced Zwitterionization of Glycine: Stabilization Mechanism and Spectral Signatures, *The journal of physical chemistry letters*, 2016, **7**, 5137-5142.
2. D. Valverde, Z. M. da Costa Ludwig, C. R. da Costa, V. Ludwig and H. C. Georg, Zwitterionization of glycine in water environment: Stabilization mechanism and NMR spectral signatures, *J Chem Phys*, 2018, **148**, 024305.
3. M. P. Bhate, J. C. Woodard and M. A. Mehta, Solvation and Hydrogen Bonding in Alanine- and Glycine-Containing Dipeptides Probed Using Solution- and Solid-State NMR Spectroscopy, *J. Am. Chem. Soc.* , 2009, **131,**, 9579–9589.
4. A. N. Alexandrova and W. L. Jorgensen, On the mechanism and rate of spontaneous decomposition of amino acids, *The journal of physical chemistry. B*, 2011, **115**, 13624-13632.
5. A. D. Becke, Density-functional thermochemistry. III. The role of exact exchange, *The Journal of Chemical Physics*, 1993, **98**, 5648-5652.
6. C. Lee, W. Yang and R. G. Parr, Development of the Colle-Salvetti correlation-energy formula into a functional of the electronic density, *Phys. Rev. B*, 1988, **37**, 785-789.
7. J. P. Perdew, K. Burke and M. Ernzerhof, Generalized Gradient Approximation Made Simple, *Physical Review Letters*, 1996, **22**, 3865-3868.
8. C. M. Aikens and M. S. Gordon, Incremental Solvation of Nonionized and Zwitterionic Glycine, *J. Am. Chem. Soc.*, 2006, **128**.
9. J. H. Jensen and M. S. Gordon, On the Number of Water Molecules Necessary To Stabilize the Glycine Zwitterion, *J. Am. Chem. Soc.*, 1995, **117**, 8159-8170.
10. M. J. Snider and R. Wolfenden, The Rate of Spontaneous Decarboxylation of Amino Acids, *J. Am. Chem. Soc.*, 2000, **122**, 11507-11508.