A Conserved Asparagine in a Ubiquitin Conjugating Enzyme Positions the Substrate for Nucleophilic Attack

Walker M. Jones, Aaron G. Davis, R. Hunter Wilson, Katherine L. Elliott and Isaiah Sumner

Department of Chemistry and Biochemistry, James Madison University, Harrisonburg, VA 22807, USA

Correspondence:
I. Sumner
901 Carrier Drive, MSC 4501
Harrisonburg, VA 22807
Phone: +1-540-568-6670
Email: sumneric@jmu.edu

Keywords: ubiquitin, E2, Ubc13, ubiquitination, molecular dynamics, QM/MM, BOMD

ABSTRACT The mechanism used by the ubiquitin conjugating enzyme, Ubc13, to catalyze lysine ubiquitination is probed with three computational techniques: Born-Oppenheimer Molecular Dynamics (BOMD), single point Quantum Mechanics/Molecular Mechanics energies (QM/MM), and classical Molecular Dynamics (MD). These simulations support a long-held hypothesis and show that Ubc13-catalyzed ubiquitination uses a step-wise, nucleophilic attack mechanism. Furthermore, they show that the first step – the formation of a tetrahedral, zwitterionic intermediate – is rate limiting. However, these simulations contradict another popular hypothesis that supposes that the intermediate is stabilized by a highly conserved asparagine (Asn79 in Ubc13). Instead, calculated reaction profiles of the N79A mutant illustrate how stabilizing the intermediate actually increases the barrier to product formation. Finally, simulations of wild-type, N79A, N79D and H77A Ubc13 show that Asn79 stabilizes the motion of the electrophile prior to the reaction, positioning it for nucleophilic attack.

1 Present address: Department of Chemistry, University of Georgia, Athens, GA 30602, USA
2 Present address: Department of Chemistry, University of Minnesota, Minneapolis, MN 55455, USA
Introduction

Cells use lysine ubiquitination – the addition of the small protein ubiquitin (Ub) to a lysine sidechain of a target protein\(^1^2\) – to regulate many processes including inflammation response, DNA repair and protein degradation\(^1^3\). Ubiquitination occurs in three enzymatic steps, catalyzed by E1 (ubiquitin activating), E2 (ubiquitin conjugating) and E3 (ubiquitin ligase) enzymes\(^1^3\). Prior to its transfer, ubiquitin is covalently bonded to the E2 via a thioester linkage between the C-terminal glycine of ubiquitin and a cysteine on the E2. The E2 then catalyzes the transfer of ubiquitin to the final protein. A RING E3 improves the efficiency of this direct transfer, but it is not necessary for all E2 enzymes\(^4^6\). (HECT and RBR E3 ligases, on the other hand, utilize a two-step process in which the ubiquitin is first transferred to an active site cysteine on the E3 before reaching its target\(^1^6\).) Ubiquitin is transferred when the amine sidechain of a lysine residue on the target protein deprotonates and attacks the thioester carbon in the E2~Ub conjugate, which leads to the cleavage of the thioester bond and the formation of an isopeptide bond between the target protein and the ubiquitin (see Scheme 1)\(^6^8\). The exact mode of catalysis employed by the RING E3 and E2 enzymes, however, remains an area of active research.

RING E3 ligases are thought to promote ubiquitin transfer through mechanical means, i.e. catalysis by proximity\(^1^4^6^9^1^6\). It is surmised that RING E3 ligases 1) bring the substrate close to the E2~Ub conjugate \(^1^6^9^1^6\), 2) put E2~Ub into the folded back position, which bends back the electrophilic thioester like a loaded spring\(^4^1^2^1^5\) and 3) reduce the conformational flexibility of both the substrate lysine and the thioester allowing them to quickly achieve a reactive configuration\(^4^6^1^4^1^5\). So, although a RING E3 ligase promotes ubiquitin transfer, it does not change the chemical mechanism used by the ubiquitin conjugating enzyme, E2.
It is commonly hypothesized that the nucleophilic attack of the E2~Ub complex proceeds through a meta-stable, zwitterionic intermediate; the thioester oxygen carries a negative charge and the amine nitrogen carries a positive charge (Scheme 1). The rate-limiting step in this mechanism is unknown. However, calculated reaction profiles of thioester aminolysis on model systems\textsuperscript{17–19} and on an E2 enzyme ternary complex indicate that forming the intermediate is rate-limiting\textsuperscript{20}. 

It is also commonly hypothesized that the sidechain of a nearby asparagine acts as an oxyanion hole and stabilizes the negative charge on the zwitterionic intermediate\textsuperscript{3,7,8,20–23}. There is evidence to support this claim. First, it is part of the HPN loop, group of residues (histidine, proline and asparagine) that are highly conserved across the entire E2 family\textsuperscript{3,7,22,24}. Second, if the asparagine is mutated, the catalytic rate and efficiency of ubiquitination decrease significantly\textsuperscript{5,8}. Third, it is the only plausible residue capable acting as an oxyanion hole\textsuperscript{20,23}. However, there have been some recent studies that have reexamined this assumption and found evidence that the asparagine hydrogen bonds to a nearby loop in the E2 enzyme\textsuperscript{5,24,25}, which may affect the loop’s dynamics. Mutations to this active site loop are also known to have a negative impact on catalysis\textsuperscript{4,5,8,26} and recent work has suggested that this loop regulates E2 activity by acting as a gate\textsuperscript{26,27}.

