THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

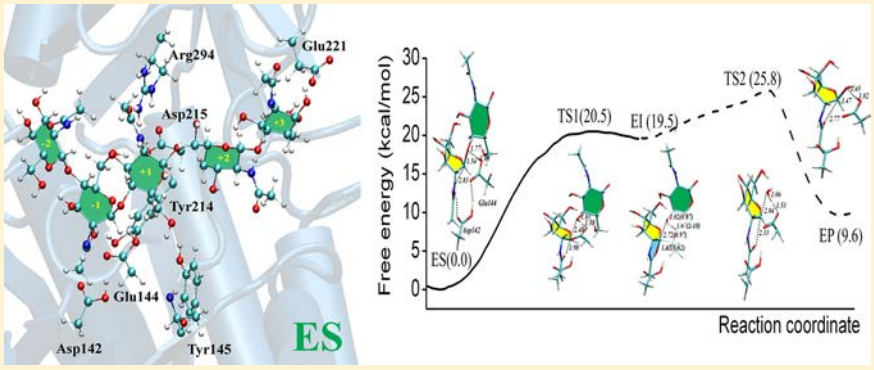
QM/MM Free-Energy Simulations of Reaction in *Serratia marcescens* Chitinase B Reveal the Protonation State of Asp142 and the Critical Role of Tyr214

Jitrayut Jitonnom,*[,†] Michael A. L. Limb,[‡] and Adrian J. Mulholland[‡]

[†]Division of Chemistry, School of Science, University of Phayao, Phayao 56000, Thailand

[‡]Centre for Computational Chemistry, School of Chemistry, University of Bristol, Bristol BS8 1TS, U.K.

Supporting Information



ABSTRACT: *Serratia marcescens* Chitinase B (ChiB), belonging to the glycosidase family 18 (GH18), catalyzes the hydrolysis of β-1,4-glycosidic bond, with retention of configuration, via an unusual substrate-assisted mechanism, in which the substrate itself acts as an intramolecular nucleophile. Here, both elementary steps (glycosylation and deglycosylation) of the ChiB-catalyzed reaction are investigated by means of combined quantum mechanics/molecular mechanics (QM/MM) umbrella sampling molecular dynamics (MD) simulations at the SCC-DFTB/CHARMM22 level of theory. We examine the influence of the Asp142 protonation state on the reaction and the role that this residue performs in the reaction. Our simulations show that reaction with a neutral Asp142 is preferred and demonstrate that this residue provides electrostatic stabilization of the oxazolinium ion intermediate formed in the reaction. Insight into the conformational itinerary ($^{1,4}B \leftrightarrow {}^{4}H_{5} \leftrightarrow {}^{4}C_{1}$) adopted by the substrate (bound in subsite −1) along the preferred reaction pathway is also provided by the simulations. The relative energies of the stationary points found along the reaction pathway calculated with SCC-DFTB and B3LYP were compared. The results suggest that SCC-DFTB is an accurate method for estimating the relative barriers for both steps of the reaction; however, it was found to overestimate the relative energy of an intermediate formed in the reaction when compared with the higher level of theory. Glycosylation is suggested to be a rate-determining step in the reaction with calculated overall reaction free-energy barrier of 20.5 kcal/mol, in a reasonable agreement with the 16.1 kcal/mol barrier derived from the experiment. The role of Tyr214 in catalysis was also investigated with the results, indicating that the residue plays a critical role in the deglycosylation step of the reaction. Simulations of the enzyme−product complex were also performed with an unbinding event suggested to have been observed, affording potential new mechanistic insight into the release of the product of ChiB.

INTRODUCTION

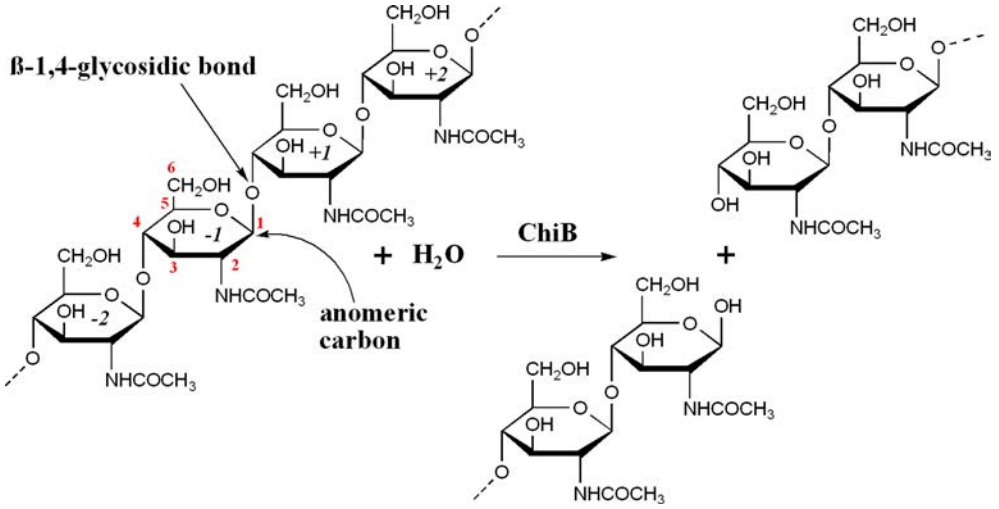
Glycoside hydrolases, or glycosidases (GHs), are the largest group of glycan-degrading enzymes in nature. They are responsible for hydrolysis of glycosidic bonds in carbohydrates and help to catalyze a wide range of important biological functions, such as glycan processing in glycoproteins, remodeling cell walls, and polysaccharide modification and degradation.[1] In addition to their numerous roles in biochemistry, they are highly efficient enzymes, accelerating reactions by as much as 10^{17} times over the spontaneous rate.[2]

GHs have been extensively studied, with >130 different families of GHs reported in the literature;[3] however, despite advances in understanding of the mechanisms underlying GH catalysis, the specific mechanistic details of many GHs remain an important challenge. The first general mechanism for GHs was proposed by Koshland[4] in late 1953: he suggested an acid−base type catalysis, which could proceed with either a retention or an inversion of the anomeric configuration of the substrate. In the ensuing years, a wealth of research (extensively reviewed elsewhere[5,6]) has provided a more detailed description of GH

Received: January 20, 2014
Revised: April 14, 2014
Published: April 15, 2014

   dx.doi.org/10.1021/jp500652x | J. Phys. Chem. B 2014, 118, 4771−4783

Scheme 1. General Representation of the Substrate and Product of the ChiB-Catalyzed Hydrolysis[a]



[a]ChiB catalyzes the cleavage of the β-1,4-glycosidic bond between the NAG residues in subsites −1 and +1 of chitin, a polysaccharide (4 units shown) consisting of *N*-acteylglucosamine (GlcNAc or NAG) units linked by β-1,4-glycosidic bonds.

function. In brief, inverting GHs operate by a single nucleophilic substitution with only one transition state, whereas retaining GHs follow a double displacement mechanism, proceeding via a stable covalent glycosyl−enzyme intermediate (glycosylation step), followed by the subsequent hydrolysis of this intermediate to form the cleaved reaction product (deglycosylation step). Inverting and retaining GHs usually require the presence of at least two key residues in the active site of the enzyme, one acting as the general acid and the other as the general base, in order for efficient catalysis to take place. In some GHs, however, suitable residues required to perform these specific roles are absent from, or poorly positioned, within the enzyme active site and therefore cannot contribute to catalysis in this proposed manner. This has led to an alternative mechanism, known as the substrate-assisted mechanism, to be proposed for these types of GHs.

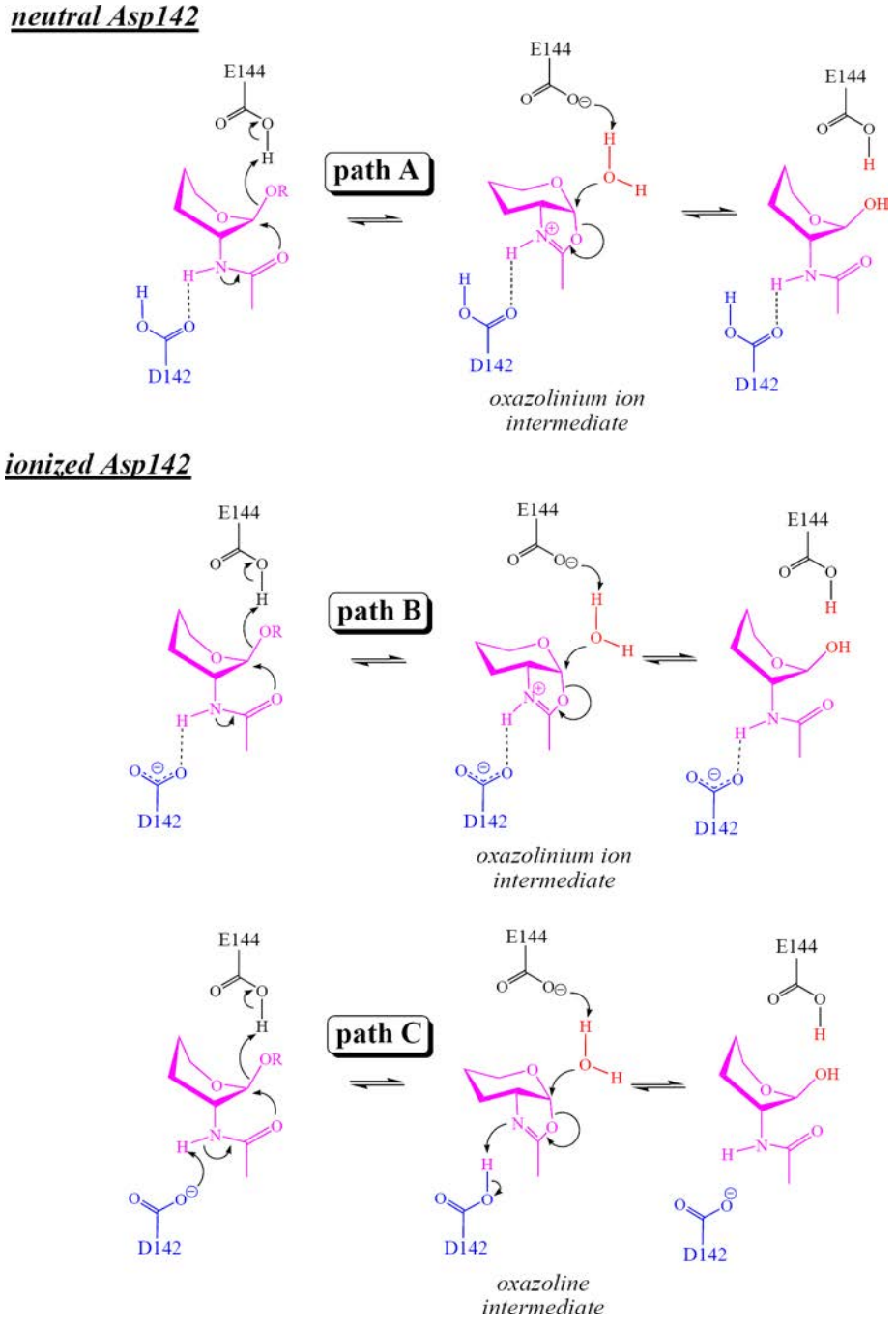
Glycosidic bond hydrolysis occurring via a substrate-assisted catalytic mechanism is an unusual case observed in some glycosidases, where the substrate itself acts as an intramolecular nucleophile instead of an enzyme residue.[7] This mechanism is thought to be used by some retaining GHs enzymes (families 18, 20, 56, 84, and 85) acting on substrates containing an *N*-acetyl (acetamido) group at the C2 position, several GH3 and GH22 enzymes,[8,9] and inverting GH19 enzymes.[10] Additionally, participation of the *N*-acetyl group has also been proposed to take place in glycosyltransferase.[11] It is believed that all these enzymes share a common mechanistic feature, where in the absence of a suitably positioned nucleophilic residue in the enzyme active site, the acetamido group on the substrate is instead directly involved in the formation of the intermediate (proposed to be either an oxazoline or oxazolinium) in the first step of the reaction.[7] Several crystal structures of GH18, GH20, GH56, GH84, and GH85 family enzymes have been solved, and in conjunction with several experimental[12−15] and computational investigations,[7,16,17] the substrate-assisted mechanism is suggested to be an accurate description of the enzyme-catalyzed reaction in these enzymes.

S. marcescens Chitinase B (ChiB), belonging to the GH18 family, degrades chitin [an insoluble linear polymer of β-(1,4)-linked *N*-acetylglucosamine $(NAG)_n$], which is the second-most abundant biopolymer in nature after cellulose. ChiB has received much attention as an attractive system for the development of new inhibitors with chemotherapeutic potential.[18,19] It has also been applied in biotechnology for conversion of insoluble polysaccharides into a commercially valuable product.[20] It is proposed that ChiB catalyzes the hydrolysis of the glycosidic bonds found in chitin via the aforementioned substrate-assisted mechanism (see Scheme 1). Previous studies on ChiB have indicated that the enzyme's catalytic function depends on a relatively large number of residues:[21] Glu144 is proposed to act as a catalytic acid/base,[15] Asp142 and Asp140 are suggested to be involved in the binding of substrates as well as catalysis,[21] and a variety of other conserved residues (e.g., Tyr10, Ser93, Tyr214, and Asp215) have also been found to influence the catalytic activity of the enzyme.[21] Of these residues, Asp142 and Tyr214 are of particular interest because their mechanistic roles are still not fully understood.

In our recent theoretical study, we modeled the glycosylation reaction of ChiB using QM/MM potential energy surface calculations, treating the Asp142 as a neutral residue.[7] Our calculations suggested that Asp142, when in its neutral form, plays a critical role in catalysis by electrostatically stabilizing the transition state and oxazolinium ion intermediate formed in the glycosylation step of the reaction. The study provided useful insight into the preferred catalytic pathway for the reaction when the Asp142 is treated as a protonated residue. However, when comparing the protonation state of the Asp142 residue in ChiB to its analogous residue in a variety of similar GHs (Asp313 in *Streptomyces plicatus* hexosaminidase, GH20,[22] and Asp174 in human *O*-GlcNAcase, GH84),[23] all of which proceed via a substrate-assisted mechanism, it is found that these similar enzymes follow a reaction mechanism where the deprotonated form of the residue is favored. Consequently, it is still not fully understood why the protonated form of the Asp142 residue is preferred in ChiB and what the fundamental impact this difference has on the role of the residue in the reaction. Therefore, in order to investigate the origin of this preference for the neutral protonation state of this vital active site residue, both possible protonation states of the residue were modeled, by simulating three possible reaction pathways for the ChiB-catalyzed reaction (Scheme 2).

dx.doi.org/10.1021/jp500652x | J. Phys. Chem. B 2014, 118, 4771−4783

Scheme 2. (A–C) Three Reaction Pathways for Hydrolysis of Chitin by ChiB Modeled in This Study Utilizing QM/MM Umbrella Sampling MD



Tyr214, in conjunction with Asp142 and Asp215, is proposed to interact with the substrate upon binding. Specifically, Tyr214 has been observed in the crystal structure to form an interaction which contributes to the distortion of the *N*-acetyl group on the −1 sugar.[15,21] Despite the presence of this apparently favorable interaction with the substrate, mutagenesis study[21] has shown that k_{cat}, and not K_m, is affected by the Y214F mutation (0.117 ± 0.011 s^{-1}), resulting in a 160-fold decrease in the k_{cat} compared with the wild-type (17.8 ± 2.3 s^{-1}) at pH 6.3. Similar perturbations to the rate have also been observed in Y390F in ChiA,[24] Y183F in hevamine,[25] and Y227F in ChiNCTU2.[26] The reason for this significant reduction in the rate and also catalytic activity upon mutation of Tyr214 remains largely unresolved. Currently it is proposed, based on evidence from experimental studies looking at the binding of reaction intermediate analogues[27,28] and our recent theoretical study,[7] that the Tyr214 plays an important role in destabilizing intermediates formed during the enzyme-catalyzed reaction and, hence, it is this that could account for the observed decrease in the k_{cat} upon mutation on the residue. In our computational study, the calculations concluded that destabilization of the oxazolinium ion intermediate is provided by a hydrogen bond formed between the Tyr214 hydroxyl group and the *N*-acetyl group of the −1 subsite NAG; however, any further insights into the role the residue plays in the later parts of the reaction was not obtained, as our calculations did

dx.doi.org/10.1021/jp500652x | J. Phys. Chem. B 2014, 118, 4771–4783

not include modeling of the deglycosylation step of the reaction. In this study, the catalytic role of the two conserved residues, Asp142 and Tyr214, was investigated by performing combined QM/MM (quantum mechanics/molecular mechanics) molecular dynamics (MD) simulations. Three different reaction pathways (see Scheme 2) were modeled to investigate the most likely protonation state of Asp142 residue in the reaction (neutral/protonated or ionized/deprotonated) and to provide a more detailed understanding of the catalytic mechanism of *S. marcescen* ChiB at the atomic level. Both the glycosylation and the deglycosylation steps of the reaction were modeled utilizing adiabatic mapping calculations.[7] For the first time, the free energy for the complete catalytic reaction for the ChiB was calculated, in order to determine the most likely pathway of the full ChiB-catalyzed reaction. Higher level energy corrections (using hybrid density functional theory) were also performed on the structures obtained to validate the computational methods used for the reaction pathway calculations. The role of Tyr214 in catalysis was examined via hydrogen bond analysis, and free-energy calculations were utilized to assess the quantitative contribution of this residue to the calculated free-energy barrier of the reaction in ChiB.

COMPUTATIONAL DETAILS

The initial structure of the enzyme−substrate (ES) complex was taken from the X-ray structure of the E144Q mutant of *S. marcescens* ChiB with a chitopentaose substrate bound (PDB code 1E6N).[15] The wild-type was recovered by manually altering Gln144 to Glu144. Two different models (neutral and negatively charged/ionized), differing in the protonation state of the catalytic residue Asp142, were prepared to test the proposed catalytic mechanisms (Scheme 2). These were generated from the wild-type structure using a standard patch in CHARMM to adjust the Asp142 protonation state. All crystallographic water molecules were retained. Hydrogen atoms were added using the HBUILD subroutine in CHARMM, and the titratable residues in the enzymes were assigned based on the pK_a estimated by PROPKA 2.0 (http://propka.ki.ku.dk)[29] at their physiological pH. The protonation state of the catalytic triad (Asp140-Asp142-Glu144) for the two models was different: in both, Asp140 was deprotonated and Glu144 was protonated, with Asp142 deprotonated in the ionized model but treated as protonated in the neutral model. All other aspartate and glutamate residues were treated as deprotonated. Histidine residues (none of which are located near the active site) were modeled in their neutral states, with their tautomeric state assigned on the basis of the hydrogen-bonding network using WHAT-IF (http://swift.cmbi.ru.nl).[30] All simulations were performed using the CHARMM program (version c27b2).[31]

For the QM/MM MD simulations, the system was partitioned into two regions: a QM region consisting of a relatively small number of atoms directly involved in the reaction and a MM region consisting of all the remaining atoms in the system in order to include the important long-range effect of protein/solvent environment in the calculation. The QM region (see Figure 2c) for the glycosylation step consisted of the side-chains of the Glu144 and Asp142 and the NAG residues at subsites −1 and +1 of the substrate (chitobiose), for a total of 72 and 73 atoms for ionized and neutral Asp142 model, respectively. The QM region for the deglycosylation step was different to the one used to model the first step of the reaction and consisted of the NAG monomer at subsite −1 of the substrate (a chitose unit), the side-chains of Glu144 and Asp142, and a catalytic water molecule, for a total of 45 and 46 atoms for ionized and neutral Asp142 model, respectively. The selection of the catalytic water molecule for the deglycosylation step was based on the product of the preceding glycosylation step, where the water observed in the QM/MM MD simulations (Figure S1 of the Supporting Information) to be best positioned for nucleophilic attack was treated with QM methods. In addition, the aglycon leaving group at subsites +1, +2, and +3 was removed from the enzyme active site and replaced by TIP3P water molecules for the deglycosylation step. The QM regions were treated using a self-consistent charge-density functional tight-binding (SCC-DFTB) method,[32] as implemented in CHARMM.[33] SCC-DFTB has successfully been applied to study reactions in several important enzymes.[34−36] The MM regions for both reaction steps, consisting of all the remaining enzyme, substrate, and solvent molecules not included in the QM region, were described with the CHARMM22 all-atom force field[37] and additional parameters for the NAG moieties, as used in our previous study.[7] Hydrogen "QQ" type link-atoms[38] were placed along covalent bonds crossing the QM/MM boundary and were situated between the C_β and C_α atoms of the enzyme residues and along the C−O bond of the glycosidic linkage for both sugar moieties in the substrate (requiring 4 H-link atoms for the glycosylation step and 2 H-link atoms for the deglycosylation step).

The systems were prepared for QM/MM MD simulations using the same protocols applied successfully in our earlier study.[7] In brief, the ES complex was solvated by a 25 Å radius sphere of pre-equilibrated TIP3P model waters[37,39] centered on the anomeric C1 carbon (see Scheme 1). A spherical deformable boundary potential[40] with a 25 Å radius was used to prevent the water from "evaporating" from the system. All atoms outside the 25 Å sphere centered on the anomeric carbon were deleted, while protein heavy atoms in the buffer zone (21−25 Å) were subject to Langevin dynamics with positional restraints using force constants scaled to increase from the inside to the outside of the buffer. All atoms within a 21 Å sphere of the reaction zone were subjected to Newtonian dynamics with no positional restraints. The ES complex for each system was thermalized in the NVT ensemble at 310 K with 1 ns of stochastic boundary QM/MM MD simulation, following the procedure described in refs 7 and 41. An integration time-step of 1 fs was used, with all of the bonds involving hydrogen atoms constrained using SHAKE.[42]

Several snapshots were taken from the equilibrated QM/MM MD simulations, 300−1000 ps, to ensure a diverse range of enzyme−substrate conformations were sampled, which has been found to be important in similar studies.[7,43] These initial geometries, extracted from MD, were subsequently energy-minimized with the Adopted Basis Newton−Raphson method until the gradient was <0.01 kcal/(mol Å), to be used as starting points for QM/MM adiabatic mapping calculations.[44] For modeling of the reaction by adiabatic mapping, the reaction coordinates for both steps of the reaction (glycosylation and deglycosylation) along the paths A−C (Scheme 2) are linear combinations of interatomic distances as defined in Figure S2 of the Supporting Information. Potential energy surfaces (PESs) were calculated, and the geometries representing the minimum energy pathway (MEP) through the surfaces were used as the putative reaction coordinates for the free energy (potential of mean force, PMF) calculations.

dx.doi.org/10.1021/jp500652x | J. Phys. Chem. B 2014, 118, 4771−4783

The free-energy profile for each pathway was computed using QM/MM umbrella sampling MD simulations requiring a series of simulations to be performed with a harmonically restrained reaction coordinate utilizing a force constant of 200 kcal/(mol Å^2). All other aspects of the umbrella sampling simulations were the same as the QM/MM MD simulations described above. Each simulation (window) consisted of 60 ps of equilibration and 40 ps of sampling dynamics. The free-energy profiles were obtained by combining the statistics from all of the simulations performed for each reaction using the weighted histogram analysis method (WHAM).[45] These approaches have been applied successfully in previous work for other enzymes,[46,47] including the GH5 member Cel5A cellulase.[34]

RESULTS AND DISCUSSION

QM/MM MD Simulations of Enzyme−Substrate Complex. Two 1 ns QM/MM MD simulations of the Michaelis complexes, with Asp142 treated as ionized and neutral, were conducted to obtain the initial average active-site structures for reaction modeling. The simulations indicate that both ES complexes stable, reaching the equilibrium after the first 300 ps of the simulations, as indicated by the root-mean-square deviations (RMSD) relative to the starting structure calculated for the simulations (Figure 1a). The ionized model (0.28 ± 0.01 Å) exhibits a slightly higher RMSD compared to the neutral model (0.24 ± 0.01 Å). Snapshots of the ES complex for both models are shown in Figure 2 (panels a and b). The substrate was found to bind within the enzyme active sites via a combination of hydrogen bonds (H-bonds) and van der Waals interactions for the majority of the simulation. These interactions maintained the distortion of the pyranose ring in the −1 subsite in the boat conformation (see angles C2−C1−O5−C5 and C3−C2−C1−O5 in Table S1 of the Supporting Information), consistent with the experimentally determined crystal structure.[15] Tyr214 and Asp142 both make major contributions to this distortion by forming tight H-bond interactions with the *N*-acetyl group of the −1 sugar (d_6 = 1.93 ± 0.13/1.68 ± 0.08 Å and d_9 = 1.70 ± 0.10/1.75 ± 0.11 Å for neutral/ionized models, respectively; Figure 2c and Table S1 of the Supporting Information), enabling the oxygen atom of the *N*-acetyl to move toward the anomeric carbon (d_4 = 2.85 ± 0.13/2.92 ± 0.14 Å for neutral/ionized models, respectively; Table S1 of the Supporting Information), thus helping to promote intramolecular nucleophilic attack by narrowing the distance between the two atoms. The other sugar subsites (−2, +2, +3) remained in a chair conformation, stabilized by the solvent-exposed aromatic residues (e.g., Trp97 and Trp220) and, to a lesser extent, by hydrogen bonding. Interestingly, a loss of the H-bond interaction between Asp142 and Glu144 (see Figure 1b) was observed in some parts of the simulation in the ionized model; this meant the Glu144 side chain could move away from the Asp142 carboxyl group and point toward Tyr145 (Figure 2b). In this new position, the side chains of both Asp142 and Glu144 were now exposed to the solvent (Figure S3 of the Supporting Information), therefore preventing them from interacting effectively with substrate. Despite this observation in the ionized model simulations, the Glu144 residue on average, in both models, was found to be well-positioned for proton donation to the glycosidic bond, as required for the first step of the reaction, with a favorable d_2 distance of ∼1.8 Å maintained over the course of both simulations.

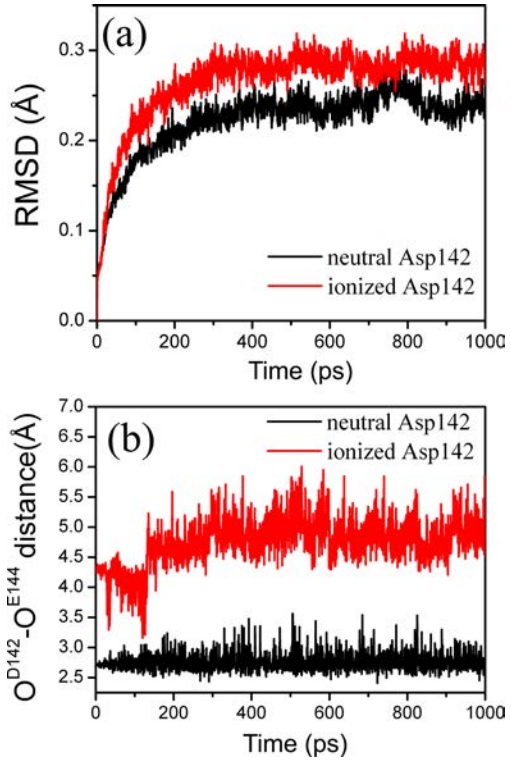


Figure 1. (a) RMSD of the protein heavy atoms and (b) Asp142 Oδ2 − Glu144 Oδ2 distance relative to the starting structure during 1 ns of QM/MM MD simulations of enzyme−substrate complex (ES) with neutral and ionized Asp142.

Protonation State of Asp142 and Catalytic Mechanism of *S. marcescens* ChiB. Two carboxyl residues (Glu144 and Asp142) in ChiB are proposed to play a crucial role during hydrolysis; while Glu144 is known to function as a catalytic acid/base, the influence of the chosen protonation state of the Asp142 residue and its impact on the residue as an electrostatic stabilizer[7] is still not fully understood. Here, three reaction pathways were modeled in order to determine which was the most likely description of the ChiB-catalyzed reaction and to investigate the effect of the protonation state of Asp142. The mechanistic details of the pathways modeled are shown in Scheme 2. In the first pathway (denoted as "path A"), the reaction proceeds via an oxazolinium ion intermediate with Asp142 treated as protonated, helping to stabilize the reaction species electrostatically,[7] as proposed by Van Aalten et al.[15] The second pathway (denoted as "path B") is similar to path A and is consistent with the mechanism proposed for GH20 β-hexosaminidase,[17] where an oxazolinium ion intermediate is formed in the reaction; however, in contrast to path A, Asp142 is treated as deprotonated and acts as a charged residue providing electrostatic stabilization of the oxazolinium intermediate. For the third pathway (denoted as "path C"), an alternative mechanism was proposed in which, Asp142, in conjunction with Glu144, acts as a proton-shuttle during hydrolysis, resulting in a mechanism which proceeds via an oxazoline intermediate similar to the one proposed to be formed in GH84 *O*-GlcNAcase[16] and GH85 endohexosaminidases.[12] The free-energy profile for each path was calculated, and the results obtained are shown in Figure 3. Representative geometries of the QM region along the reaction pathway are also displayed. Geometric details of stationary structures, as

dx.doi.org/10.1021/jp500652x | J. Phys. Chem. B 2014, 118, 4771−4783

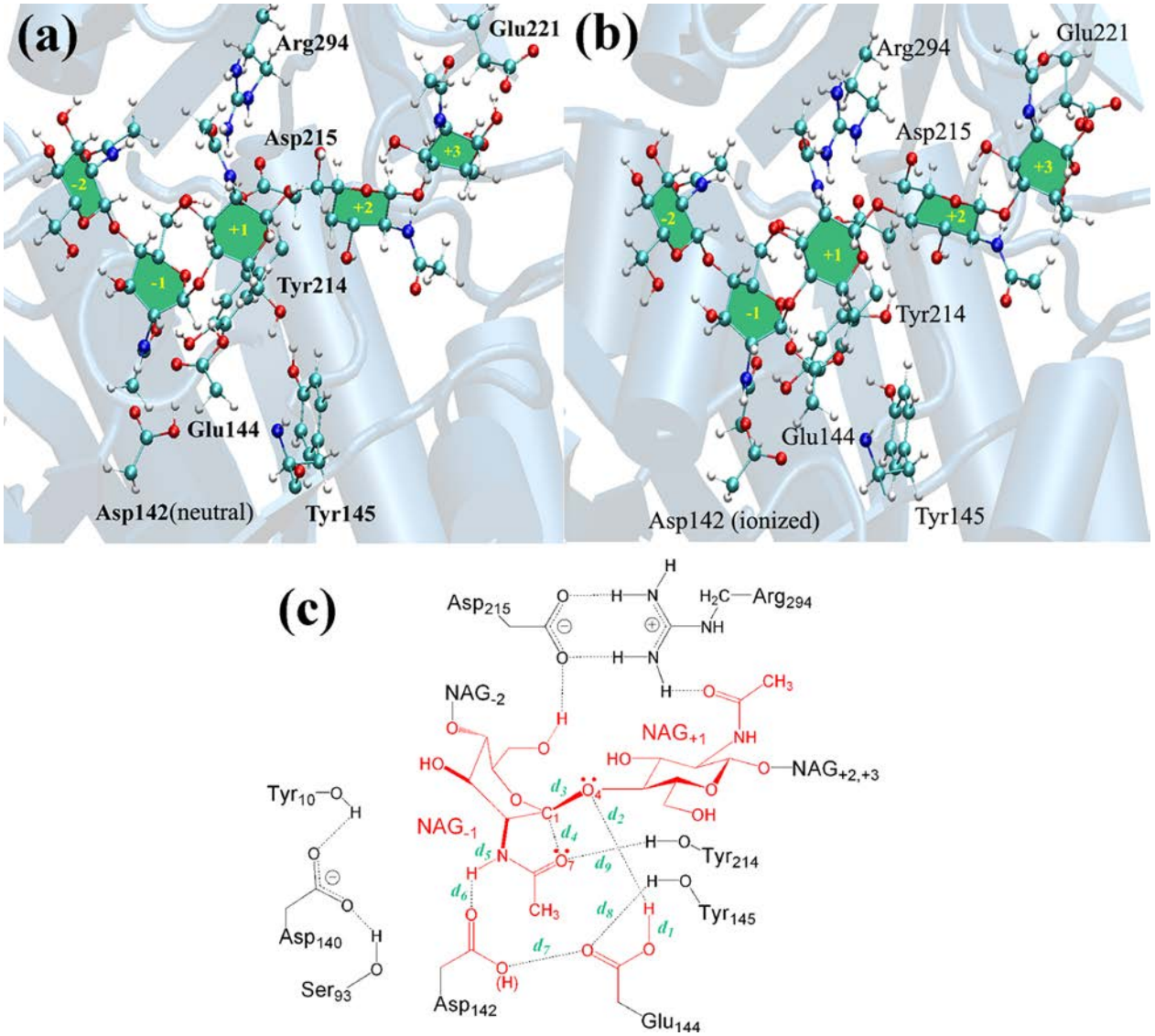


Figure 2. Snapshots of the equilibrated enzyme–substrate (ES) complex from QM/MM MD simulations containing (a) neutral and (b) ionized Asp142. (c) QM region for the glycosylation step (in red) and its structural parameters referred to in the text.

well as the conformations of the NAG moiety in subsite −1 along paths A–C, can be found in Table S2 and Figure S4 of the Supporting Information.

