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

SUPPORTING INFORMATION

The peptide bond hydrolysis in the human DPP III is regulated by stereoelectronic effect and nitrogen inversion

A. Tomić, B. Kovačević and S. Tomić

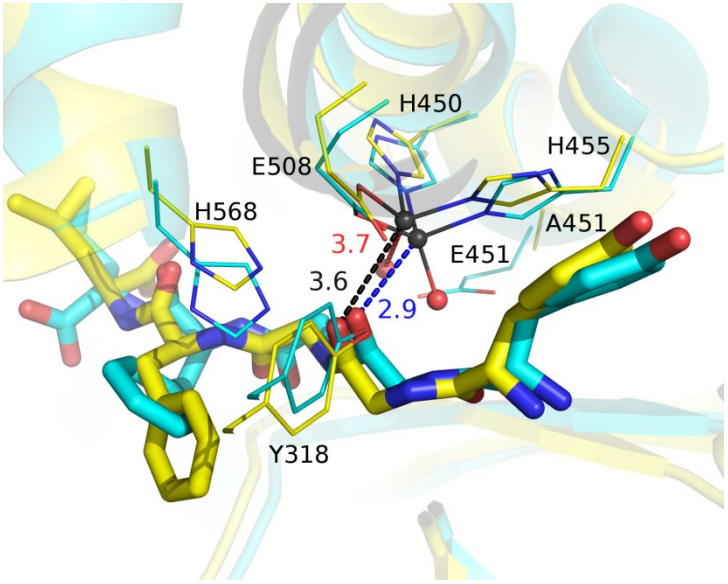
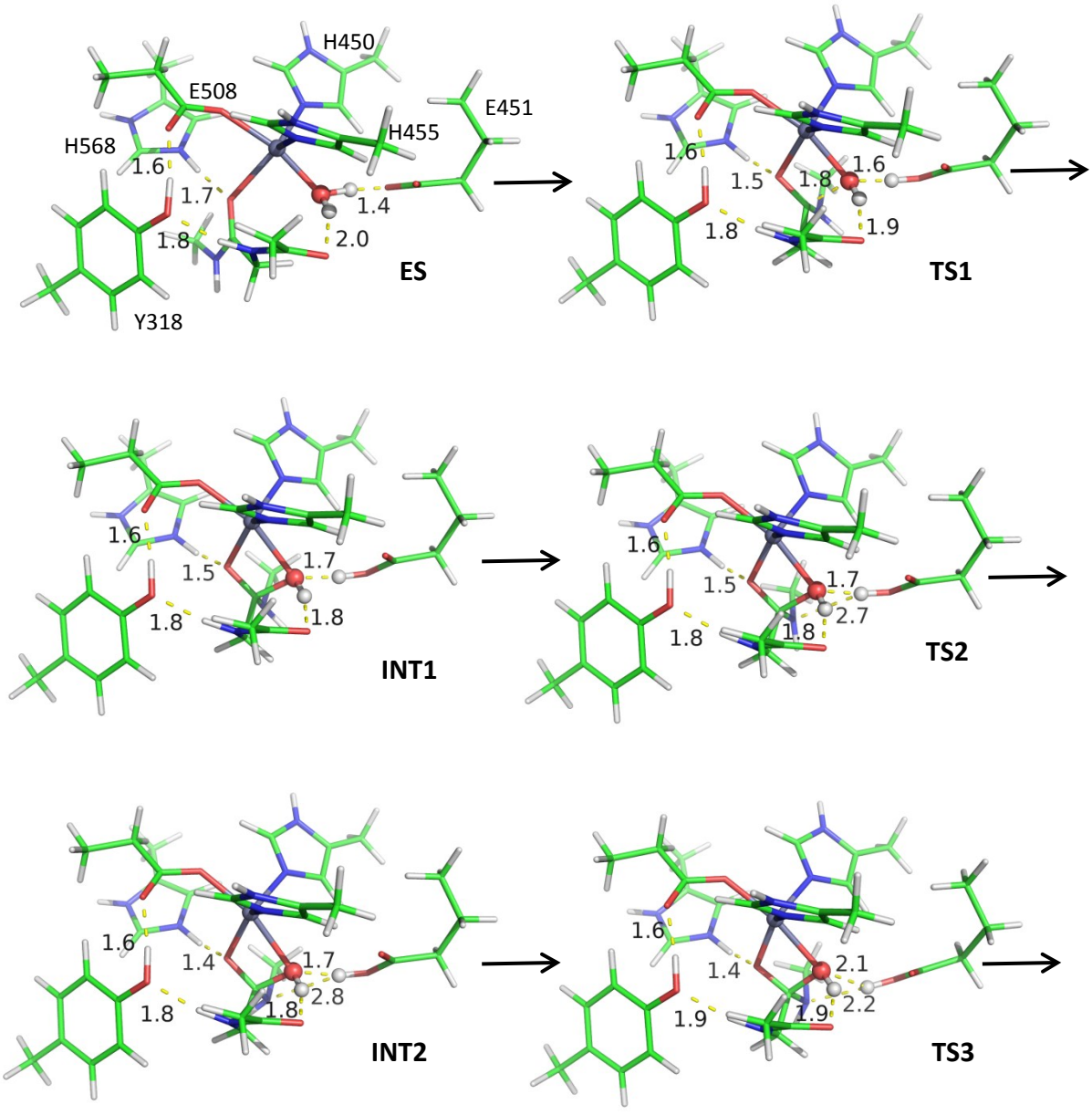


Figure S1. Overlap of the enzyme binding site in the human DPP III – Leu-enkephalin complex obtained experimentally (pdb code: 5E3A; sticks with carbon atoms coloured yellow) and by QM/MM optimization (sticks with carbon atoms coloured cyan). The major difference in ligand binding is observed for C-terminus (points in two opposite orientations). The zinc ion and water molecule are shown as grey and red spheres, respectively. Selected atom distances (in Å) are indicated.


H450
E508
E451
H455
H568
1.6
1.7
1.4
1.8
2.0
Y318
ES
1.6
1.5
1.8
1.6
1.8
1.9
TS1
1.6
1.5
1.7
1.8
1.8
INT1
1.6
1.5
1.7
2.7
1.8
1.8
TS2
1.6
1.4
1.7
2.8
1.8
1.8
INT2
1.6
1.4
2.1
2.2
1.9
1.9
TS3

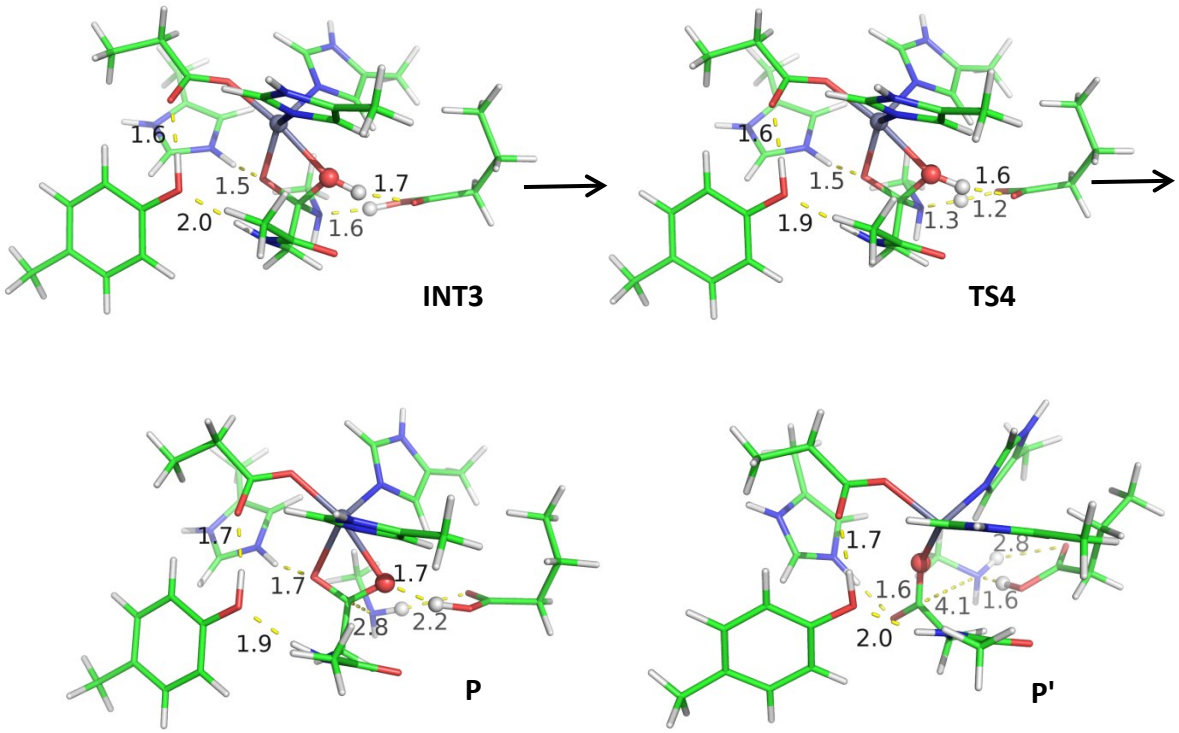


Figure S2. Optimized structures of the: model enzyme-substrate (**ES**) complex, transition states (**TS1-TS4**), intermediates (**INT1-INT3**), and products (**P** and **P'**) obtained for breakdown of the amide bond in CH_3CO-Gly-$NHCH_3$ bound to a model of the human DPP III active site with His568 in its protonated form. Calculations were performed at B97D/[6-31G(d)+LANL2DZ-ECP] level of theory. Electrostatic interactions are shown with yellow dashed lines (atom pairs distance are indicated in angstroms), and coordination bonds are shown with thick grey sticks. The zinc ion is shown as (grey) sphere, while water molecule in ball and sticks.

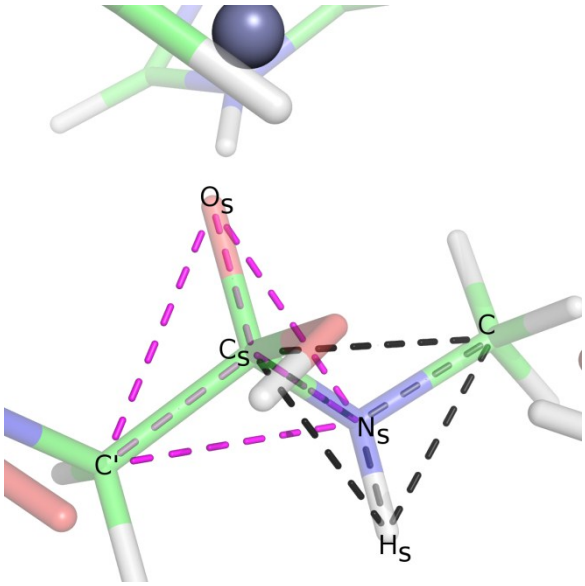
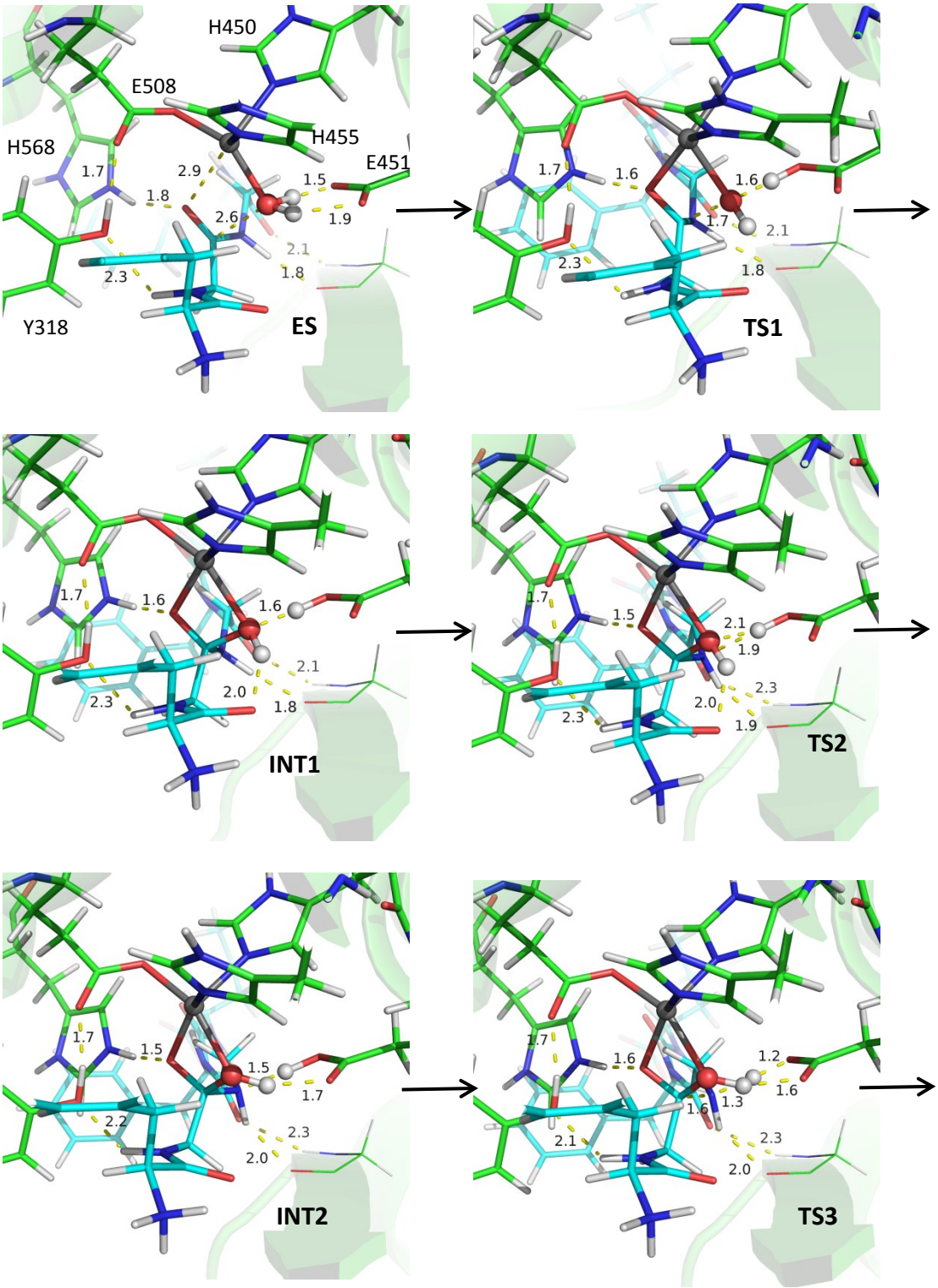