The need to re-examine the role of the conserved asparagine is also apparent from the structural data available from NMR and crystal structures of E2~Ub and E2~SUMO conjugates. Contrary to expectations, several of these structures do not show the conserved asparagine forming a hydrogen bond with the thioester carbonyl\textsuperscript{5,24,25}. (The thioester carbonyl develops the negative charge in the intermediate.) In fact, the asparagine sidechain is sometimes oriented in the opposite direction. Moreover, studies have shown that the relative position of Ub to the E2 is quite mobile and that the relative position seems to affect the hydrogen bonding partners of the conserved asparagine\textsuperscript{10,15,24,25,28}. Of the possible E2~Ub orientations, the closed or folded back position is thought to be the most active\textsuperscript{4,13–15,25}. However, a recent study suggests that this is not always the case\textsuperscript{10}.

In an attempt to understand some of these mechanistic details, we recently used classical molecular dynamics (MD) simulations to examine the role of the conserved asparagine (Asn79) in the E2 enzyme, Ubc13\textsuperscript{24}. We ran simulations of Ub conjugated with WT and mutated (N79A and
N79D) Ubc13. We found that Asn79 forms only a weak (~3 kcal mol\(^{-1}\)) hydrogen bond with the thioester carbonyl and typically hydrogen bonds to Asn116 in the active site loop. We also found that in WT Ubc13, an aspartate in the active site loop, Asp119, remains near the substrate lysine (< 7 Å away). Thus, we hypothesized that Asn79 hydrogen bonds with the active site loop in order to keep Asp119 near the substrate lysine to serve as a deprotonation site; lysine must lose two protons (see Scheme 1). However, we were unable to directly simulate reaction mechanisms since classical MD will not allow covalent bonds to break and form. Therefore, in this work, we complement our previous study by using two \textit{ab initio} methods – Born-Oppenheimer molecular dynamics (BOMD), and hybrid quantum mechanics/molecular mechanics (QM/MM) calculations – in addition to classical MD to improve our understanding of the mechanism used by Ubc13 by uncovering the rate-limiting step and by clarifying the role played by Asn79.

**Methodology**

**Structural preparation**

The preparation was the same as outlined in Ref. 24. Briefly, the initial structure of the Ubc13/ubiquitin complex was generated from the pdb code 2GMI in the RCSB protein data bank. The Mms2 coordinates were removed and the ubiquitin was rotated so that it is in the closed/folded-back position (see Fig. 1(a)). Missing hydrogen atoms were added with the LEaP program. A single lysine with a zwitterionic backbone and a neutral sidechain was then added to the structure to act as a substrate. The residues forming the thioester linkage between the Ubc13 and ubiquitin were extracted from the pdb and capped with acetyl (ACE) and N-methylamide (NME) groups. The thioester was optimized using B3LYP/6-31G(d,p) while the carbonyl carbon on ACE and the amide nitrogen on NME were fixed to their crystal positions. The optimized thioester was used to calculate RESP\(^{30}\) partial atomic charges using HF/6-31G(d). The cysteine residue within the thioester was labeled as the Amber residue CYX, the thioester glycine was labeled as CGLY, and the thioester bond between them was described using parameters from the General Amber Force Field (GAFF).\(^{31}\) HF/6-31G(d) level of theory was also used to calculate the RESP partial atomic charges of the substrate lysine and the remaining parameters were taken from a standard Amber LYN residue. The full list of parameters for the thioester and charges for the thioester and LYN are available in the SI.

**Molecular Dynamics**
All MD was performed using the Amberff12SB force field\textsuperscript{32} and the GPU-accelerated PMEMD module in the Amber14MD package\textsuperscript{29,33}. The structure was solvated in a rectilinear box ($r = 12.0$ Å) of TIP3P water molecules\textsuperscript{34} and neutralized with 12 K$^+$ and 10 Cl$^-$ ions (13 K$^+$ and 10 Cl$^-$ in N79D). The system was stabilized by an optimization, heated to 300K, and subjected to density equilibration (100 ps, with a time-step of 0.5 fs). The system was then simulated at 300K for 500 ns using a Langevin thermostat with a collision frequency of 1 ps$^{-1}$ and a time step of 2 fs. The cutoff for non-bonded interactions was 8 Å and all covalent bonds to hydrogen atoms were held fixed using the SHAKE algorithm\textsuperscript{35}. In order to tether the substrate to the active site, mimic the presence of the enzyme substrate complex, and to help it maintain a reactive geometry, restraints were added to the amine group on the substrate lysine and the thioester carbonyl; the remainder of the system was unrestrained. Specifically, if the distance between the N$_\varepsilon$ of the substrate lysine and the thioester carbon was between 3 Å and 5 Å a harmonic potential was added with a force constant of 32 kcal mol$^{-1}$ Å$^{-2}$ and if the distance was greater than 5 Å a constant force of -64 kcal mol$^{-1}$ Å$^{-1}$ was applied. To maintain the Bürgi–Dunitz angle of nucleophilic attack (107°)\textsuperscript{36} between the thioester oxygen, carbon and lysine N$_\varepsilon$, a harmonic potential with force constant 200 kcal mol$^{-1}$ rad$^{-2}$ was added if the angle was $> 122^\circ$ or $< 92^\circ$, and to maintain a tetrahedral geometry, a harmonic potential with force constant 200 kcal mol$^{-1}$ rad$^{-2}$ was added the lysine amine hydrogen, N$_\varepsilon$ and thioester carbon angles if they were $> 125^\circ$ or $< 94^\circ$.

