PCCP

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: A. Fekete and I. Komáromi, *Phys. Chem. Chem. Phys.*, 2016, DOI: 10.1039/C6CP06869C.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/pccp

View Article Online
DOI: 10.1039/C6CP06869C

Modeling the archetype cysteine protease reaction using dispersion corrected density functional methods in ONIOM-type hybrid QM/MM calculations; the proteolytic reaction of papain

Attila Fekete and István Komáromi*

Division of Clinical Laboratory Science, Department of Laboratory Medicine, Faculty of Medicine, University of Debrecen, Debrecen, Hungary

4032 Debrecen, Nagyerdei krt. 98., Hungary

*Author to whom correspondence should be addressed (komaromi@med.unideb.hu)

View Article Online
DOI: 10.1039/C6CP06869C

Abstract

A proteolytic reaction of papain with the simple peptide model N-methylacetamide substrate has been studied. Our aim was twofold: i) we propose a plausible reaction mechanism with the aid of potential energy surface scans and second geometrical derivatives calculated at the stationary points, and ii) we investigated the applicability of the dispersion corrected density functional methods in comparison with the popular hybrid generalized gradient approximations (GGA) method (B3LYP) without such correction in QM/MM calculations for this particular problem.

In the resting state of papain the ion pair and neutral forms of Cys-His catalytic dyad have approximately the same energy and they are separated by a small barrier only. Zero point vibrational energy correction shifted this equilibrium slightly into the neutral form. On the other hand, the electrostatic solvation free energy corrections, calculated by the Poisson-Boltzmann method for the structures sampled from molecular dynamics simulation trajectories, resulted in a more stable ion-pair form.

All methods we applied predicted at least a two elementary step acylation process *via* a zwitterionic tetrahedral intermediate. Using dispersion corrected DFT methods the thioester S-C bond formation and the proton transfer from histidine occur in the same elementary step, although not synchronously. The proton transfer lags behind (or at least does not precede) the S-C bond formation. The predicted transition state corresponds mainly to the S-C bond formation while the proton is still on the histidine Nδ atom. In contrast, the B3LYP method using larger basis sets predicts a transition state in which the S-C bond is almost fully formed and the transition state can be mainly featured by the Nδ(histidine) to N(amid) proton transfer. Considerably lower activation energy was predicted (especially by the B3LYP method) for the next amide bond breaking elementary step of acyl-enzyme formation. The deacylation appeared to be a single elementary step process by all methods we applied.

View Article Online
DOI: 10.1039/C6CP06869C

Introduction

Papain (EC 3.4.22.2) is the member of the C1 family of cysteine endopeptidases. In 1970, based partially on the previously revealed structural information,[1] Lowe proposed a cysteine peptidase mechanism[2] closely related to the serine protease mechanism. From the reviews that have been published on this topic so far (see, e.g. Refs 3, 4 and 5) it is clear that for all clans of these enzymes a sole and universally valid reaction mechanism cannot be expected.[3, 4, 5]

Both serine and cysteine proteases cleave peptide bonds through an acyl-enzyme (ester or thioester) intermediate.[4] The anionic side chain of Asp in the Asp102-His57-Ser195 catalytic triad of serine peptidases are replaced by a neutral group in the Asn175-His159-Cys25 catalytic triad of papain-like enzymes, where the numbering of the amino acids corresponds to the chymotrypsin and papain residue numbering, respectively (Figure 1). The Asn175 side chain contributes to the stabilization of the charged imidazolium form which appears during the enzyme reaction or even in the resting state. However, this contribution is less significant than the effect of the Asp102 in serine proteases.[6] The other considerable difference is that the side chain of the catalytic cysteine exists in a deprotonated (i.e. negatively charged thiolate) form in the nearly neutral pH range of about 4 - 8.[7-10]

Assuming that the active site Cys25 and His159 side chains are already in an ion-pair form in the resting state of papain-like cysteine peptidases, the expected role of the oxyanionic tetrahedral intermediate is less important (if it exists at all). In this case only the position of the formal charge centers are dislocated in contrast to the serine peptidase reaction in which the ion pair is formed. Indeed, the replacement of the Gln19 residue (which plays a role in the stability of the tetrahedral intermediate) influences only moderately the first (acylation) step of the peptidase reaction.[11, 12]

During the last, nearly four decades many theoretical studies have been carried out in order to reveal the fine details of the cysteine proteinase reaction.[13-28] Alternative reaction paths, compared to the originally proposed one, are also suggested.[16, 19, 20] In these paths the proton transfer from the imidazole ring of histidine to the nitrogen of the peptide bond is either a concerted process with the sulfur attack on the carbonyl oxygen, or can even precedes the latter. Others suggested[23, 26, 28] that cysteine peptidase active site cysteines can exist in S-H form as well. By means of QM/MM calculations Suhai and his co-workers[20] found that the tetrahedral intermediate structure is stable only when the amide N is also protonated from the His159 imidazolium ring.

The protonation state of the Cys-His catalytic dyad in different protein environments is also a subject of intensive studies.[29-32] In certain cases it is straightforward to assume that the surrounding water shell is explicitly involved in the proton transfer between the cysteine and histidine side chains.[32] The effect of the substrate on the proton transfer barrier and/or the equilibrium between the neutral and ion pair (zwitterionic) states has been revealed by theoretical methods as well.[30] Benchmark QM/MM calculations on the protonation site of the Cys-His dyad have also been carried out and compared to those obtained using continuum solvent approximation and/or gas phase calculations.[33]

Calculations carried out on simple model systems in solution[21, 22] showed that the hemi-thioacetal anion that corresponds to the anionic tetrahedral intermediate at serine proteases is not a stable intermediate. Whereas, Ma et al.[24] have shown that for cathepsin K on the potential energy surface of the acylation step a local minimum corresponding to the

View Article Online
DOI: 10.1039/C6CP06869C

tetrahedral intermediate exists while such intermediate was not found for the deacylation process. Then again Wei et al.[25] have found recently that the first (acylation) phase is a single elementary step process while during deacylation a tetrahedral intermediate is formed.

In the calculations on cysteine peptidases the highest level of quantum mechanical theories were the DFT B3LYP or B-LYP methods using valence double- or triple zeta quality basis with polarization and diffuse function for the energy (see e.g. references 20, 25-27).[20, 25-27] For geometry optimization usually lower levels of theories were applied. While considering dynamics in understanding enzyme catalysis has an increasingly important role,[34] in QM/MM dynamics simulations the level of theory applied for the QM subsystem, in general, should be even lower. Therefore, static QM and QM/MM calculations using sophisticated QM methods (and proper partitioning of the system) still have a great importance.

The DFT methods are ideal choices as a QM component of QM/MM methods since they can predict comparable geometries, relative energies and even activation energies to those that can be obtained from high level post SCF methods, at fractional cost. However, density functionals mentioned above are suffering from the improper treatment of the dispersion effects and the long range exchange term is also imperfectly implemented in those functionals. [35, 36, 37] While dispersive interactions at large internuclear separations is originated from electron correlation, van der Waals interactions at shorter separations (typically at "van der Waals bond" distances) have non negligible contribution from the long range exchange interaction as well.[37]

Currently, a number of DFT methods are available in which London dispersion energy is taken into account.[35, 36, 38, 39] The London energy can be simply and still satisfactorily calculated by adding pair-wise C_6/r^{-6} terms with suitable dumping function. Such terms are applied e.g. in the ωB97X-D[40] and B97-D[41] functionals. In ωB97X-D functional a long range electron exchange term is also included which can explain its excellent performance on geometry and energetics of van der Waals complexes.[37] Another way is to use functionals which are highly parameterized and the parameters are optimized on data set including large number of data from noncovalent interactions. The M06-2X[42] functional with 35 adjusted parameters is a typical example for this latter case. In this sense, even though it does not contain direct (e.g. C_6/r^{-6}) term for dispersion interaction, it can still be regarded as a (medium range) dispersion corrected density functional method. In addition, the M06-2X functional perform equally well for activation energies and transition state geometries too. [43,44,45, 46] It was shown that even a simple pair-wise empirical dispersion correction in the functional can considerably improve the calculated reaction barriers for biochemical reactions in DFT based QM/MM calculations.[47, 48] Recent papers[45, 49] have demonstrated clearly that correct reaction barriers and transition state geometries and even correctly predicted asynchronicity require careful selections of the density functional method and basis set combination.

In the present work, our main aim was to explore the applicability of a few dispersion corrected DFT methods as a QM component of ONIOM [50] type calculations and to compare them to a representative popular method (B3LYP) without such correction. We were interested in whether modern DFT theories can predict different geometries or relative energies for intermediates and transition states (i.e., alternative reaction pathways) or simply confirm those ones that were proposed from the earlier calculations. This aim can be justified by the diversity of cysteine proteinase reaction and by the fact that existence of some intermediates are not fully proven. As an example for the latter one is the "tetrahedral intermediate" since the form (anionic or zwitterionic) and even its existence on the potential energy surface are not fully proven from previous computations. In order to prove whether the

View Article Online
DOI: 10.1039/C6CP06869C

stationary points on the QM/MM (ONIOM) potential energy surface correspond to a local minimum or a first order saddle point (transition state) the second geometrical derivatives were calculated[50] which can be done in the recent Gaussian packages even at electronic embedding approximation.[51] This is still not common for such large QM/MM systems. Instead of extending the size of the QM subsystem we applied numerous levels of theories on a model system embedded in the whole protein. We also carried out potential energy surface (PES) scans along the assumingly most significant atomic movements which accompany the enzyme reaction in order to find transition states and local minima on these PESs.