Figure S3. Improper torsional angles ω_1 (C-N_s-C_s-H_s; edges of the triangular pyramid shown with dashed black lines) and ω_2 (C'-C_s-N_s-O_s; edges of the triangular pyramid shown with dashed magenta lines) defining degree of nitrogen and carbon atom pyramidalization, respectively. Angles are shown on INT1 structure obtained from of CH_3CO-Gly-$NHCH_3$ hydrolysis reaction. The zinc ion is shown as grey sphere.


H450
E508
H455
E451
H568
Y318
ES
TS1
INT1
TS2
INT2
TS3

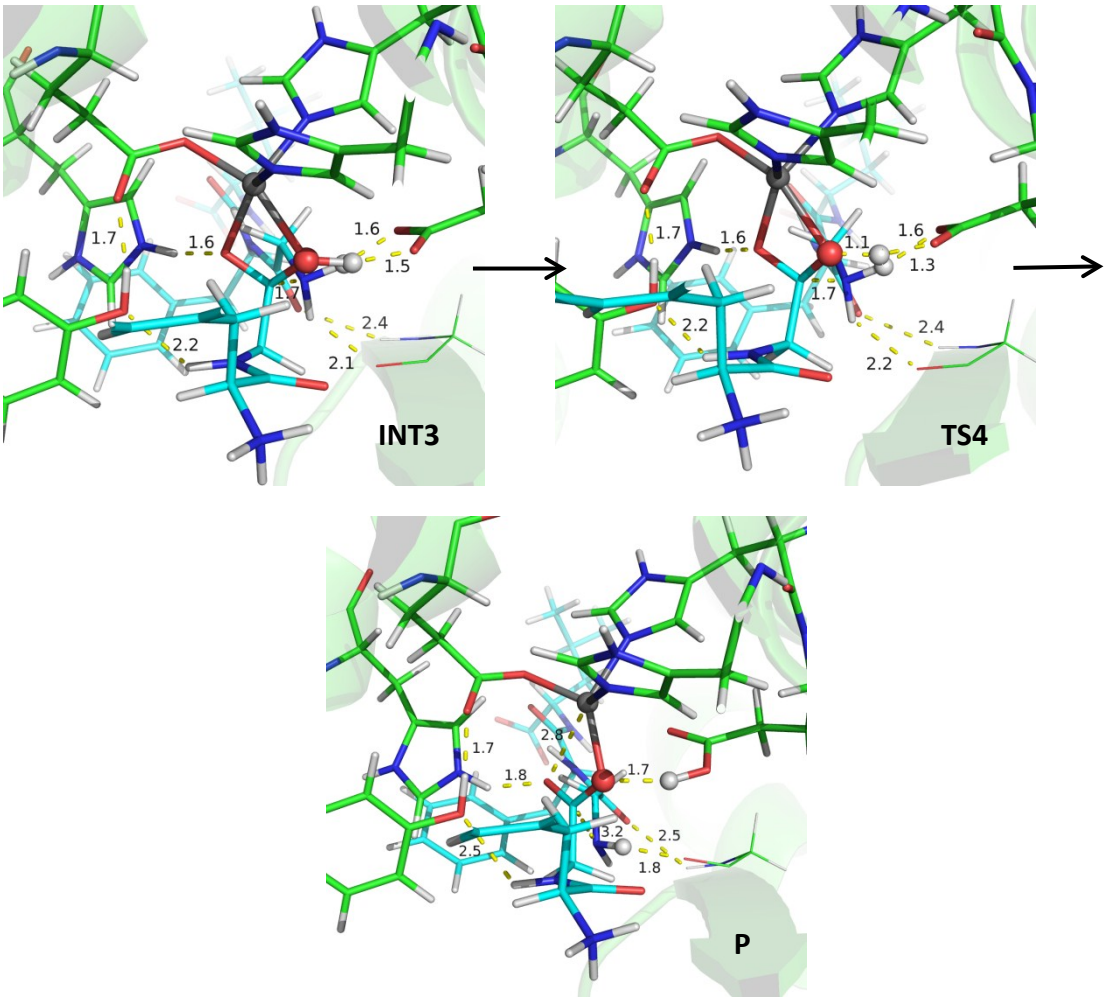


Figure S4. Optimized structures of the: enzyme-substrate (**ES**) complex, transition states (**TS1-TS4**), intermediates (**INT1-INT3**), and product (**P**) obtained by B97D/[6-31G(d) + LANL2DZ-ECP] calculations for Glu451-assisted peptide bond hydrolysis in human DPP III. His568 is in its protonated form. Amino acid residues (Y318, H450, E451, H455, E508 and H568) and substrate that are part of QM region are indicated as sticks with carbon atoms coloured green and cyan, respectively. The zinc ion is shown as grey sphere, while water molecule in ball and sticks. Amino acid Gly389 that is a part of MM region is shown in lines. Selected electrostatic interactions are shown with yellow dashed lines (atom pairs distance are indicated in angstroms), and coordination bonds are shown with thick grey sticks.

1) HYDROLYSIS IN THE MODEL SYSTEM WITH NEUTRAL HIS568

Model of the enzyme-substrate complex. In the model of the enzyme-substrate complex zinc ion is 5-coordinated. Besides by His450, His455 and Glu508, the metal ion is coordinated by water (O_w) and the substrate (O_s) oxygens (ES in Figs. S5 and S6 and Table S1). The zinc-coordinated water is polarized by negatively charged Glu451 (strong O_{e1}-H_{w1}-O_w hydrogen bond, see corresponding distances in Table S1), which is a strong base, and by the zinc cation (Zn-O_w coordination bond, Table S1). Consequently H_{w1} proton is partially shared between Glu451 and water as indicated by O_{e1}-H_{w1} and O_w-H_{w1} distances, 1.45 Å and 1.07 Å, respectively (Table S1). It should be noticed that water H_{w2} proton makes an additional hydrogen bond with the substrate terminal acetyl group (Figs. S5 and S6). In the enzyme binding site substrate is stabilized by hydrogen bonds with evolutionally conserved His568 and Tyr318, as well as by the coordinate bond with the metal ion (see ES in Fig. S6). Tyr318 is hydrogen bonded to Glu508 (Fig. S6) and this hydrogen bond is present throughout the reaction.

Glu451-assisted water addition and first proton transfer. The first reaction step, i.e. nucleophillic attack by water molecule is the rate-determine step for the substrate hydrolysis 60.8 kJ/mol (see Fig. S5 and Table S2). The deprotonation of water molecule by Glu451 is concerted with approaching of the emerging hydroxyl group to the amide carbon C_s atom (Figs. S5-6). Strong hydrogen bond between His568 and O_s oxygen in the enzyme-substrate complex facilitated nucleophile attack of the activated water molecule by dragging the negative charge away from C_s atom and positioning carbonyl group for the nucleophillic attack. The advancing formation of the oxy-anion in the TS1 structure is stabilized by substrate coordination on the metal dication (O_s-Zn distance is reduced by ~0.2 Å as compared in the enzyme-substrate complex, Table S1) and strong hydrogen bond with His568 (O_s-His568 distance equals to 1.73 Å, Table S1). Already in the TS1 structure Glu451 is protonated and planarity of the peptide bond is distorted as indicated by the pyramidalization of carbon and nitrogen atoms. The O_w-C_s bond formation (distance O_w-C_s in INT1 structure equals 1.51 Å, Table S1) system relaxes in the structure of the first intermediate (INT1 structure in Figs. S5 and S6).

The proton, H_{w1}, transfer from Glu451 to the substrate N_s atom, occurs in next two reaction steps (Fig. S5 and S6). The distance N_s-H_{w1} of 2.72 Å in INT1 structure is unusually long for a transition state of proton transfer. Therefore, rotation of Glu451 and hydroxide ion (O_w-H_{w2}) resulting with the formation of N_s-H_{w1}-O_{e2} and O_w-H_{w2}-O_{e1} hydrogen bonds are described by the second transition state (TS2) leading to INT2 structure. As a consequence, in the INT2 structure, N_s-H_{w1} distance reduces to 1.53 Å (Table S1). This facilitates H_{w1} proton transfer to amide nitrogen (N_s) and relaxation into the INT3 structure (Figs. S5 and S6). At the same time, elongation of peptide bond is observed (ionic character of amide bond in the INT3 structure is indicated by C_s-N_s distance of 1.67 Å, Table S1). Transition from INT2 to INT3 structure is accompanied by decrease in energy (Fig. S7). Beside via the hydrogen bond with His568 and bidentat coordination on the Zn^{2+} ion, INT3 structure is stabilized by two strong hydrogen bonds: O_w-H_{w2}-O_{e2} (distance H_{w2}-O_{e2} of 1.43 Å shows delocalized nature of H_{w2} atom, Table S1) and N_s-H_{w1}-O_{e1} (distance H_{w1}-O_{e1} of 1.67 Å, Table S1).

The peptide bond cleavage and second proton transfer. The collapse of INT3 into the product P takes place during three reaction steps. In the INT4 structure proton H_{w2} is transferred to the Glu451 carboxylate group and peptide bond is broken (distance C_s-N_s of 2.82 Å, Table S1). As a consequence, gradual planarization of the emerging carboxylate group, also indicated by decrease of distances O_w-C_s and O_s-C_s to 1.27 Å and 1.29 Å, respectively, can be observed (Table S1). The formed methylamine is stabilized by van der Waals interactions with the His450 and weak hydrogen bond with the Glu451 (in the INT4 structure distance H_1-O_{e1} is equal to 2.23 Å, Table S1), while the acetate part of product bidentantely coordinates the zinc ion and makes hydrogen bonds with the Glu451 and His568. The INT4 energy is 8.3 kJ/mol above the ES energy (Fig. S7 and Table S2). The next two reaction steps include the second proton (H_{w2}) transfer from the bound hydroxyl group to amine part of the product. The N_s-H_{w1} distance is 3.61 Å in the INT4 structure and this is quite long for a proton transfer. The Glu451 carboxylate group rotation towards methylamine results with formation of the strong N_s-H_{w2}-O_{e2} hydrogen bond in the TS5 structure (Figs. S5 and S6), and as a consequence the N_s-H_{w2} distance decreases to 1.58 Å in the INT5 structure (Table S1.). The hydrogen bond between the Glu451 and the acetate part is broken in this structure (Figs. S5 and S6). During the whole reaction five ligands coordinate the zinc ion, however, the 2.43 Å long O_w-Zn distance in the INT1 and INT5 structures indicates reduced contribution of O_w atom to zinc coordination. The INT5 structure has 2.9 kJ/mol higher energy than the ES complex (Fig. S7 and Table S2). In the next reaction step H_{w2} proton is transferred from Glu451 to the amine, but in the same time H_{w1} proton is transferred back to Glu451 upon the H_{w2} binding to the N_s nitrogen (with activation energy of 7.06 kJ/mol, Fig. S7 and Table S2). Formation of the neutral amine resulted with the product P that is 4.60 kJ/mol more stable than the ES complex. Regarding the INT4, INT5 and P structure energies, and the associated activation energies, any of these three structures can be considered as the final product of hydrolysis. Protonated amine (product P' show in Fig. S5 and S6) was identified in a structure where protonated amine forms two strong hydrogen bonds,

one with acetate and other with Glu451; as indicated by distances O_w-H_{w2} and H_{w1}-O_{e1} of 1.7 Å and 1.5 Å, respectively (Table S1). As a consequence H_{w1} proton is partially delocalized between Glu451 and N_s atom (indicated by N_s-H_{w1} distances of 1.13 Å), and acetate part of the product monodentatly coordinates the zinc ion (tetrahedral zinc coordination). Evidently, the N-protonated tetrahedral product is not stable.