Mechanism via Oxazolinium-Ion Intermediate. Paths A and B both involve formation of an oxazolinium cation intermediate (in glycosylation) which then collapses upon hydrolysis (deglycosylation). In Figure 3 (panels a and b), it can be seen that the ES structures from paths A and B have similar geometries, with both displaying a twisted glycosidic bond linking the two glucosyl moieties in the −1 and +1 subsites. It was observed that a H-bonding network consisting of Asp142 and Glu144 helps to position the substrate in a reactive orientation within the active site. Additionally, a H-bond formed between Tyr214 and the acetamido group of the subsite −1 NAG residue (d_9) was also found to contribute to the distortion of this acetamido group. This distortion facilitates a reduction in the distance between the two atoms involved in the intramolecular nucleophilic attack (d_4 = 2.83 ± 0.06 and 2.91 ± 0.05 Å for paths A and B, respectively) and thus helps to promote the first step of the reaction.

As shown in Figure 3a, three H-bonds are conserved throughout pathway A: Asp142 forms hydrogen bonds with the carboxylate group of Glu144 (d_7) and with the H–N bond of the acetamido group within the subsite −1 NAG (d_6) and Glu144 acts as a proton donor, forming a hydrogen bond to the scissile glycosidic linkage (d_2), with this later H-bond found to be formed at a shorter distance in path A (1.77 ± 0.09 Å) compared to the same H-bond observed in path B (1.85 ± 0.12 Å). The increased length of this H-bond, as observed in path B, coincided with the loss of the H-bond interaction between Glu144 and Asp142 (see Figure 3b). It therefore follows that the Asp142–Glu144 interaction must be present to enable the Glu144 to interact strongly with the substrate (supporting our previous conclusions made when investigating the D142N system).[7] Additionally, these findings support our previous observations from adiabatic mapping that the presence of these H-bonds and their ability to stabilize the oxazolinium cation species electrostatically contribute to a lower reaction barrier.[7]

In the glycosylation step, two chemical processes were involved in the cleavage of the glycosidic bond. The formation of EI was found to be concerted but asynchronous in both

dx.doi.org/10.1021/jp500652x | J. Phys. Chem. B 2014, 118, 4771–4783

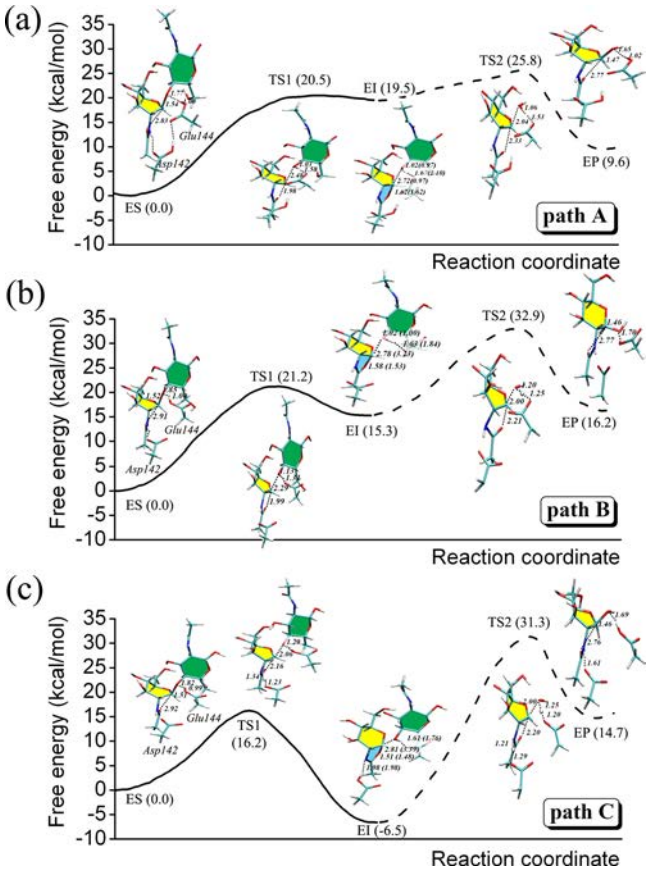


Figure 3. Free-energy profile for paths A–C calculated with QM/MM umbrella sampling MD. Representative geometries of the QM region at the stationary points (ES, TS1, EI, TS2, and EP) along the profile are displayed with important structural parameters labeled in units of Å. The unbroken line represents glycosylation, and the dashed line represents deglycosylation.

paths A and B. Starting from ES, Glu144 completely donates its proton to the glycosidic oxygen (d_2; 1.03 ± 0.03 Å for path A and 1.13 ± 0.03 Å for path B), while at the same time, the scissile glycosidic bond is cleaved to form TS1 (d_3; 2.48 ± 0.06 Å for path A and 2.29 ± 0.05 Å for path B). After this concerted step, it was observed that the covalent bond in TS1 between the *N*-acetyl nucleophile and the anomeric carbon is too large to be considered fully formed (d_4; 1.98 ± 0.06 Å for path A and 1.99 ± 0.05 Å for path B). Therefore, in an asynchronous process, intramolecular nucleophilic attack takes place, resulting in a narrowing of the *N*-acetyl oxygen and the anomeric carbon distance and the formation of the fused bicyclic ring found in the oxazolinium cation intermediate (EI). On the basis of the calculated geometries, the results indicate the reaction for paths A and B proceed first, with the transfer of the Glu144 proton onto the glycosidic oxygen before the oxazolinium intermediate is fully formed. Then, in a distinct step, nucleophilic addition of the *N*-acetyl oxygen at the anomeric center takes place, completing the formation of the intramolecular carbon–oxygen bond (d_4) and forming EI.

During the deglycosylation, the reverse process takes place. Nucleophilic attack of a water molecule at the anomeric center results in the collapse of the intermediate and the subsequent formation of the reaction product (EP). In TS2, the catalytic water was found to be positioned close to the anomeric carbon [d_3; 3.03 ± 0.07 Å (EI) → 2.04 ± 0.05 Å (TS2) for path A and 3.23 ± 0.06 Å (EI) → 2.00 ± 0.05 Å (TS2) for path B], while the fused oxazoline ring is spontaneously broken [d_4; 1.62 ± 0.07 Å (EI) → 2.33 ± 0.05 Å (TS2) for path A and 1.53 ± 0.04 Å (EI) → 2.21 ± 0.04 Å (TS2) for path B]. No proton abstraction was found at this stage in path A, as indicated by a distance of 1.53 ± 0.09 Å (d_1) between the water hydrogen and the Glu144 carboxylate oxygen, which is too large to be consistent with the formation of a bond. In contrast, path B indicates that proton abstraction between these two atoms had already taken place, with distances of 1.20 ± 0.01 Å and 1.25 ± 0.04 Å observed between the hydroxide water ($^-$OH) and its corresponding proton (d_2) and the same proton with the Glu144 carboxylate oxygen (d_1). These distances are representative of the formation of two transient bonds, indicating the transfer of a proton from the catalytic water onto the Glu144 residue. After TS2, in both paths A and B, the hydrogen abstraction was fully completed, resulting in the formation of the EP. The results of the calculation suggest that in path A the breakdown of the glycosidic linkage takes place prior to the proton abstraction of the water by Glu144, while these two steps are observed to be concerted in path B.

The calculated free-energy barriers for glycosylation and deglycosylation are 20.5 and 6.3 kcal/mol, respectively, for path A (Figure 3a) and 21.2 and 17.6 kcal/mol for path B (Figure 3b). For path A, an overall free-energy barrier of 25.8 kcal/mol was found, compared to a barrier of 32.9 kcal/mol determined for path B. These results therefore indicate that path A has a lower overall free-energy barrier to reaction. In the calculated free-energy profiles for both paths, deglycosylation was observed to be the rate-determining step in the reaction. The reactive conformation of the substrate was found to contain a glycosyl moiety at subsite −1, which was distorted away from its lowest-energy chair conformation. The conformation itinerary adopted by the NAG moiety at the subsite −1 of both pathways (Figure S4 of the Supporting Information) followed a boat ($^{1,4}B$) → half-chair $[^4H_5]^‡$ → chair ($^4C_1$) pattern for glycosylation. However, for deglycosylation the conformational changes of the moiety in the −1 subsite for path A, chair ($^4C_1$) → half-chair $[^4H_5]^‡$ → boat ($^{1,4}B$), was different from path B, chair ($^4C_1$) → boat $[^{2,5}B]^‡$ → chair ($^2C_5$), despite both pathways following a similar reaction mechanism.

The calculated energies and observed conformational itineraries for the reaction profiles showed that the major difference between paths A and B (and therefore the impact of a protonated Asp142 compared to a deprotonated Asp142) was in the deglycosylation step of the reaction. The smaller overall free-energy barrier calculated for the reaction proceeding via path A could be attributed to the reaction taking place through a lower-energy transition state (TS2) compared to the one observed in path B. This energy difference is likely due to the stronger H-bond network formed by residues Asp142, Glu144, the catalytic water, and the H–N bond of the acetamido group (d_1, d_6, d_7; Figure 2c) with these interactions providing more stabilization of the half-chair glycosyl moiety in TS2 of path A than the boat conformation formed in path B. These geometric observations are supported by the stabilization calculations performed on the system (Figure S5 of the Supporting Information) that show that the enzyme provides ∼5 kcal/mol of stabilization of the reaction in enzyme in path A compared to minimal stabilization observed in path B.

In contrast, in the glycosylation step of the reaction, path B, despite representing a less energetically favored overall pathway and greater barrier to glycosylation, indicated that the

dx.doi.org/10.1021/jp500652x | J. Phys. Chem. B 2014, 118, 4771–4783

unprotonated Asp142 stabilized the oxazolinium ion species (EI) more effectively than the protonated Asp142 in path A (presumably because of the ion−ion interaction between the negative charge of the ionized Asp142 and the developing positive charge of the oxazolinium intermediate). Consequently, a shorter d_6 distance at the EI (EI1/EI2) was observed for path B (1.41−1.58 Å, Table S2 of the Supporting Information) compared to path A (1.60−1.77 Å, Table S2 of the Supporting Information). These findings therefore provide some evidence to explain why the analogous residue to Asp142 in similar GHs (GH20) is believed to be deprotonated in the active enzyme.[17,22]

Mechanism via Oxazoline Intermediate. In contrast to paths A and B, path C consists of a mechanism that proceeds via an oxazoline intermediate with both Asp142 and Glu144 acting as the catalytic acid/base residues. The calculated energies and geometries for path C are shown in Figure 3c. The starting geometries of the substrate and active site residues in path C are almost identical to that observed in path B, as would be expected due to both systems containing Asp142 in the same protonation state. During the glycosylation, cleavage of the glycosidic bond was found to occur concomitantly with the proton abstraction by Asp142 (which acts as a general base) from the H−N bond of the acetamido group at the NAG subsite −1. This event is likely to be assisted by the H-bond polarization at the glycosidic bond by Glu144. Additionally, it was observed in TS1 that two transient bonds (d_5 =1.34 ± 0.01 and d_6 =1.23 ± 0.04 Å) between the acetamido group and the Asp142 carboxylate oxygen, and a further two transient bonds (d_3 = 2.06 ± 0.04 and d_4 = 2.16 ± 0.04 Å) between three atoms of the glycosidic oxygen, anomeric carbon, and carbonyl oxygen were formed, indicating that the Glu144 proton had not transferred fully onto the oxygen in the scissile glycosidic bond. After TS1, the Glu144 proton is fully transferred onto the cleaved glycosidic oxygen and the acetamido proton is bonded to the Asp142 carboxylate oxygen, leading to the formation of the oxazoline intermediate. In deglycosylation, the aglycon leaving group is removed and the nucleophilic attack of the oxazoline intermediate was performed. In this final step of the reaction, proton abstraction from a water molecule by Glu144 occurs concomitant with the proton transfer from Asp142 to the acetamido group, indicated by d_1, d_2, d_5, d_6 distances in the range of 1.20−1.29 Å observed in TS2 (see Figure 3c). On completion of the deglycosylation step, EP is obtained, consisting of a substrate which has retained the original stereochemistry of its anomeric carbon with respect to the reactant, as observed for paths A and B.

The calculated free-energy barriers for glycosylation and deglycosylation are 16.2 and 37.8 kcal/mol, respectively. Deglycosylation, as observed in paths A and B, was found to be the rate-determining step in the free-energy profile calculated for path C. The conformation of the ring in the −1 subsite in path C was found to adopt boat ($^{1,4}B$) → half-chair $[^4H_5]^‡$ → chair ($^5C_2$) for the glycosylation, and chair ($^5C_2$) → boat $[B_{2,5}]^‡$ → boat ($B_{2,5}$) for the deglycosylation.

Why is Path A Energetically Favored?: Comparison of SCC-DFTB and B3LYP Results. From our calculations, path A has the lowest free-energy barrier (25.8 kcal/mol) of the three pathways modeled, with paths B and C calculated to have overall barriers of 32.9 and 31.3 kcal/mol, respectively. The barriers for each step of the reaction (glycosylation and deglycosylation) differ significantly among the three pathways with deglycosylation found to have the most variation in barrier height; it is also predicted to be the likely rate-determining step of the reaction. The barrier to glycosylation was lowest in path C (16.2 kcal/mol) followed by paths A (20.5 kcal/mol) and B (21.2 kcal/mol). In contrast, the barrier to deglycosylation (calculated as the difference in energy between TS2 and EI2) was lowest in path A (6.3 kcal/mol) followed by paths B (17.6 kcal/mol) and C (37.8 kcal/mol). Path A therefore represents the most energetically favorable pathway for the ChiB-catalyzed reaction, with a free-energy barrier of 25.8 kcal/mol. This is significantly higher than the experimentally derived barrier of 16.1 kcal/mol (k_{cat} = 28 s^{-1}); this is probably due to limitations of the SCC-DFTB QM method used in the calculations.[48]

Small model, single-point (SP) calculations utilizing a higher level of theory (B3LYP) were used to estimate the relative accuracy of the energies calculated with the SCC-DFTB method. In order to perform this comparison, representative geometries of the QM region of the stationary points found along the SCC-DFTB/MM free-energy paths A to C were extracted. SP calculations were then performed on each of these structures with SCC-DFTB and B3LYP to compare the relative energies of the structures calculated with the different methods. Additionally, the basis set dependence of the B3LYP-calculated energies was assessed by performing SP B3LYP calculations with a variety of different-sized basis sets: 6-31G(d), 6-31G(d,p), and 6-311++G(d,p) (Table 1).

The SP corrections performed on path A show that the relative energies of the stationary points representing TS1 and TS2 obtained with SCC-DFTB are in good agreement with

Table 1. Comparison of Relative Potential Energies (in kcal/mol) of the QM Region at the Six Stationary Points (ES, TS1, EI1, EI2, TS2, and EP) for Both Glycosylation and Deglycosylation Among Three Reaction Pathways (A, B, and C)[a]

	glycosylation			deglycosylation		
path:methods	ES	TS1	EI1	EI2	TS2	EP
path A: US(SCC-DFTB/CHARMM)	0.0	20.5	19.5	0.0	6.3	−9.9
SP(SCC-DFTB)	0.0	19.7	13.9	0.0	11.6	−11.8
SP[B3LYP/6-31G(d)]	0.0	19.2	7.9	0.0	9.5	−15.6
SP[B3LYP/6-31G(d,p)]	0.0	17.7	5.9	0.0	10.3	−14.1
SP[B3LYP/6-311++G(d,p)]	0.0	17.1	4.7	0.0	11.4	−11.2
path B: US(SCC-DFTB/CHARMM)	0.0	21.2	15.3	0.0	17.6	0.9
SP(SCC-DFTB)	0.0	32.8	28.7	0.0	21.2	−5.2
SP[B3LYP/6-31G(d)]	0.0	29.3	22.8	0.0	21.0	−9.5
SP[B3LYP/6-31G(d,p)]	0.0	28.0	22.3	0.0	21.0	−8.1
SP[B3LYP/6-311++G(d,p)]	0.0	27.3	20.2	0.0	20.7	−6.6
path C: US(SCC-DFTB/CHARMM)	0.0	16.2	−6.5	0.0	37.8	21.2
SP(SCC-DFTB)	0.0	21.9	5.4	0.0	27.2	−3.2
SP[B3LYP/6-31G(d)]	0.0	30.0	9.5	0.0	25.0	−5.5
SP[B3LYP/6-31G(d,p)]	0.0	28.4	8.8	0.0	23.3	−5.6
SP[B3LYP/6-311++G(d,p)]	0.0	28.7	7.4	0.0	26.5	−2.1
exptl[a]	16.1					

[a]Activation energy as estimated from the experimental rate of reaction[48] by using transition state theory. [a]Values are single-point (SP) energies calculated at SCC-DFTB and B3LYP with a different basis set. Umbrella sampling (US) QM/MM free energies are also included for comparison.

dx.doi.org/10.1021/jp500652x | J. Phys. Chem. B 2014, 118, 4771−4783

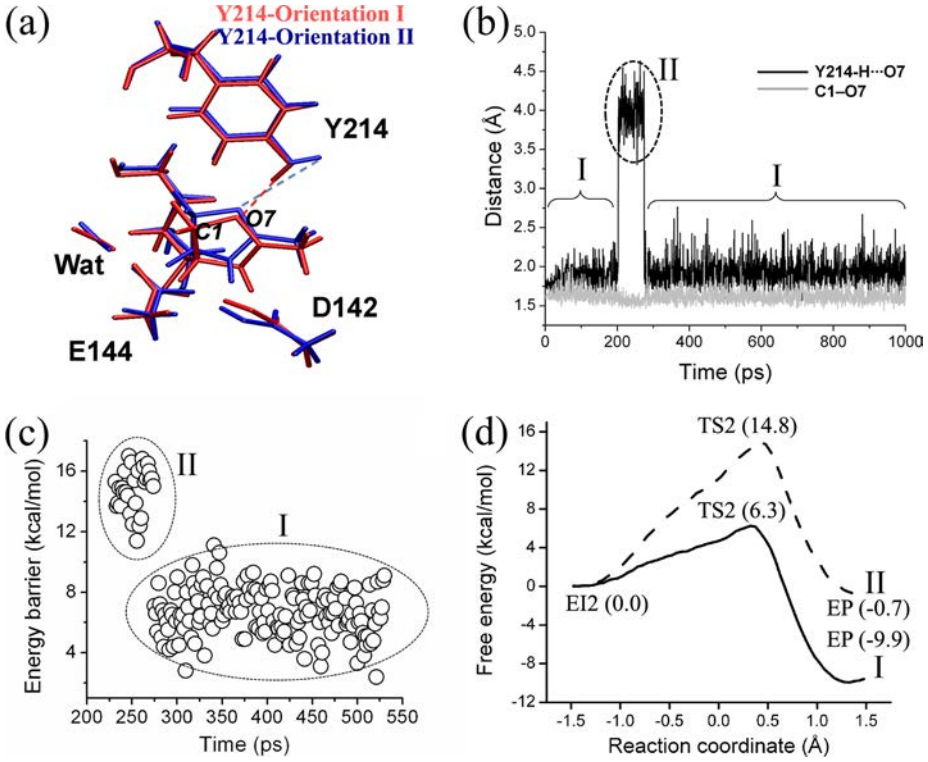


Figure 4. Analysis of Tyr214 in the deglycosylation step of ChiB catalysis. (a) Schematic representation of orientations adopted by Tyr214 with (**I**) and without (**II**) the hydrogen bond to the oxygen in the oxazolinium ring. (b) Relevant distances (Y214−H···O7 and C1−O7) related to the two orientations adopted by Tyr214 during 1 ns QM/MM MD simulation. (c) Clusters of calculated potential-energy barriers to deglycosylation (SCC-DFTB/CHARMM22) for selected snapshots extracted between 200 and 500 ps during QM/MM MD simulations. (d) Calculated free-energy profile (for deglycosylation) with Tyr214 in orientations **I** and **II**.

B3LYP/6-31G(d). SCC-DFTB was found to only slightly overestimate the energy of these two points by 0.5 and 2.1 kcal/mol for TS1 and TS2, respectively, indicating it is a reasonably accurate method for estimating the relative barriers for the glycosylation and deglycosylation steps.

For paths B and C, the B3LYP/6-31G(d) relative energies calculated for TS2 are also found to be in good agreement with SCC-DFTB (with a difference of 0.2 and 2.2 kcal/mol between the methods for paths B and C, respectively), while a larger correction was observed when comparing the energies obtained for TS1. These corrections show, first, that the SCC-DFTB calculated barriers to glycosylation for these paths are not consistent with the higher level of theory and, second, despite the difference in the relative energies of TS2 calculated by the two methods, the correction to the TS2 energy has no significant influence on the overall barrier to reaction. This is because, despite accounting for the large correction to the TS2 energy, deglycosylation is still found to be the rate-limiting step for both paths. Consequently, the overall conclusions drawn from analyzing the SCC-DFTB/MM calculated free-energy profiles are still valid.

In contrast, the SP calculations performed to correct the relative energy of EI1, formed after the glycosylation step, has significant implications in the analysis of the free-energy profile obtained for path A. The SP B3LYP energy for EI1 shows that SCC-DFTB overestimates the relative energy of the intermediate by 6.0 kcal/mol. Therefore, accounting for the fact the EI1 is calculated to be more stable with the higher level of theory but recognizing SCC-DFTB accurately calculates the relative barrier to deglycosylation, the correction shows that glycosylation, and not deglycosylation, is in fact the rate-determining step of reaction in path A. This is because correcting the SCC-DFTB energy of EI1 now lowers the energy of TS2 relative to ES, resulting in TS1 now representing a structure with the highest relative energy on the reaction pathway.

For paths B and C, despite SP calculations showing similar discrepancies for the calculated relative energies of EI1 between two methods, the paths exhibit either a significantly higher barrier for deglycosylation [path C, TS2:25.0 kcal/mol B3LYP/6-31G(d)] or a higher-energy intermediate (path B, EI1:22.8 kcal/mol). As a result, although SCC-DFTB overestimates the energy of EI1 compared with B3LYP, the correction to the SCC-DFTB energy would not be enough to stabilize TS2 such that it is lower in energy than TS1 in paths B and C. Consequently, deglycosylation remains as the rate-determining step for both of these paths with a slight increase in the height of the overall barrier to reaction for path C (by 2.1 kcal/mol).

In summary, the corrections indicate path A is still the most energetically favorable pathway but importantly indicate that glycosylation, and not deglycosylation, is in fact the rate-limiting step of the reaction due to SCC-DFTB overestimating the energy of EI1 in the free-energy profiles obtained previously. Further support for this conclusion is provided by a comparison of the free-energy profiles for the wild-type and mutant (D142N) reactions. The results (Figure S6 of the Supporting Information) from these calculations indicate that mutation affects the glycosylation step most significantly, with a relatively high free-energy barrier calculated for the glycosylation step in the mutant compared with the barrier in the wild-type. This finding is in line with experimental studies that

dx.doi.org/10.1021/jp500652x | J. Phys. Chem. B 2014, 118, 4771−4783

have shown the D142N mutant catalyzes the reaction less efficiently than the wild-type.[21]

SP calculations with different basis sets indicate that there is a small dependence of the calculated energy with respect to the size of the basis set used for the B3LYP calculation. Calculating the standard deviation of the B3LYP energies of each of the stationary points for the three paths indicates that the largest basis set dependence was observed for the energy of EI1 in path A (σ = 1.6 kcal/mol) and the lowest for energy of TS2 in path B (σ = 0.2 kcal/mol). It was also found that increasing the size of basis set used in the calculation resulted in a lowering of the overall barrier to reaction, with a decrease of 2.1, 2.9, and 0.6 kcal/mol for paths A, B, and C, respectively, when increasing the basis used for the calculation from 6-31G(d) to 6-311++G(d,p). This shows that an increase in the size of the basis set used, as well as employing a higher level theory to calculate the energy of the QM region, would likely result in a lower free-energy barrier to be calculated for the reaction, thus improving the agreement with the experimental calculated barrier of 16.1 kcal/mol.

Catalytic Importance of Tyr214. To better understand the role of Tyr214 in the reaction, further analysis was performed on the EI2 complex (with neutral Asp142) by running a 1 ns QM/MM MD simulation in order to assess its stability and examine the influence of the H-bond formed by Tyr214. The RMSD plot from the simulation, shown in Figure S7 of the Supporting Information, indicates a stable MD trajectory was obtained after 200 ps. Interestingly, during the 1 ns simulation, two different orientations (**I** and **II**) of Tyr214 were unexpectedly observed (Figure 4, panels a and b): one conformation (**I**) in which the residue could form a hydrogen bond with the oxazolinium oxygen (observed between 0−200 ps and 300−1000 ps) and the other (**II**), where its side-chain pointed away from the substrate (observed between 200−300 ps). To analyze the impact of the reorientation of the Tyr214 on the reaction energetics, several snapshots representing one of the two observed orientations, **I** and **II**, were chosen from the QM/MM MD and used as a starting structure for reaction modeling. In a similar manner as performed previously, umbrella sampling MD based on adiabatic mapping was then performed on the extracted snapshots in order to simulate reaction pathway A. **II** was not observed in any previous MD simulations on the ES complex (Figure 5a), and therefore, a restraint was imposed on the system to ensure a H-bond distance between the acetamido O7 oxygen of subsite −1 NAG and the hydroxyl group of Tyr214 could not be formed. This allowed **II** to be formed in ES and ensured that this orientation was maintained during the simulation of the glycosylation step of the reaction.

The calculations showed that different orientations of the Tyr214 resulted in significantly different reaction barriers for the reaction (Figures 4c and 5b). The deglycosylation step of the reaction was found to be most affected by this change in orientation, with a calculated free-energy barrier relative to the intermediate (EI1) of 6.3 and 14.8 kcal/mol, for **I** and **II**, respectively (Figure 4d). In contrast, the glycosylation step was affected less by the reorientation, with a calculated barrier to glycosylation of 20.5 kcal/mol compared to 16.7 kcal/mol for **I** and **II**, respectively. These calculations reveal new insight to the importance of the Tyr214 residue and specifically the influence of this H-bond with the substrate on each individual step of the enzyme-catalyzed reaction.

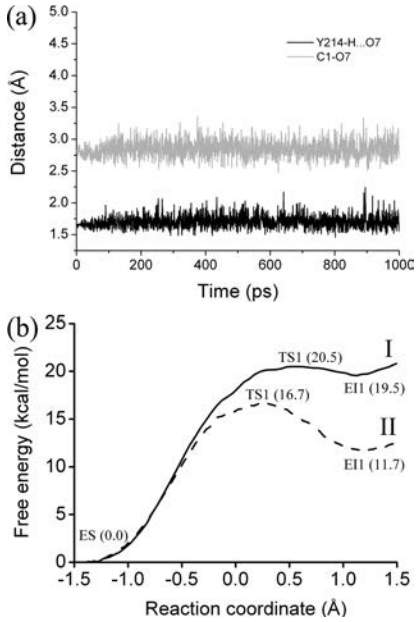


Figure 5. Analysis of Tyr214 in the glycosylation step of ChiB catalysis. (a) Relevant distances (Y214−H ···O7 and C1−O7) related to the two orientations adopted by Tyr214 during 1 ns QM/MM MD simulation. (b) Calculated free-energy profile (for glycosylation) with Tyr214 in orientations **I** and **II**.