Computational details

Starting from a structure available in the protein data bank (PDB ID: 1PPN)), short (250 ns) molecular dynamics simulations on papain were carried out. Both the neutral and ion pair forms of its catalytic dyad were considered, in order to reveal how the geometries of the catalytic sites change during simulation. In the constant particle number, constant pressure and constant temperature (NPT) simulation dodecahedral periodic box, TIP3P explicit water model[52] and 150mM ionic strength (set by Na+ and Cl- ions) were used. The FF99SB force field[53] with Berendsen barostat[54] and v-rescale thermostat[55] were applied in these calculations, which were carried out by the GROMACS software.[56] For the short range electrostatic and van der Waals energy terms 10Å cut-off distances were used, while the particle mesh Ewald (PME) method[57] was applied for long-range electrostatic energy corrections.

The initial QM/MM model was constructed also from the 1PPN X-ray structure of papain. Because no significant deviation from the starting (optimized) catalytic center geometry was observed in dynamics simulations with the exception from thermal fluctuation, the AMBER 12 sander[58] optimized structure was used as the starting geometry in all the ONIOM calculations instead of a selected frame from the dynamics trajectory. At first, the whole structure was truncated that way so that only the catalytic center and its neighborhood (Val16-Ala30, Ser60-Trp69, Val130-Ala136, Lys156-Ala162, Ile173-Gly178, Thr204-Phe207) were kept, which comprised all the amino acid residues known to have considerable influence in the enzyme reaction. During truncation always whole amino acids were kept. The truncated N- and C-terminal peptide bonds were closed by acetyl or N-methyl caps, respectively. The orientation of the N-H and C=O bonds in these "caps" corresponded to those peptide bonds that were cut previously. This way, the first model system with the N-methylacetamide (NMA) substrate consist of 780 atoms. The NMA substrate has been put to the active site by means of the xleap module of the AMBER package. In the new structure constructed this way the QM ("model") subsystem comprise the side chains and C_α atoms of His159 and Cys25 residues, as well as the substrate model NMA, when it was considered. During the optimizations and potential energy scans the Gln19-Gly-Ser-Cys-Gly-Ser-Cys25, Gly62-Cys-Asn-Gly-Gly66, Asp158-His159, Asn175-Ser-Trp177 residues and the NMA were allowed to move while all the other atoms were fixed.

First, calculations were carried out on this truncated system. When all the transition states and local minima were identified, their geometries were fitted back to the whole protein by applying the VMD[59] package, and the calculations were repeated. In this paper, only the results corresponding to the whole protein and the NMA (Figure 1) will be discussed.

For the QM ("high level") methods the B97-D [41], M06-2X [42], ωB97X-D [40] and B3LYP[60, 61] functional forms and the standard Pople style 6-31G(d,p), 6-31+G(d,p), 6-311+G(d,p) and 6-311++G(d,p) basis sets were applied. For the MM ("low level") method the AMBER force-

View Article Online
DOI: 10.1039/C6CP06869C

field implemented in the Gaussian packages was used. Electronic embedding approximation was applied throughout the ONIOM QM/MM calculations, i.e. the QM methods were polarized by the surrounding MM partial charges through the modified $1e^-$ Hamiltonian.[51] The standard link atom approach using hydrogen atoms to saturate dangling bonds was applied for covalent bonds at the QM-MM borderline region.[50, 51] At all local minima and transition states we obtained the second geometrical derivatives were calculated at the corresponding level of theory. Here we report only ONIOM type QM/MM calculations. Therefore, mentioning the QM method should be regarded as the abridgement of the full QM/MM calculations.

It is generally accepted that polar solutes (like proteins) in polar solvent (e.g., water which contains even charged particles) the electrostatic part of the solvation is the dominant one[62] that can be satisfactorily estimated by the Poisson-Boltzmann method.[63, 64] Therefore, the electrostatic solvation free energies were calculated for the protein taking into account both the neutral and the ion-pair forms of the catalytic triad. Thousand frames from the last (equilibrated) 100 ns simulations were selected equidistantly and, for each structure, the electrostatic solvation free energy were calculated by means of the DelPhi[65] software. For the calculations the Amber FF99SB force field partial charges and the standard particle sizes provided with the DelPhi package[65] were applied. The averaged solvation free energies for both the neutral and the ion-pair forms can be considered as a further correction to the zero point corrected ONIOM energy values obtained at different levels of theories.

The conversion between the PDB and Gaussian input file formats was done by the TAO package.[66] The ONIOM type QM/MM calculations were carried out by using the Gaussian 03[67] and Gaussian 09[68] software packages. The Chimera[69] software was used for molecular graphics representations. The potential energy surfaces (PESs) for the first step of acyl-enzyme formation were generated along the (Cys)Sγ…C(NMA carbonyl) *vs.* the (HisNδ)Hδ…N(NMA amide) bonds, as well as along the NMA amide N-C *vs.* both the (NMA)N-Hδ(HisNδ) and (Cys)Sγ…C(NMA carbonyl) bonds using 0.05Å adjoining grid point distances. PESs for the deacylation steps were also scanned along the (water)O… C(NMA carbonyl) *vs.* the (HisNδ)…H(water) bonds using the same grid point distances. The surface plots were generated by the DPlot[70] software. Only the B3LYP and M06-2X methods with 6-31G(d,p) and 6-31+G(d,p) basis sets in ONIOM were used for surface mapping.

Results and Discussion

The resting state of papain

The trajectories of the 250 ns molecular dynamics simulations on free papain with both the ion-pair (zwitterionic) and the neutral forms of the Cys-His catalytic dyad indicate that these systems are equilibrated during the first 100-150 ns time frame (Figure S1). The simulations also show that despite the significant fluctuation in the (Cys25)Hγ and the Nδ(His159) distance in the neutral form of papain, the shortest values (1.8 Å -2.0Å) that correspond to a usual H-bond distance are still significantly populated (Figure 2A). This suggests that the proton transfer from the (Cys25)Sγ to the Nδ(His159) can occur even without any explicitly contributing water molecule, therefore, such water molecule(s) was(were) not considered in our calculations. On the other hand, the already transferred proton also resides with high probability in a position which corresponds to the (His159Nδ)Hδ…Sγ(Cys25) hydrogen bonded distance (Figure 2B).

View Article Online
DOI: 10.1039/C6CP06869C

Analyzing the trajectories, significant differences can be noticed between them regarding the number of water molecules being in contact with the Cys25Sγ atom. While the neutral Cys25Sγ can have close contact typically with 0-3 surrounding water molecules (Figure 2C), the ion pair form Cys25Sγ can make at least 2-4 such close contacts (Figure 2D). This suggests the possibility that the solvent contributes to the stability of these forms in different degrees. Interestingly, apart from the thermal fluctuation (Figure 2A-2B), the geometry of the Cys-His dyad and its proximity, either in their (His159)Nδ…HγSγ(Cys25) or in (His159)NδHδ…Sγ(Cys25) hydrogen bonded forms, have considerable similarity to the starting X-ray structure.

The existence of the (His159)Nδ…Hγ-Sγ(Cys25) H-bond connections during the whole simulation leads to the plausible assumption that in static ONIOM calculations we can start from the optimized H-bonded structure. The results of our calculations on the substrate-free papain are summarized in the Table 1. In all cases, the ONIOM QM/MM calculations without zero point energy correction resulted in approximately the same energy values for the neutral and ion pair (thiolate and imidazolium) forms of Cys25-His159 side chains. The energy difference between the calculated energies are not greater than 0.8 kcal/mol. The ωB97X-D predicts the ion pair form the most stable one meanwhile the other methods showed mixed results. The M06-2X values are closer to those ones obtained by ωB97X-D method. Interestingly, the B97-D and B3LYP resulted in very similar energy differences between these states with the same basis-set dependence for the bases we used in these calculations. It is worth noting here that some earlier calculations showed also the ion pair form more stable than the neutral ones.[13, 19, 20]

Applying the zero point energy correction, the neutral form becomes the most stable one in all kind of calculations we carried out. Nevertheless, the corrected energy difference between the two characteristic states remain small, less than 2 kcal/mol for all cases. For the ion pair → neutral cysteine-histidine side chains conversion the methods we applied predicted lower ZPVE corrected energy for the transition state than the corresponding ion-pair endpoint. It formally means that the ion-pair spontaneously turns into its neutral form even at 0 K and without any extra (activation) energy.

These zero point corrected theoretical energy values are in conflict with experiments[7-10, 71-78] which demonstrated that the ion pair state is more populated. However, from these experiments certain (albeit significantly less than 50 %) probability could be assigned to the co-existence of the neutral form or, it is supported that, at least, such neutral forms cannot be excluded.[8, 74, 77] From this observations one can conclude that the (free) energies of the ionic and neutral forms should be close to each other.