Finally, residues Asn79 and His77 were mutated to produce N79A, N79D and H77A enzymes. Like the WT enzyme, these mutants were prepared for simulations as described in the

![Figure 1](image_url)  
**Figure 1** Ubc13-Ub complex (a) and description of collective variables (CVs) (b). In (a), Ubc13 is cyan and Ub is red. The thioester bond between Ubc13 and Ub and the active site loop (residues 116-120) in Ubc13 are highlighted in yellow. The model in (a) is used for all calculations and is based on pdb 2GMI and rotated to the closed position. The CVs used are the distance between the nucleophile and electrophile ($R_{N_\varepsilon-C}$) and the hydrogen coordination number of the lysine sidechain amine ($CN_{N_\varepsilon}$). Important residues are labeled in (b).
previous paragraphs. Five, independent, 500 ns, NVT simulations were completed for the wild-
type, N79A and N79D mutants; four simulations were completed for H77A. Trajectory analyses
were conducted using AmberTools14\textsuperscript{37}. Error bars for analyses were calculated using the standard
error, i.e., the standard deviation of the simulations divided by $\sqrt{N}$, where $N$ is the number of
simulations \textsuperscript{38}.

**Born-Oppenheimer Molecular Dynamics**

Born-Oppenheimer molecular dynamics (BOMD) was performed using CP2K (version 2.6)\textsuperscript{39,40}
and the Metadynamics (MTD) enhanced sampling technique was used to generate minimum
energy pathways (MEPs) for the reaction \textsuperscript{41,42}. Fully periodic QM/MM was used to calculate the
forces necessary to generate the trajectory; the QM images were decoupled with the GEEP
algorithm\textsuperscript{31,32}. PBE/TZV2P with GTH pseudopotentials and the Grimme D3 dispersion
corrections\textsuperscript{45} along with the GAPW \textsuperscript{46} method (280 Ry cutoff) were used for the QM region;
Amber ff12SB was used for the MM region \textsuperscript{32}. The entire system had 48,776 atoms. The QM
region contained 161 atoms and included the thioester, the substrate lysine, Asn79, the active site
loop (residues 116-120) and the 14 water molecules closest to the thioester oxygen (see Fig. 2).
Since water can move freely into and out of the QM region, the identity of the QM waters must be
allowed to change. Thus, every 2 ps, the 14 water molecules closest to the thioester were identified,
and if needed, added to or deleted from the QM region. As the reaction neared the barrier, the QM
waters were updated every 0.5 ps. Hydrogen link atoms were used to saturate the dangling bonds
and only single bonds to carbon atoms were cut (i.e., no amide bonds). Finally, the temperature
was held fixed at 300 K during the simulation using a Nosé-Hoover chain thermostat\textsuperscript{47,48} with a
time constant of 10 fs. Initial structures were taken from the MD simulations described previously
and were equilibrated for ~10 ps before starting MTD.

MTD locates MEPs by adding a Gaussian function to the potential energy surface at regular
time intervals ($\Delta t$) \textsuperscript{41,42}. Thus, at time $T = n \Delta t$, in one-dimension, the added term becomes:

$$V_{MTD}(S(R)) = \sum_{i=\Delta t, 2\Delta t, \ldots, n\Delta t} w \exp \left\{ \frac{-(S(R)-S_i)^2}{2\sigma^2} \right\}$$ \text{ Eq. (1)}

where $S(R)$ is a collective variable (CV) that depends on all atomic positions, $R$; $S_i = S(R(t = i\Delta t))$ is the value of the CV at a time $t = i\Delta t$; and $w$ and $\sigma$ are the height and width of
the Gaussian. $w = 1.0 \text{ kcal mol}^{-1}$, $\sigma =0.1$ (internal units) and $\Delta t =20 \text{ fs}$ in our simulations. Once
the system was close to the barrier, we halved \( w \) and \( \sigma \) to produce higher resolution results. Since \( V_{MTD} \) accumulates over time, it has the effect of lifting a system out of free energy minima. Thus, the system is able to sample the full reaction coordinate. Once the calculations have converged, the free energy of the system can be approximated as \( F(S(R)) \sim -V_{MTD}(S(R)) \).

In this work, we used the following two, independent CVs: \( CV_1 = R_{N_e-C} \), the distance between the nucleophile (\( N_e \) on the attacking lysine) and the electrophile (the thioester \( C \)) and \( CV_2 = CN_{N_e} \), the hydrogen coordination number of \( N_e \). (The coordination number is the number of hydrogen bound to the nitrogen.) If \( R_{N_e-C} \) exceeded 4 \( \text{Å} \), the momentum of the CV was reversed. These two CVs (shown in Fig. 1(b)) should be sufficient to describe the reaction which requires \( N_e \) on the lysine sidechain to both deprotonate and form a covalent bond to the thioester carbon. Furthermore, these CVs are general, allowing MTD to determine a reaction mechanism with any \textit{a priori} assumptions (e.g., concerted vs. stepwise, identity of proton acceptors, etc.). CP2K input files are included in the SI.