In glycosylation, the presence of the Tyr214 H-bond (**I**) results in a higher barrier for the formation of the oxazolinium ion. This is probably due to the H-bond interacting with and reducing the nucleophilicity of the O7 oxygen, as well as hindering the movement of nucleophilic oxygen as the C1−O7 NAG distance is reduced during intramolecular nucleophilic attack. Conversely, in deglycosylation, the H-bond is shown to have an opposite and greater effect on the barrier to reaction (Figure 5). In **I**, the relative energy of EI1 indicates that the Tyr214 H-bond acts to destabilize the oxazolinium ion intermediate and, as a result, this lowers the barrier to deglycosylation (by 8.5 kcal/mol). In **II**, in the absence of the Tyr214 H-bond, the destabilizing effect on EI1 is not observed and the barrier of the deglycosylation step is raised and results in deglycosylation now becoming the rate-limiting step of the reaction. These free-energy calculations performed with **I** and **II** show that, although the Tyr214 influences the efficiency of the glycosylation step, it plays a more vital role in the deglycosylation step of the reaction, reducing the barrier of this step such that it is no longer rate-determining. This subtle observation provides new evidence to explain the origin of the reduced rate observed in the Y214F mutant observed in the experiment, as well as its analogous residue in other GH enzymes, and confirms previous stabilization calculations performed on the system that showed that the Tyr214 H-bond destabilizes the transition state and enzyme intermediate in the glycosylation step of the reaction.[7]

QM/MM MD Simulations of Enzyme−Product Complex. To gain further insight into the dynamics and stability of the enzyme−product complex after the catalytic reaction, two additional QM/MM MD simulations with different semi-empirical treatments of the QM region (i.e., SCC-DFTB/CHARMM22 and AM1/CHARMM22) were performed. The RMSD plots for the heavy atoms in the enzyme−product complex indicate that both simulations are reasonably stable

dx.doi.org/10.1021/jp500652x | J. Phys. Chem. B 2014, 118, 4771−4783

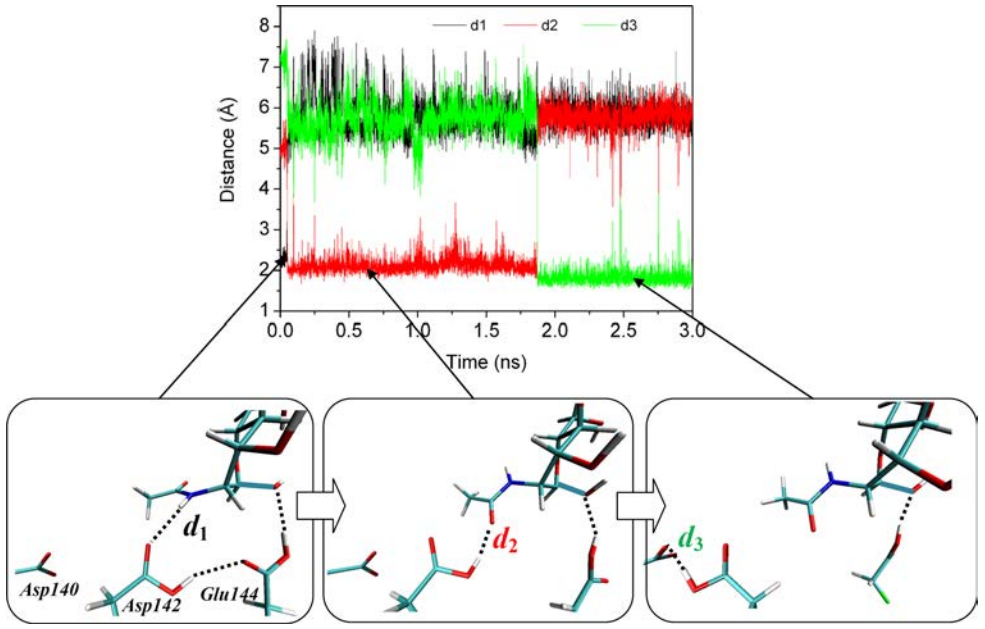


Figure 6. Snapshots showing the important conformational changes associated with the product unbinding event observed in the AM1/CHARMM22 MD simulation on the enzyme−product complex. d_1 = HN2−Oε2(Asp142), d_2 = O7−Hε2(Asp142), and d_3 = Oε2(Asp140)−Hε2(Asp142).

(Figure S8 of the Supporting Information). The simulation performed with SCC-DFTB/CHARMM22 reached an equilibrated structure faster than in the AM1/CHARMM22 MD (indicated by low RMSD, Figure S8 of the Supporting Information). A difference in the binding stability of the products was also observed. In the SCC-DFTB/CHARMM22 simulation, Asp142, Glu144, and Tyr214 were found to bind the reaction product more tightly at the −2 and −1 binding subsites of the enzyme (d_2 = 1.64 ± 0.01 Å, d_6 = 1.98 ± 0.16 Å, and d_9 = 1.73 ± 0.11 Å) compared to the binding observed in the AM1/CHARMM22 simulation (d_2 = 2.09 ± 0.01 Å, d_6 = 5.75 ± 0.61 Å, and d_9 = 4.64 ± 0.75 Å). Furthermore, the AM1/CHARMM22 simulation contained an interesting unbinding event, where the Asp142 side chain rotated in, allowing the *N*-acetyl group of subsite −1 NAG to flip its orientation, thus facilitating the release of the product (Figure 6). The boat conformation was observed at subsite −1 during the simulation. This observation provides further evidence to support the role of the −OH group on Tyr214 and Asp142 in contributing to the distortion of the *N*-acetyl group of the −1 subsite NAG and underlines the importance of their interactions in contributing to the binding of the substrate.

CONCLUSIONS

In this study, the QM/MM umbrella sampling MD free-energy simulations have been performed on the unusual substrate-assisted mechanism of *S. marcescens* ChiB. The contribution of two conserved amino acid residues, Asp142 and Tyr214, to enzyme catalysis was investigated. In particular, the role of Asp142, and the influence of its protonation state on the reaction mechanism[7] were examined through the three different reaction pathways modeled (paths A−C), treating the residue as either ionized or neutral. Our simulations show that the system containing a neutral Asp142 residue stabilizes the transition state of the deglycosylation step more effectively than in the other paths (outweighing the observed destabilization of the transition state in the glycosylation step). This resulted in the lowest overall free-energy barrier among the three pathways to be calculated for path A, supporting our previous theoretical study and in agreement with experimental findings[7,15] and underlining the important contribution of the Asp142 in the second step of the reaction. Analysis of two different possible orientations of Tyr214 enabled the important role of the H-bond formed by the residue with the substrate to be investigated. Our results showed that the interaction was necessary to facilitate efficient deglycosylation and prevent it from being the rate-limiting step in the reaction.

Small model, single point B3LYP calculations with a variety of different-sized basis sets indicated that SCC-DFTB overestimates the destabilization of the intermediate formed in the enzyme-catalyzed reaction. This emphasized the importance of validating semiempirical methods with higher level methods and provided evidence to suggest that glycosylation and not deglycosylation was the rate-determining step of the reaction.

A complex conformational change of the chitose unit in the subsite −1 along the reaction pathway was found to take place adopting boat ($^{1,4}B$) → half-chair $[^4H_5]^‡$ → chair ($^4C_1$) for the glycosylation step and the reverse conformations for deglycosylation: chair ($^4C_1$) → half-chair $[^4H_5]^‡$ → boat ($^{1,4}B$). Further QM/MM MD simulations of the enzyme−product complex were also performed in which an unbinding event is suggested to have been captured, affording new mechanistic insight into the release of the product of ChiB.

ASSOCIATED CONTENT

Supporting Information

Position of the nucleophilic water for the deglycosylation from QM/MM MD (Figure S1); definition of the reaction coordinate used in this work (Figure S2); snapshot of the ES complex when Asp142 is ionized (Figure S3); distorted conformations of NAG subsite −1 along the reaction coordinate of paths A−C (Figure S4); calculated stabilization energy profiles for paths A−C along the reaction coordinate of glycosylation and deglycosylation steps (Figure S5); free-energy

dx.doi.org/10.1021/jp500652x | J. Phys. Chem. B 2014, 118, 4771−4783

profiles for the D142N mutant (Figure S6); RMSD curves from the QM/MM simulations of the EI2 and EP complex of path A (Figures S7 and S8); geometric parameters for ES snapshots of both neutral and ionized Asp142 models (Table S1) and for stationary points along the reaction coordinates of paths A−C (Table S2). This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: jitrayut.018@gmail.com. Tel: +66 (0) 5446 6666, ext. 1834. Fax: +66 (0) 5446 6664.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was carried out within the Research Grant for New Scholar (Grant MRG5680143) that is cofunded by the Thailand Research Fund, the Office of the Higher Education Commission, and the University of Phayao, Thailand. J.J. thanks Prof. Supa Hannongbua for her kind support and guidance of this research project. We thank Dr. Joni Mujika, Dr. Julianna Oláh, and Dr. Christoph Sontag for their critical reading of the manuscript. A.J.M. is an EPSRC Leadership Fellow and thanks EPSRC for support (Grant EP/G007705/1). A.J.M. and M.A.L. thank EPSRC and BBSRC for support. This research was also funded by the University of Phayao (Grant R020056216016).

REFERENCES

(1) Carbohydrates and Glycobiology. *Science* **2001**, *291*, 2263−2502.
(2) Wolfenden, R.; Lu, X. D.; Young, G. Spontaneous Hydrolysis of Glycosides. *J. Am. Chem. Soc.* **1998**, *120*, 6814−6815.
(3) Cantarel, B. L.; Coutinho, P. M.; Rancurel, C.; Bernard, T.; Lombard, V.; Henrissat, B. The Carbohydrate-Active Enzymes Database (CAZy): An Expert Resource for Glycogenomics. *Nucleic Acids Res.* **2009**, *37*, D233−D238.
(4) Koshland, D. E. Stereochemistry and the Mechanism of Enzymatic Reactions. *Biol. Rev.* **1953**, *28*, 416−436.
(5) Vocadlo, D. J.; Davies, G. J. Mechanistic Insights into Glycosidase Chemistry. *Curr. Opin. Chem. Biol.* **2008**, *12*, 539−555.
(6) Vuong, T. V.; Wilson, D. B. Glycoside Hydrolases: Catalytic Base/Nucleophile Diversity. *Biotechnol. Bioeng.* **2010**, *107*, 195−205.
(7) Jitonnom, J.; Lee, V. S.; Nimmanpipug, P.; Rowlands, H. A.; Mulholland, A. J. Quantum Mechanics/Molecular Mechanics Modeling of Substrate-Assisted Catalysis in Family 18 Chitinases: Conformational Changes and the Role of Asp142 in Catalysis in ChiB. *Biochemistry* **2011**, *50*, 4697−4711.
(8) Vocadlo, D. J.; Davies, G. J.; Laine, R.; Withers, S. G. Catalysis by Hen Egg-White Lysozyme Proceeds Via a Covalent Intermediate. *Nature* **2001**, *412*, 835−839.
(9) Vocadlo, D. J.; Mayer, C.; He, S.; Withers, S. G. Mechanism of Action and Identification of Asp242 as the Catalytic Nucleophile of Vibrio furnisii N-Acetyl-β-D-Glucosaminidase Using 2-Acetamido-2-Deoxy-5-Fluoro-α-L-Idopyranosyl Fluoride. *Biochemistry* **2000**, *39*, 117−126.
(10) Huet, J.; Rucktooa, P.; Clantin, B.; Azarkan, M.; Looze, Y.; Villeret, V.; Wintjens, R. X-ray Structure of Papaya Chitinase Reveals the Substrate Binding Mode of Glycosyl Hydrolase Family 19 Chitinases. *Biochemistry* **2008**, *47*, 8283−8291.
(11) Tvaroska, I.; Kozmon, S.; Wimmerova, M.; Koca, J. Substrate-Assisted Catalytic Mechanism of *O*-GlcNAc Transferase Discovered by Quantum Mechanics/Molecular Mechanics Investigation. *J. Am. Chem. Soc.* **2012**, *134*, 15563−15571.
(12) Abbott, D. W.; Macauley, M. S.; Vocadlo, D. J.; Boraston, A. B. Streptococcus pneumoniae Endohexosaminidase D, Structural and Mechanistic Insight into Substrate-Assisted Catalysis in Family 85 Glycoside Hydrolases. *J. Biol. Chem.* **2009**, *284*, 11676−11689.
(13) Macauley, M. S.; Whitworth, G. E.; Debowski, A. W.; Chin, D.; Vocadlo, D. J. *O*-GlcNAcase Uses Substrate-Assisted Catalysis: Kinetic Analysis and Development of Highly Selective Mechanism-Inspired Inhibitors. *J. Biol. Chem.* **2005**, *280*, 25313−25322.
(14) Mark, B. L.; Vocadlo, D. J.; Knapp, S.; Triggs-Raine, B. L.; Withers, S. G.; James, M. N. Crystallographic Evidence for Substrate-Assisted Catalysis in a Bacterial β-Hexosaminidase. *J. Biol. Chem.* **2001**, *276*, 10330−10337.
(15) Van Aalten, D. M. F.; Komander, D.; Synstad, B.; Gaseidnes, S.; Peter, M. G.; Eijsink, V. G. H. Structural Insights into the Catalytic Mechanism of a Family 18 Exo-Chitinase. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 8979−8984.
(16) Bottoni, A.; Pietro Miscione, G.; Calvaresi, M. Computational Evidence for the Substrate-Assisted Catalytic Mechanism of *O*-GlcNAcase. A DFT Investigation. *Phys. Chem. Chem. Phys.* **2011**, *13*, 9568−9577.
(17) Passos, Ó.; Fernandes, P.; Ramos, M. Theoretical Insights into the Catalytic Mechanism of β-Hexosaminidase. *Theor. Chem. Acc.* **2011**, *129*, 119−129.
(18) Andersen, O. A.; Dixon, M. J.; Eggleston, I. M.; van Aalten, D. M. Natural Product Family 18 Chitinase Inhibitors. *Nat. Prod. Rep.* **2005**, *22*, 563−579.
(19) Houston, D. R.; Synstad, B.; Eijsink, V. G.; Stark, M. J.; Eggleston, I. M.; van Aalten, D. M. Structure-Based Exploration of Cyclic Dipeptide Chitinase Inhibitors. *J. Med. Chem.* **2004**, *47*, 5713−5720.
(20) Horn, S. J.; Sikorski, P.; Cederkvist, J. B.; Vaaje-Kolstad, G.; Soerlie, M.; Synstad, B.; Vriend, G.; Vaerum, K. M.; Eijsink, V. G. H. Costs and Benefits of Processivity in Enzymatic Degradation of Recalcitrant Polysaccharides. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 18089−18094.
(21) Synstad, B.; Gaseidnes, S.; van Aalten, D. M. F.; Vriend, G.; Nielsen, J. E.; Eijsink, V. G. H. Mutational and Computational Analysis of the Role of Conserved Residues in the Active Site of a Family 18 Chitinase. *Eur. J. Biochem.* **2004**, *271*, 253−262.
(22) Williams, S. J.; Mark, B. L.; Vocadlo, D. J.; James, M. N. G.; Withers, S. G. Aspartate 313 in the Streptomyces plicatus Hexosaminidase Plays a Critical Role in Substrate-Assisted Catalysis by Orienting the 2-Acetamido Group and Stabilizing the Transition State. *J. Biol. Chem.* **2002**, *277*, 40055−40065.
(23) Cetinbas, N.; Macauley, M. S.; Stubbs, K. A.; Drapala, R.; Vocadlo, D. J. Identification of Asp174 and Asp175 as the Key Catalytic Residues of Human *O*-GlcNAcase by Functional Analysis of Site-Directed Mutants. *Biochemistry* **2006**, *45*, 3835−3844.
(24) Papanikolau, Y.; Prag, G.; Tavlas, G.; Vorgias, C. E.; Oppenheim, A. B.; Petratos, K. High Resolution Structural Analyses of Mutant Chitinase a Complexes with Substrates Provide New Insight into the Mechanism of Catalysis. *Biochemistry* **2001**, *40*, 11338−11343.
(25) Bokma, E.; Rozeboom, H. J.; Sibbald, M.; Dijkstra, B. W.; Beintema, J. J. Expression and Characterization of Active Site Mutants of Hevamine, a Chitinase from the Rubber Tree Hevea brasiliensis. *Eur. J. Biochem.* **2002**, *269*, 893−901.
(26) Hsieh, Y. C.; Wu, Y. J.; Chiang, T. Y.; Kuo, C. Y.; Shrestha, K. L.; Chao, C. F.; Huang, Y. C.; Chuankhayan, P.; Wu, W. G.; Li, Y. K.; Chen, C. J. Crystal Structures of Bacillus cereus NCTU2 Chitinase Complexes with Chitooligomers Reveal Novel Substrate Binding for Catalysis: A Chitinase without Chitin Binding and Insertion Domains. *J. Biol. Chem.* **2010**, *285*, 31603−31615.
(27) Houston, D. R.; Shiomi, K.; Arai, N.; Omura, S.; Peter, M. G.; Turberg, A.; Synstad, B.; Eijsink, V. G.; van Aalten, D. M. High-Resolution Structures of a Chitinase Complexed with Natural Product Cyclopentapeptide Inhibitors: Mimicry of Carbohydrate Substrate. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 9127−9132.
(28) Houston, D. R.; Eggleston, I.; Synstad, B.; Eijsink, V. G. H.; van Aalten, D. M. F. The Cyclic Dipeptide Cl-4 [Cyclo-(L-Arg-D-Pro)] Inhibits Family 18 Chitinases by Structural Mimicry of a Reaction Intermediate. *Biochem. J.* **2002**, *368*, 23−27.

(29) Bas, D. C.; Rogers, D. M.; Jensen, J. H. Very Fast Prediction and Rationalization of pKa Values for Protein-Ligand Complexes. *Proteins: Struct., Funct., Bioinf.* **2008**, *73*, 765−783.
(30) Vriend, G. WHAT IF: A Molecular Modeling and Drug Design Program. *J. Mol. Graphics* **1990**, *8*, 52−56.
(31) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations. *J. Comput. Chem.* **1983**, *4*, 187−217.
(32) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. Self-Consistent-Charge Density-Functional Tight-Binding Method for Simulations of Complex Materials Properties. *Phys. Rev. B* **1998**, *58*, 7260−7268.
(33) Cui, Q.; Elstner, M.; Kaxiras, E.; Frauenheim, T.; M, K. A QM/MM Implementation of the Self-Consistent Charge Density Functional Tight Binding (SCC-DFTB) Method. *J. Phys. Chem. B* **2001**, *105*, 569−585.
(34) Liu, J.; Wang, X.; Xu, D. QM/MM Study on the Catalytic Mechanism of Cellulose Hydrolysis Catalyzed by Cellulase Cel5A from Acidothermus Cellulolyticus. *J. Phys. Chem. B* **2010**, *114*, 1462−1470.
(35) Xu, D.; Guo, H. Quantum Mechanical/Molecular Mechanical and Density Functional Theory Studies of a Prototypical Zinc Peptidase (Carboxypeptidase A) Suggest a General Acid-General Base Mechanism. *J. Am. Chem. Soc.* **2009**, *131*, 9780−9788.
(36) Xu, D.; Guo, H.; Cui, Q. Antibiotic Deactivation by a Dizinc β-Lactamase: Mechanistic Insights from QM/MM and DFT Studies. *J. Am. Chem. Soc.* **2007**, *129*, 10814−10822.
(37) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(38) Reuter, N.; Dejaegere, A.; Maigret, B.; Karplus, M. Frontier Bonds in QM/MM Methods: A Comparison of Different Approaches. *J. Phys. Chem. A* **2000**, *104*, 1720−1735.
(39) Shaw, K. E.; Woods, C. J.; Mulholland, A. J. Compatibility of Quantum Chemical Methods and Empirical (MM) Water Models in Quantum Mechanics/Molecular Mechanics Liquid Water Simulations. *J. Phys. Chem. Lett.* **2009**, *1*, 219−223.
(40) Brooks, C. L., III; Karplus, M. Deformable Stochastic Boundaries in Molecular Dynamics. *J. Chem. Phys.* **1983**, *79*, 6312−6325.
(41) Ranaghan, K. E.; Ridder, L.; Szefczyk, B.; Sokalski, W. A.; Hermann, J. C.; Mulholland, A. J. Insights into Enzyme Catalysis from QM/MM Modelling: Transition State Stabilization in Chorismate Mutase. *Mol. Phys.* **2003**, *101*, 2695−2714.
(42) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of N-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.
(43) Lodola, A.; Mor, M.; Zurek, J.; Tarzia, G.; Piomelli, D.; Harvey, J. N.; Mulholland, A. J. Conformational Effects in Enzyme Catalysis: Reaction via a High Energy Conformation in Fatty Acid Amide Hydrolase. *Biophys. J.* **2007**, *92*, L20−L22.
(44) Woodcock, H. L.; Hodoscek, M.; Brooks, B. R. Exploring SCC-DFTB Paths for Mapping QM/MM Reaction Mechanisms. *J. Phys. Chem. A* **2007**, *111*, 5720−5728.
(45) Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. The Weighted Histogram Analysis Method for Free-Energy Calculations on Biomolecules. I. The Method. *J. Comput. Chem.* **1992**, *13*, 1011−1021.
(46) Lonsdale, R.; Hoyle, S.; Grey, D. T.; Ridder, L.; Mulholland, A. J. Determinants of Reactivity and Selectivity in Soluble Epoxide Hydrolase from Quantum Mechanics/Molecular Mechanics Modeling. *Biochemistry* **2012**, *51*, 1774−1786.
(47) Mujika, J. I.; Lopez, X.; Mulholland, A. J. Mechanism of C-Terminal Intein Cleavage in Protein Splicing from QM/MM Molecular Dynamics Simulations. *Org. Biomol. Chem.* **2012**, *10*, 1207−1218.
(48) Krokeide, I. M.; Synstad, B.; Gaseidnes, S.; Horn, S. J.; Eijsink, V. G. H.; Sorlie, M. Natural Substrate Assay for Chitinases Using High-Performance Liquid Chromatography: A Comparison with Existing Assays. *Anal. Biochem.* **2007**, *363*, 128−134.

Journal of Molecular Graphics and Modelling 56 (2015) 53–59

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

A DFT study of the unusual substrate-assisted mechanism of *Serratia marcescens* chitinase B reveals the role of solvent and mutational effect on catalysis

Jitrayut Jitonnom[a,*], Chanchai Sattayanon[b], Nawee Kungwan[b], Supa Hannongbua[c]

[a] Division of Chemistry, School of Science, University of Phayao, Phayao 56000, Thailand
[b] Department of Chemistry, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand
[c] Department of Chemistry, Faculty of Science, Kasetsart University, Bangkok 10900, Thailand

ARTICLE INFO

Article history:
Accepted 8 December 2014
Available online 16 December 2014

Keywords:
DFT
Glycoside hydrolase
Chitinase B
Enzyme reaction
Substrate-assisted catalysis
Mutation

ABSTRACT

Serratia marcescens chitinase B (SmChiB) catalyzes the hydrolysis of β-1,4-glycosidic bond, via an unusual substrate-assisted mechanism, in which the substrate itself acts as an intramolecular nucleophile. In this paper, the catalytic mechanism of SmChiB has been investigated by using density functional theory. The details of two consecutive steps (glycosylation and deglycosylation), the structures and energetics along the whole catalytic reaction, and the roles of solvent molecules as well as some conserved SmChiB residues (Asp142, Tyr214, Asp215, and Arg294) during catalysis are highlighted. Our calculations show that the formation of the oxazolinium cation intermediate in the glycosylation step was found to be a rate-determining step (with a barrier of 23 kcal/mol), in line with our previous computational studies (Jitonnom et al., 2011, 2014). The solvent water molecules have a significant effect on a catalytic efficiency in the degycosylation step: the catalytic water is essentially placed in a perfect position for nucleophic attack by hydrogen bond network, lowering the barrier height of this step from 11.3 kcal/mol to 2.9 kcal/mol when more water molecules were introduced. Upon the *in silico* mutations of the four conserved residues, their mutational effects on the relative stability of the reaction intermediates and the computed energetics can be obtained by comparing with the wild-type results. Mutations of Tyr214 to Phe or Ala have shown a profound effect on the relative stability of the oxazolinium intermediate, emphasizing a direct role of this residue in destabilizing the intermediate. In line with the experiment that the D142A mutation leads to almost complete loss of SmChiB activity, this mutation greatly decreases the stability of the intermediate, resulting in a very large increase in the activation barrier up to 50 kcal/mol. The salt-bridges residues (Asp215 and Arg294) were also found to play a role in stabilizing the oxazolinium intermediate.

1. Introduction

Serratia marcescens chitinase B (SmChiB), belonging to the glycosidase family 18, degrades chitin (an insoluble linear polymer of β-(1,4)-linked *N*-acetylglucosamine $(GlcNAc)_n$) which is the second most abundant biopolymer in nature after cellulose. SmChiB has received much attention as an attractive system for the development of new inhibitors with chemotherapeutic potential [1,2]. It has also been applied in biotechnology for conversion of insoluble polysaccharides into commercially valuable product [3,4]. SmChiB has been suggested experimentally [5] and theoretically [6–8] to catalyze the hydrolysis of β-1,4-glycosidic bonds found in chitin, via an unusual substrate-assisted mechanism (see Fig. 1). In brief, SmChiB catalyzed the glycosidic bond hydrolysis via a two-step mechanism in which the formation of oxazolinium intermediate takes place in the first step (glycosylation), followed by hydrolysis of the intermediate in the second step (deglycosylation), yielding the sugar product with retention of configuration at the anomeric center. After catalytic reaction, the product is then released from the substrate binding cleft of SmChiB which is suggested to proceed via the unbinding event [6].

The catalytic function of this enzyme has been found to depend on a relatively large number of residues [9]. Asp140, Asp142 and Glu144, conserved in most family 18 chitinases, form a catalytic triad in the active site of SmChiB. Glu144 is known to act as a catalytic acid/base [5]. Asp142, together with Asp140, are suggested

* Corresponding author. Tel.: +66 05446 6666x1834; fax: +66 05446 6664.
E-mail address: jitrayut.018@gmail.com (J. Jitonnom).

http://dx.doi.org/10.1016/j.jmgm.2014.12.002
1093-3263/

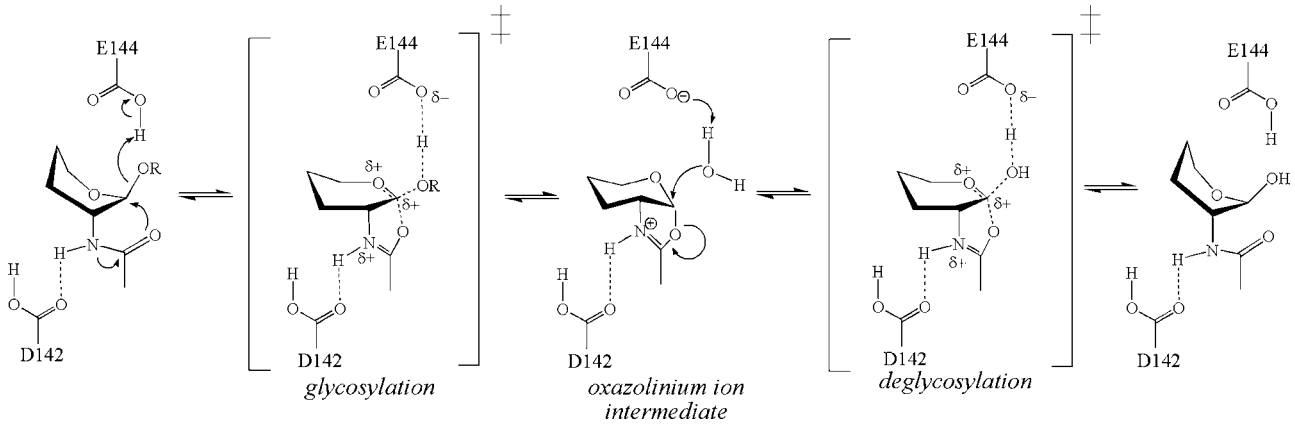


Fig. 1. Schematic presentation of substrate-assisted (retaining) mechanism of *Serratia marcescens* ChiB (SmChiB).

to be involved in binding of substrate [9] and catalysis [6,7]. Other conserved residues (e.g., Tyr10, Ser93, Tyr214, Asp215 and Arg294) have also been found to influence the catalytic activity of the enzyme [9]. Among these, Asp142, Tyr214 and a salt-bridge (Asp215 and Arg294) are of particular interest as they are located at the catalytic center of SmChiB and participate directly to the binding of substrate at subsites −1 and +1, in cooperation with other hydrophobic residues including Trp97, Trp220, and Trp403 (see Fig. 2). Moreover, Tyr214, in conjunction with Asp142 and Asp215, is proposed to interact with the substrate upon binding and has been observed to contribute to the distortion of the *N*-acetyl group on the −1 sugar [5,9]. In addition to their role in substrate binding, they have been proved to be critical for SmChiB catalysis: Asp142 has been found, in its neutral protonation state, to be important in stabilizing the transition state and the intermediate, by electrostatic interactions [6,7]. Recently, Tyr214 has been shown by our group to play a critical role in the deglycosylation step of the reaction, by destabilizing the intermediate via its H-bond formed with the substrate [6] and thus lowering the energy barrier in this step by about 8.5 kcal/mol. Similar role of analogous tyrosine was also found in GH1 beta-glucosidase [10]. Meanwhile, the role of the Asp215 and Arg294 salt-bridge is likely to contribute electrostatically to the reaction. While the catalytic role of these residues seems to be well described, it is of particular importance to probe the influence of these residues on each individual step of SmChiB reaction.

To further investigate the roles of these conserved residues (Asp142, Tyr214, Asp215 and Arg294) and in particular their fundamental impact on energetics of each individual step of the SmChiB reaction, a theoretical study on the catalytic mechanism of SmChiB with hybrid density functional theory (DFT) method has been carried out to provide structural information of intermediates and energetics along the whole catalytic reaction of wild-type SmChiB. Upon the wild-type geometry, several mutants corresponding to those four conserved residues (see Fig. 2) were *in silico* made and their computed energetics was then obtained for comparison with the wild-type results. In addition, we also modeled different number of water molecules in the second half of the reaction and have demonstrated the role of solvent during chitin hydrolysis.