We would like to emphasize that after addition of zero-point correction to electronic energies, energies of TS3 and TS4 structures are lower than the energies of INT2 and INT3 structures, respectively, and therefore, at this level of theory, they can not be considered as real transition states. Shape of the curve indicates that system more or less continuously, without energy barrier, relaxes in the INT4 structure (Fig. S7). Namely, H_{w1} proton bound to the Glu451 carboxylate group is transferred to N_s nitrogen while simultaneously C_s-N_s bond is cleavage and H_{w2} proton is transferred to Glu451.

Table S1. Selected bond lengths and dihedral angles during breakdown of the amide bond in CH_3CO-Gly-$NHCH_3$ bound into the model of the human DPP III active site with His568 modelled in its neutral state by 4-methylimidazol. Calculations were performed at the B97D/[6-31G(d) + LanL2DZ-ECP] level of theory.

	ES	TS1	INT1	TS2	INT2	TS3	INT3	TS4	INT4	TS5	INT5	TS6	P	P’
d(O_s-Zn) / Å	2.29	2.03	1.99	2.02	2.03	2.04	2.08	2.09	2.11	2.09	2.09	2.13	2.14	2.02
d(O_w-Zn) / Å	2.08	2.23	2.43	2.29	2.29	2.29	2.23	2.20	2.38	2.47	2.43	2.35	2.33	4.22
d(C_s-N_s) / Å	1.34	1.41	1.46	1.47	1.52	1.54	1.67	1.74	2.82	3.05	4.27	3.81	3.53	3.48
d(O_w-C_s) / Å	3.24	1.70	1.51	1.50	1.45	1.43	1.39	1.37	1.27	1.26	1.26	1.26	1.27	1.25
d(O_s-C_s) / Å	1.26	1.33	1.36	1.36	1.35	1.35	1.33	1.32	1.29	1.30	1.30	1.30	1.29	1.30
d(O_w-H_{w2}) / Å	0.98	0.99	1.00	0.99	1.00	1.01	1.07	1.18	1.63	2.05	3.42	3.02	2.18	1.74
d(O_w-H_{w1}) / Å	1.07	1.59	1.67	2.07	2.57	2.51	2.41	2.42	2.7	3.46	4.45	4.47	3.98	3.30
d(H_{w2}-O_{e2}) / Å	2.86	2.81	2.83	2.53	1.72	1.66	1.43	1.25	1.02	1.00	1.06	1.40	2.57	3.14
d(H_{w1}-O_{e1}) / Å	1.45	1.03	1.01	1.00	1.09	1.18	1.67	1.76	2.23	2.15	3.12	2.20	1.08	1.49
d(N_s-H_{w1}) / Å	5.12	2.71	2.72	2.14	1.53	1.37	1.07	1.06	1.02	1.02	1.02	1.03	1.54	1.13
d(N_s-H_{w2}) / Å	4.68	3.11	3.02	2.93	2.58	2.57	2.66	2.77	3.61	2.36	1.58	1.18	1.03	1.05
d(H450-Zn)[a] / Å	2.22	2.17	2.15	2.15	2.16	2.15	2.14	2.14	2.12	2.11	2.12	2.12	2.12	2.15
d(H455-Zn)[a] / Å	2.16	2.11	2.11	2.12	2.12	2.12	2.12	2.12	2.13	2.13	2.12	2.14	2.14	2.13
d(E508-Zn)[a] / Å	2.11	2.08	2.06	2.06	2.04	2.04	2.03	2.04	2.04	2.04	2.05	2.05	2.05	2.05
d(H568[Hε]-O_s) / Å	1.87	1.73	1.69	1.69	1.72	1.74	1.78	1.77	1.88	1.90	1.89	1.91	1.89	1.81
d(H568[Nε-Hε]) / Å	1.02	1.04	1.04	1.04	1.04	1.04	1.03	1.03	1.03	1.03	1.03	1.02	1.03	1.03
d(Y318-sup)[b] / Å	1.80	1.85	1.85	1.88	2.00	1.99	1.97	1.98	1.94	1.93	1.91	1.91	1.93	1.89
ω_1 (C-N_s-C_s-H_s)[c] / °	174.4	142.6	119.3	122.3	121.1	120.7	122.5	123.1						
ω_2 (C'-C_s-N_s-O_s)[c] / °	-178.7	135.8	127.0	127.6	125.4	125.1	123.3	121.8						

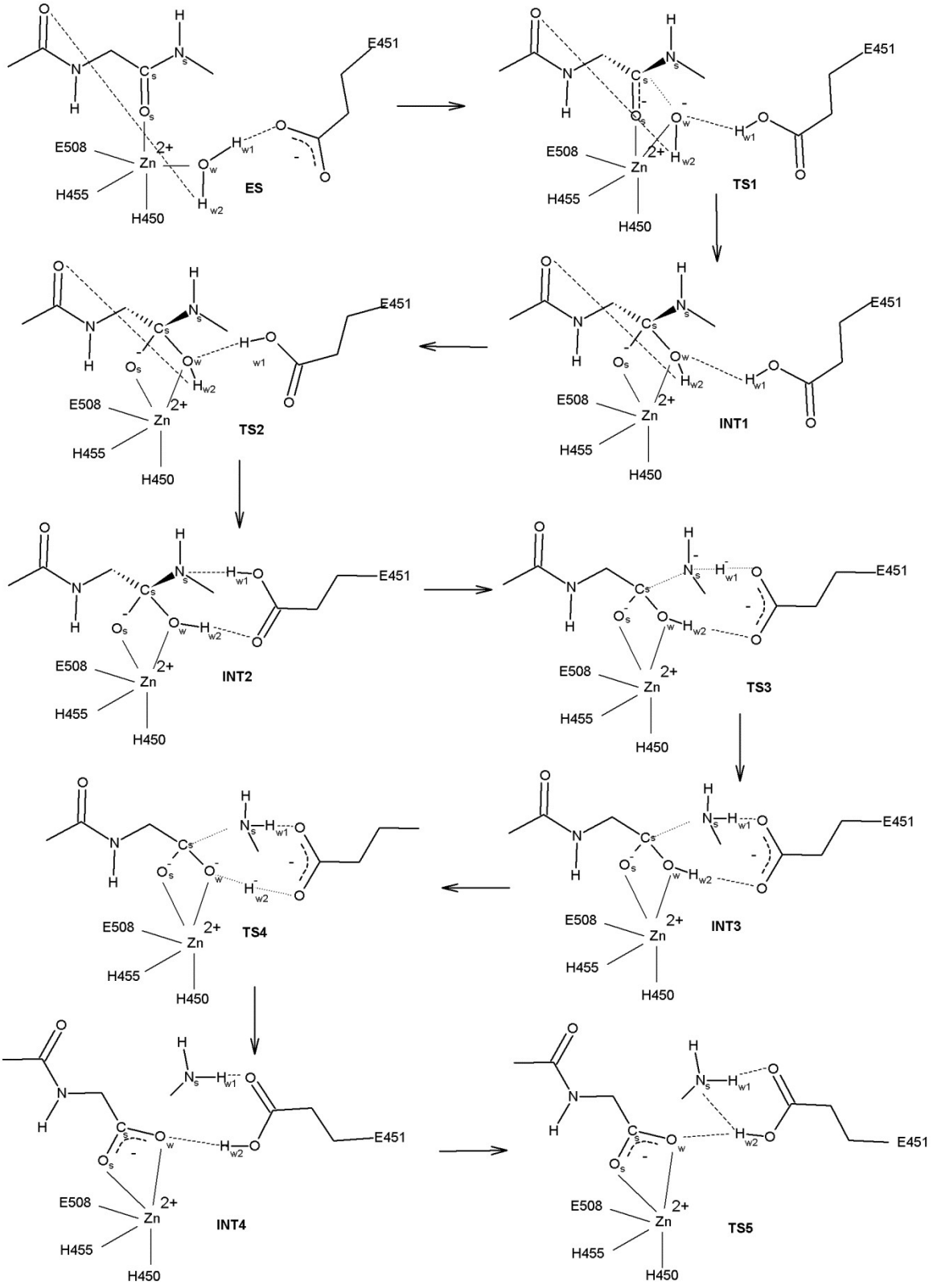
[a] distances between the zinc ion and nitrogen (Nδ) or oxygen (carboxyl) atoms from the histidine of glutamate amino acid residue, respectively

[b] distance between the oxygen atom from the Tyr318 hydroxyl group and amide hydrogen atom from the second amino acid residue from the substrate N terminus

[c] C and C’ are carbon atoms adjacent to N_s or C_s atom, respectively, while H_s is a hydrogen atom bonded to N_s

Table S2. Energy profile for breakdown of the amide bond in CH_3CO-Gly-$NHCH_3$ bound into the model of the human DPP III active site with His568 modelled in its neutral state. Calculations were performed in the presence of a solvent (dielectric constant 4) utilizing: model 1: B97D/[6-31G(d) + LanL2DZ-ECP] + $ZPVE_{\text{B97D/[6-31G(d) + LanL2DZ-ECP]}}$ and model 2: B97D/[6-311++G(d,p) + LanL2DZ-ECP]//B97D/[6-31G(d) + LanL2DZ-ECP] + $ZPVE_{\text{B97D/[6-31G(d) + LanL2DZ-ECP]}}$ levels of theory.

	ΔE / kJ mol^{-1}	
system	model 1	model 2
ES	0.0	0.0
TS1	60.8	76.5
INT1	59.9	77.1
TS2	66.5	83.7
INT2	46.8	61.0
TS3	40.9	55.4
INT3	35.9	51.1
TS4	30.0	47.1
INT4	8.3	23.5
TS5	25.2	41.1
INT5	2.9	14.3
TS6	10.0	23.1
P	−4.6	13.1
P’	−33.6	−30.2