The importance of solvent molecules on the relative stability of such an ion pair is demonstrated by sophisticated QM/MM computations that were carried out on cathepsin, which is homologous to the papain. It was shown[27] that the ion pair form was the most stable one only when their interactions with the surrounding water molecules were also considered.

Therefore, we have performed electrostatic solvation free energy calculations on both the ion-pair and neutral cysteine-histidine side chains as well in order to estimate the effect of solvation. The averaged values from the last 100 ns simulation the PB calculations predicted larger electrostatic solvation energy term for the ion pair form than for the neutral one by about 20 kT (i.e., about 11.9 kcal/mol at 300K). These solvation (free) energy values from the PB computations can be used as a further correction to the corresponding zero point corrected ONIOM energies. The correction is large enough to reverse the relative stability of ion pair

View Article Online
DOI: 10.1039/C6CP06869C

and neutral forms listed in Table 1 and predicts the previous one to be the most stable one again in all our cases. While the method we used for correction differs substantially from that one used by Mladenovich et al.,[27] our results also underline the importance of the solvent in the stabilization of the ion pair form.

The characteristic geometry parameters of the proton transfer pathway from the thiol group to the imidazol ring (i.e. the Sγ-Hγ and the Hγ-Nδ distances corresponding to d1 and d2 in Scheme 1A) are also shown in the Table 1. The data show that all the density functional methods applied in ONIOM resulted in very similar values for the S-H and imidazol neutral side chain pair. Note, however, that the B97D method predicted significantly shorter Nδ-Hγ distance and a little longer Sγ-Hγ distances, i.e., a stronger Nδ...Hγ-Sγ hydrogen bond. The importance of the size of the selected basis was negligible. It is noteworthy, that the B97D method predicted the longest S-H distances. The values obtained for the transition states spread to a little larger range.

The d1 and d2 distances in the neutral- and ion pair forms, as well as in the transition state associated with their conversion are in good agreement with the published theoretical values for the papain N-Acetyl-Phe-Gly-4-nitroanilide enzyme substrate complex.[25] It should be emphasized that the d1 and d2 distances for the ion-pair are in good agreement with the averaged values obtained from sophisticated QM/MM dynamics calculations[27] while the corresponding values for the neutral form significantly different from those in Ref 27. The latter discrepancy can be explained by the more flexible nature of the neutral form resulting in larger deviations, i.e., larger averaged values compared to the static equilibrium ones. The larger neutral (His159)Nδ…Hγ-Sγ(Cys25) H-bond distance fluctuation can be observed in our simulations as well (Figures 1A-1B)

Acyl-Enzyme formation

The main parameters of our results on the first cycle of the cysteine protease reaction of papain collected in Tables 2 and 3. According to our calculations a (zwitterionic) tetrahedral intermediate exists, therefore, the transition state energies compared to the energy of the Michaelis complex, tetrahedral intermediate as well as to the energy of products are shown in Table 2. Four characteristic distances (see, in Scheme 1B) were listed for this particular acylation process. These are shown at each of the stationary points (Michaelis complex, tetrahedral intermediate, product and the two transition states between them). In contrast to these results, carrying out B3LYP/6-31G(d) reaction path modeling which was augmented with B3LYP/6-31++G(d,p) energy calculations and pseudo-bond free energy estimation in the framework of QM/MM method, Wei et al. proposed a one elementary step mechanism for the crucial acyl enzyme formation without any stable (tetrahedral) intermediate state.[25]

Regarding the Michaelis complex geometry in Table 2, the bond length parameters demonstrated only marginal dependence on the levels of theory at which they were derived. On the other hand, the non-bonding distances represented more characteristic method dependencies. The longest values for both the Sγ-C(peptide carbonyl) and (His Nδ)H-N(peptide amide) distances were derived by B3LYP methods. The reason for the elongated distances probably is that the B3LYP method does not include proper terms for long range interactions while the other methods do. The most glaring non-bonding distance differences can be observed with the B3LYP and M06-2X methods. Comparing our characteristic distances to those that were published recently by Wei et al.[25] for the Michaelis complex of papain and N-Acetyl-Phe-Gly-4-nitroanylide, the chemical bond lengths are similar to each other, although their reported amid bond length (1.37 Å) is a little longer than ours. Similarity

View Article Online
DOI: 10.1039/C6CP06869C

between the published [25] and our present non-bonding distances can be also observed with the exception of the significantly longer (3.7 Å) Sγ-C(peptide carbonyl) separation calculated by Wei et al.

The papain-NMA Michaelis complex is stabilized by the H-bonds between the NMA carbonyl oxygen and the Gln19 side chain amid group as well as the Cys25 backbone amid H (Figure 3). In the proper orientation of the Gln19 side chain the Trp177 and Gly23 residues play a significant role. It should be mentioned that the imidazole ring of the His159 and the (Cys25)S atom are approximately in a common plane.

From Table 2 it is immediately apparent, that the tetrahedral intermediate corresponds to a formal zwitterionic structure, i.e. a single O atom is connected to the tetrahedral carbonyl C atom and the same time the amide group is protonated. This is in contrast to the anionic TI structure obtained from calculations carried out for serine proteases.[79] This implies that not only the sulfur attack on peptide carbon but also the proton transfer from the histidine to the peptide (amide) nitrogen take place in the same elementary step. However, comparing the d1 (Sγ-C(carbonyl), d3 (amide(N)-Hδ(His)) and d4 (H-Nδ(His)) distances at the transition state it is evident that the first elementary step of the acylation process is not fully synchronous. Interestingly, the B3LYP method using polarized Pople-type split shell double- and triple zeta quality basis sets augmented with diffuse functions either on heavy or on both heavy and hydrogen atoms predict different type of asynchronicity than all the other methods we applied. In the former case, the d1 (S-C) distance is close to the value that should exist in the anionic (serine proteinase analogue) TI structure and the d3 and d4 distances (i.e., the proton transfer) feature intermediate (transition state like) values. It means, that the transition state resembles to the structure which connects an already formed serin-protease-like tetrahedral intermediate to zwitterionic one. By all the other methods the transition state is mainly determined by the S-C bond formation and the H-N(imidazolium) distance is close to the value which was observed in the Michaelis complex. It should be underlined that the proton transfer neither in these cases precedes the S-C bond formation. The B97-D methods predict significantly longer S-C distances for the transition state structure than all the other levels of theories we applied (independently from the quality of basis sets we used). The same trend was observed, although in a less extent, for the tetrahedral intermediate structure as well.

The transition state(s) and the TI are also stabilized dominantly by the same H-bonds which were observed at the papain-NMA Michaelis complex. Significant difference can be observed, however, in the orientation of the imidazole ring of the His159 residue (Figure 4). In contrast to the roughly parallel orientation of the NMA amid bond and the imidazole ring in the Michaelis complex, the imidazole ring turned away from its original position and now the Nδ-H bond points to the N atom of the NMA in the TS. The imidazole ring of the His159 and the (Cys25)S atom are no longer in a common plane.

The B3LYP/6-31G(d,p) method has an exceptional feature among the methods we used in ONIOM calculations. At this level of theory an additional transition state can be localized on the pathway which connects the Michaelis complex and the zwitterionic tetrahedral intermediate (Table 3). The TS1a is similar to those transition states which were found at the dispersion corrected density functionals. However, in this case it leads to the intermediate which resembles the anionic (serine protease-like) tetrahedral intermediate. The next (separate) step on the potential energy surface is the formation of the zwitterionic tetrahedral intermediate from the anionic structure. Interestingly, the sole TS on the Michael complex → zwitterionic tetrahedral intermediate pathway found by B3LYP methods using larger basis sets roughly correspond to the anionic intermediate → zwitterionic intermediate transition.

View Article Online
DOI: 10.1039/C6CP06869C

The possible acyl-enzyme formation pathways are summarized in Scheme 2, where the middle row is valid only when B3LYP/6-31G(d,p) is the high level method in ONIOM calculation.

Regarding the energy values in Table 3 it is immediately apparent, that on the pure (i.e., not ZPVE corrected) potential energy surface the two transition states (TS1a and TS1b) and the (oxy)anionic tetrahedral intermediate between them have almost the same energies. The zwitterionic intermediate is considerably more stable than the anionic intermediate. Applying ZPVE corrections the barrier, which separates the anionic- from the zwitterionic tetrahedral intermediates disappears. It means that in our case the barrier has only theoretical importance.

From Table 2 and Table 3 one also can notice that the amide(peptide) bond breaking energy barrier (TS2), which results in the acyl-enzyme product, is predicted to be significantly lower by the B3LYP method than by all the other functional form–basis set combinations we used. On the other hand, it implies that depending on the chemical structure of the substrate, regarding the chemical environment and even the optimization methods applied, these tetrahedral intermediate/intermediates simply do not exist or can be easily missed. It can be the explanation why Wei et al. [25] found the acylation step to be a single elementary step by the B3LYP functional in QM/MM calculation, in contrast to our present results.