**Single Point QM/MM**

Single point QM/MM calculations were used to calculate the relative energies between the different steps along the reaction coordinate. The ONIOM method\(^{49-51} \) was used as implemented in the Gaussian 09 suite of programs\(^{52} \). Input files were prepared using the TAO package according to the procedures outlined in Refs. 53 and 54. The core residues of this structure were the thioester, the substrate lysine, Asn79 and the active site loop (residues 116-120). All residues within 15 \( \text{Å} \) of the core were retained in the ONIOM calculations and all residues within 10 \( \text{Å} \) of the core were flexible; the remaining atoms were held fixed. Backbones were saturated with ACE and NME residues where necessary. In total, three different structures were used for the ONIOM calculations; two were taken from MD and are presented in the main text and one was taken from MTD and is presented in the SI.

The ONIOM energy is defined as\(^{50} \)

\[
E_{\text{oniom}} = E_{\text{real}}^{\text{MM}} - E_{\text{model}}^{\text{MM}} + E_{\text{model}}^{\text{QM}} \tag{2}
\]

where the “model” system is treated at the QM level and contains the enzyme active site. In our case, the model layer varied since it depends on the atoms and residues within the active site.
Invariably, it contained the thioester, the substrate lysine and a portion of the active site loop. Water was also included, as well as the sidechain of Asn79. Our QM region contained around 90 atoms, hydrogen link atoms were used to saturate the dangling bonds and only single bonds to carbon atoms were cut (i.e., no amide bonds). The “real” system makes up the rest of the protein-enzyme complex and is calculated with more cost effective molecular mechanics. Geometry optimizations were carried out with the M06-2X/6-31G(d):ff12SB level of theory.\textsuperscript{32,55}

One drawback of ONIOM energy expression in Eq. 2, is that the model and real system only interact at the MM level; the model system calculations are done in vacuum.\textsuperscript{50} This is termed mechanical embedding. Additionally, the MM charges in the model layer do not change to reflect the changing bonding topology, which adds another source of error. To correct for the changes in charge along the reaction coordinates, we followed Ref. 53 and recalculated the MM charges in the model layer using the RESP method.\textsuperscript{30} After recalculating the charges, the geometry was reoptimized and the MM charges were recalculated. These iterations continued until the change in energy between two steps was less than 0.1 kcal mol\textsuperscript{-1} or until the new optimization finished within a single step. Occasionally, the energies oscillated and never converged. In these cases, the MM charges were taken as the average between two successive runs, the geometry was reoptimized and the process finished. Final energies were calculated at the MP2/6-311+G(d,p):ff12SB level of theory and electrostatic embedding\textsuperscript{50} was used to calculate the model layer energies in the presence of...
of the MM point charges in the real layer. We call this combination of RESP charges followed with electrostatic embedding RESP-EE.

Finally, Asn79 was mutated to generate the N79A mutant. The initial structures for the mutant reactants, intermediates and transition states were taken from the corresponding final, ONIOM-RESP-EE optimized geometries of the WT enzyme. RESP partial charges were recalculated for the QM region and the geometry was reoptimized at the M06-2X/6-31G(d):ff12SB level of theory. Final energies were calculated using electrostatic embedding and MP2/6-311+G(d,p):ff12SB. Final atomic coordinates, partial charges and connectivity data used in the Gaussian input files are available in the SI.

Results and Discussions

BOMD: A reactive configuration is required to form the TI±

We used BOMD coupled with MTD to discover the mechanism of Ubc13-catalyzed thioester aminolysis, and the role played by the residues in the active site of Ubc13. Of particular interest is the role played by Asn79, a residue that is highly conserved in the E2 family. We ran two, independent simulations – MTD1 and MTD2 – snapshots of which are shown in Fig. 2. Asn79 is in a different position in each simulation, providing us with a way to interrogate its function. In

![Diagram](image)

**Figure 3** Evolution of the collective variables and the distance between Asn79 and the thioester oxygen for two, independent simulations: MTD1 (a) and MTD2 (b). The left y-axis is distance in Å and the right y-axis is the coordination number (CN). The red curve is the CN of the substrate lysine amine nitrogen (CN\(_{N\varepsilon}\)), the blue curve is the distance between the substrate lysine amine nitrogen and the thioester carbon (R\(_{N\varepsilon-C}\)), and the black curve is the distance between the Asn79 amide sidechain and the thioester oxygen (R\(_{N\delta-O}\)). Both simulations are indicative of a stepwise mechanism; first the nucleophile attacks and then it is deprotonated. Green asterisks mark the formation of the tetrahedral, zwitterionic intermediate (TI±). A persistent hydrogen bond between Asn79 and TI± never forms.