E451
E508
Zn
2+
H455
H450
ES
TS1
INT1
TS2
INT2
TS3
INT3
TS4
INT4
TS5

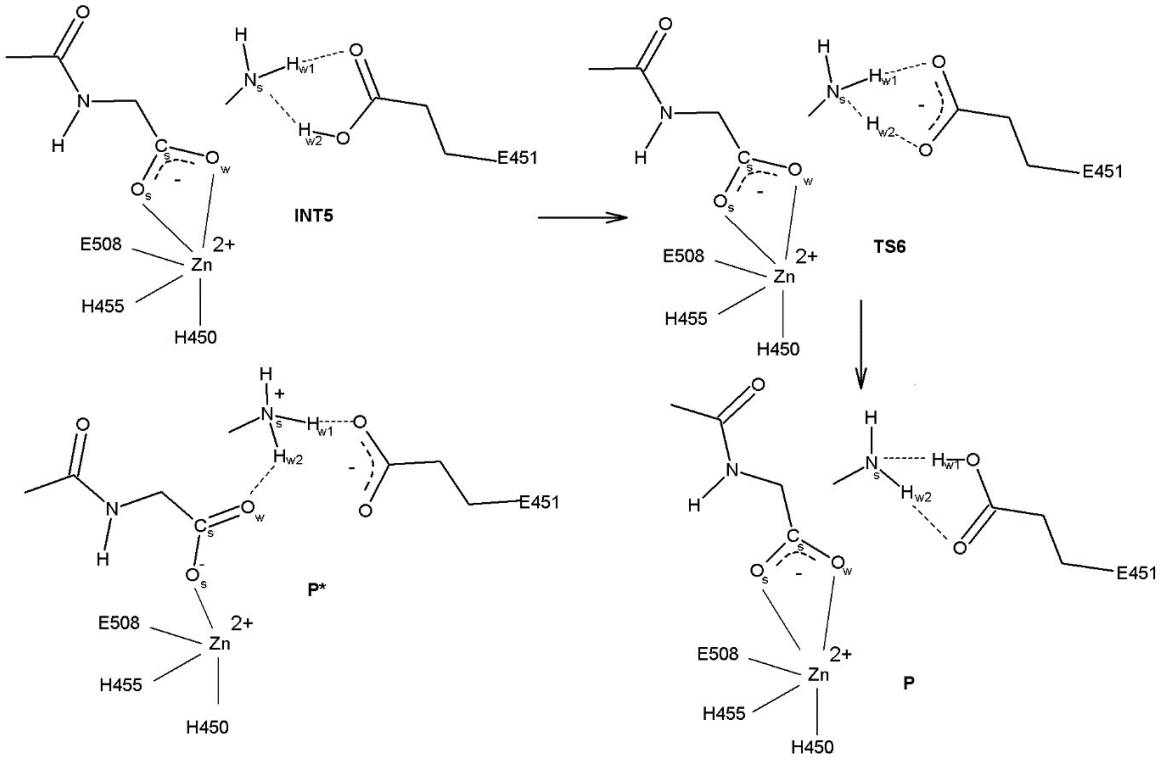
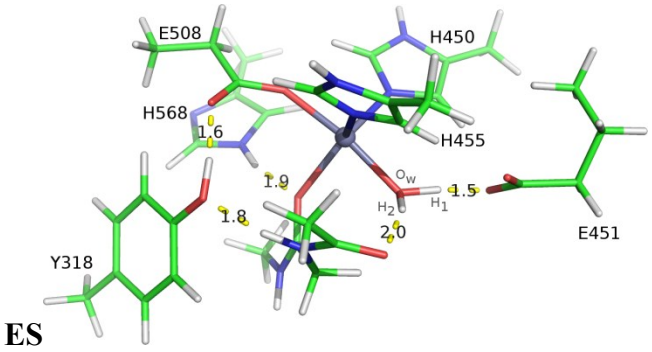
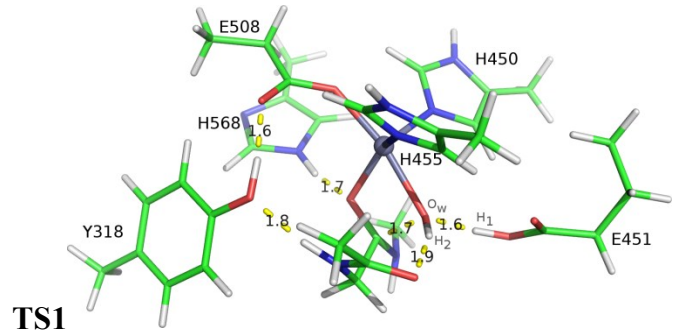


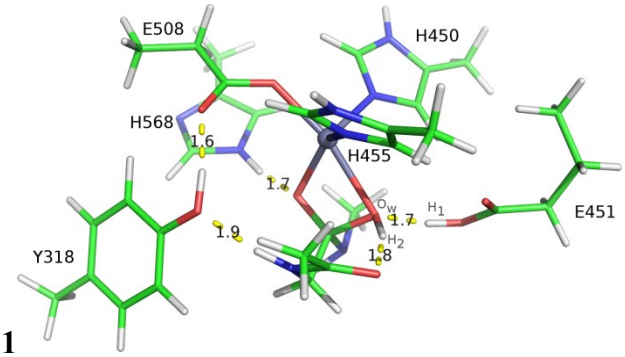
Figure S5. Reaction mechanism for breakdown of the amide bond in CH_3CO-Gly-$NHCH_3$ bound to a model of the human DPP III active site (for clarity His568 and Tyr318 are not shown) with His568 in its neutral form. Calculations were performed at the B97D/[6-31G(d) + LanL2DZ-ECP] level of theory. **ES** denotes enzyme-substrate complex, **TSn** denotes transition states 1 to 5 (TS1, TS2, TS3, TS4 and TS5), **INTn** denotes intermediates (INT1, INT2, INT3, INT4 and INT5) and **P** and **P*** denote the products. Hydrogen bonds and bonds that are either broken or formed during hydrolysis are depicted as dashed and solid black lines, respectively. Covalent and coordinative bonds are depicted as solid lines.



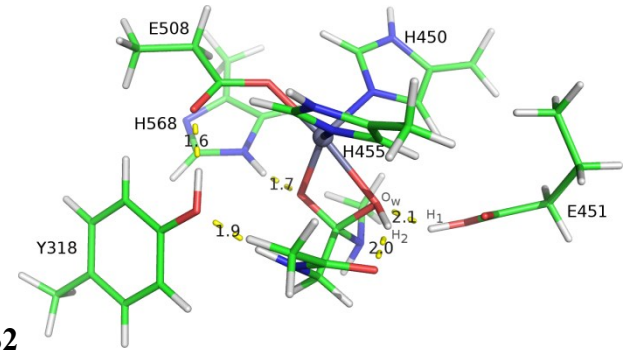
ES



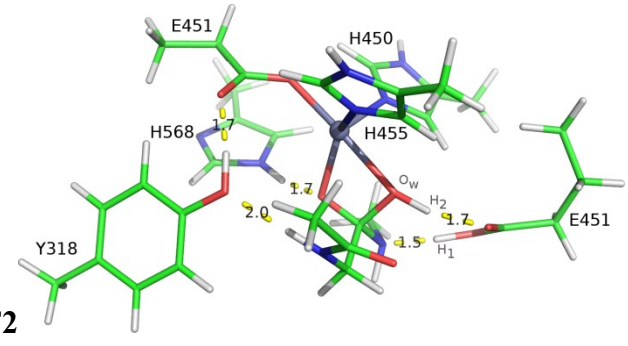
TS1



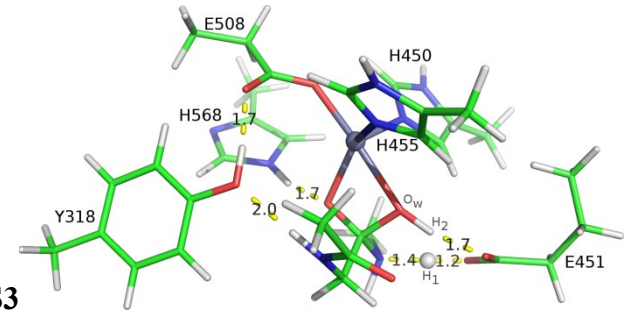
INT1



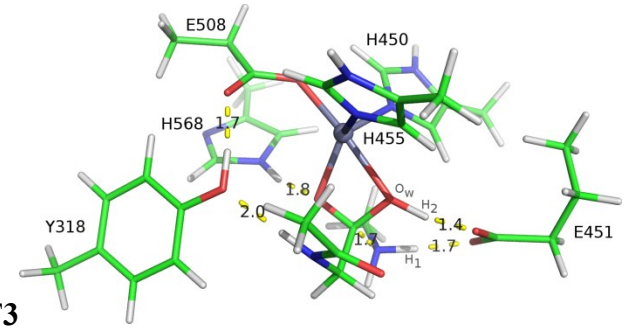
TS2



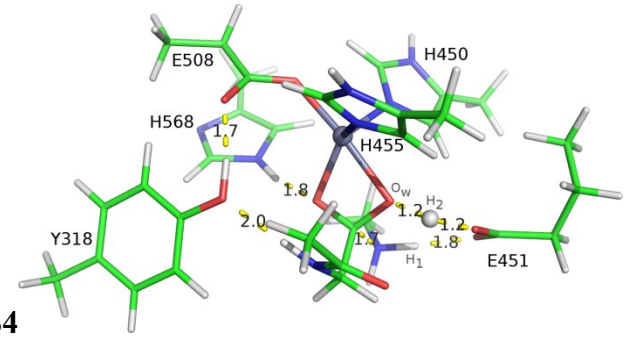
INT2



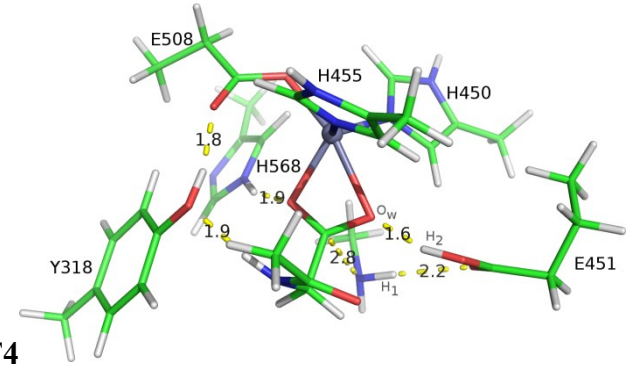
TS3



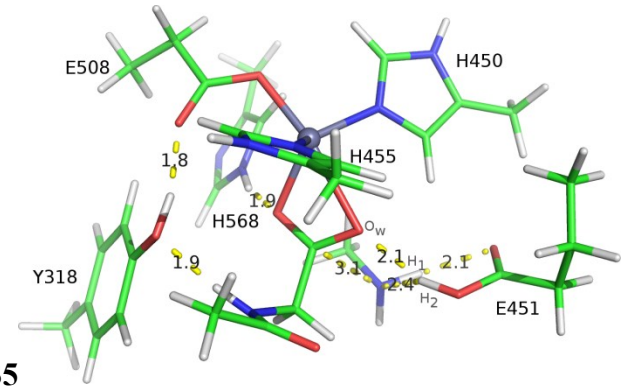
INT3



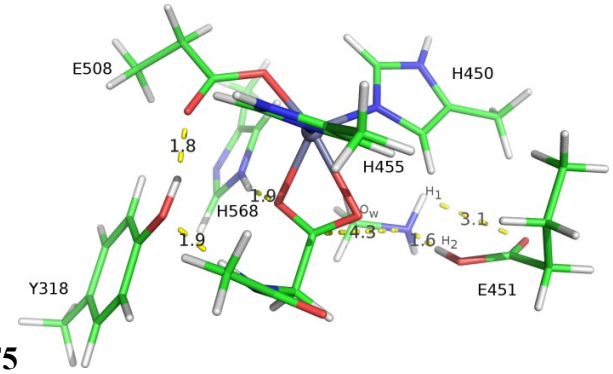
TS4



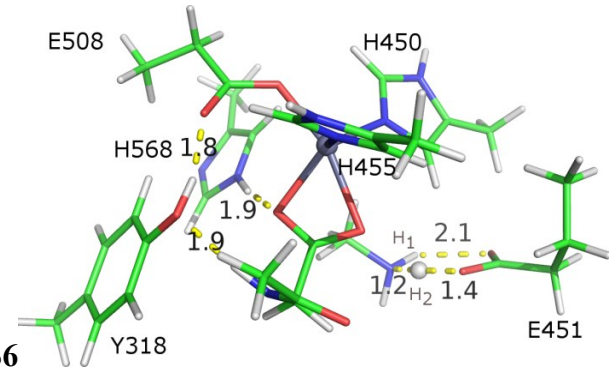
INT4


E508
H450
1.8
H455
1.9
H568
O_w
2.1
H_1
2.1
1.9
3.1
2.4
H_2
Y318
E451

TS5


E508
H450
1.8
H455
1.9
H568
O_w
H_1
3.1
1.9
4.3
1.6
H_2
Y318
E451

INT5


E508
H450
H568
1.8
H455
1.9
1.9
H_1
2.1
1.2
H_2
1.4
Y318
E451

TS6

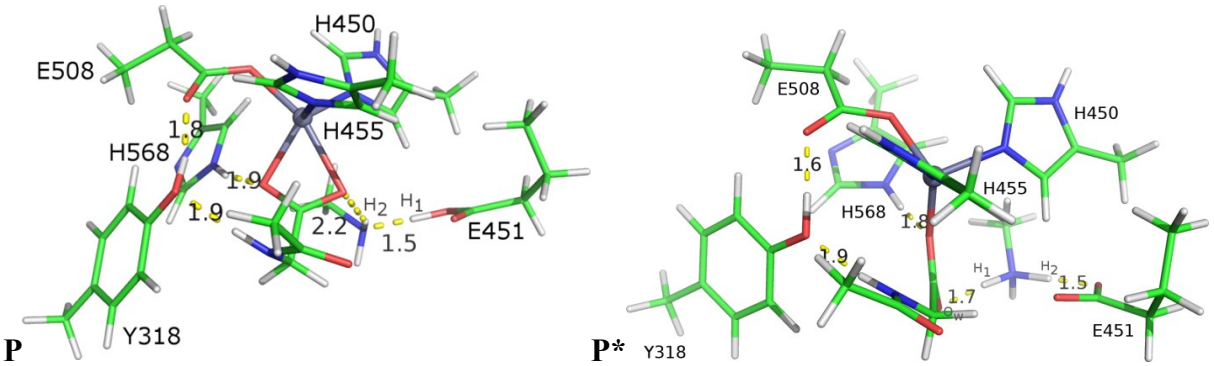


Figure S6. Optimized structures of the: model enzyme-substrate (**ES**) complex, transition states (**TS1-TS6**), intermediates (**INT1-INT5**), and products (**P** and **P***) obtained for breakdown of the amide bond in CH_3CO-Gly-$NHCH_3$ bound to a model of the human DPP III active site with His568 in its neutral form. Calculations were performed at B97D/[6-31G(d)+LANL2DZ-ECP] level of theory. Electrostatic interactions are shown with yellow dashed lines (atom pairs distance are indicated in angstroms), and coordination bonds are shown with thick gray sticks. The zinc ion is shown as grey sphere.