The potential energy surface (PES) we have derived can explain why the B3LYP/6-31G(d,p) (Figure 5A) and M06-2X/6-31G(d,p) (Figure 5B) methods predict different TS structures and also why it was extraordinarily difficult to find the transition state on the B3LYP/6-31G(d,p) (and all the other B3LYP) surfaces. On the B3LYP/6-31G(d,p) PES there are two transition states, TS1a at d1≈2.7Å and d3≈1.8Å, and TS1b at d1≈2.26Å and d3≈1.47Å. The first one connects the Michaelis complex to the "classic" anionic intermediate, while the TS1b connects the anionic intermediate to the zwitterionic intermediate. On the extremely flat B3LYP PESs region, the position and even the existence of transition state(s) can be influenced by small geometry perturbations or changes in the basis sets. In contrast, the M06-2X PES (Figure 5B) has a much more characteristic saddle point. Because geometry optimization is performed at each grid point with the exception of the grid variables, only the transition states and their proximity were mapped and, e.g., the Michaelis complexes can be found outside of the map. Nevertheless, unconstrained geometry optimization from the grid points leads either to the Michaelis complex or to the product with the exception of the B3LYP/6-31G(d,p) PES as expounded above.

It should be stressed that there was no spontaneous C-N (amide) bond breaking in the substrate during the PES scan. However, during the potential energy scan (see, details above) only the (amid)N…Hδ(His159) and the (Cys25)Sγ…C(amid carbonyl) distances are scanned and the C-N (amid) bond was only optimized. In order to examine how other choices of the variables can influence the potential energy, similar scans along the C-N (amid) and the (Cys25)Sγ…C(amid carbonyl), as well as the (amid)N…Hδ(His159) distances also were carried out. The results of these mappings are shown in the supplementary Figure (Figure 2S). Interestingly, during the (amid)N…Hδ(His159) and C-N (amid) scan the (Cys25)Sγ…C(amid carbonyl) bond spontaneously forms at short (amid)N…Hδ(His159) distances and the transition state (amid)N…Hδ(His159) geometry parameters estimated from these surfaces (Figures S2C and S2D) are approximately the same as we obtained from the transition state search based on the surface scan expounded previously. While this is not an exact and unambiguous proof of the acylation reaction mechanism(s) we proposed above, it can be regarded as a further support for them. It should be also mentioned that the C-N (amid) and (Cys25)Sγ…C(amid carbonyl) as well as the (amid)N…Hδ(His159) and C-N distances scans

View Article Online
DOI: 10.1039/C6CP06869C

show substantially weakened C-N (amid) bond at the B3LYP/6-31G(d,p) level compared to those we obtained at the M06-2X/6-31G(d,p) level (Figures S2A and S2C *vs.* S2B and S2D, respectively). It is also a possible explanation of a single elementary step for acyl enzyme formation observed by Wei et al.[25] from QM/MM molecular dynamics simulation.

Hybrid QM/MM calculations carried out by Harrison et al.[19] showed much perfect synchronicity between the Sγ attack and the proton transfer. Meanwhile Wei et al.[25] obtained a TS where the proton transfer, the Sγ-C(peptide carbonyl) bond formation and the peptide bond breaking of the substrates are approximately synchronous processes. The discrepancy between their results and the results presented here can be related to the different levels of theories, different computational models and the different QM/MM partitioning schemes. As explained above, all these factors can essentially influence the position and even the existence of the transition state at the B3LYP level of theory Wei et al. used in QM/MM computation.[25] The second transition state structure which connects the tetrahedral intermediate state to the product (i.e., acyl-enzyme structure with already cleaved amide bond) is featured a considerably elongated C-N distance (with ~2Å or even longer bond length) for the scissile amide bond. Notable exceptions can be observed using the B97-D method with all but the smallest (6-31G(d,p)) basis set. It should also be underlined that a significantly smaller activation energy barrier predicted for the zwitterionic tetrahedral intermediate → acyl-enzyme reaction step by the B3LYP methods than all the other methods applied in our ONIOM calculations.

For the product (i.e., the acyl-enzyme plus the methyl-amine (cleaved model "C-terminal part of the peptide") all methods we applied predict very similar results for the bond lengths. The notable exception is the slightly longer S-C distances predicted by B97-D methods (independently from the basis sets we applied) compared to all the other levels of theory. In general, the B3LYP method predicts the largest values for the interatomic separations examined here while the other DFT methods, especially the M06-2X, predict significantly smaller values. The reason probably is that, as mentioned above, the exchange-correlation functional applied in B3LYP method is not a proper one for the long-range electron-electron exchange interaction, and it does not have a proper long range correlation term. Comparing the values listed in Table 2 to those reported by Wei et al. [25] a remarkable similarity in bond lengths is observed. Nevertheless, Wei et al. [25] obtained significantly shorter d4 distances (2.09 Å) which indicates a stronger H-bond between the N-terminal amide hydrogen with histidine Nδ atom.

Regarding the activation energies, the general conclusion derived from our results is that for the acyl enzyme formation the first elementary step (i.e., the formation of the zwitterionic tetrahedral intermediate) is the rate determining step. The calculated activation (ZPVE-corrected) energies are in the range of ~10-13 kcal/mol. All methods predict the zwitterionic TI less stable than the Michaelis complex. From the tetrahedral intermediate state, the second transition state (leading to the acyl-enzyme product) can be much more easily accessed than the first one leading back to the Michaelis complex. The calculated total energy of the system consisted of the cleaved product and the papain was found to be higher than the papain Michaelis complex comparing the (TS2-TI)-(TS2-AE) *vs.* (TS1-TI)-(TS1-MC) energy differences from Table 2. This suggests that the reaction is shifted toward the reactant. It should be considered, however, that in the real reaction, the amine group (i.e., the N-terminal-end of the released peptide/protein sequence in real cases) can dissociate from the papain and can be protonated which should shift the reaction equilibrium into the product direction.

View Article Online
DOI: 10.1039/C6CP06869C

Owing to the quite extended set of parameters used to parameterize the exchange and correlation functionals of the B97-D, ωB97X-D and M06-2X methods it is difficult to explain the origin of differences observed between the calculated energies and geometries. However, the general observation is that for this particular reaction the B3LYP method predicts remarkably longer non-bonding interatomic separation which is probably caused by the missing long-range terms in the exchange–correlation functional of B3LYP method. Nevertheless, all the corresponding calculated distance values, especially those that represent "chemical bonds" have remarkable similarities.

Acyl-enzyme hydrolysis

The next process on the amide (peptide) hydrolysis is the acyl hydrolysis reaction. We modeled it by adding a simple water molecule to the acyl-enzyme structure. The specific distances that can be used to follow the reaction are the S-C (d1), O(water)-C (d2), H(water)-O(water) (d3) and H(water)-N(imidazol) (d4) distance (Scheme 1C). In the acyl-enzyme water complex each method resulted in distance parameters that are very similar to each other (Table 4). Only the slightly longer S-C and C-O distances calculated by the B3LYP and B97-D methods are worth mentioning. Note that longer S-C distances were already observed at the product site of acylation step. None of these parameters showed any significant basis set dependency in the applied 6-31G(d,p) to 6-311++G(d,p) range.

In the transition state, which connects the acyl-enzyme water complex to the acetic acid and papain complex, the results are essentially the same: all levels of theories we applied predicted very similar transition state geometries (Table 4). For the calculated transition state structures, the d2 distances are predicted to be slightly longer by the B3LYP and B97-D methods. Regarding the breaking d3 and the newly forming d4 bond lengths, they are almost the same in the transition state. Only the ωB97X-D method predicted the breaking bond to be a little shorter than the newly forming one.

The PESs calculated at the B3LYP/6-31+G(d,p) and M06-2X/6-31+G(d,p) levels of theory showed rather synchronous reaction regarding the new d2 and d4 bond distances (Figures 6A-B). However, at the TS geometries the S-C (d1) distances are only marginally longer than they were in the reactants (Table 4). The S-C (d1) bond breaks spontaneously at shorter d2 and d4 values and geometry optimizations starting from the pre- and post TS region lead to the reactant (acyl-enzyme plus water) and product (enzyme plus acetic acid) geometries, respectively

Based on these data, an asynchronous but still a single elementary step reaction mechanism can be proposed for the deacylation process in which the O(water)-C(carbonyl) and H(water)-Nδ(histidine) bond formation and O(water)-H(water) bond breaking precede the S-C bond breaking. The product geometries were also very similar to each other with the exceptions of slightly longer S-C and remarkably longer O(carboxylate)-Hδ(histidine) (formerly O-H (water) non-bonding distances predicted by the B3LYP method.

Interestingly, Wei et al.[25] proposed a two distinct step mechanism for this particular step of papain hydrolysis. However, the activation (free) energy required for the second elementary reaction step was extremely small. Comparing the transition state structures for the first elementary step that Wei et al.[25] obtained to the sole transition state that we calculated, a remarkable similarity can be recognized between them. In their paper the transition state Ow-C(carbonyl), Ow-Hw and Nδ-Hw distances were 1.68Å, 1.16Å and 1.32, respectively, which are in good agreement with our d2,d3 and d4 values listed in the Table 4. The second

View Article Online
DOI: 10.1039/C6CP06869C

elementary step that Wei et al.[25] proposed corresponds mainly to an S-C bond breaking which is peculiar to the second phase in the single elementary reaction step in our calculations.