2. Computational details

The initial structure of the enzyme–substrate (ES) complex was taken from the X-ray structure (PDB entry 1E6N) [5] of the E144Q mutant of *S. marcescens* ChiB with a chitopentaose $(GlcNAc)_5$ substrate bound along subsites −2 to +3 (see Fig. 2). The wild-type was recovered by manually altering Gln144 to Glu144. The enzyme active site was truncated via Cβ carbon atoms, generating a quantum cluster model of ES for the first (glycosylation) step that consists of side-chain atoms of Asp142, Glu144, Tyr214, Asp215 and Arg294 and one chitobiose molecule (occupying at GlcNAc subsites −1 and +1). In the next (deglycosylation) step, another cluster

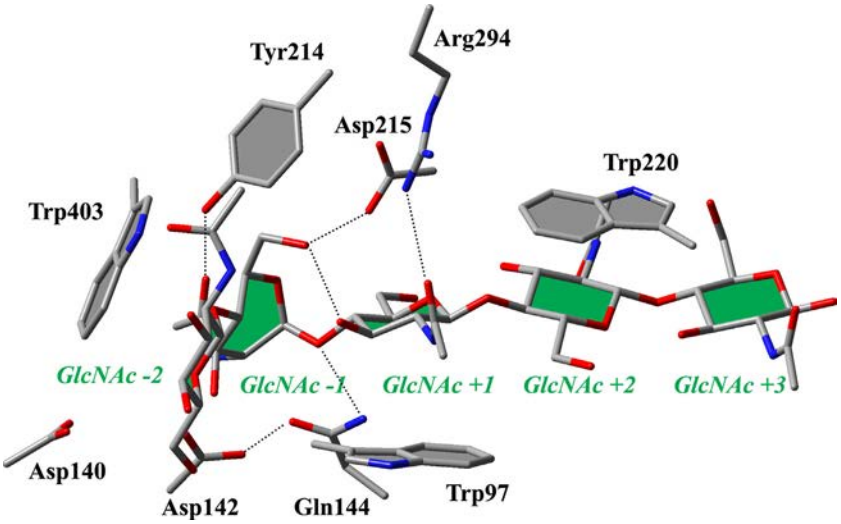


Fig. 2. X-ray structure of the active site of SmChiB complexes with *N*-acteylglucosamine substrate, $(GlcNAc)_5$, (coordinates taken from PDB entry 1E6N).

model (denoted as EI2) was built based on the final geometry of the last step of glycosylation (denoted as EI1) with one chitose molecule occupying at GlcNAc subsite +1 removed. One water molecule was added at a position of leaving group oxygen (O_{gly}), generating a nucleophilic water model (**1wat**). To test the solvent effect on the catalysis, additional water molecules were explicitly added into the EI2 model by gradually placing a water molecule around the (optimized) catalytic water and computing their configurations and energetic. Note that six water molecules were added but only four could maximally be modeled with stationary point found during the geometry optimization. The failure to find the optimized geometry of the higher number of water molecules (i.e., five and six waters) is probably due to the small size of the EI2 model which may not large enough to contain those waters. Finally, four different number of water models were studied and denoted as **1wat**, **2wat**, **3wat** and **4wat**, respectively. All cluster models were subjected to geometry optimizations with zero net charge. The Becke-3-Lee-Yang-Parr (B3LYP) exchange correlation functional and a standard basis set (6–31G(d)) [11] were used. Using quantum chemical cluster calculation, a computational method that has been widely applied in modeling enzymatic reaction [12], the geometries of stationary points were obtained with truncated atoms (Cβ atoms) constraint to their corresponding positions from the X-ray structure, keeping the optimized structures close to those obtained experimentally. Frequencies were calculated at the same level of theory as the geometry optimizations to confirm the nature of the stationary points and also to obtain zero-point energies (ZPE). Since some atoms were frozen to their crystallographic positions, a few small negative eigenvalues usually appear, in this case all below 50i cm^{-1}. These frequencies do not contribute significantly to the ZPE and can be ignored [13,14]. Effects of implicit solvent and protein environment (with the conductor-like polarized continuum model) [15] as well as energy correction at larger basis set (B3LYP/6-311++G(3df,3pd)//B3LYP/6-31G(d)) were also tested on the optimized structures, which give a similar result in terms of relative energetics and mechanism (see Table S1; Supporting Information). The energy profile of glycosylation was calculated using the ES as reference while that of deglycosylation was calculated with respect to the EI2. All calculations reported in this work were performed using Gaussian 09 [16].

3. Results and discussion

In order to understand the mutational effects of the conserved residues on the catalysis of the wild-type SmChiB, all structural and energetic information for the wild-type reaction need to be established as a reference. As mentioned above, the first step (glycosylation) of the reaction involves formation of an oxazolinium intermediate which then collapses upon hydrolysis in the second step (deglycosylation), yielding the sugar product. The detailed information for the two consecutive steps is described as follows.

3.1. Enzyme–substrate complex

The structures and energetic along the glycosylation and deglycosylation paths as calculated by the DFT cluster method at the B3LYP level are shown in Fig. 3. Key structural parameters are summarized in Table 1. Let consider the geometry at the ES complex shown in Fig. 3A. As seen, the quantum cluster model of ES generally represents the geometry observed for the X-ray structure of mutant (E144Q) SmChiB (see Fig. 2 and Table 1). In particular, a twisted glycosidic bond within the $(GlcNAc)_2$ substrate could be maintained through seven hydrogen bonds (**a**–**g**), assisted mainly by several enzyme active site residues, and thus supporting the role of these conserved residues in the substrate binding. Asp142, Glu144 and Tyr214 are fully responsible for the substrate distortion: the two carboxyl residues form a strong hydrogen-bonding network (see distances **a**, **b** and **g**) that essentially places the substrate in a reactive orientation while Tyr214 helps the two residues in twisting the *N*-acetyl group of the −1 GlcNAc via a hydrogen bond (distance **c**). This distortion facilitates a decrease in the C1—O_{NAc} distance involved in the intramolecular nucleophilic attack and thus helps to promote the formation of oxazolinium intermediate in the first step (glycosylation) of the SmChiB reaction.

3.2. Glycosylation path

As shown schematically in Fig. 1, the glycosylation step is the formation of an oxazolinium cation intermediate where two chemical processes are involved, that is, the proton transfer from Glu144 proton to the scissile glycosidic oxygen (see ***a***; O_{gly}—HO(Glu144)) and the nucleophilic addition of the *N*-acetyl oxygen at the anomeric C1 carbon (C1—O_{NAc}). Our DFT models indicate that these two processes occur simultaneously in an asynchronous manner. Starting from ES to TS1, Glu144 completely donates its proton to the glycosidic oxygen (see ***a***; O_{gly}—HO(Glu144) = 1.10 Å), activating the cleavage of glycosidic bond. At the same time, the scissile glycosidic bond (O_{gly}—C1) is elongated (1.48 Å at ES to 2.27 Å at TS1) and the C1—O_{NAc} distance between the anomeric carbon (C1) and the *N*-acetyl oxygen (O_{NAc}) becomes shorter (3.08 Å at ES to 2.29 Å at TS1). At this stage, the −1 GlcNAc ring was found to adopt a coplanar geometry (0.6°) between atoms C3, C2, C1 and O_{ring}, featuring a $^4H_5$ half-chair conformation, which is in agreement to our previous observations [6,7]. Upon the cleavage of the glycosidic bond, the intramolecular nucleophilic attack takes place in a synchronous process, resulting in a narrowing of the *N*-acetyl oxygen and the anomeric carbon distance and the formation of the bicyclic structure of the oxazolinium intermediate, which is characterized by an intramolecular covalent bond formed (C1—O_{NAc} = 1.57 Å) inside the sugar at the EI1. The glycosylation step is endothermic by 11 kcal/mol with a calculated activation energy of 23.0 kcal/mol (see Fig. 3B). The conformation itinerary adopted by the GlcNAc moiety at the subsite−1 of this step is also depicted in Fig. 3B, which is found to follow a boat $(^{1,4}B) \rightarrow$ a half-chair $(^4H_5)^{\ddagger} \rightarrow$ chair $(^4C_1)$ pattern.

3.3. Deglycosylation path

The second step (deglycosylation) of the reaction requires an essential water molecule near the anomeric carbon (C1) to act as a nucleophile during the breakdown of the intermediate formed in the first step. In our previous study, we found that there are at least three water molecules participating in a near-attack position close to the C1 atom during the 1 ns QM/MM MD simulation of enzyme-intermediate complex (EI2) [6]. However, only one water molecule having the best positioning for nucleophilic attack was chosen to be treated quantum mechanically in the QM/MM reaction modeling. Here we included more water molecules explicitly in our cluster model. The higher number of water molecules was modeled to understand the role of the solvent on the catalytic efficiency of the SmChiB enzyme, which is often ignored in most theoretical studies of glycosidase reaction [6,7,10,17–23]. Representative structures and energetics along the deglycosylation path with one to three water molecules modeled (**1wat**, **2wat** and **3wat**) are shown in Fig. 4. This step is represented by the **2wat** model as this is a minimum model that could produce a water nucleophile in a perfect position for reaction and its geometric parameter is included in Table 1. Additional geometric parameters for the remaining water models are summarized in the Supporting Information (Table S2).

In an opposite direction to the first step, the second step is initiated by nucleophilic attack of the catalytic water at the anomeric

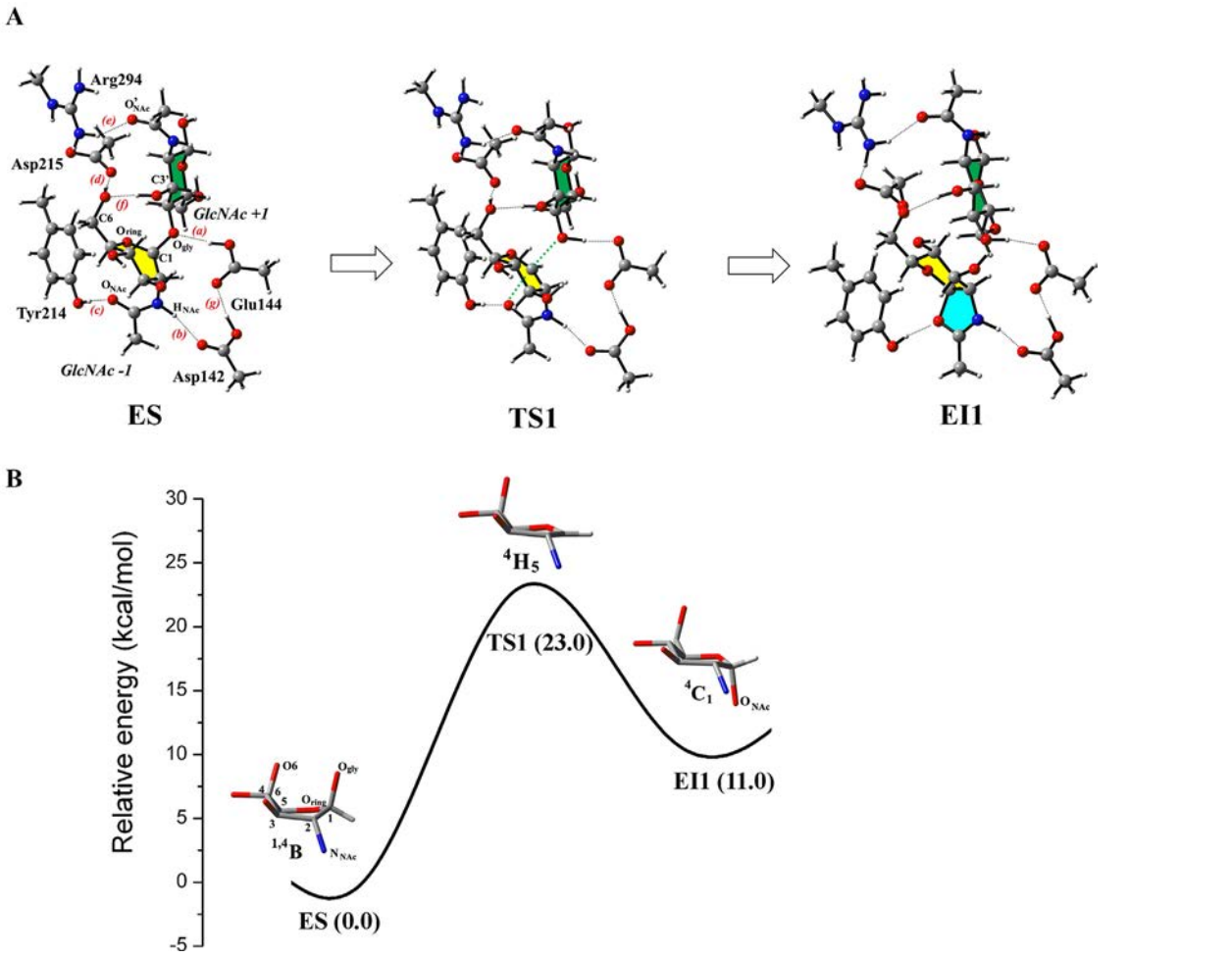


Fig. 3. Structures (A) and energetic (B) along the glycosylation path of SmChiB reaction. Conformation of the −1 GlcNAc is also included with some hydrogen atoms are omitted for clarity.

center, following by the collapse of the intermediate (EI2) and the subsequent formation of the reaction product (EP) as shown in Fig. 4A. At TS2, the catalytic water was found to be positioned close to the anomeric carbon (C1—O_{wat}; 2.88 Å (EI2) → 2.42 Å (TS2)) while the fused oxazoline ring is spontaneously broken (C1—O_{NAc}; 1.55 Å (EI2) → 2.14 Å (TS2)). A fully bonded O_{wat}—H_{wat} distance of 1.01 Å at TS2 indicates that no proton abstraction occurs at this stage, that is, the transfer of a proton from the catalytic water onto the Glu144 residue does not occur, in agreement to our previous observation [6]. After TS2, the proton abstraction was achieved with the fully H_{wat}—O(Glu144) bonded distance of 1.00 Å, resulting in the formation of the EP and a subsequent release of the product, which is

Table 1
Selected structural parameters (interatomic distances and angles) of the stationary points on the potential energies of glycosylation and deglycosylation paths.

Distance (Å), angle (°)	Expt[a]	Glycosylation			Deglycosylation (**2wat**)		
		ES	TS1	EI1	EI2	TS2	EP
O_{gly}—C1	1.39	1.48	2.27	2.87	–	–	–
C1—O_{NAc}	2.98	3.08	2.29	1.57	1.56	2.14	2.96
O_{gly}—HO(Glu144) (a)	3.19[b]	1.90	1.10	0.99	–	–	–
C1—O_{ring}	1.45	1.38	1.28	1.34	1.34	1.28	1.38
O_{wat}—C1	–	–	–	–	2.92	2.42	1.46
H_{wat}—O(Glu144)	–	–	–	–	1.78	1.59	0.99
O_{wat}—H_{wat}	–	–	–	–	0.98	1.02	1.86
H_{NAc}—O(Asp142) (b)	2.89[c]	2.01	1.79	1.55	1.54	1.84	2.08
O_{NAc}—HO(Tyr214) (c)	2.66[b]	1.80	1.91	2.13	2.04	1.86	1.77
C6OH—O(Asp215) (d)	2.64[b]	1.61	1.59	1.65	1.72	1.64	1.67
O'_{NAc}—HN(Arg294) (e)	3.09[c]	1.95	1.90	1.93	–	–	–
C6:HO—HO:C3′ (f)	2.61[b]	1.71	1.79	1.81	–	–	–
O(Glu144)—HO(Asp142) (g)	2.71[b]	1.74	1.43	1.05	1.50	1.46	1.74
C3—C2—C1—O_{ring}	−41.0	−33.0	0.6	26.4	26.5	12.1	−16.3

[a] Values taken from the experimentally determined X-ray structure of E144Q mutant ChiB (PDB entry 1E6N, chain A) from Ref. [5].
[b] O—O distance.
[c] N—O distance.

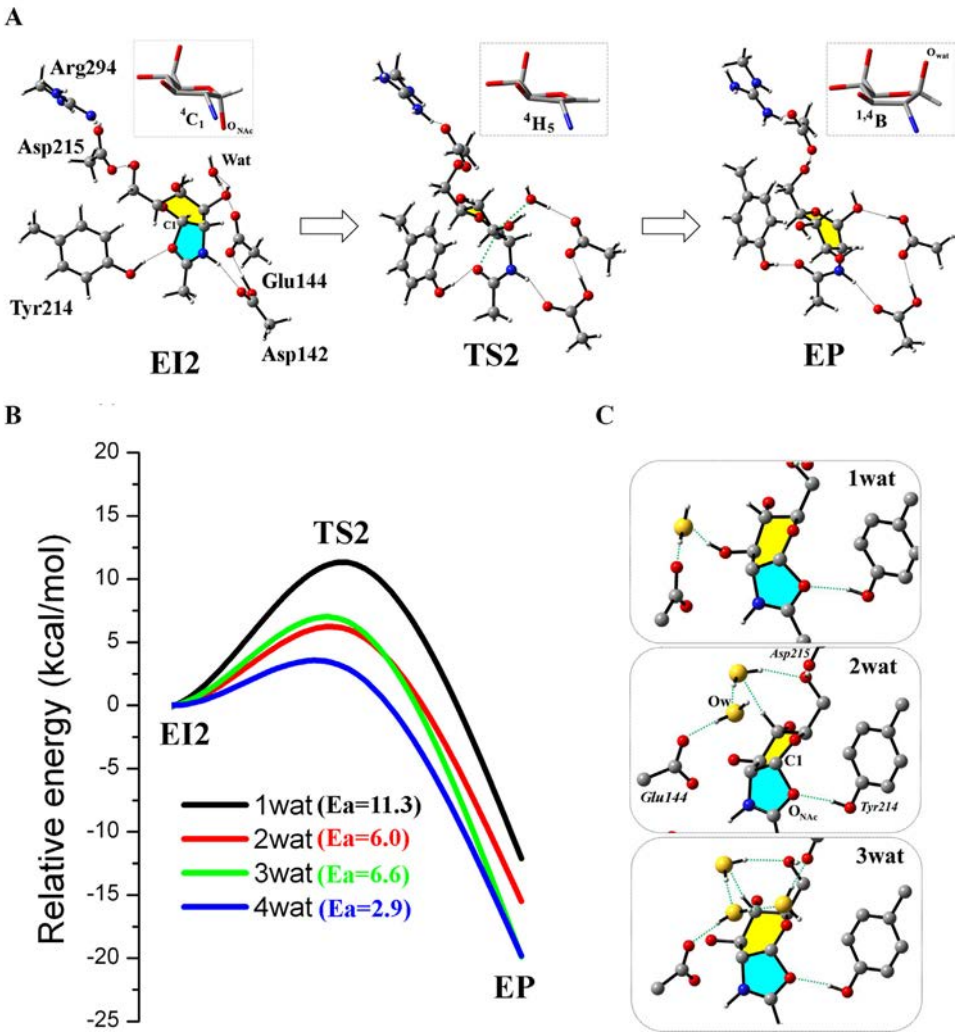


Fig. 4. Representative structures (A) and energetic (B) along the deglycosylation path of SmChiB reaction with different number of water molecules modeled near C1 atom. Initial optimized structures of EI2 (C) for **1wat**, **2wat** and **3wat** are demonstrated with oxygen atoms colored in yellow. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

proposed to be triggered by the rotation of Asp142 side chain in concomitant with the flip of the *N*-acetyl group of subsite −1 NAG [6].

In this deglycosylation, a reverse conformations of the −1 GlcNAc sugar was found as chair ($^4C_1$) → half-chair [$^4H_5$]‡ → boat ($^{1,4}B$).

As shown in Fig. 4B, the barrier height of the second step was significantly decreased (from 11.3 kcal/mol to 2.9 kcal/mol) when more number of water molecules were introduced in the model system, emphasizing the essential role of water molecules for efficient catalysis. While all model systems seem to have a similar geometry, a minor difference for the position of the water nucleophile is still seen at the EI2 (see Fig. 4C) where a hydrogen bond (H-bond) network playing a role. As shown, a large barrier (11.3 kcal/mol) observed at TS2 of **1wat** is observed due to a weak H-bond stabilization provided by water molecules. On the other hand, increasing number of water molecules (for either **2wat** or **3wat** models) appears to create more H-bond network and helps in positioning the nucleophile water in a perfect position for reaction, as evident by the shorter C1—O_{wat} distances of 2.92 Å (**2wat**) and 2.87 Å (**3wat**) compared to that (4.45 Å) in **1wat**. Thus, the more water molecules would provide a stronger H-bond stabilization that leads to a lower energy barrier in the deglycosylation step. In addition, they also contribute to the product stability (see Fig. S1 in the Supporting Information).

The lower energy barriers for the second step (in range of 11.3–2.9 kcal/mol relative to the EI2) comparing to that (23.0 kcal/mol) for the first step (Fig. 3B) suggest that the glycosylation path is the rate-determining step of the overall reaction of SmChiB.

3.4. Energetic analysis of mutants

To further understand the mutational effect on the energetic along both glycosylation and deglycosylation pathways, several *in silico* mutations were made based on the geometry of the optimized structure of the wild-type SmChiB. We selected the **2wat** system as a representative model of the second step for energetic analysis of mutants.

3.4.1. Y214F

Our previous simulation [6] has shown that the Tyr214 residue could adopt two orientations, corresponding to the different H-bond interaction formed between the OH-group of Tyr214 and the oxazoline ring within the intermediate, and we found that this H-bond is essential for an efficient deglycosylation step in the SmChiB

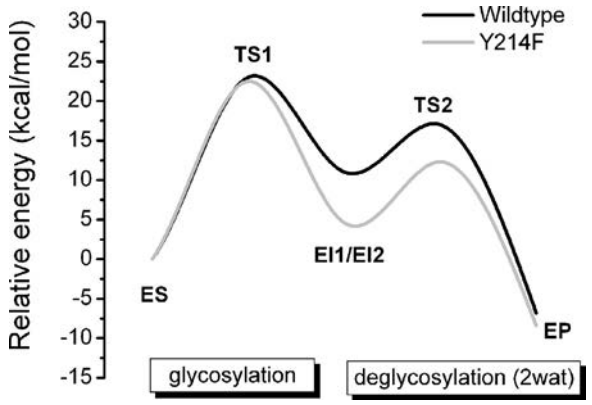


Fig. 5. Relative potential energy profiles for overall reaction in wild-type (black) and Y214F mutant (grey) SmChiB based on the two waters model (**2wat**) system.

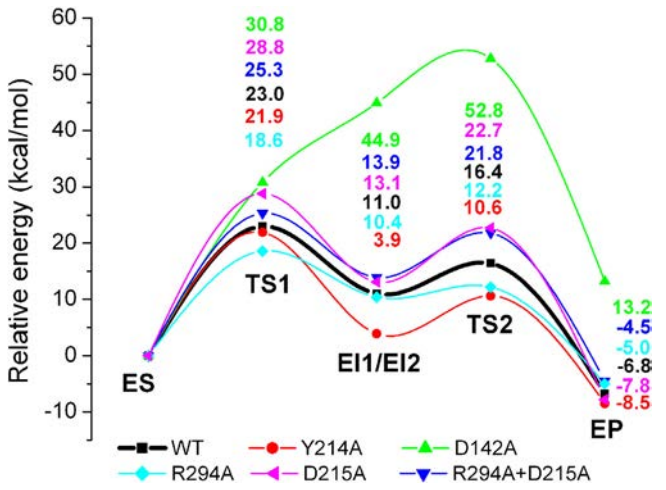


Fig. 6. Relative potential energy profiles for both glycosylation and deglycosylation paths of SmChiB reaction for wild-type (black) and other alanine mutations.

reaction, which provides explanation to the origin of the reduced rate observed in the Y214F mutant observed in the experiment, as well as its analogous residue in other GH enzymes [24–26]. However, this conclusion was drawn based on energetic analysis of the wild-type, not the Y214F mutant, that does not account for the electronic effect of the Tyr214 residue (as this residue was treated classically during the simulation), which provides only an indirect implication to the reduced rate of Y214F mutant. Here, the electronic effect of Tyr214 is fully described and the Y214F mutant was generated based on *in silico* mutation of the optimized geometry of the wild-type SmChiB. Calculated reaction energetics for each individual step of the mutant is shown in Fig. 5, in comparison with the energy profile of the wild-type. Clearly, the Y214F mutant causes only a small effect on the glycosylation step with a slightly lower barrier at TS1 (22.5 kcal/mol). On the other hand, the mutant has a profound effect on the deglycosylation step, and in particular the relative stability of oxazolinium intermediate (EI1/EI2) formed in the glycosylation step: it stabilizes the intermediate more effectively as indicated by a large decrease in the EI1/EI2 energy from 11.0 kcal/mol in wild-type to 4.1 kcal/mol in Y214F. This result also implies that Tyr214 exhibits a stronger destabilization to the intermediate compared to the Phe residue, as the tyrosine residue polarizes, via its H-bond, the oxazoline ring formed in the glycosylation step, thus facilitating an efficient deglycosylation. More interestingly, the reduction of the EI energy leads to an increase in a barrier height of the deglycosylation step from 6.0 to 8.8 kcal/mol in wild-type and Y214F, respectively. This result was also found in other water models (see Figs. S2 and S3 in the Supporting Information). A similar change of the energy profile was also found in the case of Y214A mutation, as described below. This energetic analysis indicate clearly that the Y214F mutation has a dominant effect on the second step of the reaction, not on the first step, further supporting our recent study of the vital role of Tyr214 in deglycosylation [6] in which the residue plays a role in destabilizing the intermediate via hydrogen bond. This information would also explain why mutations of similar analogous residue (Y217F and Y228A) decreased hydrolytic activities with improved transglycosylation ability in endo-beta-N-acetylglucosaminidases [27] and *Serratia proteamaculans* ChiD [28], respectively.

3.4.2. Other mutants

To further probe the reaction energetic of other mutants, several alanine mutations on residues surrounding the −1 and +1 GlcNAc were performed and their corresponding energetics along the reaction were examined, as shown in Fig. 6. The mutations were involved two single mutations of Asp142 and Tyr214 and a double mutation of Arg294 and Asp215. We found that the Y214A mutation exhibited a similar reaction energy profile to that of the Y214F. Interestingly, D142A was found to be most affected on both energetics and mechanism. Instead of a two-step mechanism the mutant proceed the reaction, via one transition state, that required a very large energy barrier of 52.8 kcal/mol. This result supports our previous observations [7], emphasizing a necessary of the Asp142–Glu144 interaction and their ability to stabilize the oxazolinium cation species electrostatically, leading to a lower reaction barrier. The large barrier generated by the D142A mutation not only reflects a decreased rate of SmChiB hydrolysis, but also provides evidence to support the improved transglycosylation activity observed in variants of Asp-142/313, analogous residues in SmChiB and SmChiA, respectively [27,29]. The salt-bridge residues, Arg294 and Asp215 are likely to give an opposite role. By comparing to that of wild-type reaction, the D215A was attributed to the higher energy barrier for glycosylation and deglycosylation whereas R294A was found to lower the barriers of both steps to 18.6 and 1.8 kcal/mol, respectively. The role of these salt-bridge residues can be further seen from the double mutation which clearly showed an increased barrier for both reaction steps, with the deglycosylation having the most effect. Like Asp142, this salt-bridge was also found to play a role in stabilizing the oxazolinium intermediate as this double mutation gives a higher energy of the EI compared to that of the wild-type.

4. Conclusions

In this study, quantum cluster models with density functional theory are successfully used not only to describe the unusual substrate-assisted mechanism at a molecular level but also to demonstrate the influence of four conserved residues (Asp142, Tyr214, Asp215 and Arg294) at the catalytic center of SmChiB on catalysis. Structural and energetic information of the reaction in wild-type SmChiB were provided by the models, supporting the two-step mechanism involving the formation of an oxazolinium intermediate and the subsequent collapse of the intermediate upon hydrolysis yielding the sugar product. The first step was found to be the rate-determining step of the overall reaction with the calculated activation barrier of 23.0 kcal/mol, in agreement to our previous studies. In the second step, the reaction barrier was found to be lower (from 11.3 kcal/mol to 2.9 kcal/mol) when more water molecules were introduced in the model systems, emphasizing the important role of solvent in catalytic efficiency. By comparing with the wild-type results, the mutational effect of four conserved residues in the SmChiB active site on reaction energetics of each individual step was allowed to be studied. Mutations of Tyr214 to

Phe or Ala have shown a profound effect on the relative stability of the oxazolinium intermediate, which exhibits the decreased EI energy that turns out to raise the energy barrier in the deglycosylation step. This observation has confirmed another important role of Tyr214 residue in intermediate destabilization and deglycosylation activity, besides its role in substrate distortion. The D142A mutation greatly decreases the stability of the intermediate, resulting in a very large increase in the activation barrier up to 50 kcal/mol. The salt-bridges residues (Asp215 and Arg294) were also found to play a role in stabilizing the oxazolinium intermediate. Our modeling approach may be useful as a computational probe to understand the influence of key conserved residues on the individual step of enzyme reaction, complementary to the site-direct mutagenesis experiment.

Acknowledgements

This work was carried out within the Research Grant for New Scholar (Grant no. MRG5680143) that is co-funded by the Thailand Research Fund, the Office of the Higher Education Commission, the University of Phayao. N.K. thanks Chiang Mai University for financial support. J.J. is funded by the University of Phayao (Grants nos. R020056216016 and R020057316004) and thanks the National e-Science Infrastructure Consortium for providing computing resources (http://www.e-science.in.th).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jmgm.2014.12.002.