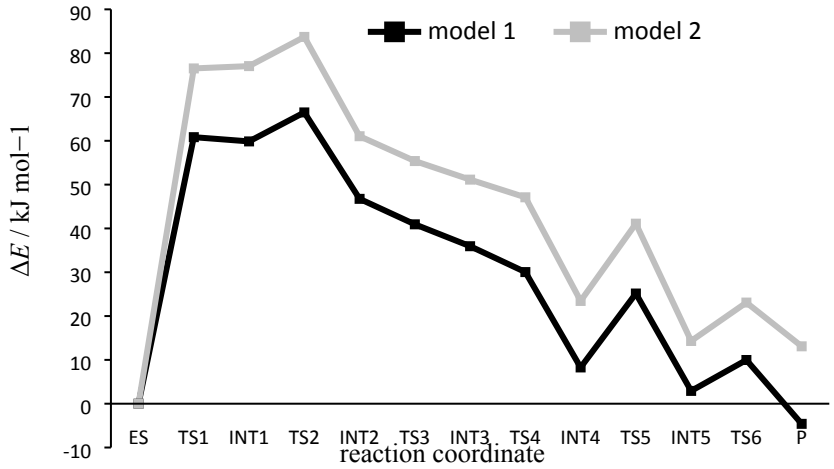


Figure S7. Energy profile for breakdown of the amide bond in CH_3CO-Gly-$NHCH_3$ bound into the model of the human DPP III active site with His568 in its neutral form. Energies are obtained using: model 1: B97D/[6-31G(d) + LanL2DZ-ECP] + $ZPVE_{\text{B97D/[6-31G(d) + LanL2DZ-ECP]}}$ (black squares and line) and model 2: B97D/[6-311++G(d,p) + LanL2DZ-ECP]//B97D/[6-31G(d) + LanL2DZ-ECP] + $ZPVE_{\text{B97D/[6-31G(d) + LanL2DZ-ECP]}}$ (gray squares and line) levels of theory. All calculations were performed in the presence of a solvent (dielectric constant 4) utilizing the polarizable continuum model (CPCM).

2) ENZYMATIC HYDROLYSIS IN THE SYSTEM WITH NEUTRAL HIS568

Two initial enzyme – Leu-enkephaline complex structure in which histidine residue (His568) was in its neutral state were built. After optimization the zinc coordination number in one of them was 5 and in the other 4 (Fig. S1 a and b). Both structures were used as initial structures for unconstrained energy surface scan, where O_w-C_s distance was decreased by 0.05 Å (with 5 optimization steps per scan) in a way to establish new chemical bond (Fig.S1 c and d). Structures corresponding to energy maxima or minima on scan profiles were optimized either in a way to find the transition state or local minimum, respectively. However all optimization procedures ended with initial enzyme-substrate complex structure.

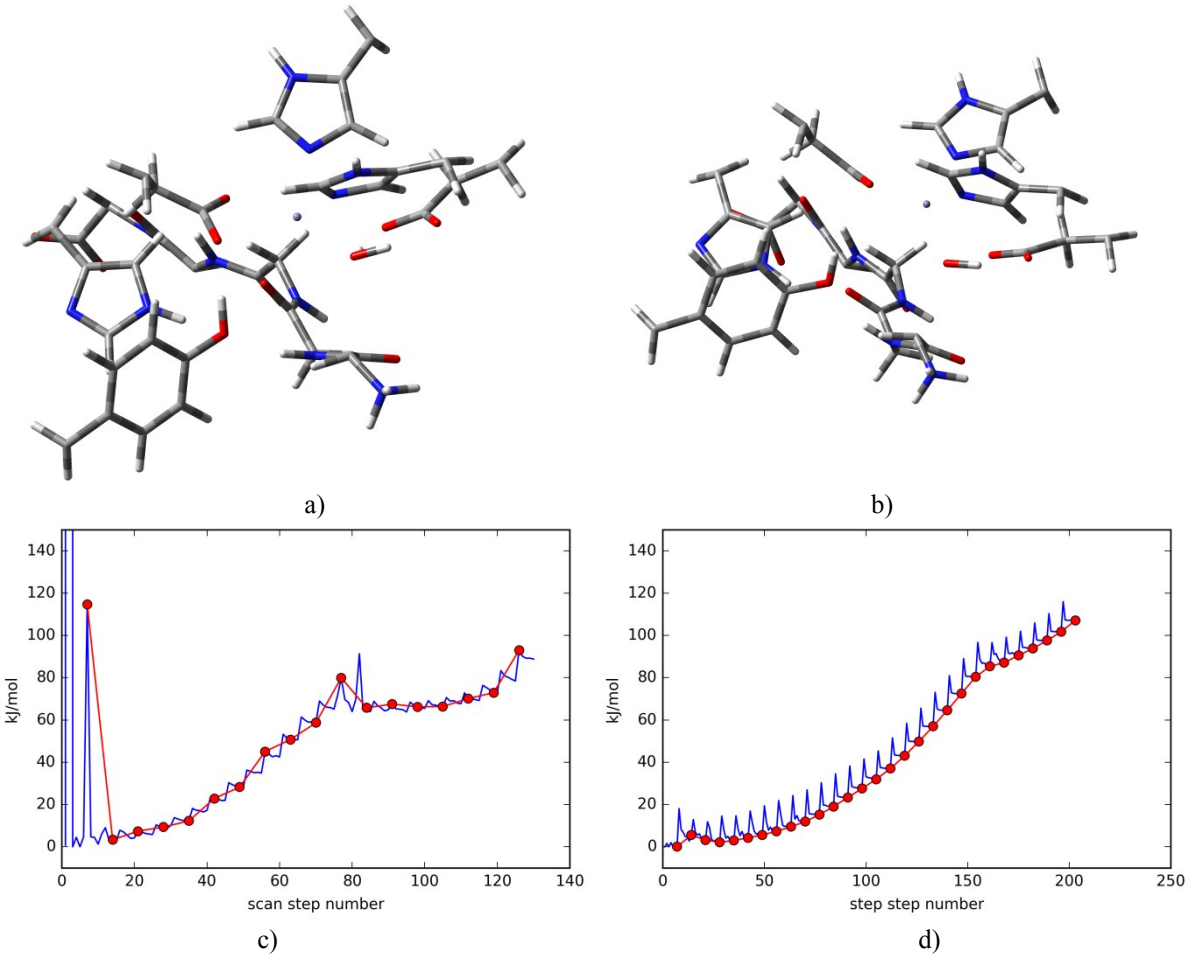


Figure S8. QM part of optimized DPP III – Leu-enkephalin complex structure with His568 in its neutral state. In both structures the zinc ion is coordinated by His450, His455, Glu508 and water molecule, while substrate molecule either: a) coordinates it, CN(Zn^{2+})=5, or b) interacts electrostatically with it, CN(Zn^{2+})=4. Energy surface scan profiles for O_w-C_s distance reduction by 0.05 Å (5 optimization steps per scan) in a way to establish new chemical bond, in which initial enzyme-substrate structure was the one shown in: c) a and d) b.

3) ENZYMATIC HYDROLYSIS IN THE SYSTEM OBTAINED DIRECTLY FROM THE MD SIMULATIONS

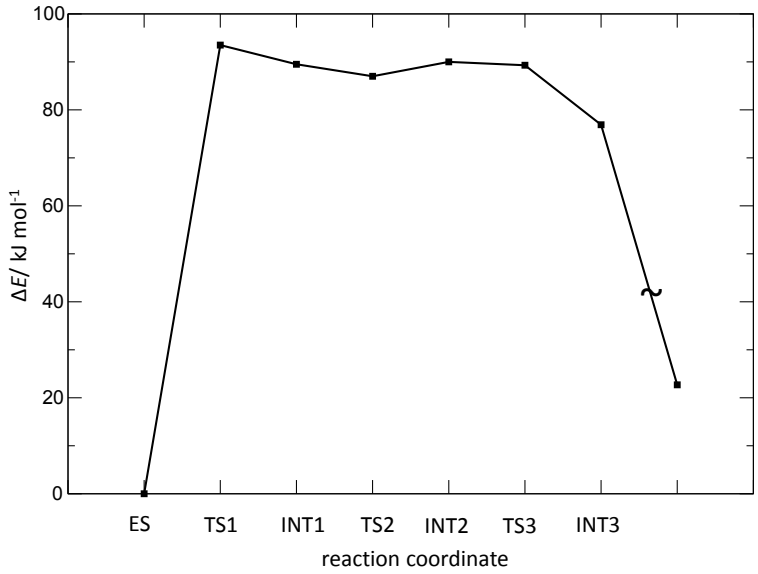
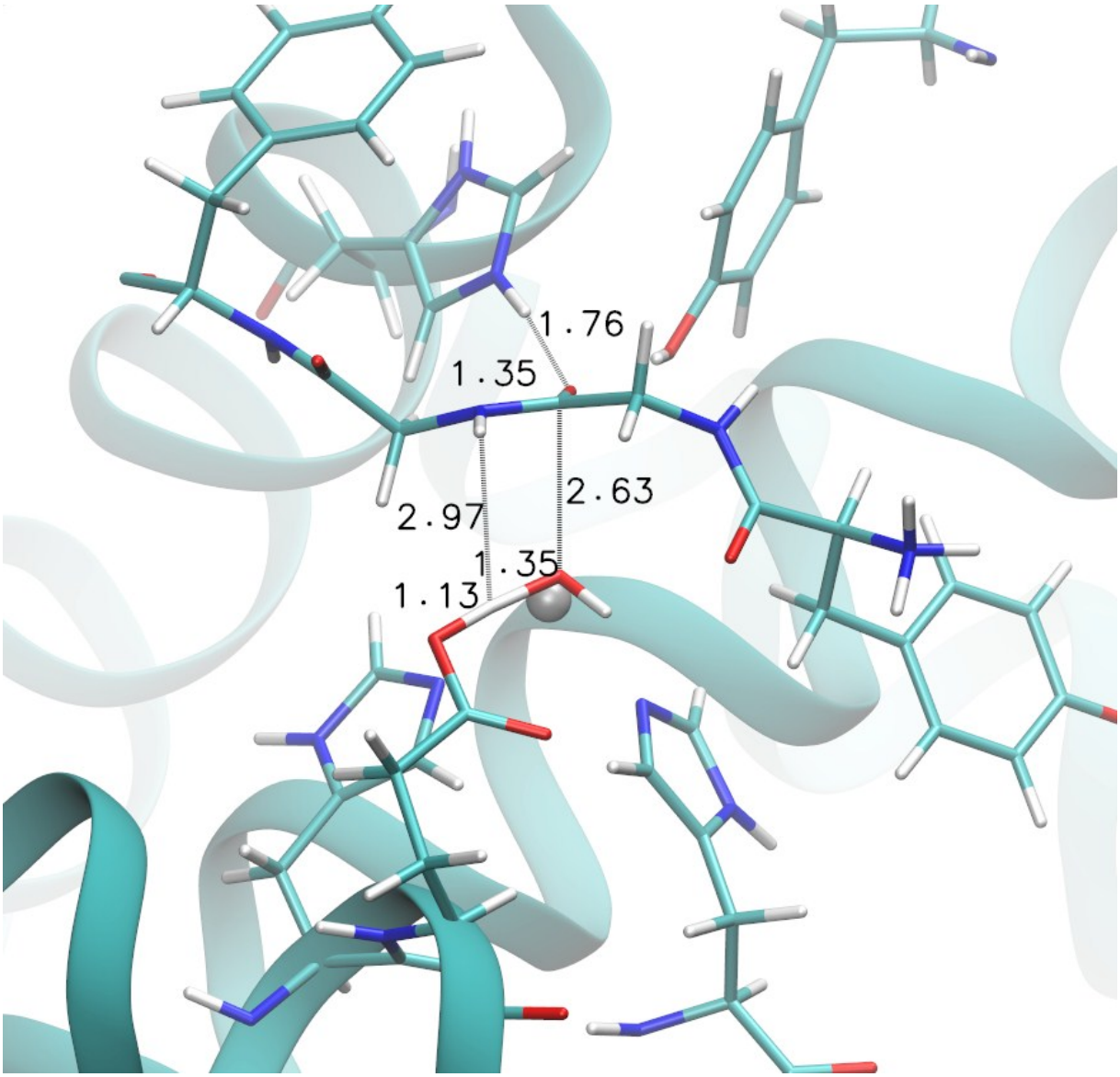
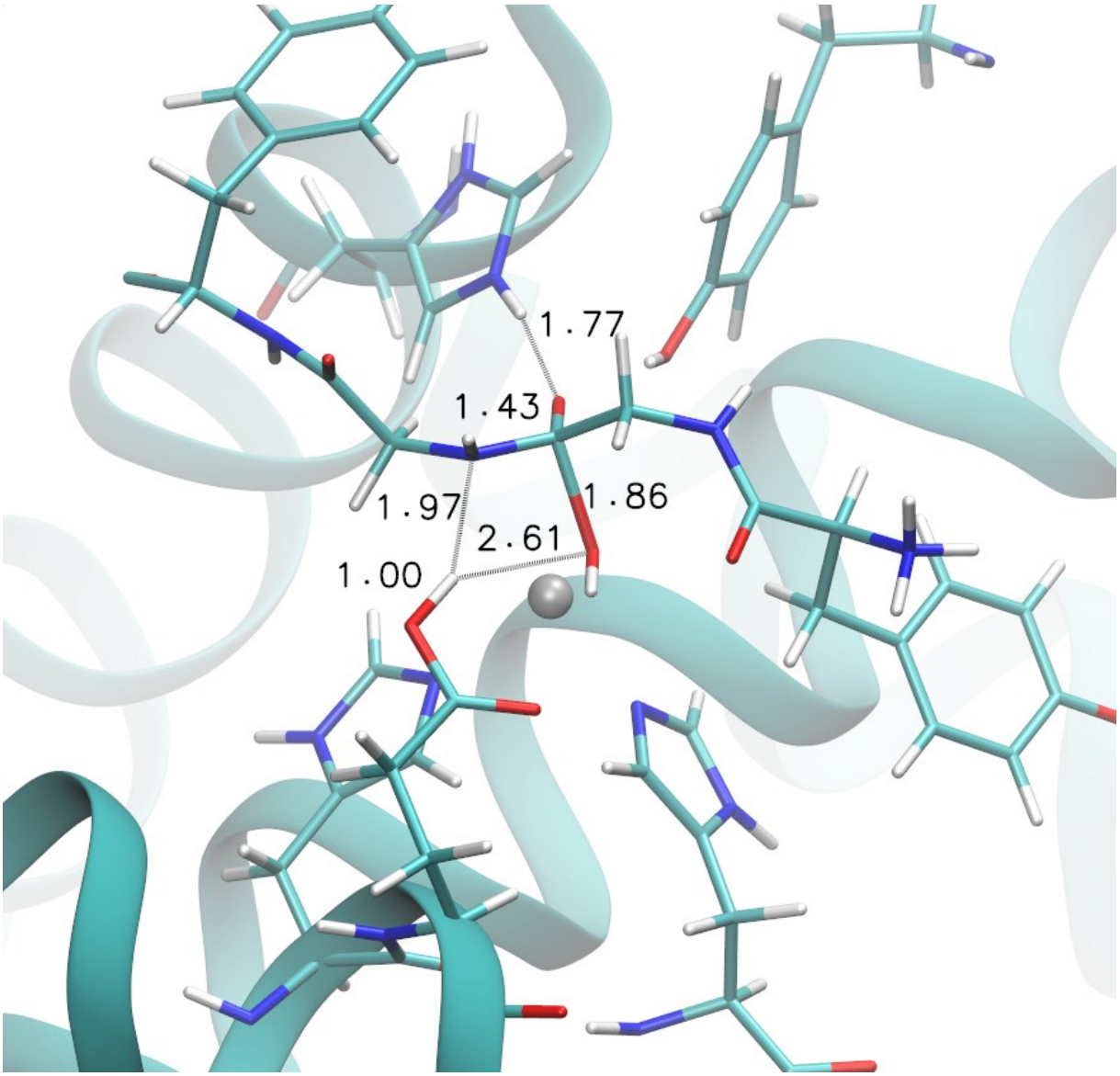