Regarding the transition state energies, all the density functional methods predict values which are in the ~10-14 kcal/mol range. In general, the more complete the basis set the higher the activation energy calculated. It should be also recognized that the energy level of the product site is always higher than that of the corresponding acyl-enzyme reactant. As it was emphasized in the case of acyl enzyme formation, neither the solvation of the product site nor the leaving of the acetic acid (model peptide fragment) from the papain site have been considered.

Comparing the two distinct chemical processes (acyl-enzyme formation and acyl-enzyme hydrolysis), the rate limiting steps have comparable activation energies. Nevertheless, the second (deacylation) reaction step can be featured with slightly higher activation energies at most of the methods we applied. However, it is generally accepted that the rate determining step for amide bond hydrolysis is the acylation, in contrast to the ester hydrolysis where it is the deacylation step.[3,4] The available few theoretical calculations on the whole cysteine protease reaction demonstrated comparable reaction barriers for these distinct steps too, with a little larger barrier for the acylation step.[24, 25] It should be remembered, however, that the most simple peptide model was used in our calculations and different reference points for the acyl-enzyme formation and the deacylation processes (the papain – NMA and the acyl-enzyme – (single molecule) water complexes, respectively) were applied.

Applying the Eyring-Polanyi equation for the k_{cat} values[80-82] (which are typically in the range of 1-50 s^{-1}, depending on the substrate, temperature and other conditions) at 310 K and taking the transmission coefficient to be unity, the activation Gibbs free energies could be calculated to be in the range of 15.8-18.1 kcal/mol. Assuming that neither the entropy nor the $p\Delta V$ term of enthalpy has crucial contribution, these values should be close to the activation energies. The available activation energies or Gibbs free energies derived from reaction kinetics parameters are in the range ~5 kcal/mol[83] to ~18 kcal/mol.[25] These values can be derived also from the k_{cat} rate constants using the classic Arrhenius equation and the equation given for conventional transition state theory,[84] respectively. Therefore, in spite of the obvious simplifications in our model calculations, they still predict reliable activation energy values.

Conclusions

A proteolytic reaction of papain has been modeled by ONIOM type QM/MM methods using the simple peptide model substrate, N-methylacetamide. The applicability of a hybrid GGA (the popular B3LYP) method to a few of the more recent DFT methods with long range correction term or suitable parameters for such interaction, as the QM component in the ONIOM QM/MM calculations has been evaluated.

Our calculations show that in the resting state of papain the ion pair and neutral forms of His Cys side chains of the catalytic dyad have approximately the same energy levels that are separated by a small barrier. Zero point correction shifts this equilibrium slightly into the direction of the neutral form while implicit solvent model correcting the ONIOM method with PB computations predicts the ion pair form to be the most populated one, in good agreement with the available experimental data.

View Article Online
DOI: 10.1039/C6CP06869C

Regarding the enzyme mechanism, all the dispersion corrected DFT methods applied, as well as the B3LYP method using larger (6-31+G(d,p) to 6-311++G(d,p)) basis sets predict two elementary steps for the acylation phase and a single elementary deacylation step. According to these calculations, in the acylation phase a zwitterionic tetrahedral intermediate exists where the carbonyl carbon became tetrahedral and holds a negatively charged oxygen and, simultaneously, the amide nitrogen is protonated as well. The activation energies we derived are in the range that can be found in the literature or can be derived from the available kinetic constants. Although all the density functional methods applied in this work predict an asynchronous rate determining first step for the acylation reaction, it should be emphasized that ONIOM type QM/MM computations using long range corrected DFT methods resulted in a significantly different transition state compared to those that can be obtained by the B3LYP method using larger (6-31+G(d,p) to 6-311++G(d,p)) basis sets. Nonetheless, the proton transfer lags behind (or at least does not precede) the S-C bond formation. Interestingly, the B3LYP/6-31G(d,p) method in ONIOM predicts three elementary steps for the first (acylation) step through an anionic and a subsequent zwitterionic intermediate. It should be emphasized also that the activation energy between these two intermediates was found to be extremely small. Since the B3LYP functional in ONIOM predicted very small barrier for the amid bond breaking (independently from the basis sets), it increases the possibility that, depending on the chemical environment and the method one applies for transition search, different, one to three elementary step mechanisms are possible for the acyl enzyme formation.

An additional conclusion which can be drawn also from our present work is that, for the cysteine protease reaction, the usage of dispersion corrected DFT methods are strongly advised both in pure QM and QM/MM calculations, because it might qualitatively influence the computationally derived reaction mechanism, in contrast to the serine protease mechanism.

Acknowledgments: This work was supported by the Hungarian Scientific Research Fund (OTKA K-106294 and OTKA K-105459) and by the European Union and the State of Hungary, co-financed by the European Social Fund in the framework of TAMOP-4.2.2.C-11/1/KONV-2012-0010.1. The authors thank the National Information Infrastructure Development Institute (Grant No. NIIF-1057) for computer time. A.F is grateful for scholarships by the “Talentum Foundation” of the Gedeon Richter PLC and by Human Capacities Grant Management Office (Grant No. NTP-EFÖ-P-15-0610-A). We are thankful to Dr. Arpad Somogyi (CCIC Mass Spectrometry and Proteomics Facility, Ohio State University, Columbus, OH) for his useful comments and suggestions.

View Article Online
DOI: 10.1039/C6CP06869C

Legends to Schemes and Figures:

Scheme 1.: (A): The Cys25-His159 catalytic dyad of papain and the (Cys)Sγ-H and H-Nδ(His) distances (d1 and d2, respectively) used to follow the proton transfer between Cys25-His159 residues. (B): The N-methylacetamide (NMA) and the Cys25-His159 catalytic dyad of papain. The specific (Cys)Sγ-C(NMA-carbonyl), (NMA carbonyl)C-N(NMA amide), (His Nδ)Hδ-N(NMA amide) and (His Nδ)Hδ-Nδ(His) distances (d1, d2, d3 and d4, respectively) are used to follow the acyl-enzyme formation reaction. (C): The acyl-enzyme (acetylated papain) and a water molecule participating in the acyl-enzyme hydrolysis. The specific (Cys)Sγ-C(acetyl carbonyl), (acetyl carbonyl)C-O(water), (water)O-H(water) and (water)H-Nδ(His) distances (d1, d2, d3 and d4, respectively) are used to follow the deacylation reaction.

Scheme 2.: The schematic summary of the possible reaction pathways for the acyl-enzyme formation part of the whole peptidase reaction derived based on our calculations. The middle row corresponds to that three elementary reaction pathway that was resulted only in the B3LYP/6-31G(d,p):AMBER ONIOM calculations without zero point energy correction.

Figure 1.: The graphics representation of the Michaelis complex of papain and N-methylacetamide (NMA). The NMA and the Cys25, His159 and Asn175 are shown by ball and stick representation. For the papain the "cartoon" representation coloring is based on residue position from blue to red (from N- to C-terminal, respectively), while the transparent surface is colored based on the surface electrostatic potential from positive to negative ones (from blue to red, respectively).

Figure 2.: The minimum (Cys25)Hγ and (His159)Hδ proton distance from their bridgehead (Cys25)Sγ and (His159)Nδ atoms (A,B) and the number of water molecules - Sγ(Cys25) contacts (C,D) during simulations carried out on papain with neutral (A,C) and ion pair (B,D) form of the catalytic dyad.

Figure 3.: The active site structure of the papain-N-methylacetamide (NMA) Michaelis complex from M06-2X/6-311++G(d,p):AMBER ONIOM calculation. The H-bonded network playing a role in positioning and stabilizing the NMA substrate is shown by orange lines. The interactions that can have a substantial contribution to the reaction coordinate are labeled by dashed red sticks.

Figure 4.: The transition state structure of the first (acyl-enzyme formation) step of the hypothetical papain-NMA reaction predicted by the M06-2X/6-311++G(d,p):AMBER ONIOM calculation. The most important atomic movements during this reaction step are labeled by dashed red sticks.

Figure 5.: Relaxed potential energy surfaces (PESs) for the system used for modeling the formation of acyl-enzyme structure as the function of two internal coordinates (Sγ-C(peptide carbonyl) and (imidazolium Nδ)Hδ-N(peptide amide)) which presumably play key role in the corresponding reaction coordinate calculated at the ONIOM(B3LYP/6-31G(d,p):AMBER) (A) and ONIOM(M06-2X/6-31G(d,p):AMBER) (B) levels of theories.

Figure 6.: Relaxed potential energy surfaces (PESs) for the deacylation step of the model acyl-enzyme system as the function of two internal coordinates (O(water)-C(acyl carbonyl) and H(water)-Nδ(imidazol)), which presumably play a key role in the corresponding reaction

View Article Online
DOI: 10.1039/C6CP06869C

coordinate calculated at (A) ONIOM(B3LYP/6-31+G(d,p):AMBER) (B) ONIOM(M06-2X/6-31G(d,p):AMBER) levels of theories.