MTD1, Asn79 remains less than 5 Å from the thioester oxygen, whereas in MTD2, Asn79 drifts and is ~ 8 Å away during the reaction. (The tertiary structures of MTD1 and MTD2 are displayed...
in Fig. S4 in the SI.) The analyses of these two simulations reveal three important insights: 1) the reaction occurs in a step-wise fashion, 2) the formation of the zwitterionic, tetrahedral intermediate (TI±) has the largest energetic barrier and, 3) the reaction occurs faster when Asn79 is near the thioester oxygen (i.e., in MTD1). The third point has some subtleties. Specifically, although Asn79 remains near the TI± in MTD1 (< 5 Å in Fig. 3), it never forms a strong, persistent hydrogen bond, even after the TI± forms. Therefore, we hypothesize that the increased rate seen in MTD1 is not due to Asn79 stabilizing TI±. Instead, we attribute the increased reaction rate in MTD1 to the fact that a reactive configuration of the nucleophile and electrophile occurs simultaneously, whereas this configuration takes longer to construct in MTD2. The data in Figs 3 and 4 illustrate this point.

Figure 3 shows the time evolution of the distance between Asn79 and the thioester oxygen (RNεO) as well as the CVs in both simulations. The reaction in both MTD1 and MTD2 is step-wise, as previously hypothesized.3,7 Figure 3 shows that the amine nucleophile attacks first and is deprotonated second (RNεC decreases, then CNNe goes from two to one). Both simulations also show that ubiquitination proceeds through a zwitterionic, tetrahedral intermediate (TI±): the thioester oxygen develops a formal negative charge and the lysine amine develops a formal positive charge. In Fig. 3, the TI± is located at RNεC ~ 1.2 Å and CNNe ~ 2 and is marked by a green asterisk.

![Figure 4](image)

**Figure 4** Evolution of the nucleophilic attack during independent simulations MTD1 (a) and MTD2 (b). The left y-axis is distance in Å and the right y-axis is angle in degrees. The blue curve is the distance between the substrate lysine amine nitrogen and the thioester carbon (RNεC), and the black curve is the angle between the substrate lysine amine nitrogen, the thioester carbon and the thioester oxygen (θNCO). The red line is set at 107°, the Bürgi-Dunitz angle of nucleophilic attack. In both simulations, the TI± only forms if RNεC < 2 Å and θNCO > 107°. Both criteria are met simultaneously in MTD1, but not MTD2.

Although the simulations utilize the same mechanism, the reaction occurs on different time-scales. MTD1 reacts in 13 ps and MTD2 reacts in 21 ps. Although these timings differ by only 8 ps, the free energy barrier for MTD2 is 12 kcal mol⁻¹ higher than MTD1 (see Fig. S4). We
further note that these reactions are greatly accelerated in comparison to in situ kinetics due to the use of the MTD algorithm. The difference in reaction times is mainly due to the way the nucleophiles and electrophiles are oriented during the simulations. In order to form the TI±, not only does the nucleophile have to approach the carbonyl carbon, but the orientation has to be conducive to a reaction. This means that the angle between the substrate lysine amine nitrogen, the thioester carbon and the thioester oxygen (θNCO) has to approach the Bürgi-Dunitz angle of nucleophilic attack (107°) 36. These two parameters just prior to TI± formation are plotted versus time in Fig. 4. In MTD1, both events happen simultaneously at ~10.5 ps in Fig. 4(a). In MTD2, there are instances where the nucleophile/electrophile distance is short, but the angle is too low. For example, at 19.5 ps in Fig. 4(b), RNe-C < 2 Å, but θNCO ~ 100°, so the reaction does not proceed. Thus, we hypothesize that Asn79 helps promote the proper reaction geometry.

ONIOM: Stabilizing the TI± increases the reaction barrier

Figure 5 The reaction profile (left) and proton transport transition states (TS2) (right) for active site Geometry 1, paths A (red) and B (blue). The energies in each path are relative to one of its RCs, so both paths have one RC with an energy of 0 kcal mol⁻¹. Here, Asn79 hydrogen bonds to the thioester oxygen. The bonds breaking during the reaction are shown in red, dotted lines and the transferring proton is indicated with a green arrow. Proton transfer has the largest barrier (TS2) and is 23.0 kcal mol⁻¹ for Path A and 23.2 kcal mol⁻¹ for Path B calculated at the MP2/6-311+G(d,p):ff12SB RESP-EE level of theory. The multiple barrier heights in Path B are due to different configurations of water molecules in the active site.

Because it is challenging to determine whether the MTD free energy calculations have converged, we also computed detailed energy profiles using the ONIOM QM/MM method. We
pulled two different frames from our MD data, each with different Ubc13 active site geometries. In Geometry 1, Asn79 hydrogen bonds to the thioester oxygen and in Geometry 2, water hydrogen bonds instead. Furthermore, we explored two different proton transport (PT) paths; in Path A, PT occurs directly from the lysine $N_\varepsilon$ to the cysteine sulfur and in Path B, PT occurs through a water molecule (see Fig. 5). (A third path where an active site aspartate accepts the proton is detailed in the SI.) Each path has a reactant complex (RC), a TI±, a product complex (PC) and two transition states (TS1, TS2), where TS2 is the PT step. The ONIOM calculations uncover two valuable insights: 1) they confirm that Ubc13-catalyzed ubiquitination proceeds with a stepwise mechanism and 2) they show greater TI± stability leads to a larger barrier to product formation.