References

[1] O.A. Andersen, M.J. Dixon, I.M. Eggleston, D.M. van Aalten, Natural product family 18 chitinase inhibitors, Nat. Prod. Rep. 22 (2005) 563–579.
[2] D.R. Houston, B. Synstad, V.G. Eijsink, M.J. Stark, I.M. Eggleston, D.M. van Aalten, Structure-based exploration of cyclic dipeptide chitinase inhibitors, J. Med. Chem. 47 (2004) 5713–5720.
[3] S.J. Horn, P. Sikorski, J.B. Cederkvist, G. Vaaje-Kolstad, M. Soerlie, B. Synstad, et al., Costs and benefits of processivity in enzymatic degradation of recalcitrant polysaccharides, Proc. Natl. Acad. Sci. U. S. A. 103 (2006) 18089–18094.
[4] G. Vaaje-Kolstad, S.J. Horn, M. Sorlie, V.G. Eijsink, The chitinolytic machinery of *Serratia marcescens* – a model system for enzymatic degradation of recalcitrant polysaccharides, FEBS J. 280 (2013) 3028–3049.
[5] D.M.F. Van Aalten, D. Komander, B. Synstad, S. Gaseidnes, M.G. Peter, V.G.H. Eijsink, Structural insights into the catalytic mechanism of a family 18 exo-chitinase, Proc. Natl. Acad. Sci. U. S. A. 98 (2001) 8979–8984.
[6] J. Jitonnom, M.A. Limb, A.J. Mulholland, QM/MM free-energy simulations of reaction in *Serratia marcescens* chitinase B reveal the protonation state of Asp142 and the critical role of Tyr214, J. Phys. Chem. B 118 (2014) 4771–4783.
[7] J. Jitonnom, V.S. Lee, P. Nimmanpipug, H.A. Rowlands, A.J. Mulholland, Quantum mechanics/molecular mechanics modeling of substrate-assisted catalysis in family 18 chitinases: conformational changes and the role of Asp142 in catalysis in ChiB, Biochemistry 50 (2011) 4697–4711.
[8] C.M. Payne, J. Baban, S.J. Horn, P.H. Backe, A.S. Arvai, B. Dalhus, et al., Hallmarks of processivity in glycoside hydrolases from crystallographic and computational studies of the *Serratia marcescens* chitinases, J. Biol. Chem. 287 (2012) 36322–36330.
[9] B. Synstad, S. Gaseidnes, D.M.F. van Aalten, G. Vriend, J.E. Nielsen, V.G.H. Eijsink, Mutational and computational analysis of the role of conserved residues in the active site of a family 18 chitinase, Eur. J. Biochem. 271 (2004) 253–262.
[10] S. Badieyan, D.R. Bevan, C. Zhang, Probing the active site chemistry of beta-glucosidases along the hydrolysis reaction pathway, Biochemistry 51 (2012) 8907–8918.
[11] A.D. Becke, Density-functional thermochemistry III. The role of exact exchange, J. Chem. Phys. 98 (1993) 5648–5652.
[12] P.E. Siegbahn, F. Himo, Recent developments of the quantum chemical cluster approach for modeling enzyme reactions, J. Biol. Inorg. Chem. 14 (2009) 643–651.
[13] K. Wang, Q. Hou, Y. Liu, Insight into the mechanism of aminomutase reaction: a case study of phenylalanine aminomutase by computational approach, J. Mol. Graph. Model. 46 (2013) 65–73.
[14] X. Sheng, Y. Liu, C. Liu, Theoretical studies on the common catalytic mechanism of transketolase by using simplified models, J. Mol. Graph. Model. 39 (2013) 23–28.
[15] M. Cossi, N. Rega, G. Scalmani, V. Barone, Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model, J. Comput. Chem. 24 (2003) 669–681.
[16] M.J.G.W.T. Frisch, H.B. Schlegel, G.E. Scuseria, M.A.J.R.C. Robb, G. Scalmani, V. Barone, B. Mennucci, G.A.H.N. Petersson, M. Caricato, X. Li, H.P. Hratchian, A.F.J.B. Izmaylov, G. Zheng, J.L. Sonnenberg, M. Hada, M.K.T. Ehara, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y.O.K. Honda, H. Nakai, T. Vreven, J.A. Montgomery Jr., et al., Gaussian 09, Revision D.01, 2013.
[17] X. Biarnes, J. Nieto, A. Planas, C. Rovira, Substrate distortion in the Michaelis complex of Bacillus 1,3-1,4-b-glucanase: insight from first principles molecular dynamics simulations, J. Biol. Chem. 281 (2006) 1432–1441.
[18] N.F. Bras, P.A. Fernandes, M.J. Ramos, QM/MM studies on the beta-galactosidase catalytic mechanism: hydrolysis and transglycosylation reactions, J. Chem. Theory Comput. 6 (2010) 421–433.
[19] J. Liu, X. Wang, D. Xu, QM/MM study on the catalytic mechanism of cellulose hydrolysis catalyzed by cellulase Cel5A from Acidothermus cellulolyticus, J. Phys. Chem. B 114 (2010) 1462–1470.
[20] L. Petersen, A. Ardevol, C. Rovira, P.J. Reilly, Mechanism of cellulose hydrolysis by inverting GH8 endoglucanases: a QM/MM metadynamics study, J. Phys. Chem. B 113 (2009) 7331–7339.
[21] M.E.S. Soliman, J.J.R. Pernia, I.R. Greig, I.H. Williams, Mechanism of glycoside hydrolysis: a comparative QM/MM molecular dynamics analysis for wild type and Y69F mutant retaining xylanases, Org. Biomol. Chem. 7 (2009) 5236–5244.
[22] A. Bottoni, G. Pietro Miscione, M. Calvaresi, Computational evidence for the substrate-assisted catalytic mechanism of O-GlcNAcase. A DFT investigation, Phys. Chem. Chem. Phys. 13 (2011) 9568–9577.
[23] Ó. Passos, P. Fernandes, M. Ramos, Theoretical insights into the catalytic mechanism of β-hexosaminidase, Theor. Chem. Acc. 129 (2011) 119–129.
[24] E. Bokma, H.J. Rozeboom, M. Sibbald, B.W. Dijkstra, J.J. Beintema, Expression and characterization of active site mutants of hevamine, a chitinase from the rubber tree *Hevea brasiliensis*, Eur. J. Biochem. 269 (2002) 893–901.
[25] Y.C. Hsieh, Y.J. Wu, T.Y. Chiang, C.Y. Kuo, K.L. Shrestha, C.F. Chao, et al., Crystal structures of Bacillus cereus NCTU2 chitinase complexes with chitooligomers reveal novel substrate binding for catalysis: a chitinase without chitin binding and insertion domains, J. Biol. Chem. 285 (2010) 31603–31615.
[26] Y. Papanikolau, G. Prag, G. Tavlas, C.E. Vorgias, A.B. Oppenheim, K. Petratos, High resolution structural analyses of mutant chitinase A complexes with substrates provide new insight into the mechanism of catalysis, Biochemistry 40 (2001) 11338–11343.
[27] M. Umekawa, W. Huang, B. Li, K. Fujita, H. Ashida, L.-X. Wang, et al., Mutants of *Mucor hiemalis* endo-β-N-acetylglucosaminidase show enhanced transglycosylation and glycosynthase-like activities, J. Biol. Chem. 283 (2008) 4469–4479.
[28] J. Madhuprakash, K. Tanneeru, P. Purushotham, L. Guruprasad, A.R. Podile, Transglycosylation by chitinase D from *Serratia proteamaculans* improved through altered substrate interactions, J. Biol. Chem. 287 (2012) 44619–44627.
[29] H. Zakariassen, M.C. Hansen, M. Joranli, V.G. Eijsink, M. Sorlie, Mutational effects on transglycosylating activity of family 18 chitinases and construction of a hypertransglycosylating mutant, Biochemistry 50 (2011) 5693–5703.

Computational and Theoretical Chemistry 1044 (2014) 29–35

Contents lists available at ScienceDirect

Computational and Theoretical Chemistry

journal homepage: www.elsevier.com/locate/comptc

Theoretical study on the mechanism and kinetics of ring-opening polymerization of cyclic esters initiated by tin(II) *n*-butoxide

Chanchai Sattayanon [a], Watit Sontising [a], Jitrayut Jitonnom [b], Puttinan Meepowpan [a,c], Winita Punyodom [a,c], Nawee Kungwan [a,c,*]

[a] Department of Chemistry, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand
[b] Division of Chemistry, School of Science, University of Phayao, Phayao 56000, Thailand
[c] Biomedical Polymers Technology Unit, Department of Chemistry, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand

ARTICLE INFO

Article history:
Received 22 March 2014
Received in revised form 10 June 2014
Accepted 11 June 2014
Available online 19 June 2014

Keywords:
Ring-opening polymerization
Coordination–insertion mechanism
Cyclic esters
Tin(II) *n*-butoxide
Density functional theory
Transition state theory

ABSTRACT

The coordination–insertion mechanisms of ring-opening polymerizations (ROP) of γ-valerolactone (**GVL**), δ-valerolactone (**DVL**) and L-lactide (**LL**) initiated by tin(II) *n*-butoxide were investigated using hybrid density functional theory at the B3LYP level with a mixed basis set. Analysis of geometries and corresponding energies revealed six intermediate states (including two transition states) observed in all three ROP reactions, which confirmed a coordination–insertion mechanism. Calculated energy profiles demonstrated exothermicity for each ROP reaction with a rate-determining step found for the first transition state that involved the attraction of an oxygen atom of the monomer to the metal atom of the initiator. By using the resulting energy barriers and transition state theory, the thermal rate constants could be obtained with the predicted rates of **GVL** and **LL** found to be equal and faster than that of **DVL**. In addition, by comparisons of the ring-strain energies and the barrier heights, we found that the effect of ring strain was at most only a partial driving force for the ROP reactions.

1. Introduction

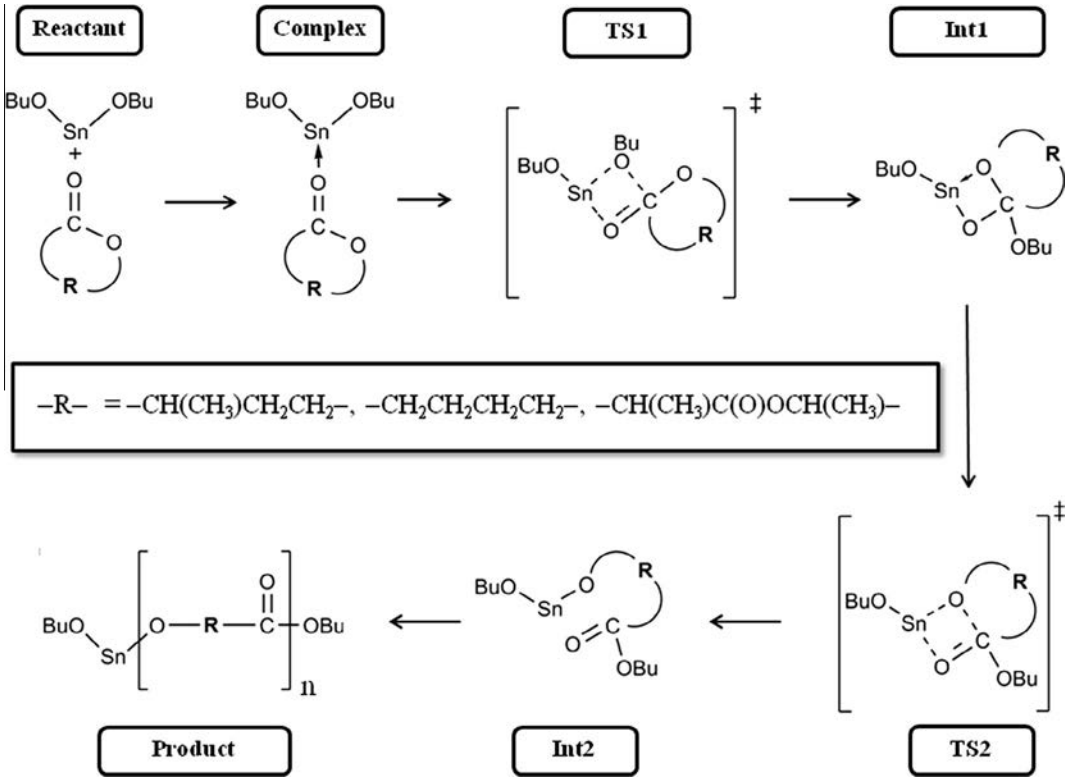
Polyester was first approved for use as biodegradable suture in the 1960s. Since then, various products based on cyclic ester have been produced as biocompatible materials [1–3]. A cyclic ester is the most potentially useful monomer for preparing biodegradable polymer in industries. Compared to a hydroxyl carboxylic acid as a monomer used in polycondensation reactions to directly produce polyester, the cyclic ester does not give any water as a by-product in the reaction. Without producing any by-product, high molecular weight of desirable polyester could be achieved. Ring opening polymerization (ROP) is one of the most widely used and effective technique for obtaining polyesters with high molecular weight [4–6] and metal alkoxides has been widely used as effective initiators for ROP of cyclic esters. Various organometallics derivatives of metals with *d*-orbitals of favorable energy such as Al, Sn, Y, Nd, Yb, Sm, La, Fe, Zn, Zr, Ca, Ti and Mg have been developed in the ROP of cyclic esters [5–7]. The ROP of cyclic esters is proposed by Kricheldorf et al. [8] to proceed via a coordination–insertion mechanism that involved in an insertion of monomer into the metal–oxygen bond of initiator (see Scheme 1).

For decades, $Sn(Oct)_2$, a well-known metal alkoxide, has been used in both academia and industries to synthesize polyesters from the ROP of cyclic esters [4,9–14] but its own drawbacks in terms of long induction time and uncontrollable molecular weight of polymer makes it unfeasible for current use in industries. On the other hand, tin(II) alkoxide, $Sn(OR)_2$, initiators have increasingly been used in the ROP of cyclic ester [15,16].

In our previous work, we studied kinetics model with computational tools to predict theoretical rate constants for the ROP of ε-caprolactone (CL) and tin(II) alkoxides [17]. The thermal rate constants were calculated, which give the predicted values in excellent agreement with the experiment, and we found that tin(II) *n*-butoxide ($Sn(OBu)_2$) is the most effective initiator in the $Sn(OR)_2$ series studied. Here, we extend our previous theoretical study to gain more understanding of the ring strain effect on ROP of cyclic monomers using the best $Sn(OBu)_2$ initiator. Three cyclic ester monomers, γ-valerolactone (**GVL**) [18], δ-valerolactone (**DVL**) [19], and L-lactide (**LL**) [20], which cover both steric and ring strain effect were chosen for this investigation, as shown in Fig. 1. Usually, the ROP of cyclic monomer is thought to be driven mainly by the ring strain of monomer; however, there are some cases

* Corresponding author at: Department of Chemistry, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand.
E-mail address: naweekung@gmail.com (N. Kungwan).

http://dx.doi.org/10.1016/j.comptc.2014.06.008
2210-271X/



Scheme 1. Coordination–insertion mechanism of ROP of cyclic monomers initiated by $Sn(OBu)_2$.

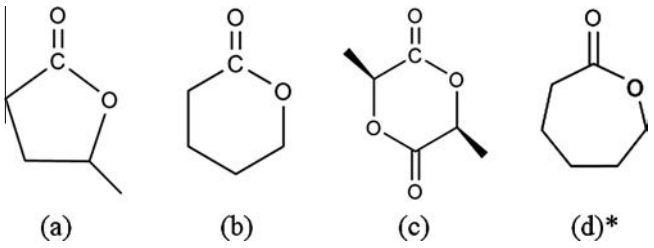


Fig. 1. Three cyclic ester monomers used in this study: (a) γ-valerolactone (**GVL**), (b) δ-valerolactone (**DVL**), (c) L-lactide (**LL**) and one monomer is, (d)* ε-caprolactone (**CL**) as the reference from our previous work [17].

where ROPs might not be driven by this effect. For instance, the rate constant of basic hydrolysis of γ-butyrolactone and CL reported by Ya and Leonid [21] at comparable conditions are close to each other (1.5×10^{-4} and 2.6×10^{-4} mol^{-1} L s^{-1}, respectively), although their corresponding ring strains were found to be energetically different (5 vs −29 kJ mol^{-1}) [22–24]. These experimental observations emphasize the need for further sophisticated studies in order to fully understand the ring strain effect on the ROP of cyclic esters.

In the present work, we provide a comparative theoretical investigation that demonstrates such important effect through following calculations. The mechanism for ROP of three cyclic ester monomers (**GVL**, **DVL** and **LL**) initiated by $Sn(OBu)_2$ were initially modeled with quantum chemical methods following the coordination–insertion mechanism, as depicted in Scheme 1. Geometries, energies and vibrational frequencies of all stationary points (reactant, transition state, intermediate and product) were then characterized and identified along the reaction energy profiles. Using the activation barriers obtained from the energy profiles, the thermal rate constants for ROP of three monomers (**GVL**, **DVL** and **LL**) initiated by $Sn(OBu)_2$ were estimated by using transition state theory.

2. Computational details

Hybrid density functional theory (DFT) at B3LYP level [25] with a mixed basis set, which is a popular and computationally cost effective method [26–28], was used to investigate the ROP mechanism of three cyclic ester monomers (**GVL**, **DVL**, **LL**) using $Sn(OBu)_2$ as initiator. For the metal atom, a double-ζ-valence quality basis set LANL2DZ was assigned for a Sn atom, which is relativistic electron core potential (ECP) developed by Hay and Wadt replaced the Sn core electron [29,30]. For non-metal atoms, a valence triple zeta with polarization function (VTZ2P) at cc-pVTZ was assigned for C, H, and O atoms. This B3LYP method with mixed basis set has been successfully used in our previous study on the ROP of CL and tin(II) alkoxides [17]. Geometry optimizations were carried out without any symmetry or geometric constraints, and energies of all stationary points (reactant, complex, transition state, intermediate and product) along the reaction path were computed to obtain the relative energy profiles. The nature of the minima (intermediate) and the extrema (transition states) were confirmed with analytical frequency calculations [26,31]. Furthermore, the pathway that connects reactants, transition state and products was checked with the assistance of the intrinsic reaction coordinate (IRC) [32]. The default IRC procedure using the hessian-based predictor–corrector integrator (HPC) [32,33] was carried out using the step size of 20 and maxpoint of 10 both forward and reverse directions which means that reaction coordinate was calculated every 0.2 $amu^{1/2}$ bohr. The energy barrier heights of all reactions were corrected by including the zero-point energy corrections [34]. All calculations were performed with the Gaussian03 software package [35]. The electronic density was analyzed using the natural bond orbital (NBO) analysis. The degree of charge separation between

positive and negative charges of all points along the reaction path was provided using the dipole moment data.

The obtained information from the quantum chemical calculations was employed to determine thermal rate constants of the reactions. These rate constants in a temperature range of 100–120 °C were calculated using a conventional transition state theory (TST) method [36] by University of Utah's web-based kinetics module within the Computational Science and Engineering Online suite (CSEOnline) [37]. This TST method was developed as a predictive tool for thermal rate constants based on the potential energy surface with a statistical representation of classical dynamics. TST is a well-known developed formula for obtaining thermal rate constants of chemical reaction [38]. Let us consider a general reaction for any initiation step of the polymerization reaction, where the reactant monomer (M) is being initiated by the initiator (I) to form the activated complex (MI) or the transition state (TS) and after that the product (P) is achieved.

$$M + I \leftrightarrow MI \rightarrow P$$

The thermal rate constant, $k(T)$, for the initiation step is given by Eq. (1) below

$$k(T) = \frac{k_B T}{h} \frac{Q_{MI}(T)}{Q_M(T) Q_I(T)} e^{\{-\Delta V^{\neq}/k_B T\}} \quad (1)$$

where k_B is the Boltzmann's constant, T is the temperature, h is the Planck's constant. $\Delta V^{\neq}$ is the classical barrier height; that is, a different energy between the M and TS. Q_M and Q_I are the total partition functions of the M and I, respectively. Q_{MI} is the MI or TS partition function [38]. The partition functions of the M and TS can be computed by providing the frequency information into TheRate [39] and the thermal rate constants of all ROP with different ring strain monomer initiated with $Sn(OBu)_2$ are calculated.

3. Results and discussion

3.1. Coordination–insertion mechanism of ROP

The ROP mechanisms for **GVL**, **DVL** and **LL** monomers initiated by $Sn(OBu)_2$ were investigated using DFT(B3LYP) and mix basis set method. The optimized geometries and corresponding energies for each step (shown in Scheme 1) of the **GVL** reaction are depicted in Fig. 2, respectively. Since all three ROPs share the same mechanism with six intermediates observed, we thus discussed in detail only on the ROP of **GVL** as a representative case for other cyclic ester monomers (**DVL** and **LL**). More detailed information for the mechanism of **DVL** and **LL** are given in Figs. S1 and S2 of the Supplementary data, respectively. In general, the ROP mechanism shown in Fig. 2 is described below.

The ROP reaction is initiated by a direct insertion of a carbonyl oxygen (O_2) of **GVL** to the metal center ($Sn—O_2$ = 2.67 Å) which leads to formation of a stable complexation (**Complex**) between **GVL** and the initiator as indicated by a lower energy (-7.14 kcal mol^{-1}) compared to the reactant energy. Then, the **Complex** is transformed into **Int1** via the coordination–insertion mechanism (see **TS1** → **Int1**), which involved a geometric rearrangement between the butoxy group (O_3—Bu) and the $Sn—O_2$ bond, forming a coordinated geometry featuring a spiral bicyclic species having the two oxygen atoms (O_2 and O_3) bridged between the Sn and C_1 atoms. A single negative (imaginary) frequency (158i cm^{-1}) was observed at **TS1** (see Table S1) with a smooth IRC curve connecting **Complex** and **Int1** (Fig. S3 in the Supplementary data). The $Sn—O_2$/$Sn—O_3$ distances are around 2.2 Å representing the forming/breaking bonds, respectively. A shortening of $C_1—O_3$ distance ($C_1—O_3$ = 3.47 Å) at **Complex** to 2.01 Å at **TS1** indicates that this bond is formed. Such rearrangement resulting in **TS1** requires at least 13.83 kcal mol^{-1} of energy before it can proceed to the intermediate species (**Int1**) with $C_1—O_3$ bond.

During the formation of **Int1**, rotation of **GVL** ring around the $C_1—O_2$ bond is required to make the $Sn—O_1$ distance shorter (3.77 Å at **TS1** to 2.42 Å at **Int1**), allowing the $Sn—O_1$ bond to be formed at **Int2** (Fig. 2). The newly formed $C_1—O_3$ bond is 1.41 Å with a sp^3 hybridization observed on the C_1 atom (changing from sp^2 in **Complex**). The energy of **Int1** is 11.23 kcal mol^{-1} above that of **Complex**. The next step is the formation of **TS2**, which is achieved through the insertion of the oxygen atom (O_1) of **GVL** to the Sn atom that leads to another coordinated geometry, similar to that observed at **TS1** (see angles in Table S5). This step requires energy of 1.34 kcal mol^{-1} for **GVL**. The **TS2** structure features a distorted four-membered ring fused with the five-membered ring of monomer having the $Sn—O_1$/$Sn—O_2$ distances of 2.19 Å. The true structure of **TS2** was confirmed, by vibration frequency and IRC calculations, with an imaginary frequency of 216i cm^{-1} and the appearance of a saddle connecting **Int1** and **Int2**, respectively (see Table S1 and Fig. S4 in the Supplementary data). A highly constraint geometry of **GVL** was observed in **TS2** as indicated by a fast elongation of $C_1—O_1$ bond as well as shortening of $Sn—O_1$/$C_1—O_2$ bonds. As a result, ring-opening of the monomer occurred via $C_1—O_1$ bond, producing **Int2**. Finally, the C_1 atom of carbonyl group regains the sp^2 hybridization at **Product**.

3.2. Natural bond orbital analysis

Changes in NBO charges on some selected atoms of all intermediates found along the ROP of **GVL** with $Sn(OBu)_2$ are illustrated in Fig. 3. This analysis provides a more qualitative way to understand the chemical changes occurring at atomic level, as demonstrated by previous studies [34,40] which showed a good correlation between the charges of nucleophilic center and reactivity (i.e., the more negative the charge, the higher the reactivity). Here, only the O_1, O_2, O_3, C_1, and Sn atoms are considered as the nucleophilic center. As the reaction progresses, all three oxygen atoms (O_1, O_2 and O_3) show a large change in charge distribution with the O_1 and O_3 charges found to change in an opposite trend: the O_1 increases in negative charge while the O_3 shows a decrease in negative charge. This change could be expected as the two oxygen atoms (O_1 and O_3) have evolved during the bond breaking and forming processes, respectively. For the O_2 atom, the increasing and decreasing curve can be seen as it evolves in both processes, that is, bond forming (**Complex** → **Int1**) and bond breaking (**Int1** → **Int2**). A slight change in the Sn and C_1 charges was observed, which is probably due to the charge compensation from the O_2 and O_3 to Sn at **Complex** and **TS1** (Fig. 3), which was also observed in our previous study [17]. Such changes in the NBO charges are also similar for **DVL** and **LL** (Fig. S6 in the Supplementary data).

3.3. Dipole moment

In addition to the NBO analysis, dipole moment can be used as an indicator for the degree of separation between positive and negative charges in a molecule. As plotted in Fig. 4, the calculated dipole moments of **Int1** and **Int2** are smaller than that of **Complex**. Thus, a solvent with low-polarity (nonpolar solvent) could stabilize **Int1** and **Int2** and a fast monomer insertion reaction could be expected, which was reported by experimental studies [41,42]. In addition, ROP mechanism of **DVL** and **LL** initiated by $Sn(OBu)_2$ are generally similar to that of **GVL** in terms of bond lengths and NBO charges found on the reaction intermediates (**Complex**, **TS1**, **Int1**, **TS2**, **Int2**). More information on vibrational frequencies and IRC calculations along the reaction intermediates are given in the supplementary data.

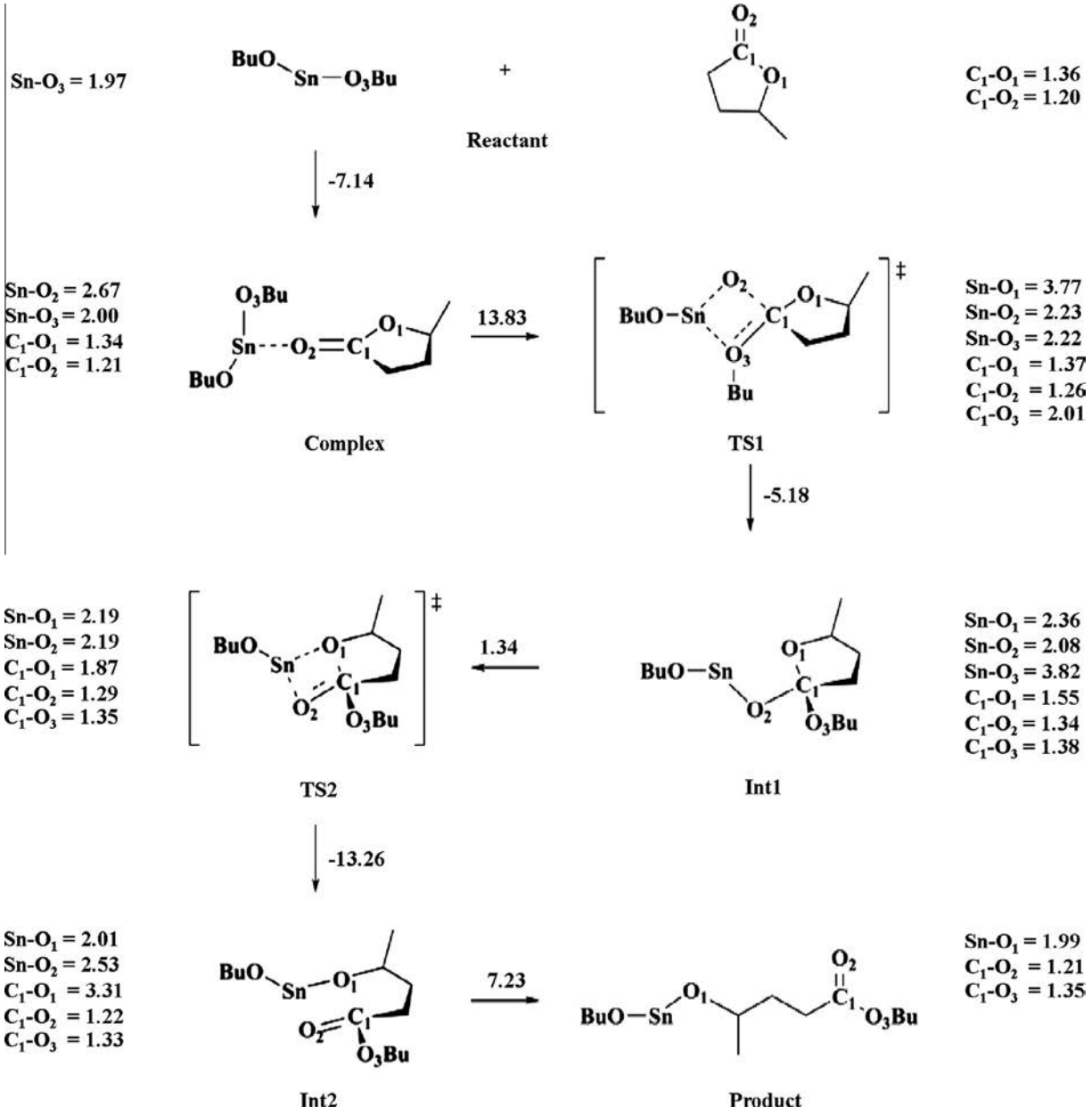


Fig. 2. ROP coordination–insertion mechanism of **GVL** initiated with $Sn(OBu)_2$. Bond lengths are noted in Å and formation energies are reported in kcal mol^{-1}.

3.4. Comparison of the barrier height, effect of ring strain and thermal rate constants

Calculated relative energies of key intermediates observed in the ROP of **GVL**, **DVL**, **LL** and **CL** initiated by $Sn(OBu)_2$ based on DFT calculations were summarized in Table 1 and plotted in Fig. 5. As shown, the stability of **Complex** is found in the following order: **DVL** > **GVL** > **LL**. The formation of **TS1** for all monomers requires a moderate energy of around 5–10 kcal mol^{-1}: **DVL** (10.69) > **GVL** (6.69) > **LL** (5.49). However, while the formation of **TS2** requires a lower energy than that of **TS1**, the **TS2** energies share the same order as **TS1**: **DVL** (4.91) > **GVL** (2.85) > **LL** (2.10). All monomers indicate the formation of **TS1** to be the rate-determining step and **LL** was found to have the fastest rate constants with the lowest barrier height. More detailed information on the comparison of rate constants calculated for the ROP of these monomers will be discussed in the next paragraph.

Theoretically, the trend of barrier heights should be in the following order: **CL** > **DVL** > **LL** > **GVL** according to ring strain of cyclic monomers [43,44], in which the ring size of each monomer is: CL = 7, LL = 6, DVL = 6, and GVL = 5. However, as shown in Fig. 5, the trend of our calculated barrier heights (**DVL** > **CL** > **GVL** > **LL**; the lower barrier, the faster rate) does not follow the theoretical prediction. Therefore, the ring strain of cyclic monomer is not the main driving force for the ring-opening polymerization. So the detail analysis on the stability of transition state formation is needed to clarify the main driving force. Many researchers [45] have suggested that the polymerizability of cyclic monomers can be evaluated by its free energy of polymerization (ΔGp). However, a complete sampling of the configurational space of the polymer is required to obtain the ΔGp, which is very difficult by the computational techniques available at present time. In this work, DFT calculations were employed to obtain the internal energy difference between **reactants** (initiator plus cyclic monomer) and their

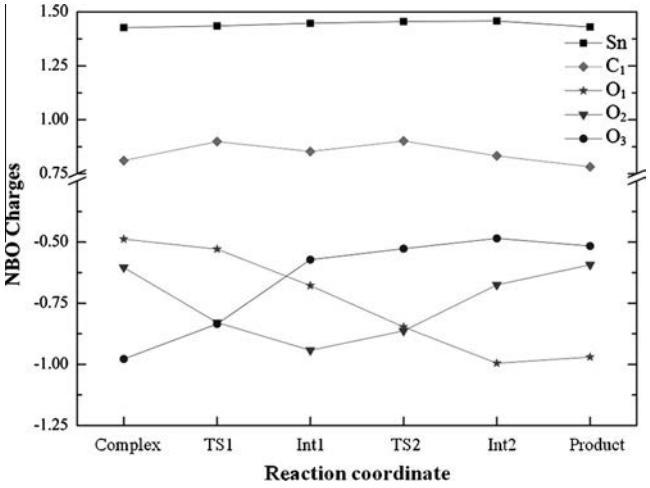


Fig. 3. Changes in NBO charges (e) of selected atoms for the reaction intermediates observed in the ROP of **GVL**.

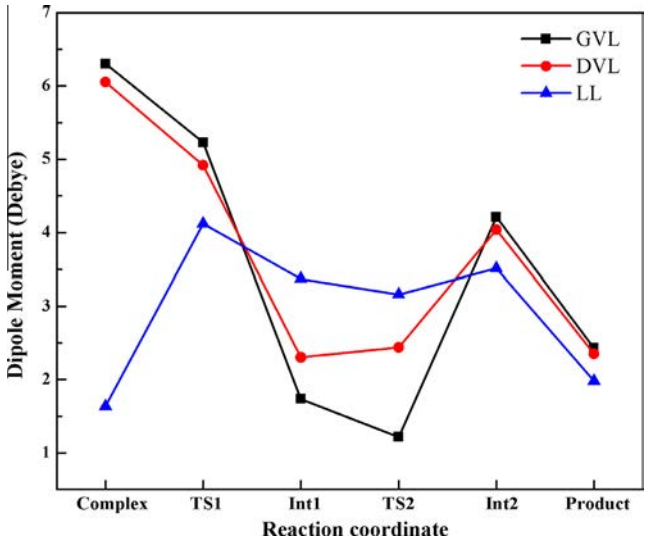


Fig. 4. Calculated dipole moments of the reaction intermediates observed in the ROP reaction of **GVL**, **DVL** and **LL**.

Table 1
Calculated relative energies (in kcal mol^{-1}) of key intermediates observed in the ROP of **GVL**, **DVL**, **LL** and **CL** initiated by $Sn(OBu)_2$.