Supplementary Figures:

Figure S1.: Root mean square deviation of the non-hydrogen atoms of papain from their starting positions during simulations (A) and the radius of gyration calculated for the papain from the simulation trajectory (B). Both the neutral and ion pair (zwitterionic) form of the catalytic dyad were considered and they marked by black and red, respectively.

Figure S2.: ONIOM potential energy surfaces (PESs) calculated along the (NMe)N-C(Ace) and the (Cys25)Sγ…C(amid carbonyl) (A,B), as well as along the (NMe)N-C(Ace) and the (amid)N…Hδ(His159) (C,D) internal distance pairs. The PESs were calculated using both B3LYP/6-31G(d,p):AMBER (A,C) and the M06-2X/6-31G(d,p):AMBER (B,D) methods.

View Article Online
DOI: 10.1039/C6CP06869C

References

1. J. Drenth, J. N. Jansonius, R. Koekoek, H. M. Swen and B. G. Wolthers, *Nature*, 1968, **218**, 929-932.
2. G. Lowe, *Philos Trans R Soc Lond B Biol Sci*, 1970, **257**, 237-248.
3. L. Polgar, in *Handbook of Proteolytic Enzymes*, eds. A. J. Barrett, N. D. Rawlings and W. J. Fred, Academic Press, 3rd edn., 2013, vol. 2, ch. 405, pp. 1773–1784.
4. L. Polgar and P. Halasz, *Biochem J*, 1982, **207**, 1-10.
5. L. Polgar, *International Journal of Biochemistry*, 1977, **8**, 171-176.
6. T. Vernet, D. C. Tessier, J. Chatellier, C. Plouffe, T. S. Lee, D. Y. Thomas, A. C. Storer and R. Menard, *J Biol Chem*, 1995, **270**, 16645-16652.
7. L. Polgar, *Eur J Biochem*, 1973, **33**, 104-109.
8. L. Polgar, *FEBS Lett*, 1974, **47**, 15-18.
9. L. A. Sluyterman and M. J. de Graaf, *Biochim Biophys Acta*, 1970, **200**, 595-597.
10. F. A. Johnson, S. D. Lewis and J. A. Shafer, *Biochemistry*, 1981, **20**, 44-48.
11. R. Menard, J. Carriere, P. Laflamme, C. Plouffe, H. E. Khouri, T. Vernet, D. C. Tessier, D. Y. Thomas and A. C. Storer, *Biochemistry*, 1991, **30**, 8924-8928.
12. R. Menard, C. Plouffe, P. Laflamme, T. Vernet, D. C. Tessier, D. Y. Thomas and A. C. Storer, *Biochemistry*, 1995, **34**, 464-471.
13. P. T. VanDuijnen, B. T. Thole, R. Broer and W. C. Nieuwpoort, *International Journal of Quantum Chemistry*, 1980, **17**, 651-671.
14. P. T. van Duijnen, B. T. Thole and W. G. Hol, *Biophys Chem*, 1979, **9**, 273-280.
15. J. A. Rullmann, M. N. Bellido and P. T. van Duijnen, *J Mol Biol*, 1989, **206**, 101-118.
16. D. Arad, R. Langridge and P. A. Kollman, *Journal of the American Chemical Society*, 1990, **112**, 491-502.
17. Y. Lin and W. J. Welsh, *J Mol Graph*, 1996, **14**, 62-72, 92-63.
18. W. J. Welsh and Y. Lin, *Theochem-Journal of Molecular Structure*, 1997, **401**, 315-326.
19. M. J. Harrison, N. A. Burton and I. H. Hillier, *Journal of the American Chemical Society*, 1997, **119**, 12285-12291.
20. W. G. Han, E. Tajkhorshid and S. Suhai, *Journal of Biomolecular Structure & Dynamics*, 1999, **16**, 1019-1032.
21. M. Strajbl, J. Florian and A. Warshel, *Journal of Physical Chemistry B*, 2001, **105**, 4471-4484.
22. K. Byun and J. L. Gao, *Journal of Molecular Graphics & Modelling*, 2000, **18**, 50-55.
23. G. P. Miscione, M. Calvaresi and A. Bottoni, *Journal of Physical Chemistry B*, 2010, **114**, 4637-4645.
24. S. Ma, L. S. Devi-Kesavan and J. Gao, *J Am Chem Soc*, 2007, **129**, 13633-13645.
25. D. Wei, X. Huang, J. Liu, M. Tang and C. G. Zhan, *Biochemistry*, 2013, **52**, 5145-5154.
26. M. Shokhen, N. Khazanov and A. Albeck, *Proteins*, 2009, **77**, 916-926.
27. M. Mladenovic, R. F. Fink, W. Thiel, T. Schirmeister and B. Engels, *Journal of the American Chemical Society*, 2008, **130**, 8696-8705.
28. Z. Ke, Y. Zhou, P. Hu, S. Wang, D. Xie and Y. Zhang, *J Phys Chem B*, 2009, **113**, 12750-12758.
29. W. Lee and B. Engels, *Biochemistry*, 2014, **53**, 919-931.
30. A. Paasche, A. Zipper, S. Schafer, J. Ziebuhr, T. Schirmeister and B. Engels, *Biochemistry*, 2014, **53**, 5930-5946.

View Article Online
DOI: 10.1039/C6CP06869C

31. W. Lee and B. Engels, *J Phys Chem B*, 2013, **117**, 8095-8104.
32. W. Lee, S. R. Luckner, C. Kisker, P. J. Tonge and B. Engels, *Biochemistry*, 2011, **50**, 5743-5756.
33. A. Paasche, T. Schirmeister and B. Engels, *J Chem Theory Comput*, 2013, **9**, 1765-1777.
34. S. C. Kamerlin and A. Warshel, *Proteins*, 2010, **78**, 1339-1375.
35. S. Grimme, *Wiley Interdisciplinary Reviews-Computational Molecular Science*, 2011, **1**, 211-228.
36. N. Marom, A. Tkatchenko, M. Rossi, V. V. Gobre, O. Hod, M. Scheffler and L. Kronik, *Journal of Chemical Theory and Computation*, 2011, **7**, 3944-3951.
37. T. Tsuneda and K. Hirao, *Wiley Interdisciplinary Reviews-Computational Molecular Science*, 2014, **4**, 375-390.
38. D. Roy, M. Marianski, N. T. Maitra and J. J. Dannenberg, *Journal of Chemical Physics*, 2012, **137**.
39. M. Marianski, A. Asensio and J. J. Dannenberg, *Journal of Chemical Physics*, 2012, **137**.
40. J. D. Chai and M. Head-Gordon, *Physical Chemistry Chemical Physics*, 2008, **10**, 6615-6620.
41. S. Grimme, *J Comput Chem*, 2006, **27**, 1787-1799.
42. Y. Zhao and D. G. Truhlar, *Theoretical Chemistry Accounts*, 2008, **120**, 215-241.
43. Y. Zhao and D. G. Truhlar, *Journal of Chemical Theory and Computation*, 2011, **7**, 669-676.
44. Y. Zhao and D. G. Truhlar, *Chemical Physics Letters*, 2011, **502**, 1-13.
45. T. A. Rokob, A. Hamza and I. Papai, *Org Lett*, 2007, **9**, 4279-4282.
46. M. Walker, A. J. Harvey, A. Sen and C. E. Dessent, *J Phys Chem A*, 2013, **117**, 12590-12600.
47. R. Lonsdale, J. N. Harvey and A. J. Mulholland, *Journal of Chemical Theory and Computation*, 2012, **8**, 4637-4645.
48. L. Toth, L. Muszbek and I. Komaromi, *Journal of Molecular Graphics & Modelling*, 2013, **40**, 99-109.
49. M. Linder and T. Brinck, *Phys Chem Chem Phys*, **15**, 5108-5114.
50. S. Dapprich, I. Komaromi, K. S. Byun, K. Morokuma and M. J. Frisch, *Journal of Molecular Structure-Theochem*, 1999, **462**, 1-21.
51. T. Vreven, K. S. Byun, I. Komaromi, S. Dapprich, J. A. Montgomery, K. Morokuma and M. J. Frisch, *J Chem Theory Comput*, 2006, **2**, 815-826.
52. W. L. Jorgensen, Chandrasekhar, J., Madura, J. D., Impey, R.W., Klein, M. L., *J Chem Phys*, 1983, **79**, 926-935.
53. V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins-Structure Function and Bioinformatics*, 2006, **65**, 712-725.
54. H. J. C. Berendsen, J. P. M. Postma, W. F. Vangunsteren, A. Dinola and J. R. Haak, *Journal of Chemical Physics*, 1984, **81**, 3684-3690.
55. G. Bussi, D. Donadio and M. Parrinello, *J Chem Phys*, 2007, **126**, 014101.
56. S. Pronk, S. Pall, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess and E. Lindahl, *Bioinformatics*, 2013, **29**, 845-854.
57. T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089–10092.
58. D. A. Case, T. A. Darden, I. Cheatham, T.E. , C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, R. C. Walker, W. Zhang, K. M. Merz, B. Roberts, S. Hayik, A. Roitberg, G. Seabra, J. Swails, A. W. Götz, I. Kolossváry, K. F. Wong, F. Paesani, J. Vanicek, R. M. Wolf, J. Liu, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, Q. Cai, X. Ye, J.