Energy profiles for each path in Geometry 1 are shown in Fig. 5. Dotted lines signify if local minima were connected to a transition state via an intrinsic reaction coordinate (IRC) calculation or a relaxed scan along a reaction coordinate. The energy barrier for the direct proton transfer is 23.0 kcal mol$^{-1}$ and the barrier for the water mediated transfer is similar (23.2 kcal mol$^{-1}$). However, due to the rearrangement of active site water, a TI± with slightly higher relative energy was found for Path B, which could reduce the PT barrier to 19.3 kcal mol$^{-1}$.

The mechanistic pathways found for Geometry 2, where water is the hydrogen bond donor, are shown in Fig 6. As also seen in Fig. 5, different configurations of active site waters provide the possibility of several barrier heights. Thus, energy barriers for Path A are 26.4 and 47.3 kcal mol$^{-1}$ while the barriers for Path B are 29.7 and 37.3 kcal mol$^{-1}$. These barriers are significantly higher than those seen in Geometry 1. This implies that water stabilizes the negative charge on

![Figure 6](image-url)

*Figure 6* The reaction profile for active site Geometry 2, paths A (red) and B (blue). The energies in each path are relative to one of its RCs, so both paths have one RC with an energy of 0 kcal mol$^{-1}$. Here, water hydrogen bonds to the thioester oxygen. There are two distinct reaction profiles for each pathway calculated at the MP2/6-311+G(d,p):ff12SB RESP-EE level of theory. Proton transfer has the largest barrier and is 47.3 and 26.4 kcal mol$^{-1}$ in path A and 37.3 and 29.7 kcal mol$^{-1}$ in path B. The multiple barrier heights in both paths are due to different configurations of water molecules in the active site.
TI\textsuperscript{±} better than Asn79, lowering its energy relative to TS2. In other words, these reaction profiles refute the hypothesis that stabilizing TI\textsuperscript{±} increases the reaction rate.

To reinforce this point, we mutated Asn79 to alanine in Geometry 1 and reoptimized the RC, TI\textsuperscript{±} and TS2 geometries. This mutation should destabilize the TI\textsuperscript{±}, since alanine is not a hydrogen bond donor. This is indeed the case as shown in Fig. 7. The N79A TI\textsuperscript{±} is 1.5 kcal mol\textsuperscript{-1} higher in energy than the N79A RC. In the WT enzyme the TI\textsuperscript{±} is 2.0 kcal mol\textsuperscript{-1} lower in energy than the WT RC. However, the mutant TS2 is only 17.4 kcal mol\textsuperscript{-1} higher in energy than the mutant RC, whereas the WT TS2 is 21.2 kcal mol\textsuperscript{-1} higher in energy than the WT RC. Therefore, the mutation reduces the barrier by 7.3 kcal mol\textsuperscript{-1}. Although only one route is shown for clarity, the N79A reaction profiles for all routes (see SI) show a similar, barrier lowering effect. The barrier in the other route in Path B is lowered by 6.5 kcal mol\textsuperscript{-1} and the barrier in Path A is lowered by 5.5 kcal mol\textsuperscript{-1} (see Fig. S6). However, it is known that the N79A mutation decreases the rate of Ubc13\textsuperscript{5} (i.e. raises the energetic barrier). The juxtaposition of our computational data with the experimental rates, therefore, indicates that the final, proton transfer step is not rate-limiting and it is therefore unlikely that Asn79 stabilizes the TI\textsuperscript{±} as part of Ubc13’s catalytic strategy. The contention that stabilizing the TI\textsuperscript{±} increases the barrier is further illustrated in Fig. S7 in the SI, where we recalculate the energies of some of the N79A structures in the presence of a negative or positive point charge placed near the thioester carbonyl. Positive, stabilizing charges increase the barrier, whereas negative, destabilizing charges decrease it.
Finally, we note that our conclusion that the final, proton transfer step is not rate-limiting seems directly contradicted by our ONIOM calculations, since the ONIOM data shows that this step has the largest barrier. However, our BOMD simulations (correctly) indicate that the RC to TI± step is rate-limiting. The difference exists because we placed the nucleophile and electrophile in attack positions in our ONIOM calculations. (The average $R_{Ne-C}$ in our ONIOM calculations is 3.01 Å and the average $\theta_{NCO}$ is 98.0°, a similar configuration to the BOMD data immediately prior to TI± formation (see Fig. 4).) Thus, we have neglected the contribution of conformational entropy required to attain a reactive RC. The data presented in the next section suggest that this is a key component to understanding how Ubc13 works.