Reaction coordinate	Relative energies[a]			
	GVL	**DVL**	**LL**	**CL**[b]
Reactant	0.00	0.00	0.00	0.00
Complex	−7.14	−8.14	−6.85	−7.53
TS1	6.69	10.49	5.49	7.05
Int1	1.51	3.19	0.90	1.67
TS2	2.85	4.91	2.10	11.29
Int2	−10.41	−11.28	−6.21	−3.45
Product	−3.18	−8.91	−11.80	−9.37

[a] All energies were calculated in relative to the reactant energy.
[b] Ref. [16].

TS1 (rate-determining step) (Table S4), which provided an estimation of the enthalpy of polymerization (ΔH_p). Generally, the ΔH_p is directly related to the ring strain, which is released through the polymerization process to provide the main driving force for the polymerization of cyclic monomers [46–49]. The experimental ΔH_p energy of different cyclic lactones was reported [44] and their

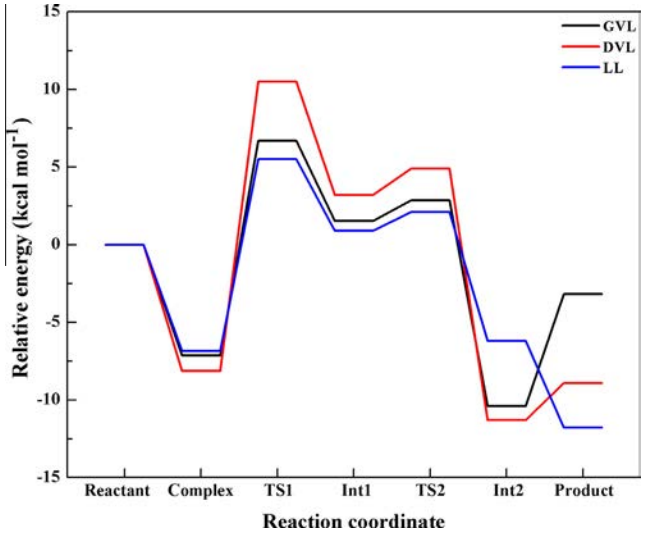


Fig. 5. Relative energy profiles for the ROP of **GVL** (black), **DVL** (red) and **LL** (blue) initiated by $Sn(OBu)_2$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

corresponding values are in the following order: **CL** (−6.89) > **DVL** (−6.6) > **LL** (−5.4) > **GVL** (1.2 kcal mol^{-1}). This order is in well agreement with the calculated ΔH_p by Refs. [43,46,49]. Moreover, the estimated energies of ring strain for different cyclic monomers were reported to be following this order: **CL** (11.7) > **DVL** (9.9) > **GVL** (7.4) > **LL** (5.5 kcal mol^{-1}) [6,49]. From these estimated values, the thermal rate constants of ROP of **CL** should be the fastest and that of **LL** should be the slowest compared to other cyclic monomers. However, these values are solely from free cyclic monomers which are only applied for self-ring opening polymerization but not for the ROP with catalysts. Therefore, taken only ring strain effect into account for the prediction of thermal rate constants is not enough, other factors need to be considered in order to explain the relative heights of our calculated reaction barriers (**TS**s).

Since the **TS1** is the rate-determining step, we analyzed the important bond angles relevant to their original cyclic esters ($\angle O_1$—C_1—C (see Fig. 2), or angle a in supplementary data,) and the values of all three systems are given in Table S4 of the Supplementary data. It is found that angles at position $\angle O_1$—C_1—C are ranked in this order: **CL** (118°) [17] > **DVL** (116°) > **LL** (113°) > **GVL** (108°). These values suggest that **GVL** has the highest ring strain and theoretically may lead to the highest rate constants. Moreover, the ΔH values of the formation of **TS1** are found to be 7.45, 6.70 [17], 6.61, 5.20 kcal mol^{-1} for **DVL**, **CL**, **GVL**, and **LL**, respectively. These values imply that **LL** should have the fastest among the ROP of cyclic monomer initiated with $Sn(OBu)_2$.

The calculated rate constants of all there system are summarized in Table 2. The order of the calculated rate constants is **GVL** = **LL** > **CL** > **DVL** which is in accordance with the trend of barrier height and the ΔH of **TS1**. The agreement between the calculated rate constant of **CL** and **DVL** with their corresponding ring-strains energies (**CL** (11.7) > **DVL** (9.9 kcal mol^{-1}) is observed, following the theoretical prediction with the higher ring strain, the faster rate constant. However, the ring-strain energies of **GVL** and **LL** are 7.4 and 5.5 kcal mol^{-1}, respectively, and our computed angles at position $\angle O_1$—C_1—C shows that the **TS1** of **GVL** (108°) is more strain than that of **LL** (113°). If the driving force for the polymerization process is dependent only on the ring strains of the **TS1**, the calculated rate constants for **GVL** with $Sn(OBu)_2$ should be larger than that of **LL** with $Sn(OBu)_2$. However, the calculated rate constants of **GVL** are almost equal to those of **LL**, which is caused from the lower barrier height of **LL**. These findings suggest that

Table 2
Calculated rate constants for the ROP of **GVL**, **DVL**, **LL**, and **CL** initiated by $Sn(OBu)_2$ under various temperatures.

Temperature (°C)	Thermal rate constants (L mol^{-1} min^{-1})			
	GVL	**DVL**	**LL**	**CL**[a]
100	421.88	0.61	436.69	33.57
102	447.89	0.66	458.36	–
104	474.98	0.72	481.48	–
106	503.51	0.78	504.96	–
108	533.13	0.85	529.52	–
110	564.56	0.92	555.16	45.01
112	597.42	1.00	581.89	–
114	632.10	1.08	609.34	–
116	668.22	1.17	638.24	–
118	706.15	1.26	667.86	–
120	745.88	1.36	698.92	59.56

[a] Ref. [16].

the principal driving force for the ROP of cyclic monomers is not only from the ring strain but also from the energy required to overcome the rate-determining step.

4. Conclusion

DFT calculations on coordination–insertion mechanism and relative energy profile were carried out for the ROP of three different cyclic monomers (γ-valerolactone (**GVL**), δ-valerolactone (**DVL**) and L-lactide (**LL**)) initiated by tin(II) butoxide, $Sn(OBu)_2$. Results indicate that the **TS1** for all three systems is the rate-determining step, which is similar to that of ε-caprolactone (**CL**) with tin(II) alkoxides [17]. The order of barrier heights is found to follow this order: **DVL** (10.69) > **CL** (7.05) > **GVL** (6.69) > **LL** (5.49 kcal mol^{-1}), which is not the same order as their corresponding ring-strain energy **CL** (11.7) > **DVL** (9.9) > **GVL** (7.4) > **LL** (5.5 kcal mol^{-1}) [6,49]. The transition state theory was used to estimate the thermal rate constants for all systems in which the order of the rate constants are found to be: **GVL** = **LL** > **CL** > **DVL**, which is with the trend of their corresponding barrier heights. The similar rate constants of **GVL** and **LL** as well as the disagreement in term of the order in the ring-strain energy and barrier heights are explained by the partial driving force of ring strain for the polymerization process.

Acknowledgements

The authors wish to thank the National Research University Project under Thailand's Office of the Higher Education Commission for financial support. J.J. is a TRF-CHE Young Research Scholar (MRG5680143), and thanks the University of Phayao for support (R020056216016). Department of Chemistry, Faculty of Science and The Graduate School, Chiang Mai University is also acknowledged.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.comptc.2014.06.008.

References

[1] R.K. Kulkarni, K.C. Pani, C.C. Neuman, F.F. Leonard, Polylactic acid for surgical implants, Arch. Surg. 93 (1966) 839–843.
[2] E.E. Schmitt, R.A. Polistina, Polyglycolic Acid Prosthetic Devices, American Cyanamid Co., Continuation-in-part of U.S. 3297033, 1969, pp. 10
[3] E.I. du Pont de Nemours Co, Surgical Sutures, 1966, pp. 18.
[4] A.-C. Albertsson, I.K. Varma, Recent developments in ring opening polymerization of lactones for biomedical applications, Biomacromolecules 4 (2003) 1466–1486.
[5] C. Jérôme, P. Lecomte, Recent advances in the synthesis of aliphatic polyesters by ring-opening polymerization, Adv. Drug Deliv. Rev. 60 (2008) 1056–1076.
[6] O. Dechy-Cabaret, B. Martin-Vaca, D. Bourissou, Controlled ring-opening polymerization of lactide and glycolide, Chem. Rev. 104 (2004) 6147–6176.
[7] J. Wu, T.-L. Yu, C.-T. Chen, C.-C. Lin, Recent developments in main group metal complexes catalyzed/initiated polymerization of lactides and related cyclic esters, Coord. Chem. Rev. 250 (2006) 602–626.
[8] H.R. Kricheldorf, T. Mang, J.M. Jonté, Polylactones, 2 copolymerization of glycolide with β-propiolactone, γ-butyrolactone or δ-valerolactone, Die Makromol. Chem. 186 (1985) 955–976.
[9] A. Kowalski, A. Duda, S. Penczek, Mechanism of cyclic ester polymerization initiated with tin(II) octoate. 2. Macromolecules fitted with tin(II) alkoxide species observed directly in MALDI-TOF spectra, Macromolecules 33 (2000) 689–695.
[10] H.R. Kricheldorf, K. Bornhorst, H. Hachmann-Thiessen, Bismuth(III) n-hexanoate and tin(II) 2-ethylhexanoate initiated copolymerizations of ε-caprolactone and L-lactide, Macromolecules 38 (2005) 5017–5024.
[11] A. Kowalski, J. Libiszowski, T. Biela, M. Cypryk, A. Duda, S. Penczek, Kinetics and mechanism of cyclic esters polymerization initiated with tin(II) octoate. Polymerization of ε-caprolactone and L,L-lactide co-initiated with primary amines, Macromolecules 38 (2005) 8170–8176.
[12] M. Sobczak, W. Kolodziejski, Polymerization of cyclic esters initiated by carnitine and tin(II) octoate, Molecules 14 (2009) 621–632.
[13] M. Sobczak, Ring-opening polymerization of cyclic esters in the presence of choline/SnOct2 catalytic system, Polym. Bull. 68 (2012) 2219–2228.
[14] J. Fernández, E. Meaurio, A. Chaos, A. Etxeberria, A. Alonso-Varona, J.R. Sarasua, Synthesis and characterization of poly (L-lactide/ε-caprolactone) statistical copolymers with well resolved chain microstructures, Polymer 54 (2013) 2621–2631.
[15] A. Kleawkla, Controlled Ring-Opening Polymerization of Cyclic Esters, Department of Chemistry, Faculty of Science, Chiang Mai University, Chiang Mai (Thailand), 2008.
[16] M. Dumklang, N. Pattawong, W. Punyodom, P. Meepowpan, R. Molloy, M. Hoffman, Novel tin(II) butoxides for use as initiators in the ring-opening polymerisation of e-caprolactone, Chiang Mai J. Sci. 36 (2009) 136–148.
[17] C. Sattayanon, N. Kungwan, W. Punyodom, P. Meepowpan, S. Jungsuttiwong, Theoretical investigation on the mechanism and kinetics of the ring-opening polymerization of ε-caprolactone initiated by tin(II) alkoxides, J. Mol. Model. 19 (2013) 5377–5385.
[18] C.W. Lee, R. Urakawa, Y. Kimura, Copolymerization of γ-valerolactone and β-butyrolactone, Eur. Polymer J. 34 (1998) 117–122.
[19] Y. Mahha, A. Atlamsani, J.-C. Blais, M. Tessier, J.-M. Brégeault, L. Salles, Oligomerization of ε-caprolactone and δ-valerolactone using heteropolyacid initiators and vanadium or molybdenum complexes, J. Mol. Catal. A: Chem. 234 (2005) 63–73.
[20] E. Dawes, N. Rushton, Response of macrophages to poly(L-lactide) particulates which have undergone various degrees of artificial degradation, Biomaterials 18 (1997) 1615–1623.
[21] I.G.d. Ya, I.B.k. Leonid, Strain and reactivity in monocyclic systems, Russian Chemical Reviews 29 (1960) 214.
[22] A. Duda, S. Penczek, P. Dubois, D. Mecerreyes, R. Jérôme, Oligomerization and copolymerization of γ-butyrolactone — a monomer known as unable to homopolymerize, 1. Copolymerization with ε-caprolactone, Macromol. Chem. Phys. 197 (1996) 1273–1283.
[23] B.V. Lebedev, A.A. Evstropov, N.K. Lebedev, E.A. Karpova, E.B. Lyudvig, B.G. Belen'kaya, Thermodynamics of ε-caprolactone, poly-ε-caprolactone and the ε-caprolactone polymerization process in the region 0–350 K, Vysokomol. Soedin., Ser. A 20 (1978) 1974–1980.
[24] A.A. Evstropov, B.V. Lebedev, E.G. Kiparisova, V.A. Alekseev, G.A. Stashina, Thermodynamic parameters of the transformation of γ-butyrolactone to poly-γ-butyrolactone under normal pressures at 0–400 K, Vysokomol. Soedin., Ser. A 22 (1980) 2450–2456.
[25] P.J. Stephens, F.J. Devlin, C.F. Chabalowski, M.J. Frisch, Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields, J. Phys. Chem. 98 (1994) 11623–11627.
[26] H. von Schenck, M. Ryner, A.-C. Albertsson, M. Svensson, Ring-opening polymerization of lactones and lactides with Sn(IV) and Al(III) initiators, Macromolecules 35 (2002) 1556–1562.
[27] J. Liu, J. Ling, X. Li, Z. Shen, Monomer insertion mechanism of ring-opening polymerization of [varepsilon]-caprolactone with yttrium alkoxide intermediate: a DFT study, J. Mol. Catal. A: Chem. 300 (2009) 59–64.
[28] M. Ryner, K. Stridsberg, A.-C. Albertsson, H. von Schenck, M. Svensson, Mechanism of ring-opening polymerization of 1,5-dioxepan-2-one and L-lactide with stannous 2-ethylhexanoate. A theoretical study, Macromolecules 34 (2001) 3877–3881.
[29] P.J. Hay, W.R. Wadt, Ab initio effective core potentials for molecular calculations. Potentials for the transition metal atoms Sc to Hg, J. Chem. Phys. 82 (1985) 270–283.
[30] W.R. Wadt, P.J. Hay, Ab initio effective core potentials for molecular calculations. Potentials for main group elements Na to Bi, J. Chem. Phys. 82 (1985) 284–298.
[31] J.L. Eguiburu, M.J. Fernandez-Berridi, F.P. Cossio, J.S. Roman, Ring-opening polymerization of L-lactide initiated by (2-Methacryloxy)ethyloxy−aluminum trialkoxides. 1. Kinetics, Macromolecules 32 (1999) 8252–8258.
[32] H.P. Hratchian, H.B. Schlegel, Accurate reaction paths using a Hessian based predictor–corrector integrator, J. Chem. Phys. 120 (2004) 9918–9924.

[33] H.P. Hratchian, H.B. Schlegel, Chapter 10 – finding minima, transition states, and following reaction pathways on ab initio potential energy surfaces, in: C.E. Dykstra, G. Frenking, K.S. Kim, G.E. Scuseria (Eds.), Theory and Applications of Computational Chemistry, Elsevier, Amsterdam, 2005, pp. 195–249.
[34] R. Zhu, R. Wang, D. Zhang, C. Liu, A density functional theory study on the ring-opening polymerization of D-lactide catalyzed by a bifunctional-thiourea catalyst, Aust. J. Chem. 62 (2009) 157–164.
[35] G.W.T.M.J. Frisch, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, J.A. Montgomery Jr., 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, J.A. Pople, Gaussian 03 (Revision E.01), Gaussian Inc, Wallingford CT, 2004.
[36] A. Khanna, Y. Sudha, S. Pillai, S. Rath, Molecular modeling studies of poly lactic acid initiation mechanisms, J. Mol. Model. 14 (2008) 367–374.
[37] T.N. Truong, S. Zhang, VKLab Version 1.0, University of Utah, Salt Lake (UT), 2001.
[38] K.J. Laidler, M.C. King, Development of transition-state theory, J. Phys. Chem. 87 (1983) 2657–2664.
[39] W.T. Duncan, R.L. Bell, T.N. Truong, TheRate: program for ab initio direct dynamics calculations of thermal and vibrational-state-selected rate constants, J. Comput. Chem. 19 (1998) 1039–1052.
[40] J. Ling, J. Shen, T.E. Hogen-Esch, A density functional theory study of the mechanisms of scandium-alkoxide initiated coordination–insertion ring-opening polymerization of cyclic esters, Polymer 50 (2009) 3575–3581.
[41] N. Ropson, P. Dubois, R. Jerome, P. Teyssie, Macromolecular engineering of polylactones and polylactides. 14. A carbon-13 and aluminum-27 NMR study of the effect of .gamma.-butyrolactone on the structure of aluminum isopropoxide in toluene, Macromolecules 26 (1993) 6378–6385.
[42] H. Ma, J. Okuda, Kinetics and mechanism of L-lactide polymerization by rare earth metal silylamido complexes: effect of alcohol addition, Macromolecules 38 (2005) 2665–2673.
[43] W. Saiyasombat, R. Molloy, T.M. Nicholson, A.F. Johnson, I.M. Ward, S. Poshyachinda, Ring strain and polymerizability of cyclic esters, Polymer 39 (1998) 5581–5585.
[44] A. Duda, A. Kowalski, Thermodynamics and Kinetics of Ring-Opening Polymerization, Handbook of Ring-Opening Polymerization, Wiley-VCH Verlag GmbH & Co. KGaA, 2009.
[45] G.C. Eastmond, A. Ledwith, S. Russo, P. Sigwalt (Eds.), Comprehensive Polymer Science: The Synthesis, Characterization, Reactions, and Applications of Polymers, Chain Polymerization, Part. I, vol. 3, Pergamon Press, 1989.
[46] K.N. Houk, A. Jabbari, H.K. Hall, C. Aleman, Why δ-valerolactone polymerizes and γ-butyrolactone does not, J. Org. Chem. 73 (2008) 2674–2678.
[47] H.R. Allcock, Ring-chain equilibria, J. Macromol. Sci., Part C 4 (1970) 149–189.
[48] H. Sawada, Thermodynamics of Polymerization, Dekker, 1976.
[49] T. Chen, Z. Qin, Y. Qi, T. Deng, X. Ge, J. Wang, X. Hou, Degradable polymers from ring-opening polymerization of [small alpha]-angelica lactone, a five-membered unsaturated lactone, Polym. Chem. 2 (2011) 1190–1194.

Chiang Mai J. Sci. 2014; 41(5.2) : 1220-1229
http://epg.science.cmu.ac.th/ejournal/
Contributed Paper

Effects of Silicon-Bridge and π-Ligands on the Electronic Structures and Related Properties of Dimethyl Zirconocene Polymerization Catalysts: A Comparative Theoretical Study

Jitrayut Jitonnom*[a] and Wijitra Meelua [a,b]

[a] Division of Chemistry, School of Science and [b] Demonstration School, University of Phayao, Phayao 56000, Thailand.

*Author for correspondence; e-mail: jitrayut.018@gmail.com

Received: 2 June 2013
Accepted: 27 July 2013

ABSTRACT

Group 4 metallocene has gained more attention recently as a potential catalyst for ring-opening polymerization of lactones. Introduction of a silicon-bridge and/or modifications of biscyclopentadiene (Cp_2) ligand with different π-ligands such as bisindene (Ind_2), mixed indene/cyclopentadiene (CpInd) or mixed uorene/cyclopentadiene (CpFlu) have been found to significantly help improve catalyst activities, polymer yields and molecular weight. In order to relate the silicon bridge (B) and π-ligand (L) effects with the catalyst activities, molecular structures and properties of those bridged dimethylzirconocene complexes as well as their insertion behaviors with ε-caprolactone were investigated by using the density functional theory method. Calculations showed that different π-ligands had a profound effect on the electronic geometries and related properties of the title complexes. All bridged catalysts were geometric constraints compared to the unbridged catalysts. The HOMO-LUMO gap and dipole moments were decreased as the large π-ligands were introduced (Cp_2 > CpInd > Ind_2 > CpFlu). Only Si and Zr atoms were observed to carry a positive charge with Si having the most. For the monomer-inserted complexes, most catalyst structures shared the same orientation of ε-caprolactone with its ester group pointing away from a methyl group of Zr atom, except the catalyst with Ind_2 ligand due to steric hindrance. A planar conformation of the monomer was observed in all catalysts and, hence, is required for each insertion. The data provide a structural basis for both inactivated and activated bridged-dimethylzirconocene complex catalysts with different π-ligands, which will be helpful for further studies on the lactone polymerization initiated by group 4 metallocene.

Keywords: silicon-bridge, zirconium, metallocene, density functional theory

1. INTRODUCTION

Ring-opening polymerization (ROP) of lactones represents a growing field of polymer research because the produced polyesters are biodegradable and biocompatible materials potentially used for biomedical and pharmaceutical applications as well as

environmentally friendly materials [1-3]. Several metal complexes have been well established for the ROP of lactones, such as alkoxides of Li [4], Al [5], Sn [6], Zn [7], Ln [8] and Ti [9]. Apart from these metal alkoxides, group 4 metallocenes have drawn increased attention, as an interesting group of transition metal complexes potentially applied for the ROP of lactones.

Previous studies have reported the use of dimethylzirconocene complexes, as pre-catalyst for the ROP of lactones with various substituted cyclopentadienyl (Cp), indenyl (Ind) and uorenyl (Flu) ligands and a bridge in the catalytic systems. Hayakawa et al. [10] was one of the first to report the ROP of lactones initiated by zirconocene-borate complex catalysts, which was found to proceed via cationic mechanism (Figure 1). Their catalytic system used is based on bis(cyclopentadiene), Cp_2ZrMe_2 and $CpCp^*ZrMe_2$ (Cp*: pentamethyl cyclopentadiene) with high molecular weights and low polydispersities of polylactone obtained. Kostakis et al. [11] reported the polymerization of ε-CL and δ-VL using various catalytic systems consisting of a combination of three C_{2v} zirconocene complexes and three borate cocatalysts. They introduced the bis(indene) ligand, Ind_2ZrMe_2, and also the Cp_2ZrMe_2, $(tBuCp)_2ZrMe_2$ in the zirconocene catalytic systems to initiate the polymerizations, producing polymers with relatively high molecular weights and narrow molecular weight distributions. Recently, Villaseñor et al. [12] have first introduced the *ansa*-zirconocene complex with silicon methyl (Me_2Si) as a bridge for the ROP of ε-CL. They successfully synthesized mixed indenyl/cyclopentadienyl zirconocene systems, $CpIndZrMe_2$, with and without the silicon bridge and used them in the polymerizations, yielding the resulting polymer with impressive quantitative outcome without cocatalyst (medium molecular weight polymers with moderate to broad polydispersities). Based on those studies, modifications of the π-ligands and the bridge in the catalyst are key factors to improve the catalytic activity, and thus guiding a development of group 4 metallocene complexes for the use as potential catalysts of the ROP of lactones. This has inspired us for a requirement of fundamental understanding of electronic effect of the π-ligands and the bridge on the efficiency of *ansa*-zirconocene catalysts in the ROP of ε-CL.

In this work, the electronic structures and related properties of four *ansa*-dimethylzirconocene complexes (**I-IV**) bearing Cp, Ind or Flu ligands (see Figure 2) were comparatively investigated by means of density functional theory. We analyzed the electronic geometries of the catalysts when the *ansa*-bridge and the ε-ligands are introduced to understand their structural basis. We continued analyzing the regioselectivity of the ε-CL monomer to the four catalysts during the monomer-insertion step by finding the most stable structure of the monomer-inserted complexes for each catalyst. To the best of our knowledge, the comparisons of the electronic structures and properties of catalysts **I-IV** have not been investigated theoretically.

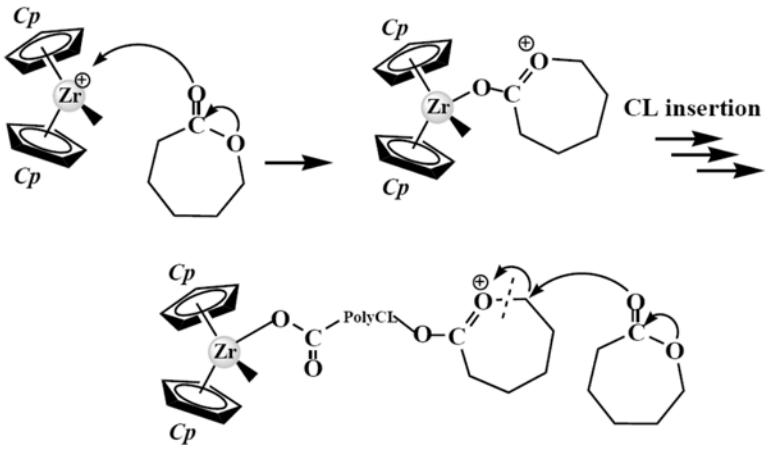


Figure 1. Cationic ROP mechanism of lactones initiated by dimethylzirconocene catalyst, Cp_2ZrMe_2.

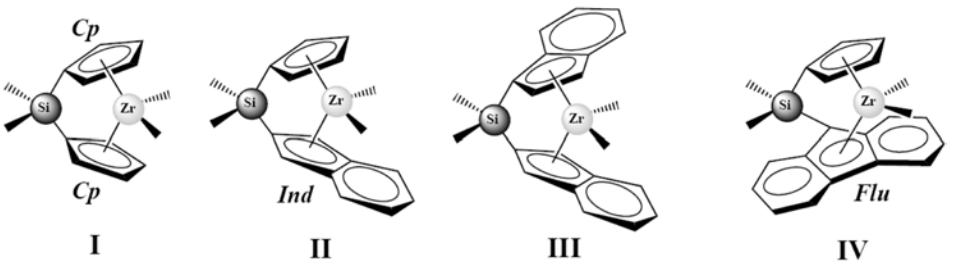


Figure 2. Generic structure of silicon-bridged dimethylzirconocene complex with different π-ligands.

2. MATERIAL AND METHODS

Geometries at the standard state (298 K and 1 atm) of all title complexes were optimized using Hartree Fock (HF) and Density Functional Theory (DFT) at the B3LYP level with the Gaussian09 program [13]. This popular and computationally cheap method predicted reliable geometries and quantum chemical results. We used a double-ζ basis set (6-31G*) for all non-metal atoms and the effective core potential double-ζ basis set (LANL2DZ) [14] for zirconium and silicon atoms. This mixed basis set was created through the use of the GEN keyword in Gaussian09. Both of these basis sets have been widely used along with DFT methods for the study of transition metal containing systems [14, 15]. Note that the chosen DFT methods were used only for describing the gas-phase structures of catalysts, which does not account for the dynamics of monomer and/or catalyst in solution.

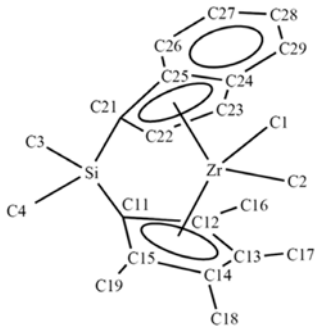
3. RESULTS AND DISCUSSION

3.1 Comparison of the Computed Structure with the X-ray Structure

Although the popular DFT method is widely used for locating the stable geometries of transition metal complexes [15], it is important to compare the DFT optimized structure of the investigated system with the available crystal structure in order to test the reliability of the method. To this end, catalyst **II** was chosen as a test model to compute its geometry at HF/LANL2DZ and B3LYP/LANL2DZ methods. The results are summarized in Table 1, along

with the experimental observation from the literature [12]. In general, the optimized geometry of catalyst **II** from the B3LYP/LANL2DZ method is more reasonably agreement with the X-ray crystal structure than that from the HF/LANL2DZ method, as evidenced by a relatively low mean absolute error (0.48 for B3LYP/LANL2DZ and 0.64 for HF LANL2DZ). The HF method gave an overestimation for most bond lengths compared to those structural parameters from the B3LYP method, for example, the distances of Zr-Cen1 and Zr-Cen2 bonds are calculated at 2.326/2.359 Å (HF) and 2.289 2.320 Å (B3LYP), higher than the X-ray observation (2.222/2.264 Å), respectively. The Cen1-Zr-Cen2 angles are calculated at 124.8° (HF) and 126.3° (B3LYP), which is close to the X-ray structure (128.1°). This result represents a case where the exchange and correlation functional is necessary, supporting the benefit of the B3LYP method over the HF method for studying the transition metal complex.

Table 1. Experimental and calculated bond lengths and bond angles for catalyst **II.**



		Calculation			
Bond length, angle	X-ray[a]	HF	error	B3LYP	error
Zr–C1	2.267	2.285	0.018	2.271	0.004
Zr–C2	2.271	2.295	0.024	2.282	0.011
Zr–C21	2.485	2.552	0.067	2.536	0.051
Zr–C22	2.500	2.561	0.061	2.537	0.037
Zr–C23	2.600	2.685	0.085	2.661	0.061
Zr–C24	2.670	2.79	0.120	2.777	0.107
Zr–C25	2.580	2.703	0.123	2.686	0.106
Si–C3	1.860	1.895	0.035	1.897	0.037
Si–C4	1.866	1.899	0.033	1.901	0.035
Si–C11	1.880	1.911	0.031	1.908	0.028
Si–C21	1.867	1.893	0.026	1.894	0.027
Zr-Cen1	2.222	2.326	0.104	2.289	0.067
Zr-Cen2	2.264	2.359	0.095	2.320	0.056
av Zr–C(Cen1)[b]	2.531	2.599	0.068	2.588	0.057
av Zr–C(Cen2)[b]	2.567	2.658	0.091	2.639	0.072
C1–Zr–C2	96.4	96.0	0.4	98.5	2.1
C3–Si–C4	107.4	104.9	2.5	105.5	1.9
C11–Si–C21	97.1	96.6	0.5	96.6	0.5
Cen1–C11–Si	160.6	163.9	3.3	162.5	1.9
Cen2–C21–Si	161.5	164.0	2.5	162.7	1.2
Cen1–Zr–Cen2	128.1	124.8	3.3	126.3	1.8
Mean absolute error			0.64		0.48

[a] Taken from Ref [11]. [b] Refers to the average bond length between Zr and the carbon atoms of the C5 ring of the corresponding cyclopentadienyl moiety. Cen1 and Cen2 are the centroids of C11–C15 and C21–C25, respectively. For clarity, the figure representing a chemical structure of catalyst **II** is included without the hydrogen atoms.

3.2 The Bridging Effect on the Electronic Geometries of Catalysts I-IV

The optimized structures of catalysts **I-IV** calculated at the B3LYP/LANL2DZ level of theory are illustrated in Figure 3 and their corresponding geometric parameters are listed in Table 2. It appears that silicon bridge (Me_2Si) has a significant effect to the catalyst structure compared to the unbridge one: it reduces the angle (α) formed between the metal and the two π-ligands and increases the angle (β) formed between the plane of π-ligands: for catalyst **I** the α and β values are calculated at 124.9° and 59.9°, respectively, for the bridge but are 133.8° and 46.1° for the unbridge. Similar observation was also found for other catalysts (see Table 2). More compact structures were observed for all bridging catalysts, as evidenced by the decreased distances between two centroids of π-ligands (Cen-Cen) and between the metal and the centroid (M-Cen) shown in Table 2. With a constrained geometry in the catalysts, the metal must arrange itself to fit with the packed structure, leading to increase the D distance in all four catalysts, especially in the catalyst **III** (with the maximum extent of D = ~0.4 Å). This metal extension may facilitate the cocatalyst activation and/or monomer insertion.