Figure S9. Energy profile for breakdown of the amide bond in Leu-enkephaline bound into the human DPP III active site. The system is gleaned through MD simulation of ES complex. Energies are obtained using: B97D/[6-31G(d) + LanL2DZ-ECP] + $ZPVE_{\text{B97D/[6-31G(d) + LanL2DZ-ECP]}}$.

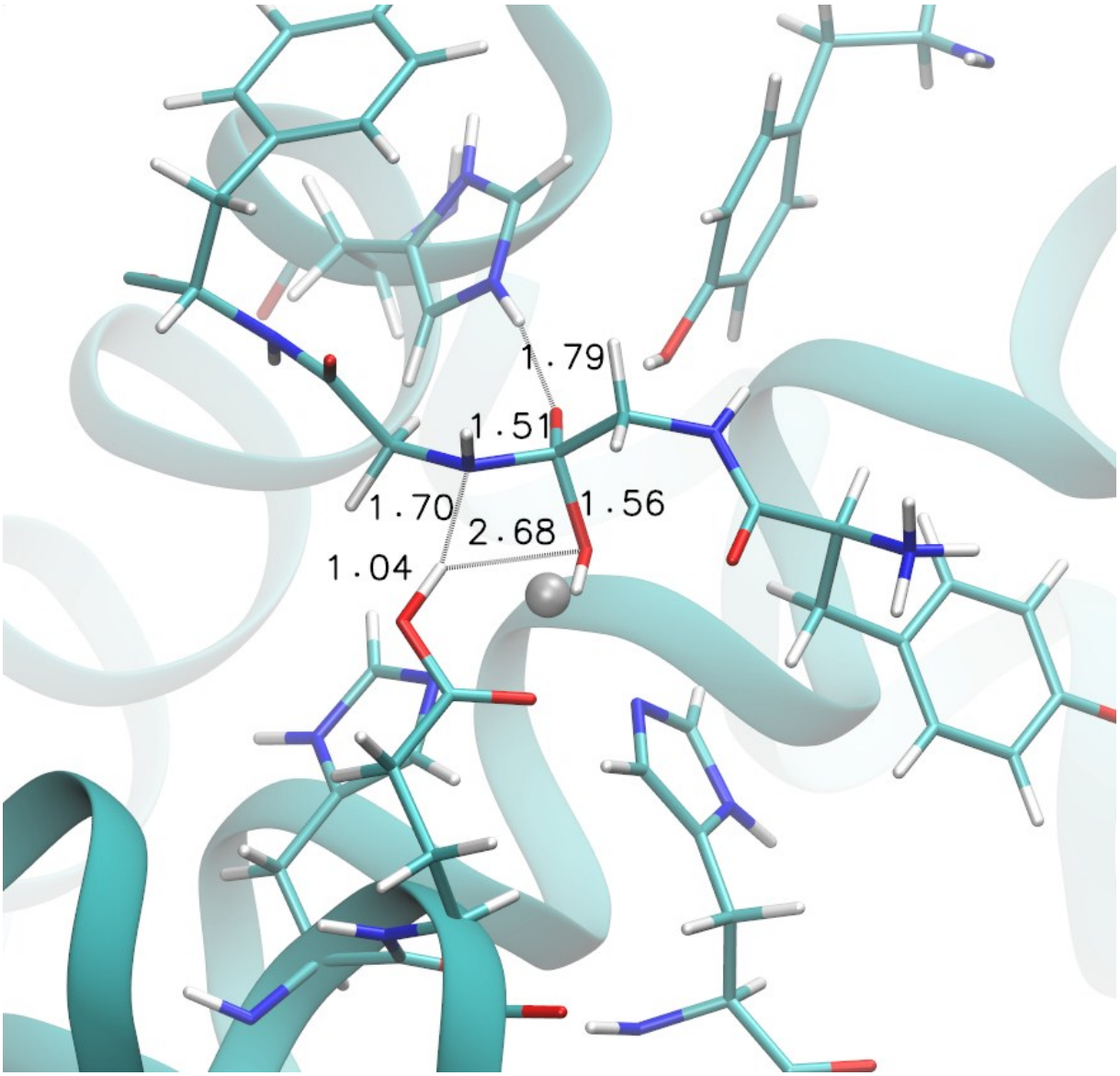
In the ES system constructed on the basis of experimental data, the first step in the reaction is concerted deprotonation of water molecule by Glu451 and nucleophilic attack of the emerging hydroxyl group on the amide carbon (C_s). The second step is N_s inversion associated with formation of H-bond with **Glu451** (*vide infra*). In the present system, which is obtained directly from MD simulations, these two reaction steps appear as a concerted process (TS1). Following intermedium (INT1) is a minimum on PES where OH is attached to peptide bond, and nitrogen atom of peptide bond has a pyramidal configuration with lone pair directed toward the **Glu451**. The transition state (TS2) involves proton transfer from **Glu451** towards N_s atom of peptide bond. The proton transfer is associated with low energy barrier which disappears upon adding zero point vibration energy. Therefore, we can assume that all three processes (addition of OH group to the peptide bond, N-inversion and proton transfer) are concerted. INT2 represents the structure where the peptide bond is almost cleaved (N_s-C_s distance is 1.67Å). Transition structure TS3 correspond to N_s-C_s bond cleavage with N_s-C_s distance of 1.76 Å. In the intermediate INT3 the peptide bond is finally cleaved considering the N_s-C_s bond distance of 2.24 Å