**MD: Fluctuations of the thioester increase upon mutation**

Based on our BOMD simulations – which show that the reaction time depends on the configuration of the nucleophile and electrophile – and our ONIOM calculations – which show that the PT barrier increases if there is a strong hydrogen bond to the electrophile – we hypothesize that Asn79 forms a weak hydrogen bond to stabilize the fluctuations of the electrophile, i.e. the thioester carbon. To test this hypothesis, we ran MD simulations of WT and mutated Ubc13 and examined the root mean square fluctuation (RMSF) of the thioester. We studied N79A and N79D mutants, since both mutations eliminate the possibility of hydrogen bonding to the thioester. (In our study, the Asp was deprotonated; its experimental protonation state is unclear.) We also studied the H77A mutant. In the WT enzyme, the imidazole ring in His77 accepts a hydrogen bond from the backbone of Asn79, holding its backbone in place. The H77A mutation removes this interaction and Asn79 moves out of position, completely eliminating the hydrogen bond to the thioester (see Fig. S11). This and other data regarding the H77A mutant are presented in the SI. The effect of mutations on the thioester RMSF is shown in Fig. 8.
Each mutant shows an increase in thioester fluctuations over the WT. In the MD simulations, the WT thioester RMSF is $0.99 \pm 0.09$ Å, whereas the N79D thioester RMSF is $1.61 \pm 0.20$ Å, N79A is $1.26 \pm 0.14$ Å, and H77A is $1.21 \pm 0.09$ Å. Each mutant is known to decrease the enzymatic efficiency and rate\textsuperscript{5,7,8,22}. Our simulation data suggests this decrease may be caused by the increased difficulty that the mutants have in getting the nucleophile and electrophile in the proper configuration. This hypothesis is further supported by analyzing the MTD data (blue and green columns in Fig. 8). The thioester RMSF is higher in MTD2 than MTD1 (1.15 Å vs 0.74 Å) and MTD2 takes longer to form the product (see Fig. 2).

![Figure 8](image)

**Figure 8** The root mean square fluctuation (RMSF) of the thioester for the wild type (WT) and mutant simulations. Both BOMD (MTD) data and classical MD data are shown for the WT enzyme. The protein backbone was aligned prior to RMSF calculation. All three mutants show an increase in thioester fluctuations, indicating that the WT enzyme keeps the thioester more rigid. When N79 is near the thioester carbonyl (MTD1) the fluctuation decreases. The error bars on the MD data is the standard error over the independent simulations; the error bars are omitted for the MTD data due to the short simulation length (< 25 ps).

**Conclusions**

We used three different types of simulations – BOMD, single point QM/MM energy calculations, and classical MD – to uncover the mechanism used by the E2 enzyme, Ubc13, to catalyze the direct transfer of ubiquitin to a substrate lysine. Our data clearly show that this thioester aminolysis reaction is stepwise. The nucleophilic attack is rate-limiting and leads to a stable, zwitterionic, tetrahedral intermediate (TI\textpm). This is followed by proton transport (PT) from the TI\textpm to the thioester sulfur leading to the product. Our simulations also shed light on the role of a highly conserved asparagine residue in the E2 enzyme family, Asn79 in Ubc13. Based on our calculations, we hypothesize that Asn79 helps position the thioester to promote the formation of a reactive geometry.

First, we used BOMD coupled with the rare-events sampling method MTD to locate minimum energy reaction pathways. We ran two independent simulations, each with different
Asn79 positions; in MTD1, Asn79 was near the thioester and in MTD2, Asn79 was distant. (We note that these two configurations are consistent with previously published experimental and computational data.) In both BOMD simulations, the first step – the formation of the TI± – took much longer than the second step, the formation of the product. However, MTD1 reacted faster than MTD2. The data show that this difference in rate results from the fact that it took MTD2 longer to achieve a reactive configuration.

Second, we used QM/MM electronic structure calculations (ONIOM) to generate detailed reaction profiles. Again, we used two different geometries, each with two proton transport pathways. In Geometry 1, Asn79 hydrogen bonded to the thioester and in Geometry 2, water hydrogen bonded to the thioester. These calculations confirmed that Ubc13 uses a stepwise mechanism. However, the PT barrier was higher in Geometry 2, even though water is a stronger hydrogen bonding donor than asparagine and can therefore stabilize the developing negative charge on TI± better than Asn79. This discrepancy occurs because water energetically stabilizes TI± more than the transition state, thereby increasing the barrier. These data contradict the hypothesis that asparagine stabilizes this charge, since charge stabilization leads to larger reaction barriers. We elaborated on this point by showing that the PT barrier in the N79A mutant is lower than the WT barrier by 5.5 – 7.3 kcal mol⁻¹. Alanine cannot donate a hydrogen bond and therefore destabilizes TI±, which lowers the relative reaction barrier. This data implies that the N79A mutant is more efficient than the WT, directly contradicting experimental evidence. Therefore, the PT step cannot be rate-limiting and the main role of Asn79 cannot be to stabilize the negative charge on the reaction intermediate.