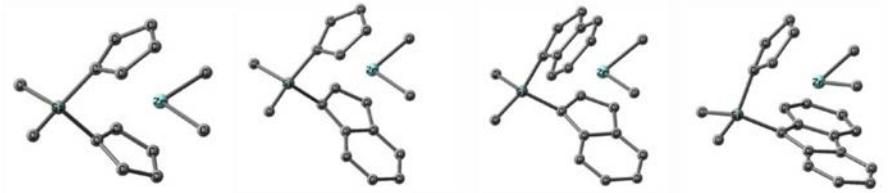
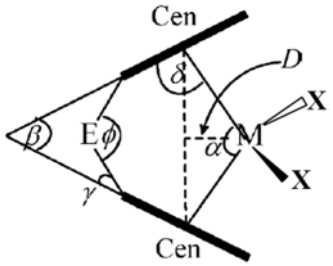


Figure 3. Optimized structures of catalysts **I-IV** calculated at the B3LYP/LANL2DZ level of theory.

Table 2. Geometric parameters for the optimized structures of catalysts **I-IV** calculated at the B3LYP/LANL2DZ level of theory. The geometric structures of the unbridged silicon of catalysts **I-IV** are also indicated in parenthesis.



Catalyst	Φ	Cen-Cen (Å)	D (Å)	α (°)	β (°)	γ (°)	δ (°)	M-Cen (Å)	M-X (Å)	X-M-X (°)	E-M (Å)
I	81.98	4.05	1.05	124.9	59.9	17.9	87.5	2.284	2.280	39.9	3.383
		(4.20)	(0.89)	(133.8)	(46.1)		(88.7)	(2.286)	(2.287)	(99.3)	
II	81.95	4.14	1.03	128.8	60.4	10.0	87.0	2.270	2.275	40.1	3.359
		(4.22)	(0.99)	(132.1)	(51.9)		(85.0)	(2.337)	(2.275)	(98.1)	
III	89.83	4.15	1.30	117.9	68.7	12.0	84.1	2.522	2.265	40.0	3.351
		(4.90)	(0.95)	(140.3)	(43.1)		(85.3)	(2.643)	(2.286)	(41.0)	
IV	96.74	4.20	1.04	127.2	65.9	13.5	82.7	2.267	2.267	98.1	3.325
		(4.32)	(0.92)	(133.8)	(53.3)		(83.1)	(2.270)	(2.270)	(98.3)	

3.3 The π-Ligands Effect on the Electronic Geometries and Properties of Catalysts I-IV

Considering the size of the π-ligands in catalysts **I-IV**, the number of aromatic ring is in the order of **I** < **II** < **III** = **IV**. This increment in the size of the π-ligands within the catalysts was found to have a significant effect on the electronic geometries of the catalysts. We observed a decrease in the angle (δ) formed between the metal and the centroid changing from 87.5° (**I**) → 87.0° (**II**) → 84.1° (**III**) → 82.7° (**IV**). This is due to an increase in the van der Waals (VDW) interaction between the aromatic ring of the π-ligand and the methyl groups at the Zr atom. The lowest angle (δ) observed in catalyst **IV** is due to the VDW interaction between two methyl groups and two benzene rings of the fluoryl ligand. With the strong VDW in catalyst **IV**, the X-M-X angle was found to be two-fold large (~98°) compared to that for other catalysts (**I-III**) (~40°), as shown in Table 2. Among the four catalysts, catalyst **III** has the most steric hindrance between one methyl group of Zr atom and the two benzene rings of the indene ligands. That is why the highest value of both β and M-Cen was observed in catalyst **III**.

The π-ligand introduction has a significant effect on the electronic properties. We found a decrease in the HOMO-LUMO energy gap for the four catalysts in the order of **I** > **II** > **III** > **IV**, as shown in Figure 4, associated with the decreased dipole moment (Table 3). We also observed a positive charge on the Si and Zr in all catalysts, with Si carrying the most.

Table 3. Energy, frontier orbital (HOMO,LUMO), dipole moment and Muliken charges (*e*) of catalysts **I-IV**.

Catalyst	Energy (a.u.)	LUMO (eV)	HOMO (eV)	Δ(LUMO-HOMO)	Dipole moment	Mulliken charges (*e*) of Si/Zr
I	-596.040	-1.225	-6.177	4.95	2.01	1.228/0.765
II	-749.672	-1.197	-5.470	4.27	1.78	1.242/0.785
III	-903.303	-1.252	-5.306	4.05	1.62	1.252/0.788
IV	-903.306	-1.361	-5.170	3.81	1.47	1.239/0.805

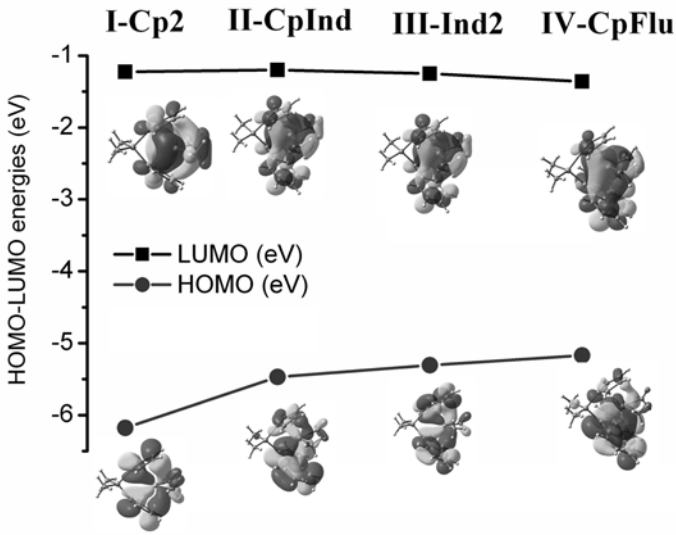


Figure 4. HOMO-LUMO energy gaps and frontier orbitals of catalysts **I-IV**.

3.4 Analysis of Regioselectivity of ε-CL from the Monomer-Inserted Complex Structure

To understand a regioselectivity for the monomer-insertion step of the ROP of ε-CL using those title catalysts (**I-IV**), the electronic structures of the monomer-inserted complex with different monomer insertions were analyzed. The optimized structures for the monomer complexes (**I-IV**) were calculated at the B3LYP/LANL2DZ level of theory and the results are illustrated in Figures 5-8. We found that the possible number of monomer-inserted complex depends on the symmetry of the catalyst. Two possible orientations of the monomer are observed for catalyst **I** and **IV** due to the C_{2v} symmetry, as shown in Figures 5 and 8. On the other hand, catalysts **II** and **III** with the C_1 symmetry gave four possible orientations of monomer. Excluding of catalyst **III**, the most energetically stable structure for each catalyst was found to share the same conformation with the ester group of the monomer pointing away from the methyl group at Zr atom (see complexes **I-a**, **II-a**, and **IV-a** in Figures 5, 6 and 8). The orientation of the catalysts was found to correspond well with the shortest distance of Zr-O1 bond (~2.18 Å- ~2.19 Å). In case of catalyst **III**, the most stable structure was found in complex **III-c** in which the ester group still keeps pointing to the catalyst but its monomer is situated at the other side, due to a steric influence from the π-ligand. A planar conformation in the lactone ring was also observed in all catalysts, with the C1-C2-O2-C3 torsion angle of -4.41° - 3.45°, suggesting that the geometric rearrangement is required for each insertion

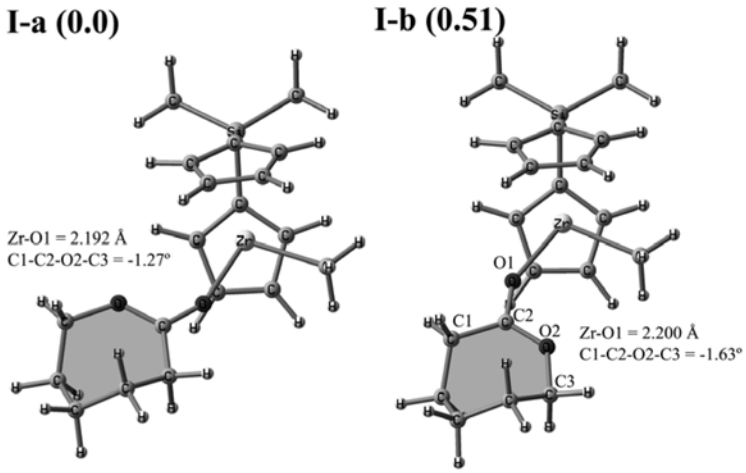


Figure 5. Activated complex structure of catalyst **I** with different ε-caprolactone insertions calculated at the B3LYP/LANL2DZ level of theory. Relative energies are indicated in parenthesis.

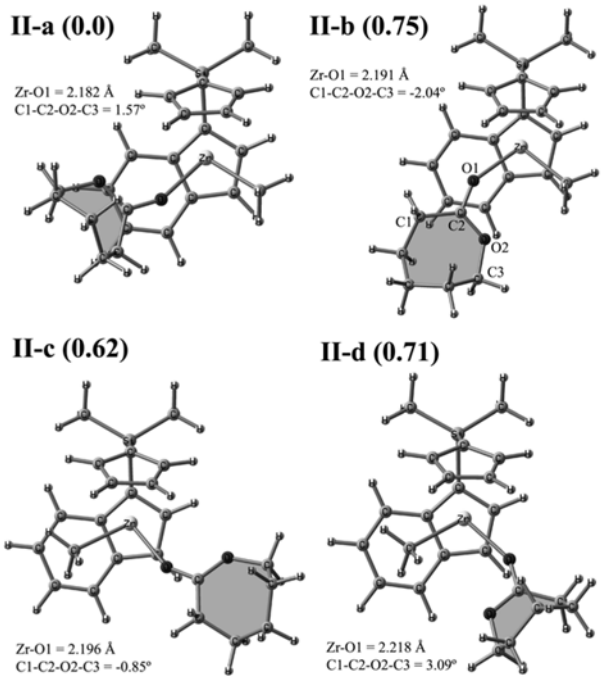


Figure 6. Activated complex structure of catalyst **II** with different ε-caprolactone insertions calculated at the B3LYP/LANL2DZ level of theory. Relative energies are indicated in parenthesis.

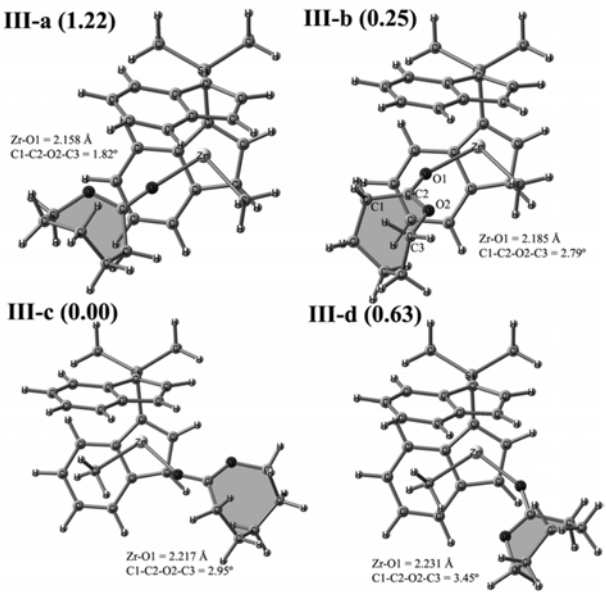


Figure 7. Activated complex structure of catalyst **III** with different ε-caprolactone insertions calculated at the B3LYP/LANL2DZ level of theory. Relative energies are indicated in parenthesis.

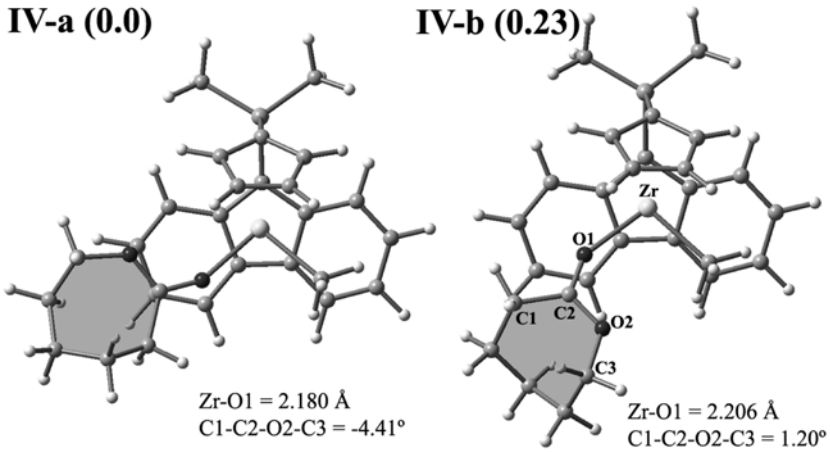


Figure 8. Activated complex structure of catalyst **IV** with different ε-caprolactone insertions calculated at the B3LYP/LANL2DZ level of theory. Relative energies are indicated in parenthesis.

4. CONCLUSIONS

In this paper, we present the computational analyses for the electronic structures and properties of four bridged dimethylzirconocene catalysts (**I-IV**) bearing different π-ligands such as cyclopentadienyl, indenyl or fluorenyl ligands. The introduction of the silicon-bridge and the π-ligands had a profound effect on the electronic geometries and properties in all four catalysts. In comparison with the unbridged catalysts, we observed a constrained geometry for all bridged catalyst with the Zr atom moving away from the catalyst center to some extent. The energy intervals between LUMO and HOMO and the dipole moments for the four catalysts are in the order of **I** > **II** > **III** > **IV**. The positive charges are observed only on Si and Zr for each of the four complexes. We then examined the monomer-inserted complexes of all catalysts and found that, excluding of catalyst **III**, most energetic favorable structures of the complexes have the shortest Zr-O distance with a stable conformation having the ester group of the monomer pointing to the catalyst. A steric hindrance is dominant in catalyst **III**, thus giving the different complex structure. In addition, a planar conformation in the lactone ring was also observed in all catalysts. These data provide a structural basis of dimethyl zirconocene and dimethyl zirconocene-monomer complex, which will be helpful for further experimental and theoretical investigations on the ROP of ε-CL initiated by group 4 metallocene catalyst.

ACKNOWLEDGEMENTS

This work is supported by grants from the University of Phayao (R020056216016) and the Thailand Research Fund (MRG-5680143). We thank Dr. Jon I. Mujika for helpful comments and Dr. Banjong Chairinkham for editorial assistance. J.J. would like to thank National e-Science Infrastructure Consortium for providing computing resources that have contributed to the research results reported within this paper (URL: http://www.e-science.in.th).

REFERENCES

[1] Seyednejad H., Ghassemi A.H., van Nostrum C.F., Vermonden T. and Hennink W.E., Functional aliphatic polyesters for biomedical and pharmaceutical applications, *J. Control Release*, 2011; **152**: 168-176.

[2] Khan J.H., Schue F. and George G.A., Heterogeneous ring-opening polymerization of lactones for biomedical applications, *Polym. Inter.*, 2009; **58**: 296-301.

[3] Albertsson A.C. and Varma I.K., Recent developments in ring opening polymerization of lactones for biomedical applications, *Biomacromolecules*, 2003; **4**: 1466-1486.

[4] Meelua W., Bua-Own V., Molloy R. and Punyodom W., Comparison of metal alkoxide initiators in the ring-opening polymerization of caprolactone, *Adv. Mater. Res.*, 2012; **506**: 142-145.

[5] Gao A., Mu Y., Zhang J. and Yao W., Aluminum complexes with bidentate N,N-dialkylaniline-arylamido ligands: synthesis, structures, and catalytic properties for efficient ring-opening polymerization of e-caprolactone, *Eur. J. Inorg. Chem.*, 2009; **2009**: 3613-3621.

[6] Dumklang M., Pattawong N., Punyodom W., Meepowpan P., Molloy R. and Hoffman M., Novel tin(II) butoxides for use as initiators in the ring-opening polymerisation of ε-caprolactone, *Chiang Mai J. Sci.*, 2009; **36**: 136-148.

[7] Huang B.H., Lin C.N., Hsueh M.L., Athar T. and Lin C.C., Well-defined sterically hindered zinc aryloxides: Excellent catalysts for ring-opening polymerization of [var epsilon]-caprolactone and l-lactide, *Polymer*, 2006; **47**: 6622-6629.

[8] Sheng H., Li J., Zhang Y., Yao Y. and Shen Q., $LnNa_8[OC(CH_3)_3]_{10}(OH)$ alkoxide clusters: Highly active single-component catalysts for the homopolymerization and copolymerization of ε-caprolactone and trimethylene carbonate, *J. Appl. Polym. Sci.*, 2009; **112**: 454-460.

[9] Meelua W., Molloy R., Meepowpan P. and Punyodom W., Isoconversional kinetic analysis of ring-opening polymerization of ε-caprolactone: Steric influence of titanium(IV) alkoxides as initiators, *J. Polym. Res.*, 2012; **19**: 9799-9810.

[10] Hayakawa M., Mitani M., Yamada T. and Mukaiyama T., Living ring-opening polymerization of lactones using cationic zirconocene complex catalysts, *Macromol. Chem. Phys.*, 1997; **198**: 1305-1317.

[11] Kostakis K., Mourmouris S., Karanikolopoulos G., Pitsikalis M. and Hadjichristidis N., Ring-opening polymerization of lactones using zirconocene catalytic systems: block copolymerization with methyl methacrylate, *J. Polym Sci. Part A: Polym. Chem.*, 2007; **45**: 3524-3537.

[12] Villaseñor E., Gutierrez-Gonzalez R., Carrillo-Hermosilla F., Fernández-Galán R., López-Solera I., Fernández-Pacheco A.R. and Antiñolo A., Neutral dimethylzirconocene complexes as initiators for the ring-opening polymerization of ε-caprolactone, *Eur. J. Inorg. Chem.*, 2013; **2013**: 1184-1196.

[13] Gaussian 09, Revision B.01, Gaussian Inc., Wallingford CT, USA, 2010.

[14] Yang Y., Weaver M.N. and Merz K.M., Jr., Assessment of the "6-31+G** + LANL2DZ" mixed basis set coupled with density functional theory methods and the effective core potential: Prediction of heats of formation and ionization potentials for first-row-transition-metal complexes, *J. Phys. Chem. A.*, 2009; **113**: 9843-9851.

[15] Cramer C.J. and Truhlar D.G., Density functional theory for transition metals and transition metal chemistry, *Phys. Chem. Chem. Phys.*, 2009; **11**: 10757-10816.

Chiang Mai J. Sci. 2014; 41(X) : 1-11
http://epg.science.cmu.ac.th/ejournal/
Contributed Paper

Catalytic Oxidation of Glucose with Hydrogen Peroxide and Colloidal Gold as Pseudo-Homogenous Catalyst: A Combined Experimental and Theoretical Investigation

Jitrayut Jitonnom [a] and Christoph Sontag *[a]
[a] Division of Chemistry, School of Science, University of Phayao, Phayao 56000, Thailand.
*Author for correspondence; e-mail: c.sontag@web.de

Received: 10 July 2013
Accepted: 25 April 2014

ABSTRACT

Gold nanoparticles have been proved to act as oxidation catalyst for glucose oxidation, offering a "chemical" synthetic route to gluconic acid and gluconates - nowadays commercially produced by an enzyme catalyzed oxidation. Our investigations of the gold catalyzed oxidation route showed that gold nanoparticles produced by a modified Turkevich method have a high activity for this pseudo-homogenous catalytic reaction. Under mild reaction conditions, glucose could be oxidized in good yields (~70%) and the resulting gluconate could be isolated by column chromatography and precipitation as calcium salt. The catalytic oxidation reaction was found to follow the first-order kinetic with a rate constant of 4.95 h^{-1}, in good agreement with previous finding. The underlying reaction mechanism is discussed, assuming that the formation of a gold-glucose cluster intermediate is a key catalytic step. Several structures of the gold-glucose intermediates were examined using density functional theory methods. The molecular behavior of glucose adsorption in gold colloid solution is present.

Keywords: colloidal gold catalyst, glucose oxidation, gold nanoparticle, gluconic acid, DFT

1. INTRODUCTION

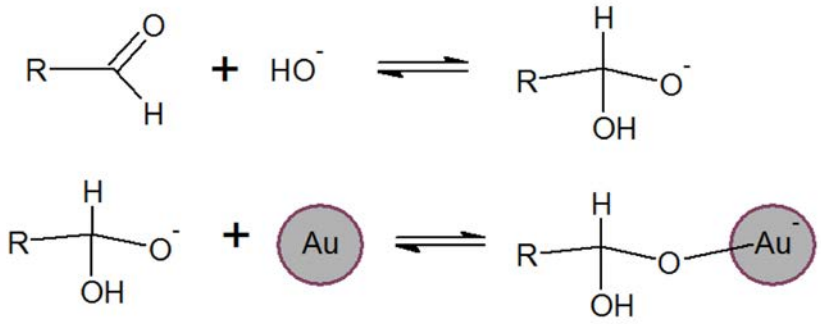
Gold nanoparticles (Au NPs) as highly effective green catalysts for oxidative processes have become a research topic of wide-spread interest in recent years [1-5], representing a bridge between homogeneous and heterogeneous catalysis. Au NPs are among the most popular metal nanoparticles due to their stability and fascinating properties. They have applications in very important fields such as catalysis and medicine [5,3,6].

Catalysis by Au NPs has been emerged by the discovery of gold nanoparticles as highly active supported catalysts for oxidation of CO into CO_2 by atmospheric air [7]. Since then, much effort has been tried to understand the catalytic activity of Au NPs [8-10]. With high potency of Au NPs, they have attracted many industries utilizing metal-catalyzed chemical processes. Among those reactions, the gold-catalyzed oxidation of glucose to

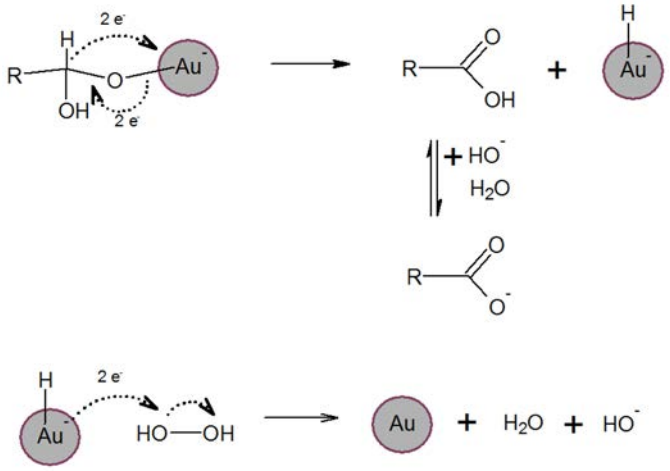
gluconic acid by Au NPs has received much attention in recent years [11-14,10,15]. In addition, the glucose-gluconic acid oxidation reaction have shown to be a promising "chemical" synthetic route to sodium gluconate [14], which is commercially produced by a fermentation process [16], making a catalytic chemical reaction of high industrial interest. The product gluconic acid is an important compound with cation complexing abilities and is used as additive with many applications in the pharmaceutical, food and detergent industry as mild, non-toxic acid [16].

The glucose oxidation can be efficiently catalyzed by using supported and unsupported Au NPs. The carbon-supported gold catalysts have been widely used for the liquid-phase oxidation of glucose due to their high activity compared to conventional catalysts. Biella et al. [15] firstly used the carbon-supported Au NPs as catalysts for aerobic oxidation of D-glucose, exhibiting a much higher activity than the conventional platinum and paladium based catalysts. Further work by Önal et al. [14] found that D-glucose oxidation to D-gluconic acid over Au/C catalysts is structure sensitive. The kinetics of glucose oxidation to gluconic acid with Au/Al_2O_3 catalyst has also been reported by Prüβe and co-workers [11] with a modified oxidative dehydrogenated mechanism proposed. Despite the rare use of unsupported colloidal gold catalyst compared to those supported catalysts, it was found that catalysis by colloidal gold showed a surprisingly high activity in the aerobic oxidation of glucose, comparable with the supported Au/C catalyst and an enzymatic system [10,13]. Presently, unsupported Au NPs as catalyst for glucose oxidation with hydrogen peroxide have been examined yet only by Comotti et al. [12], showing good results in terms of reaction time and yield.

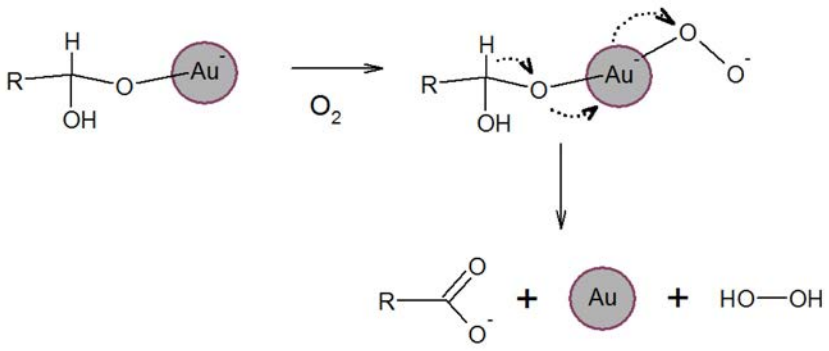
The reaction mechanism for the glucose oxidation with Au NPs catalysts has been proposed recently [11]. The first reaction step is the formation of the Au-glucose intermediate, which is initiated by the nucleophilic attack of a hydroxide ion on glucose followed by adsorption of the resulting alkoxy group to a gold particle. As a result, the relative high electro-negativity of gold is generated (Scheme 1). The oxidation step could occur via either gold hydride anion [11] or peroxygen [12] depending on the present of oxygen. In hydride anion pathway, the oxidation proceeds by transferring the carbon-bound hydrogen together with two electrons to the gold particle, releasing the gluconate anion. Subsequently, the transferred hydrogen could be transferred to hydrogen peroxide as hydride anion, reducing it to water and hydroxyl (Scheme 2). In peroxygen pathway, the addition of oxygen (from the decomposition of hydrogen peroxide) to the Au-glucose adduct was proposed (Scheme 3). It is interesting to note that both pathways share common specie of Au-glucose intermediate, which is thought to be critical during catalysis [11,12]. At present, no structural information of the key intermediate has been reported and, thus, the complete understanding for the glucose oxidation mechanism is limited.



Scheme 1. The formation of the Au-glucose intermediate



Scheme 2. The decomposition of the Au-glucose intermediate via hydride anion if the oxygen is absent



Scheme 3. The decomposition of the Au-glucose intermediate via peroxygen if the oxygen is present

In this study, the conditions for a high-yield oxidation of glucose by hydrogen peroxide using unsupported Au NPs were examined, together with a theoretical model to explain the structure of the key Au-glucose intermediate as well as its molecular behavior of glucose adsorption in gold colloid solution. Our study indicates that, even in very low concentration (about 30 ppm), gold colloid is an effective catalyst for the oxidation of glucose to gluconate salts with hydrogen peroxide under mild, alkaline conditions. The underlying mechanism for the glucose absorption on gold nanoparticle is also discussed.

2. EXPERIMENTAL

2.1 Chemicals and Equipment

D(+)-glucose monohydrate (for biochemical use), sodium hydroxide p.a. grade (Aldrich), hydrochloric acid conc. (Aldrich), silica gel for chromatography (Aldrich) and hydrogen peroxide solution 50% (commercial grade) was used. As gold precursor, gold leaves of 96.5% purity ("Thai gold") were used. For the recording of the VIS spectra, a Jasco V-530 UV/VIS spectrometer was used.

2.2 Preparation of the Nano Gold Catalyst

A small gold leave (weight 3 mg) was dissolved in ca. 0.5 ml conc. HCl and 0.5 ml hydrogen peroxide (50%) by gentle heating in a test tube. The resulting clear yellow solution was evaporated to dryness in a water bath to eliminate all traces of HCl. The residue was dissolved in 5 ml of distilled water. In an Erlenmeyer, 100 ml distilled water with 5 ml 1% trisodium citrate solution was heated to boiling under magnetic stirring. Compared to the classical Turkevich method [18,19], a higher water volume showed to produce smaller nanoparticles with better reproducibility. The above solution of chloroauric acid was added quickly under vigorous stirring. After 10 minutes of boiling and stirring, the solution was allowed to cool to room temperature and a VIS-spectrum was taken to confirm the particle size of the gold nanoparticles. The reaction mixture was used as catalyst without further treatment.

2.3 Oxidation of Glucose

1.8g (10 mmol) of glucose and 2 ml (26 mmol) of hydrogen peroxide (50%) was added to the nano gold suspension and the mixture was heated under stirring to 50°C. With the use of a pH meter, the pH was continually adjusted to 9 by drop wise addition of 0.2M NaOH. The progress of the reaction:

$$R\text{-}C(=O)H + H_2O_2 + OH^- \rightarrow R\text{-}C(=O)O^- + 2\ H_2O$$

could be estimated by the amount of sodium hydroxide - to obtain 100 % yield, a total of 50 ml is required (10 mmol hydroxide).

2.4 Isolation of the Product

In order to separate the product (sodium gluconate) from remaining glucose and catalyst, the reaction mixture was concentrated to about 20 ml by rotation evaporation. A column with diameter of 3-4 cm was filled with slurry of silica gel in iso-propanol (about 50 ml). The reaction mixture was passed slowly through the column using iso-propanol as solvent. The out-coming fractions were tested for product by TLC (using iso-propanol/water 4:1). The product containing fractions showed an oily layer which is supposed to be sodium gluconate dissolved in iso-propanol. This gel-like product could not be crystallized, so it was

transferred to the calcium salt: the product fractions were evaporated to dryness resulting in a light brown gel, which was dissolved in about 10 ml distilled water. About 1 g of calcium chloride was added forming a clear solution. Slow addition of iso-propanol resulted in a white precipitate, which could be filtrated to give 1.6 g of calcium-gluconate (3.65 mmol), corresponding to 73% yield.