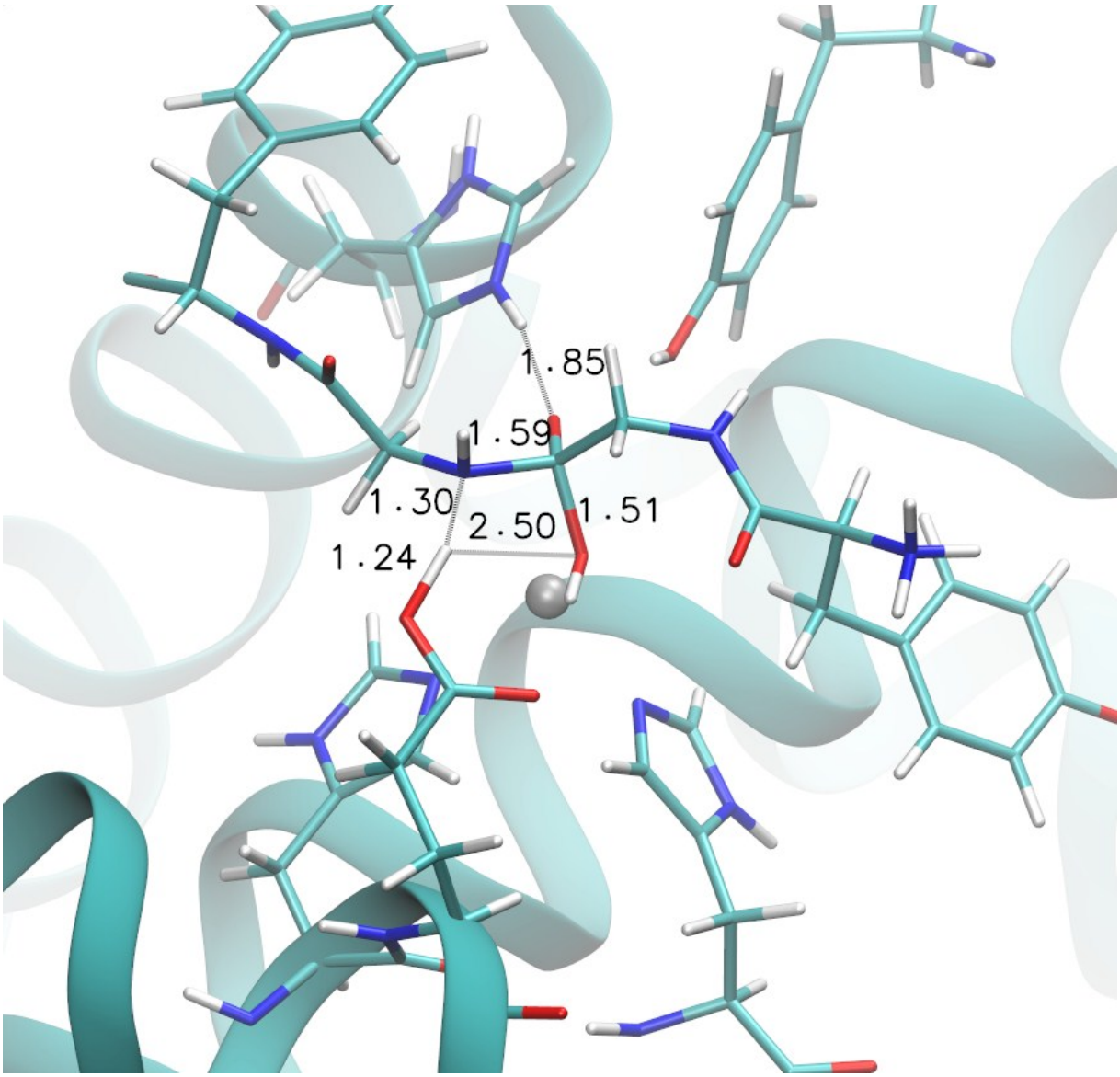
ES



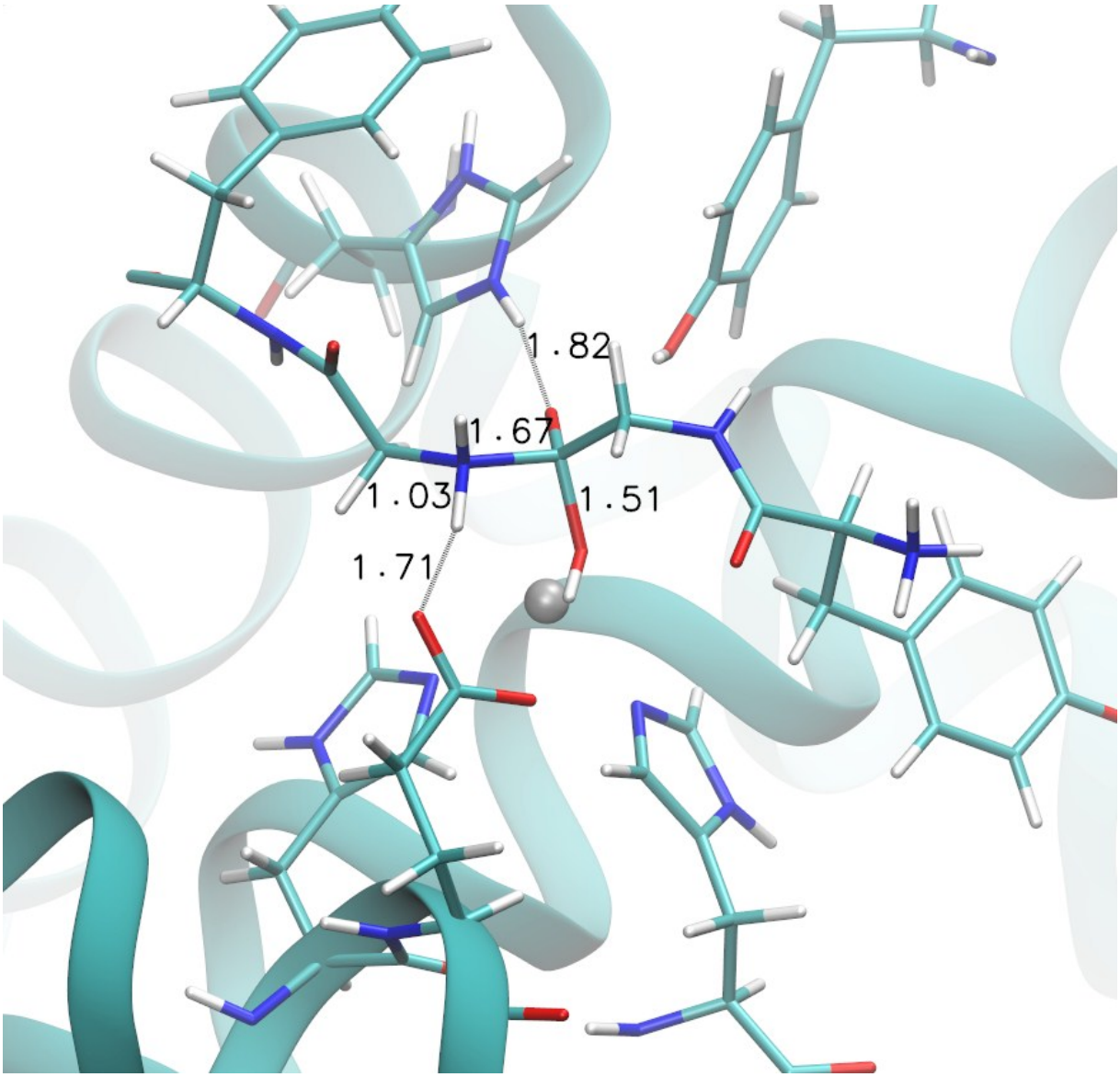
TS1



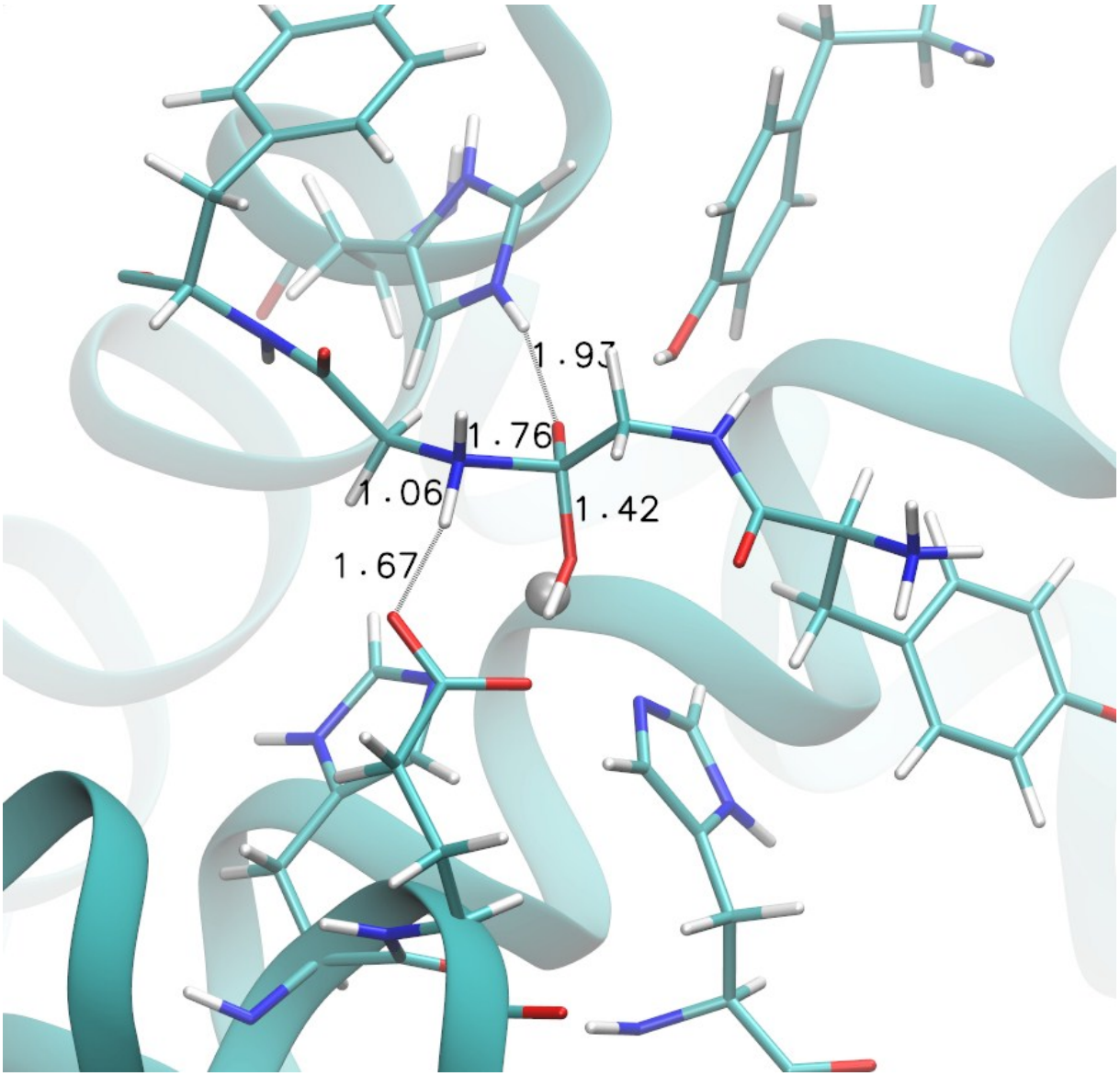
INT1



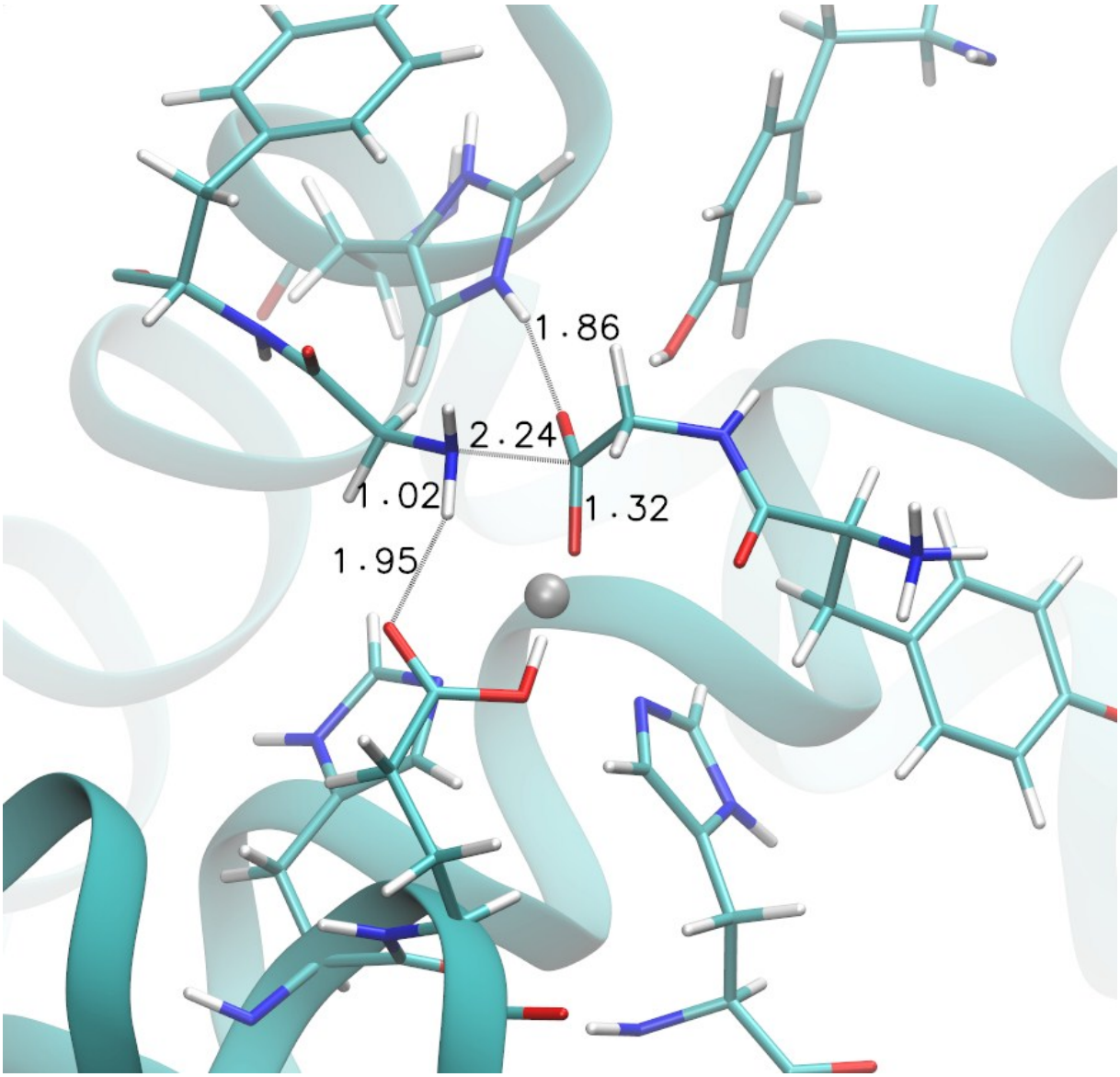
TS2



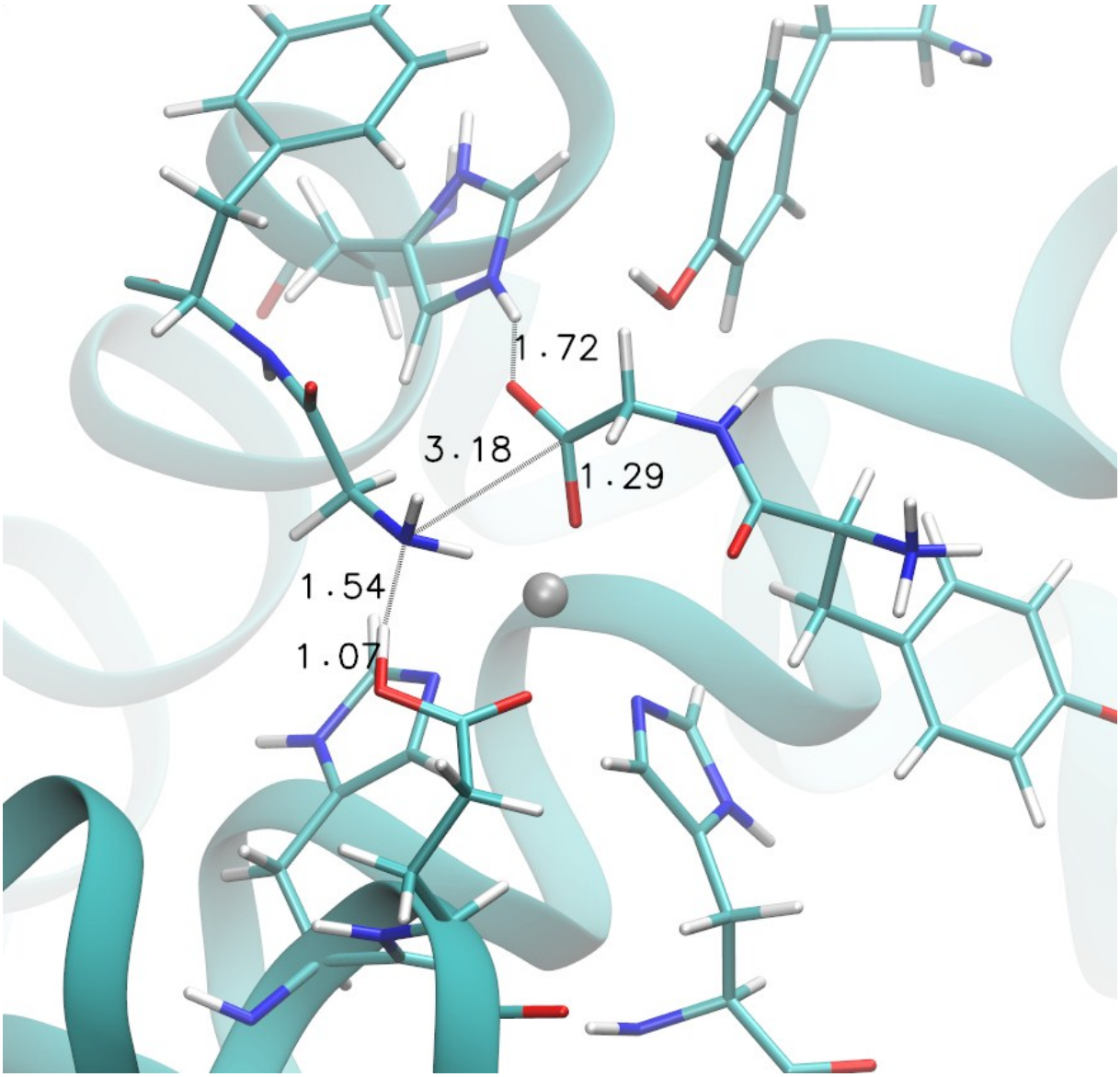
INT2



TS3



INT3

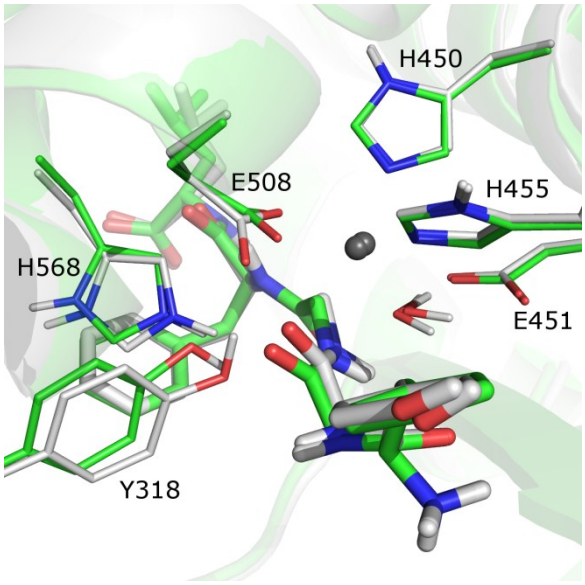


P

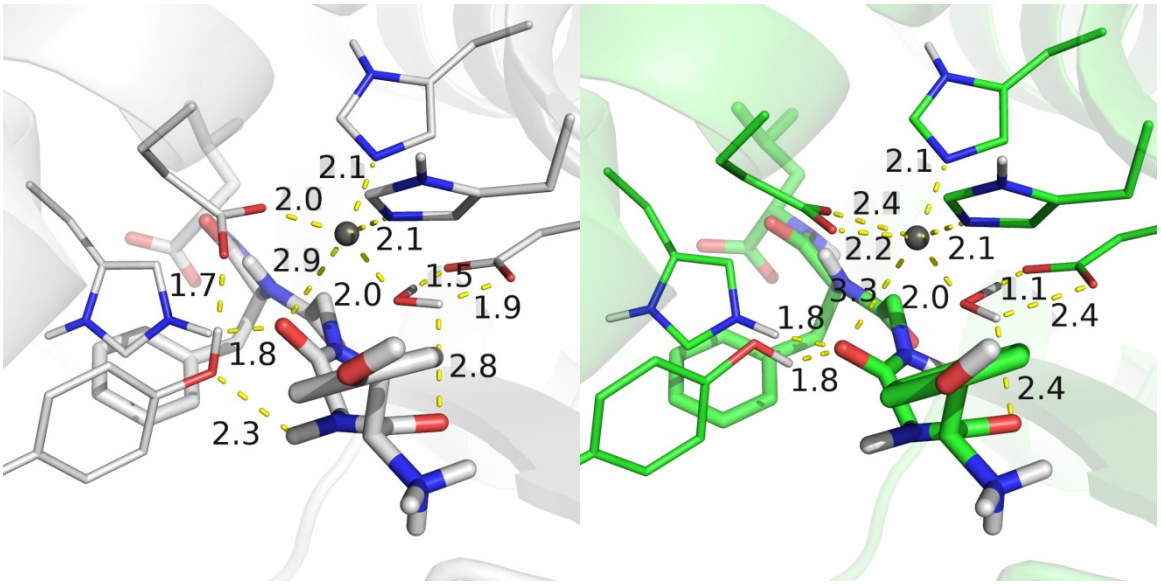
Figure S10. Optimized structures of the: enzyme-substrate (**ES**) complex, transition states (**TS1-TS3**), intermediates (**INT1-INT3**), and product (**P**) obtained by B97D/[6-31G(d) + LANL2DZ-ECP] calculations for peptide bond hydrolysis in enzyme – Leu-enkephalin complex structure obtained directly from MD simulations. His568 is in its protonated form. Amino acid residues (Y318, H450, E451, H455, E508 and H568), water and substrate that are part of QM region are indicated as sticks with carbon atoms coloured cyan. The zinc ion is shown as grey sphere. Selected atom pairs distance are indicated in angstroms

Table S3. Energy profile for breakdown of the amide bond in Leu-enkephaline bound into the human DPP III active site. bound into the human DPP III active site with His568 in its protonated form. Energies are obtained using: B97D/[6-31G(d) + LanL2DZ-ECP] + $ZPVE_{\text{B97D/[6-31G(d) + LanL2DZ-ECP]}}$.