View Article Online
DOI: 10.1039/C6CP06869C

Wang, M.-J. Hsieh, G. Cui, D. R. Roe, D. H. Mathews, M. G. Seetin, R. Salomon-Ferrer, C. Sagui, V. Babin, T. Luchko, S. Gusarov, A. Kovalenko and K. P.A., *University of California, San Francisco.*, 2012.

59. W. Humphrey, Dalke, A. and Schulten, K.,, *J. Molec. Graphics*, 1996, **14**, 33-38.
60. A. D. Becke, *Journal of Chemical Physics*, 1993, **98**, 5648-5652.
61. C. T. Lee, W. T. Yang and R. G. Parr, *Physical Review B*, 1988, **37**, 785-789.
62. A. Leach, *Molecular Modelling: Principles and Applications*, Prentice Hall, Harlow, 2nd Edition edn., 2001.
63. L. Li, C. Li, S. Sarkar, J. Zhang, S. Witham, Z. Zhang, L. Wang, N. Smith, M. Petukh and E. Alexov, *BMC Biophys*, 2012, **5**, 9.
64. W. Rocchia, E. Alexov and B. Honig, *Journal of Physical Chemistry B*, 2001, **105**, 6507-6514.
65. C. Li, M. Petukh, L. Li and E. Alexov, *Journal of Computational Chemistry*, 2013, **34**, 1949-1960.
66. P. Tao and H. B. Schlegel, *Journal of Computational Chemistry*, 2010, **31**, 2363-2369.
67. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. Montgomery, J. A., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G. A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, P. Y. Ayala, K. Morokuma, G. A. Voth, P. Salvador, J. J. Dannenberg, V. G. Zakrzewski, S. Dapprich, A. D. Daniels, M. C. Strain, O. Farkas, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. V. Ortiz, Q. Cui, A. G. Baboul, S. Clifford, J. Cioslowski, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, M. Challacombe, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, *Gaussian, Inc., Wallingford CT*, 2004.
68. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. Montgomery, J. A., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian, Inc., Wallingford CT*, 2009.
69. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin, *J Comput Chem* 2004, **25**, 1605-1612. .
70. *Journal*.
71. G. Lowe and A. S. Whitworth, *Biochem J*, 1974, **141**, 503-515.
72. L. A. Sluyterman and J. Wijdenes, *Eur J Biochem*, 1976, **71**, 383-391.
73. S. D. Lewis, F. A. Johnson and J. A. Shafer, *Biochemistry*, 1981, **20**, 48-51.
74. S. D. Lewis, F. A. Johnson and J. A. Shafer, *Biochemistry*, 1976, **15**, 5009-5017.
75. P. Halasz and L. Polgar, *Eur J Biochem*, 1977, **79**, 491-494.

View Article Online
DOI: 10.1039/C6CP06869C

76. L. Polgar, *Eur J Biochem*, 1979, **98**, 369-374.
77. D. J. Creighton, M. S. Gessouroun and J. M. Heapes, *FEBS Lett*, 1980, **110**, 319-322.
78. D. J. Creighton and D. J. Schamp, *FEBS Lett*, 1980, **110**, 313-318.
79. M. Topf, P. Varnai and W. G. Richards, *J Am Chem Soc*, 2002, **124**, 14780-14788.
80. G. Lowe and Y. Yuthavong, *Biochem J*, 1971, **124**, 107-115.
81. F. Gauthier, T. Moreau, G. Lalmanach, M. Brillard-Bourdet, M. Ferrer-Di Martino and L. Juliano, *Arch Biochem Biophys*, 1993, **306**, 304-308.
82. E. M. Papamichael, M. K. Roustas and J. G. Bieth, *Brazilian Archives of Biology and Technology*, 1999, **42**, 277-280.
83. J. F. Shaw, R. C. Chang and Y. J. Wang, *Botanical Bulletin of Academia Sinica*, 1987, **28**, 131-138.
84. J. R. Alvarez-Idaboy, A. Galano, G. Bravo-Perez and M. E. Ruiz, *J Am Chem Soc*, 2001, **123**, 8387-8395.

Table 1. Selected geometry parameters in Å s ((Cys)Sγ-H and H-Nδ(His) distances, d1 and d2, respectively, see on Scheme 1A) for the neutral (N) and zwitterionic (deprotonated Cys25 and doubly protonated His159 side chains) forms of the catalytic dyad of papain, as well as for the transition state (TS) between them. The relative ONIOM TS energies compared to the corresponding neutral and ion-pair energy values (TS-N and TS-Z, respectively, in kcal/mol units), as well as the zwitterionic center energies compared to the neutral one are also shown. In each case the zero point vibrational energy corrected (ZPVEc) energies are given as well.

Methods	Basis sets	Zwitterionic center (Z)		TS		Neutral center (N)		TS - Zwitterionic center		TS - Neutral center		Zwitterionic center - Neutral Center	
		d1	d2	d1	d2	d1	d2	ONIOM extrapolated	ZPVEc	ONIOM extrapolated	ZPVEc	ONIOM extrapolated	ZPVEc
B3LYP	6-31G(d,p)	2.008	1.096	1.630	1.339	1.390	1.882	1.967	-0.946	2.262	0.842	0.295	1.788
	6-31+G(d,p)	2.045	1.085	1.618	1.351	1.384	1.925	2.587	-0.521	2.405	0.944	-0.182	1.465
	6-311+G(d,p)	2.052	1.083	1.626	1.346	1.382	1.940	2.581	-0.560	2.828	1.373	0.247	1.933
	6-311++G(d,p)	2.055	1.082	1.625	1.347	1.383	1.938	2.650	-0.504	2.788	1.328	0.137	1.832
B97-D	6-31G(d,p)	1.936	1.130	1.662	1.314	1.414	1.774	0.814	-1.379	1.476	0.282	0.662	1.662
	6-31+G(d,p)	1.990	1.108	1.636	1.340	1.409	1.799	1.421	-1.123	1.374	0.092	-0.047	1.215
	6-311+G(d,p)	1.999	1.105	1.641	1.337	1.405	1.822	1.461	-1.111	1.641	0.341	0.181	1.451
	6-311++G(d,p)	2.006	1.102	1.637	1.341	1.406	1.820	1.562	-1.051	1.592	0.287	0.030	1.338
ωB97X-D	6-31G(d,p)	2.013	1.085	1.596	1.358	1.381	1.888	2.625	-0.517	2.213	0.983	-0.413	1.500
	6-31+G(d,p)	2.035	1.078	1.590	1.364	1.376	1.925	3.081	-0.244	2.446	1.197	-0.635	1.441
	6-311+G(d,p)	2.041	1.077	1.599	1.357	1.374	1.947	3.033	-0.343	2.842	1.530	-0.191	1.873
	6-311++G(d,p)	2.042	1.076	1.598	1.358	1.375	1.944	3.062	-0.311	2.813	1.513	-0.249	1.824
M06-2X	6-31G(d,p)	1.985	1.096	1.578	1.388	1.377	1.899	2.151	-0.789	1.618	0.259	-0.533	1.049
	6-31+G(d,p)	2.008	1.088	1.572	1.395	1.367	1.987	2.615	-0.428	1.874	0.630	-0.740	1.058
	6-311+G(d,p)	2.014	1.086	1.581	1.384	1.363	2.033	2.863	-0.480	2.388	1.231	-0.475	1.711
	6-311++G(d,p)	2.016	1.085	1.581	1.385	1.363	2.030	2.155	-0.449	2.353	1.189	0.198	1.638

Table 2: Selected geometry parameters in Å̊s ((Cys)Sγ-C(NMA-carbonyl), (NMA carbonyl)C-N(NMA amide), (His Nd)Hδ-N(NMA amide) and (His Nδ)Hδ-Nδ(His) distances, d1, d2, d3 and d4, respectively, see, Scheme 1B) at the characteristic points (Michaelis complex (MC), zwitterionic tetrahedral intermediate (ZI) and acyl enzyme (AE) as well as the transitional states (TS1 and TS2) between them) of the acyl-enzyme formation process for the papain and N-methylacetamid hypothetical reaction, calculated at different levels of theory. The zero point vibrational energy corrected (ZPVEc) transition state energy values compared to the corresponding MC, ZI and AE state energies are also shown in kcal/mol units.