Third, we used classical MD to examine the behavior of WT and mutant Ubc13. In particular, we tested the hypothesis that Asn79 reduces the motion of the thioester. We found this to be true. Thioester fluctuations increased in mutants without Asn79 (N79A and N79D) and in a mutant that disrupted the hydrogen bond network that provides Asn79 with structural stability (H77A). Upon reexamination of the BOMD data, we also found that electrophile fluctuations in MTD1 are smaller than in slow-reacting MTD2. Therefore, we hypothesize that Asn79 may promote a reactive configuration by keeping the electrophile in place. This is similar to the mechanism that the RING-E3 ligase is hypothesized to use. In fact, we recently surveyed 16 crystal structures of E2~Ub conjugates and found that six out of nine structures with a RING-E3
ligase have the conserved asparagine within hydrogen bonding distance of the thioester. Of the four structures with no E3 present, only one has the asparagine in this position. (The remaining two structures also do not have the asparagine in a hydrogen-bond configuration, but were bound to HECT and RBR-E3 ligases which use a different mechanism than the RING-E3 ligase. Thus, it seems possible that the presence of the E3 may increase the likelihood that the asparagine forms a hydrogen bond to the thioester, partially explaining its catalytic effect. We are currently running molecular dynamics simulations based on Ubc13~Ub conjugates crystallized with RING-E3 ligases (pdbs 5AIT/5AIU)\(^1\)\(^4\) to test this hypothesis\(^5\)\(^8\). It is also possible to experimentally test the fluctuation hypothesis by using Arrhenius kinetics\(^5\)\(^9\),\(^6\)\(^0\). Since the Arrhenius prefactor gives a measure of the entropic contribution\(^6\)\(^0\), a comparison of combinations of mutant and WT Ubc13 with and without a RING-E3 ligase can provide insight into the role of fluctuations in Ubc13’s catalytic strategy.

Our calculations also do not examine the overall stabilizing effect that Asn79 and the enzyme environment may have on the developing negative charge on the transition state as compared to the same reaction in water. This type of comparison is crucial to understand the catalytic effect of the enzyme\(^1\)^\(^9\),\(^2\)\(^6\),\(^6\)\(^1\)–\(^6\)\(^3\). Recent estimates suggest that the lysine ubiquitination is \(10^8\)-fold faster in the E2 enzyme as compared to water\(^1\)^\(^9\),\(^2\)\(^6\). In fact, even mutant E2 enzymes have a catalytic effect of \(10^5\) – \(10^7\)-fold over water\(^5\),\(^8\),\(^2\)\(^6\) meaning the active site has a catalytic effect even in the absence of Asn79. The addition of an E3 seems to improve the rate by only a factor of ten\(^5\). We argue that in order maximize its efficiency, Ubc13 strikes a balance between reducing the motion of the thioester and over-stabilizing the TI±. This balancing act is observed in MTD1. Asn79 never forms a strong hydrogen bond with the thioester, even when a negative charge forms on the oxygen. It remains between 3 and 5 Å away from the oxygen on TI± (see Fig. 3). A similar effect (intermediate destabilization) was comprehensively studied in Refs \(6\)\(^4\) and \(6\)\(^5\) for other enzymes that proceed through an oxyanion intermediate, although not without controversy\(^6\)\(^3\).

Finally, we note that even if Asn79 positions the thioester for nucleophilic attack, it can still stabilize the active site loop\(^5\),\(^2\)\(^4\). The asparagine sidechain is capable of donating two hydrogen bonds and we have seen that it can simultaneously donate a hydrogen bond to the active site loop (Asn116 in particular) and remain oriented towards the thioester (see Fig. S9 in the SI). Furthermore, we present a reaction profile in the SI (Fig. S8) showing that Asp119 can deprotonate
the lysine, consistent with claims made in Refs. 8, 24, 66 and 67. A picture of Ubc13 is slowly emerging that shows how a network of hydrogen bonds of differing strengths – His77 to Asn79, Asn79 to Asn116, and Asn79 to the thioester – promote ubiquitin transfer by providing structural stability to the active site and substrate.

Author Contributions

IS conceived and supervised the study and designed the experiments. IS, AGD, WMJ, RHW and KLE conducted the experiments. IS and WMJ wrote the manuscript.

Acknowledgements

This work was supported by an NSF-REU CHE-1461175 & CHE-1062629; Extreme Science and Engineering Discovery Environment (XSEDE) NSF ACI-1548562; The MERCURY consortium (http://mercuryconsortium.org/) under NSF grants CHE-1229354 and CHE-1662030; The Thomas F. and Kate Miller Jeffress Memorial Trust, Bank of America, N.A., Trustee; and a Research Corporation Department Development Award #7957. I. S. thanks Dr. C. E. Berndsen for many stimulating conversations.

Supporting Information Additional simulation information about the H77A MD simulations including the hydrogen bonding in its active site, backbone RMSDs; reaction profiles for PT from the substrate lysine to Asp119; custom force field charges and parameters; and input files for the ONIOM and BOMD simulations, including partial atomic charges, geometries and connectivities.

References


(7414), 115–120.


(9) Schiffer, J. M.; Malmstrom, R. D.; Parnell, J.; Ramirez-Sarmiento, C.; Reyes, J.; Amaro, R. E.; Komives, E. A. Model of the Ankyrin and SOCS Box Protein, ASB9, E3 Ligase Reveals a Mechanism for Dynamic Ubiquitin Transfer. Structure 2016, 24 (8), 1248–1256.

(10) Magala, P.; Bocik, W. E.; Majumdar, A.; Tolman, J. R. Conformational Dynamics Modulate Activation of the Ubiquitin Conjugating Enzyme Ube2g2. ACS Omega 2017, 2 (8), 4581–4592.


(33) Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit


(39) CP2K-v2.6 (www.cp2k.org).


(44) Laino, T.; Mohamed, F.; Laio, A.; Parrinello, M. An Efficient Linear-Scaling Electrostatic