system	ΔE / kJ mol^{-1}
ES	0.0
TS1	93.5
INT1	89.5
TS2	87.0
INT2	90.0
TS3	89.3
INT3	76.9
P	22.7



a)



b) c)

Figure S11. Overlap (a) of the Michaelis complex structures obtain from: b) the "hybrid" enzyme-substrate structure, and c) the structure taken directly from MD simulations. Leu-enkephalin is shown with thicker, while amino acid residues that constitute QM treated part of the enzyme thinner (only side chain atoms are shown), sticks (only polar hydrogens are shown). The zinc ion is show as a grey sphere. Selected bond lengths are indicated in Å.

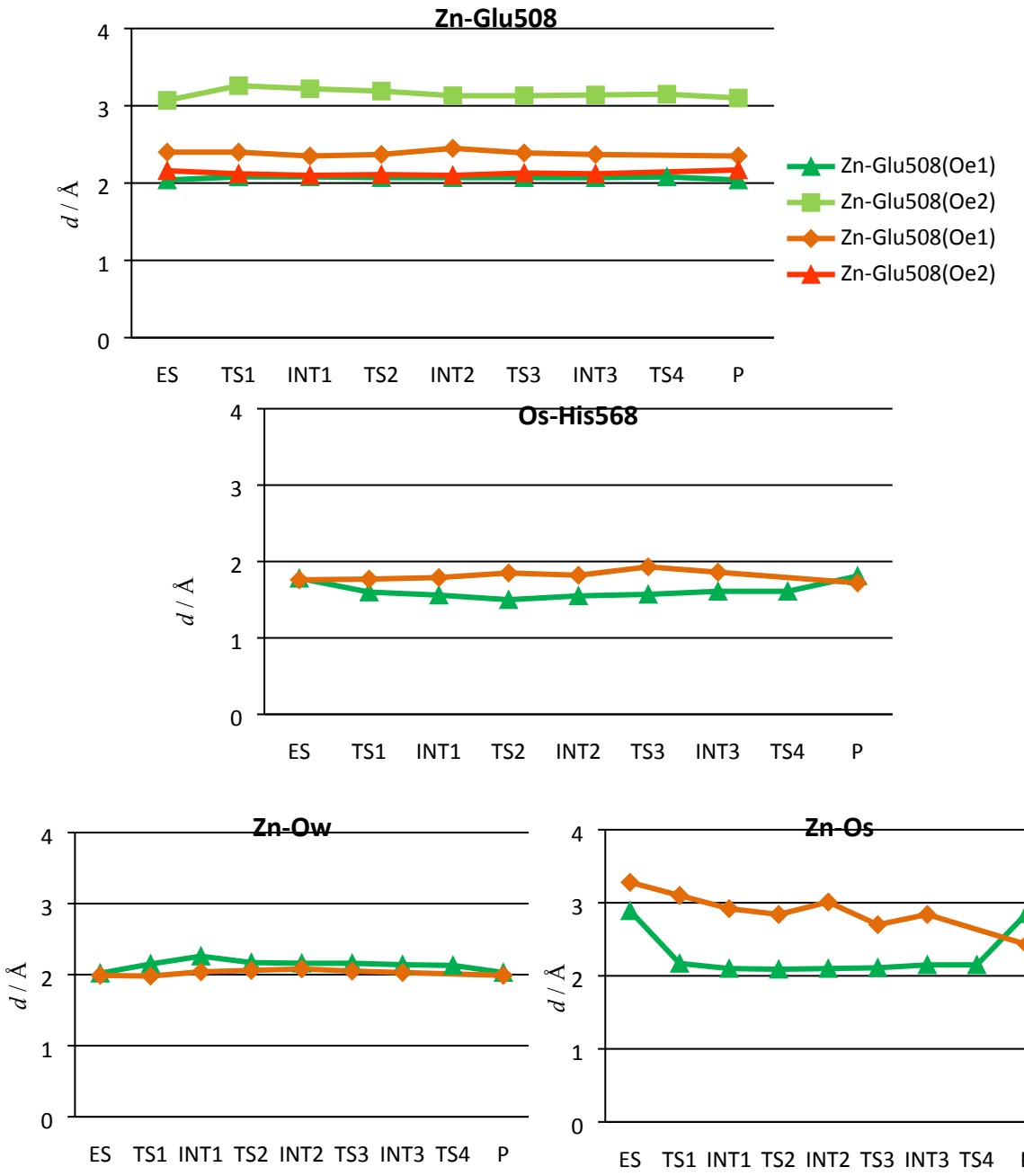


Figure S12. Selected bond lengths during peptide bond cleavage in the real systems obtained using: a) the "hybrid" enzyme-substrate structure (green) and b) the structure taken directly from MD simulations (orange), as a starting point for QM/MM calculations. One can notice that during the QM/MM calculations performed from the "MD structure" Glu508 bidentately coordinates the zinc ion (Zn-Glu508 distance) and that substrate molecule is never coordinated to the metal ion (Zn-O_s distance).