Methods	Basis sets	Michaelis Complex (MC)				TS1				Tetrahedral Intermed. (ZI)				TS2				Product (AE)				TS1-MC	TS - ZI	TS2-ZI	TS2-AE
		d1	d2	d3	d4	d1	d2	d3	d4	d1	d2	d3	d4	d1	d2	d3	d4	d1	d2	d3	d4	ZPVEc relative energies			
B3LYP	6-31+G(d,p)	3.469	1.357	3.726	1.091	2.173	1.520	1.437	1.194	2.009	1.638	1.057	1.814	1.883	2.078	1.031	2.008	1.799	3.755	1.019	2.803	12.290	3.933	0.800	4.235
	6-311+G(d,p)	3.451	1.356	3.708	1.090	2.195	1.520	1.423	1.203	2.016	1.646	1.054	1.824	1.900	2.019	1.032	1.981	1.801	3.755	1.017	2.803	12.836	4.226	0.341	4.821
	6-311++G(d,p)	3.450	1.356	3.711	1.088	2.197	1.520	1.420	1.205	2.017	1.645	1.054	1.823	1.900	2.021	1.032	1.981	1.801	3.756	1.017	2.803	12.941	4.192	0.351	4.840
B97-D	6-31G(d,p)	3.375	1.367	3.337	1.137	2.852	1.435	1.812	1.059	2.072	1.704	1.067	1.741	1.969	1.926	1.047	1.819	1.821	3.503	1.024	2.360	10.212	6.793	1.538	1.851
	6-31+G(d,p)	3.364	1.364	3.387	1.111	2.786	1.441	1.763	1.065	2.070	1.662	1.071	1.731	2.053	1.683	1.068	1.731	1.813	3.555	1.023	2.440	10.261	5.440	1.726	1.724
	6-311+G(d,p)	3.355	1.363	3.380	1.109	2.746	1.447	1.739	1.069	2.077	1.670	1.067	1.744	2.057	1.696	1.064	1.745	1.815	3.552	1.021	2.428	10.385	5.243	1.626	2.503
	6-311++G(d,p)	3.354	1.363	3.385	1.105	2.711	1.453	1.714	1.075	2.079	1.669	1.067	1.744	2.059	1.695	1.064	1.744	1.815	3.553	1.021	2.429	10.201	4.900	1.621	2.484
ωB97X-D	6-31G(d,p)	3.117	1.354	3.366	1.095	2.523	1.456	1.670	1.077	1.989	1.614	1.058	1.755	1.853	2.138	1.028	1.955	1.790	3.374	1.020	2.267	11.933	7.057	2.061	1.984
	6-31+G(d,p)	3.124	1.351	3.383	1.084	2.470	1.454	1.688	1.073	1.971	1.598	1.058	1.758	1.843	2.157	1.027	1.982	1.784	3.435	1.019	2.337	12.840	7.171	3.149	2.159
	6-311+G(d,p)	3.128	1.350	3.375	1.082	2.434	1.459	1.689	1.073	1.978	1.600	1.055	1.763	1.853	2.119	1.026	1.965	1.785	3.424	1.018	2.324	13.340	7.247	2.782	2.654
	6-311++G(d,p)	3.128	1.350	3.375	1.081	2.426	1.459	1.687	1.073	1.978	1.600	1.055	1.764	1.853	2.120	1.026	1.966	1.785	3.424	1.018	2.325	13.489	7.206	2.805	2.681
M06-2X	6-31G(d,p)	3.044	1.354	3.314	1.110	2.599	1.442	1.718	1.071	1.985	1.612	1.063	1.740	1.835	2.216	1.025	2.013	1.791	3.047	1.021	2.176	11.541	6.972	2.175	0.635
	6-31+G(d,p)	3.058	1.351	3.299	1.096	2.555	1.443	1.714	1.071	1.969	1.597	1.063	1.743	1.831	2.218	1.024	2.034	1.794	2.657	1.019	2.154	12.113	6.856	3.490	1.164
	6-311+G(d,p)	3.055	1.351	3.274	1.093	2.516	1.449	1.706	1.072	1.974	1.602	1.057	1.764	1.841	2.169	1.024	2.015	1.795	2.702	1.018	2.160	12.406	6.759	2.669	1.593
	6-311++G(d,p)	3.056	1.351	3.279	1.092	2.503	1.451	1.698	1.074	1.974	1.602	1.057	1.765	1.841	2.172	1.024	2.017	1.795	2.704	1.018	2.161	12.308	6.633	2.676	1.579

Table 3. Selected geometry parameters in Ås ((Cys)Sγ-C(NMA-carbonyl), (NMA carbonyl)C-N(NMA amide), (His Nd)Hδ-N(NMA amide) and (His Nδ)Hδ-Nδ(His) distances, d1, d2, d3 and d4, respectively, see, Scheme 1B) at the characteristic points (Michaelis complex (MC), oxyanionic tetrahedral intermediate (OI), zwitterionic tetrahedral intermediate (ZI) and the acyl-enzyme product (AE), as well as the transitional states (TS1a, TS1b and TS2) between them) of the acyl-enzyme formation process for the papain and N-methylacetamide hypothetical reaction, calculated by ONIOM method at the B3LYP/6-31G(d,p):AMBER level. The ONIOM extrapolated transition state energy values as well as the zero point vibrational energy corrected (ZPVEc) transition state energy values compared to the corresponding MC, OI, ZI and AE state energies are also shown in kcal/mol units.

	d1	d2	d3	d4
Complex MC	3.470	1.360	3.690	1.105
TS1a	2.704	1.447	1.779	1.063
Oxianionic Intermediate (OI)	2.353	1.493	1.602	1.110
TS1b	2.262	1.513	1.467	1.174
Zwitterionic Intermediate (ZI)	2.021	1.667	1.056	1.812
TS2	1.900	2.036	1.033	1.962
Product (AE)	1.806	3.634	1.020	2.534

	TS1a - MC	TS1a - OI	TS1b - OI	TS1b - ZI	TS2 - ZI	TS2 - AE
ONIOM extrapolated	12.519	0.275	0.069	6.826	0.917	2.965
ZPVEc	13.558	0.597	-1.195	4.671	0.102	3.689

Table 4. Selected geometry parameters in Ås ((Cys)Sγ-C(acetyl carbonyl), (acetyl carbonyl)C-O(water), (water)O-H(water) and (water)H-Nδ(His) distances (d1, d2, d3 and d4, respectively) see on Scheme 1C) at the characteristic points (acyl-enzyme - water complex (C), enzyme – acetic acid product (P) and the transition state (TS)) of the deacylation process for the papain and N-methylacetamide hypothetical reaction, calculated at different levels of theory. The zero point vibrational energy corrected (ZPVEc) transition state energy values compared to the corresponding C and P state energies are also shown in kcal/mol units.

	Basis sets	Complex (C)				TS				Product (P)				TS - C	TS - P
		d1	d2	d3	d4	d1	d2	d3	d4	d1	d2	d3	d4	ZPVEc TS energies	
B3LYP	6-31G(d,p)	1.808	2.741	0.979	1.965	2.008	1.703	1.235	1.268	3.036	1.354	3.149	1.085	12.639	10.378
	6-31+G(d,p)	1.802	2.803	0.981	1.978	1.964	1.752	1.238	1.264	3.060	1.354	3.273	1.076	13.199	10.634
	6-311+G(d,p)	1.804	2.814	0.978	1.986	1.976	1.757	1.255	1.247	3.060	1.352	3.264	1.075	14.536	10.897
	6-311++G(d,p)	1.804	2.814	0.978	1.987	1.977	1.758	1.255	1.246	3.061	1.353	3.269	1.073	14.501	11.039
B97-D	6-31G(d,p)	1.819	2.801	0.984	1.938	2.079	1.712	1.263	1.249	2.969	1.369	2.851	1.096	10.397	7.432
	6-31+G(d,p)	1.813	2.791	0.986	1.936	2.024	1.756	1.271	1.240	2.979	1.369	2.984	1.085	10.878	7.773
	6-311+G(d,p)	1.815	2.808	0.982	1.946	2.038	1.756	1.282	1.229	2.981	1.367	2.976	1.082	11.928	8.045
	6-311++G(d,p)	1.815	2.808	0.982	1.947	2.038	1.757	1.281	1.230	2.985	1.367	2.981	1.080	11.894	8.203
ωB97X-D	6-31G(d,p)	1.794	2.668	0.975	1.947	1.965	1.678	1.198	1.290	2.993	1.347	2.892	1.074	12.571	10.863
	6-31+G(d,p)	1.788	2.721	0.976	1.951	1.932	1.710	1.189	1.299	3.004	1.346	2.979	1.067	13.055	10.727
	6-311+G(d,p)	1.790	2.733	0.973	1.956	1.941	1.711	1.202	1.282	3.003	1.344	2.968	1.065	14.506	10.981
	6-311++G(d,p)	1.790	2.734	0.974	1.956	1.941	1.711	1.202	1.283	3.004	1.344	2.969	1.065	14.439	11.087
M06-2X	6-31G(d,p)	1.794	2.625	0.975	1.985	1.977	1.637	1.185	1.310	2.945	1.348	2.780	1.080	11.741	9.767
	6-31+G(d,p)	1.790	2.655	0.975	1.998	1.945	1.664	1.173	1.325	2.959	1.347	2.856	1.071	12.347	10.026
	6-311+G(d,p)	1.792	2.663	0.973	1.993	1.956	1.662	1.189	1.303	2.959	1.347	2.837	1.065	14.005	10.097
	6-311++G(d,p)	1.792	2.664	0.973	1.993	1.956	1.662	1.189	1.304	2.960	1.347	2.837	1.064	13.983	10.243