3. THEORETICAL

Previous studies [20-24] have shown that the DFT calculations are quite useful for understanding the adsorption orientation of a molecule on a metal nanoparticle at molecular level. Here, two models of glucose (Glc) and glucose substituted with hydroxide from metal alkaline (GlcOH) absorbed on the gold atom were chosen to represent the formation of key gold-glucose intermediate. Two orientations of the glucose were considered in which an aldehyde or a C6-OH end-group of glucose is preferred to adsorb with the gold atom. A total of four models were thus modeled in this study. The D-glucose isomer was considered in the models. A single Au atom was used as a representative surface of gold nanoparticle. Despite the gold model used is very simple, it has also been applied in several literatures [21-23], which is enough for the purpose of the present study. All system structures were fully optimized using the Density Functional Theory (DFT) method with the Becke three parameter hybrids functional with Lee-Yan-Parr correlation function (B3LYP) level of theory. An effective core potential basis set LANL2DZ (Los Alamos set of double-zeta) [25] was used for a description of Au atom. The energetic of the adsorption process were calculated at the same level as geometry optimization using following equation:

$$E_{ads} = E_{complex} - E_{glucose} - E_{Au}$$

where $E_{complex}$, $E_{glucose}$ and E_{Au} are the total ground state (electronic) energies of the Au–glucose complex, the isolated glucose (Glc or GlcOH) and the isolated Au, respectively. The more negative the adsorption energy, the stronger the adsorption. All calculations were performed using Gaussian09 program [26].

4. RESULTS AND DISCUSSION

4.1 Characterizations of Gold Nanoparticle Size, Optimum pH and Temperature for the Reaction and the Gluconate Product

Figure 1 shows the VIS spectrum of the gold NP suspension after preparation with sodium citrate and after the completion of reaction with glucose and hydrogen peroxide at pH 9. It is clearly shown that addition of glucose significantly changes in both color of the gold suspension and the spectrum. The color of the suspension changed during one hour from wine-red to dark-blue (data not shown) and a part of the particles agglomerated giving rise to a second mai peak around 600 nm, as would be expected in an alkaline solution with sodium anions [27]. Haiss et al. [28] have demonstrated that there is a good correlation between the medium particle size of gold NP and the wavelength of the absorption maximum in visible light. According to them, the particle size can be estimated using the formula: $d \sim \ln[(\lambda_{max} - 512)/6.53]/0.0216$ nm. With an absorption maximum at about 520 nm, it can be concluded that the particle sizes are in the magnitude of 10 nm.

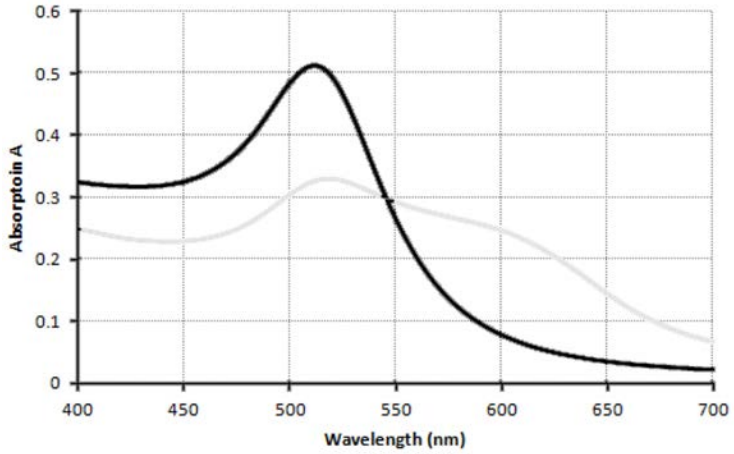


Figure 1. VIS spectrum of gold NP suspension directly after the preparation with sodium citrate (black line) and after one-hour-reaction with glucose and hydrogen peroxide at pH 9 (grey line)

To find the optimum conditions for the oxidation reaction, it has to be ensured that the pH and temperature are well below the critical limits when caramelization of glucose occurs. This phenomenon [17] could be qualitatively followed by observing the color of a solution of 10 mmol glucose in diluted sodium hydroxide at different pH. Several solutions were prepared with different pH values and heated at 60°C, which, according to [11], could be the maximum reaction temperature. It could be detected that glucose starts to decompose into brown-colored by-products at pH >= 10 at T >= 55°C within 1 hour. From this finding, a pH of 9 and 50°C were chosen to avoid this decomposition reaction during the oxidation with hydrogen peroxide.

To characterize the products after completed reaction and after re-crystallization as calcium salt, standard TLC plates with silica on aluminum were used. A mixture of i-propanol with water (4:1) is a suitable solvent. To develop the TLC, the plates were heated on a hot plate where the products appear as brown spots. Figure 2 shows a TLC of the sodium gluconate from the reaction mixture, which has a spot at significant lower R_f than glucose (0.24 vs. 0.64). The final product, the calcium salt, does barely move on the plate and has an R_f close to zero, shown in the right part of Figure 2.

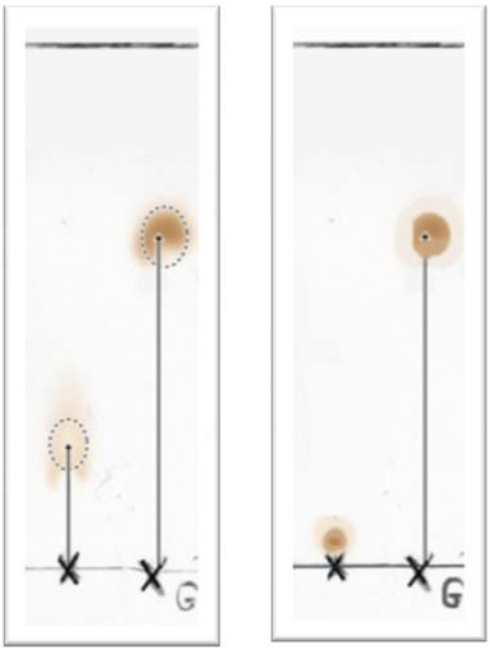


Figure 2. TLC of sodium gluconate (left) and calcium gluconate (right) vs. glucose in i-propanol/water 4:1

4.2. Reaction Rate

Following the addition of sodium hydroxide to the reaction mixture of glucose and hydrogen peroxide in gold NP suspension, a theoretical yield of nearly 90% could be detected after 30 minutes. The amount of added sodium hydroxide to maintain pH 9 declines during the reaction, indicating a first order kinetics with respect to glucose. The remaining glucose concentration in mol/L can be calculated from the consumption of NaOH: 1 ml of 0.2M NaOH corresponds to the loss of 0.2 mmol of glucose, corresponding to 2 mmol/L. For a first order kinetics, the natural logarithm of the glucose concentration over time should give a straight line with the slope representing the reaction constant. The result is shown in Figure 3. A rate constant for the first-order reaction was found to be 4.95 h^{-1}, in good agreement with the kinetic observation of Comotti et al. [12], who estimated a rate constant of 4 h^{-1} for the product formation.

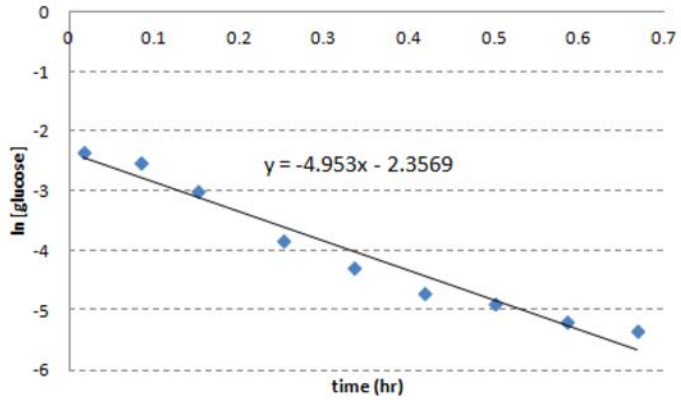


Figure 3. Plot of the logarithm of the remaining glucose concentration during the oxidation reaction in alkaline H_2O_2 over the reaction time.

4.3 Proposed Glucose Adsorption on Gold Nanoparticle

In order to understand the molecular behavior of glucose adsorption in gold colloid solution such as adsorption orientation and interaction, four models of glucose absorbed on gold nanoparticles were calculated using the DFT calculations with LANL2DZ basis set. The optimized structures (**Model I-IV**) of the Au-Glc and Au-GlcOH complexes with one Au atom absorbed on the oxygen atoms of the aldehyde and C6-hydroxyl end-group of glucose are illustrated in Figure 4. The calculated adsorption energies for the four models are listed in Table 1. For the Au-Glc complexes, the Au atom is preferred to adsorb with the C6-hydroxyl oxygen of glucose end-group (see **Model II**; Figure 4b, right), as suggested by the energetically favorable structure and a shorter Au-O distance of 2.54 Å compared to those from **Model I** (Figure 4a, left). For the Au-GlcOH complexes, the GlcOH molecule is most likely to adsorb on the Au atom through the carbonyl oxygen of the aldehyde end-group of glucose, shown in **Model III** (Figure 4b, left), with the lowest energy structure found among the four models (Table 1) and a short Au-O distance of 2.25 Å. This orientation was found to give the most favorable absorption with the lowest adsorption energy of -22.56 kcal/mol, indicating that the glucose adsorption is an exothermic process.

Based on the obtained calculations, the glucose may weakly interact with the gold nanoparticles through the CH_2OH end-group then it may react with the NaOH prior to adsorb with the gold nanoparticles as shown in Figure 5. However, it is also possible that the glucose may react with NaOH to generate the glucose anion and then adsorb on the gold nanoparticles.

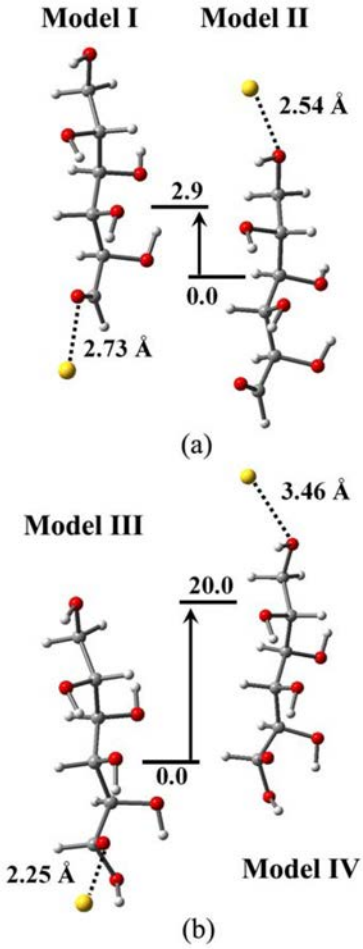


Figure 4. Optimized structures of four models (**I-IV**) representing the Au-Glc (a) and Au-GlcOH (b) complexes with Au absorbed on the oxygen atoms from the aldehyde (left) and C6-OH (right) end-group of glucose. The unit of energy is in kcal/mol**.**

Table 1. The adsorption energies for the four models (**I-IV**) from Figure 4.

Model	total energies/relative energies(kcal/mol)	adsorption energy(kcal/mol)
I	-516168 / 0.0	4.3
II	-516171 / 2.9	1.4
III	-563801 / 0.0	-22.6
IV	-563781 / 20.0	-2.5

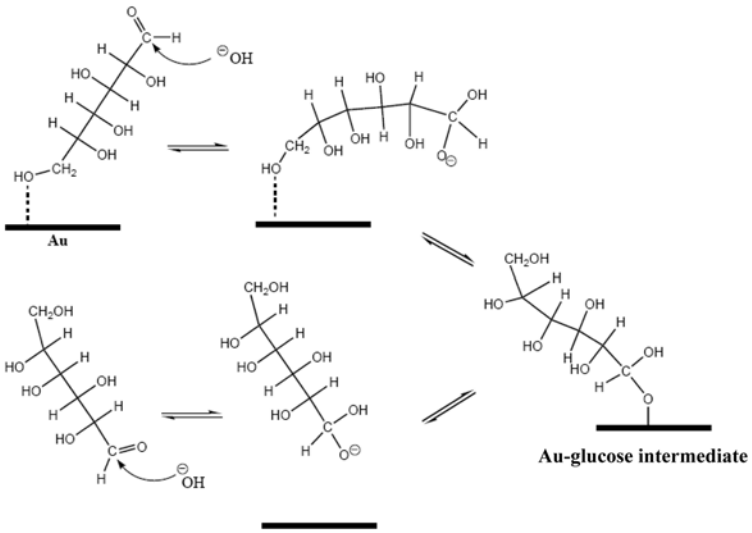


Figure 5. Possible glucose adsorption on gold nanoparticles during the formation of key Au-glucose intermediate as suggested by the present calculations.

5. CONCLUSIONS

Our investigations of the gold catalyzed oxidation route showed that the gold nanoparticles produced by a modified Turkevich method have a high activity for this reaction, which then becomes a pseudo-homogenous catalysis. It was found that nanoparticles in the magnitude of around 10 nm could be produced in higher dilution and citrate concentrations as by the classic method. Under mild reaction conditions, glucose could be oxidized in good yields (72%) and the resulting sodium gluconate could be isolated by column chromatography with iso-propanol on silica gel. Since the underlying reaction mechanism for the glucose oxidation is most likely based on the formation of Au-glucose intermediates, four possible reaction intermediates of Au-glucose complexes were examined based on DFT methods and the calculations showed that the GlcOH molecule gives a stronger adsorption to the gold atom compared to the Glc molecule, emphasizing the requirement of alkaline conditions in the glucose oxidation. Based on the present calculations, the glucose adsorption on gold nanoparticles during the formation of Au-glucose intermediate was explained for the first time, demonstrating that the aggregation of the alkoxy-group to a gold-cluster is indeed thermodynamically favored.

ACKNOWLEDGMENTS

The authors would like thank Ms. Benjawan Auanwijit and Ms. Chitchanok Manian for their help with the experiments. We thank the National e-Science Infrastructure

Consortium for providing computing resources that have contributed to the research results reported within this paper (URL: http://www.e-science.in.th). J.J. is a Thailand Research Fund–Commission on Higher Education (TRF-CHE) Young Research Scholar (MRG5680143), and thanks the University of Phayao for financial support (R020056216016).

REFERENCES

[1] Astruc D., Lu F. and Aranzaes J.R., Nanoparticles as recyclable catalysts: the frontier between homogeneous and heterogeneous catalysis, *Angew. Chem. Int. Ed.*, 2005; **44**: 7852-7872.

[2] Haruta M. and Masakazu D., Advances in the catalysis of Au nanoparticles, *Appl. Catal. A*, 2001; **222**: 427-437.

[3] Haruta M., Gold as a novel catalyst in the 21st century: Preparation, working mechanism and applications, *Gold Bulletin*, 2004; **37**: 27-36.

[4] Hutchings G.J., Catalysis by gold, *Catal. Today*, 2005; **100**: 55-61.

[5] Thompson D.T., Using gold nanoparticles for catalysis, *Nano Today*, 2007; **2**: 40-43.

[6] Liu A. and Ye B., Application of gold nanoparticles in biomedical researches and diagnosis, *Clin. Lab*, 2013; **59**: 23-36.

[7] Haruta M., Yamada N., Kobayashi T. and Iijima S., Gold catalysts prepared by coprecipitation for low-temperature oxidation of hydrogen and of carbon monoxide, *J. Catal.*, 1989; **115**: 301-309.

[8] Hvolbæk B., Janssens T.V.W., Clausen B.S., Falsig H., Christensen C.H. and Nørskov J.K., Catalytic activity of Au nanoparticles, *Nano Today*, 2007; **2**: 14-18.

[9] Mikami Y., Dhakshinamoorthy A., Alvaro M. and Garcia H., Catalytic activity of unsupported gold nanoparticles, *Catal. Sci. Technol.*, 2013; **3**: 58-69.

[10] Comotti M., Della Pina C., Matarrese R. and Rossi M., The catalytic activity of "naked" gold particles, *Angew. Chem. Int. Ed.*, 2004; **43**: 5812-5815.

[11] Prüβe U., Heidinger S. and Baatz C., Catalytic conversion of renewables: kinetic and mechanistic aspects of the gold-catalyzed liquid-phase glucose oxidation, *Agric. Forestry Res.*, 2011; **3**: 261-272.

[12] Comotti M., Della Pina C., Falletta E. and Rossi M., Aerobic oxidation of glucose with gold catalyst: Hydrogen peroxide as intermediate and reagent, *Adv. Synth. Catal.*, 2006; **348**: 313-316.

[13] Beltrame P., Comotti M., Della Pina C. and Rossi M., Aerobic oxidation of glucose: II. Catalysis by colloidal gold, *Appl. Catal. A*, 2006; **297**: 1-7.

[14] Önal Y., Schimpf S. and Claus P., Structure sensitivity and kinetics of d-glucose oxidation to d-gluconic acid over carbon-supported gold catalysts, *J. Catal.*, 2004; **223**: 122-133.

[15] Biella S., Prati L. and Rossi M., Selective oxidation of D-glucose on gold catalyst, *J. Catal.*, 2002; **206**: 242-247.

[16] Ramachandran S., Fontanille P., Pandey A. and Larroche C., Gluconic acid: properties, applications and microbial production, *Food Technol. Biotechnol.*, 2006; **44**: 185-195.

[17] Ajandouz E.H., Tchiakpe L.S., Ore F.D., Benajiba A. and Puigserver A., Effects of pH on caramelization and maillard reaction kinetics in fructose-lysine model systems, *J. Food Sci.*, 2001; **66**: 926-931.

[18] Kimling J., Maier M., Okenve B., Kotaidis V., Ballot H. and Plech A., Turkevich method for gold nanoparticle synthesis revisited, *J. Phys. Chem. B*, 2006; **110**: 15700-15707.

[19] Turkevich J., Stevenson P.C. and Hillier J., A study of the nucleation and growth processes in the synthesis of colloidal gold, *Discuss Faraday Soc.*, 1951; **11**: 55-75.

[20] Sajan D., Bena Jothy V., Kuruvilla T. and Hubert Joe I., NIR-FT Raman, FT-IR and surface-enhanced Raman scattering and DFT based theoretical studies on the adsorption behaviour of (S)-phenylsuccinic acid on silver nanoparticles, *J. Chem. Sci.*, 2010; **122**: 511-519.

[21] Maniu D., Chis V., Baia M., Toderas F. and Astilean S., Density functional theory investigation of p-aminothiophenol molecules adsorbed on gold nanoparticles, *J. Optoelectron. Adv. Mater.*, 2007; **9**: 733-736.

[22] Ma W. and Fang Y., Experimental (SERS) and theoretical (DFT) studies on the adsorption of p-, m-, and o-nitroaniline on gold nanoparticles, *J. Colloid Interf. Sci.*, 2006; **303**: 1-8.

[23] Ma W. and Fang Y., Experimental (FT-IR) and theoretical (DFT) studies on the adsorption behavior of p-nitroaniline (PNA) on gold nanoparticales, *J. Nanopart. Res.*, 2006; **8**: 761-767.

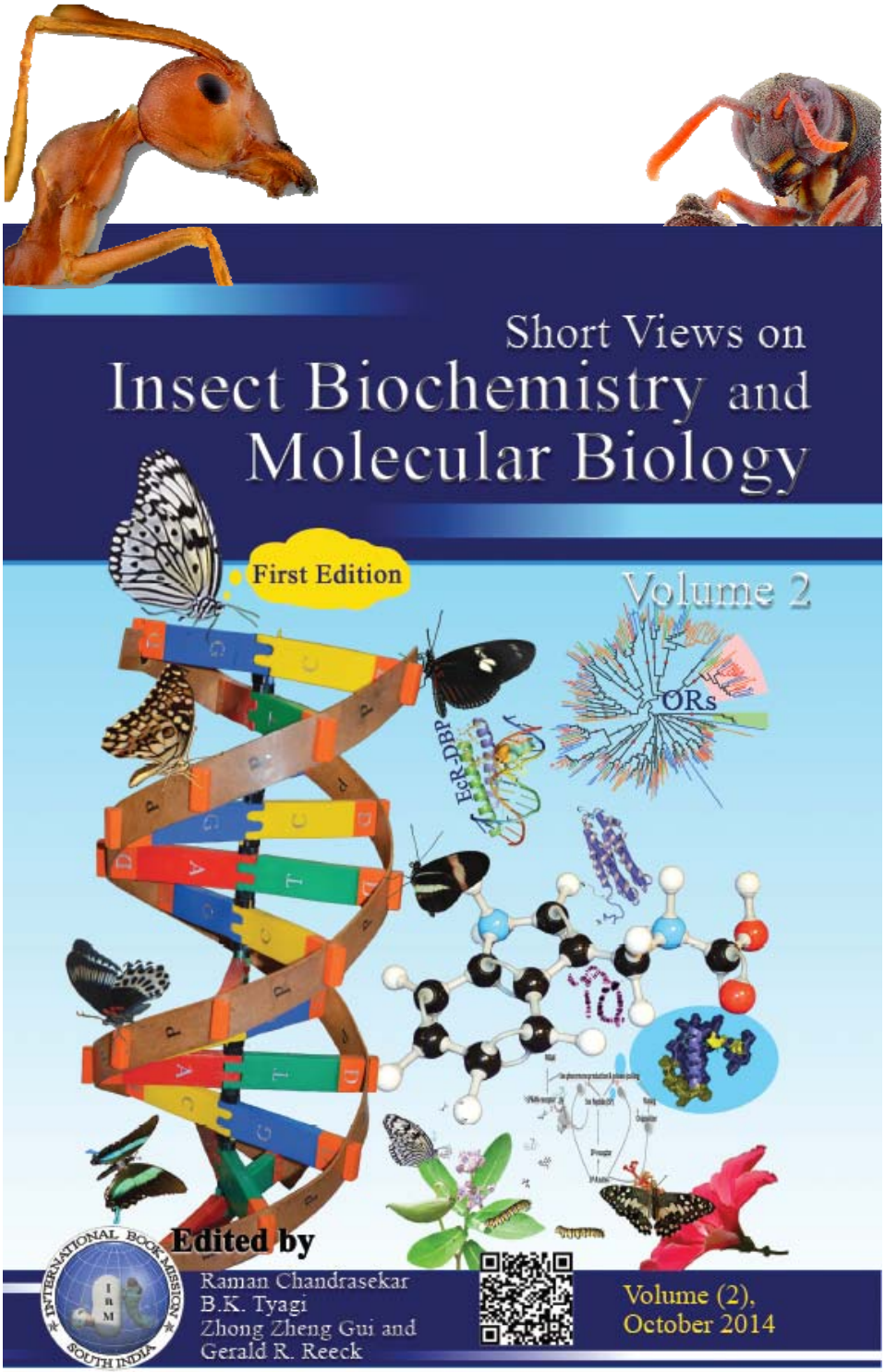
[24] Pérez L.A., López-Lozano X. and Garzóna I.L., Density functional study of the cysteine adsorption on Au nanoclusters, *Eur. Phys. J. D*, 2009; **52**: 123-126.

[25] Hay P.J. and Wadt W.R., Ab initio effective core potentials for molecular calculations. Potentials for the transition metal atoms Sc to Hg, *J. Chem. Phys.*, 1985; **82**: 270-283.

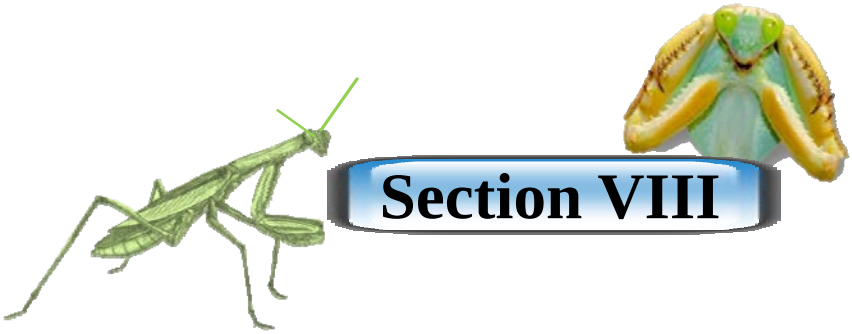
[26] Gaussian 09, Revision B.01, Gaussian Inc., Wallingford CT, USA, 2010.

[27] McFarland, A.D., Haynes, C. L., Mirkin, C.A., Van Duyne, R.P. and Godwin, H.A., Color my nanoworld, *J. Chem. Educ.*, 2004; **81**: 544a.

[28] Haiss W., Thanh N.T.K, Aveyard J. and Fernig D.G., Determination of size and concentration of gold nanoparticles from UV-VIS spectra, *Anal. Chem.*, 2007; **79**: 4215-4221.



Printed in the Unitated States of America, 2014
ISBN : 978-1-63315-205-2



Section VIII

Insect Bioinformatics

Printed in the Unitated States of America, 2014
ISBN : 978-1-63315-205-2

Short Views on Insect Biochemistry and Molecular Biology *Invited Review*
© Vol. (2) 685 – 707, 2014

Chapter –30

Computer-aided pesticide design: A short review

Jitrayut Jitonnom

Division of Chemistry, School of Science, University of Phayao, Phayao 56000, Thailand.

Abstract

With a large resistance globally to available insecticides used today, a new, effective, safer strategy for insect control is highly demand for agrochemical industry. For this reason, a rational approach that could be used to produce more selective and harmless insecticides is urgently needed. Computer-aided molecular design has been shown for many years to play essential roles in the field of insecticide discovery, providing useful information that could guide the design and development of new and better insecticidal agents. A short overview of recent applications of computational methods on the rational design of insecticides is presented through several insecticide-targeting receptors, including acetylcholinesterase, voltage-gated sodium channel, cytochrome P450 monooxygenase and nicotinic acetylcholine receptor.

***Key words*:** insecticides, docking, 3D-QSAR, acetylcholinesterase, voltage-gated sodium channels, cytochrome P450 monooxygenases, nicotinic acetylcholine receptor, computer-aided molecular design

**For Correspondence* (*email:* jitrayut.018@gmail.com)

1. Introduction

With the human suffering and financial losses worldwide caused by insect pests, new and environmentally safe insecticides are urgently needed to cope with these serious problems. Insecticides are commonly used as chemical agents (e.g., organophosphate or carbamate) or biological origin (e.g., biological controls or plant derived inhibitors) to control insects. Controlling the insect may occur from killing the insect or otherwise preventing it from engaging in behaviors deemed destructive. They

Printed in the United States of America, 2014
ISBN: 978-1-63315-205-2

Overview

have been used for a long time as a major strategy of choices for insect control in agricultural fields. However, the use of insecticide can cause not only the ecological damage to non-target organisms but also resistance to other non-target insects. This results in a large increase of insecticide resistance in many insects and other organisms, causing a serious global problem of insect control in agriculture (1,2). DDT and related organochlorines are, for example, the best-known insecticides with massive use, causing a great impact to human and environmental health. Consequently, these insecticides were banned in 1970s and have progressively been replaced by alternative, more specific and less toxic chemicals partly because of the emergence of insecticide resistance in the target species (3). Currently, over 500 species of insects and mites have been reported to develop resistance to a wide range of insecticides (4). Many of these cases involve two major mechanisms of insecticide resistance i.e., either modification of target sites and/or enhancement of metabolic and detoxification processes. For target site resistance, the resistance can occur from direct changes to the protein active sites that normally bind to the insecticides. Modification of this binding site can lead to insensitivity toward a variety of insecticides. For example, mutations in acetylcholinesterase, an important enzyme target to organophosphates and carbamates, have been well documented in many insect species (1-3). Other similar examples of target site modifications have been reported for cases of knockdown resistance to pyrethroids, DDT resistance through reduced sodium channel sensitivity (5,6), resistance to spinosad and neonicotinoids through modification of nicotinic acetylcholine receptors (7-10) and modification of the γ-aminobutyric acid-regulated chloride channel leading to resistance to dieldrin and fipronil (11,12). For metabolic resistance, this resistance involves the enhanced detoxification. There are three groups of enzymes, the cytochrome P450, the carboxylesterases and the glutathione-S transferases that have been identified as potentially being largely responsible for this metabolic resistance (13). Much effort has been directed to understand the resistance mechanism of these enzymes (1,14-17).

Since increasing numbers of insects have evolved resistance to a wide range of pesticides, greater efforts are thus being directed towards the development of new insecticides with the ability to minimize resistance and preserve the utility of

Printed in the United States of America, 2014
ISBN: 978-1-63315-205-2

insecticides. More importantly, the newly designed insecticides should have optimal insecticide activities with minimal toxicity to humans and environment.

In the past, designing insecticides as lead discovery is often difficult with random hit and thus a better and efficient approach in lead screening is needed. At present, it is clear that this challenging task become as a routine job with an increase of hit rate. The change of such history may be appreciated by the highly development of combinatorial chemistry (18) and high-throughput screening (19). This advance progress, however, is still limited in industrial company where the useful data is hardly accessible and does not provide as public data (20).

Computer-aided molecular design is a rational approach that is often used in drug design (21-23) as an essential tool for lead was screening and optimization and the design of new potent inhibitors including insecticidal agents (24). This approach has increasingly been applied during the past years as a result of the current advance of biochemistry and structural biology. It has also been expected to be more powerful for development of better and effective insecticides when it incorporates with the development of a global mechanism of action model including high-throughput screening and/or a predictive quantitative structure–activity relationship model (Fig.1). Considering a rapid rate of lead hitting of new natural and synthetic compounds with high insecticidal activity and low toxicity (25-32) and the advance of the computational methods, one might expect a dramatic growth for a number of insecticides in the market with new, presumably safer, modes of action in the future.

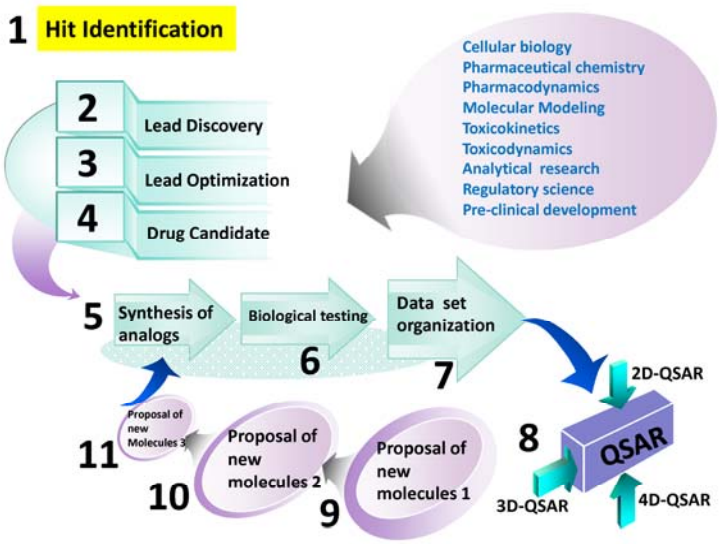


Fig.1 A schematic diagram illustrating the rational drug design process.

Printed in the United States of America, 2014
ISBN: 978-1-63315-205-2

This short review will provide some recent applications of computational methods in guiding the design and development of new insecticidal agents. In this context, four insect enzyme targets including acetylcholinesterase, voltage-gated sodium channel, cytochrome P450 monooxygenases and nicotinic acetylcholine receptor will be described as case studies, highlighting the growing field of computer-aided insecticide discovery.

2. Computational Methods

With a rapidly growing field of molecular biology and structural biology, computer-aided molecular design become as a promising rational approach in the field of agrochemistry (33,34), allowing one to rationale the experimental data and guide the design of new and better insecticides. Computer-aided molecular design of insecticide is sometimes referred as structure-based design (SBD) of insecticide (20,35). This SBD method has been successfully used in many insecticides based on protein structure of insect receptors. When the structural data is limited, ligand-based approaches (36) become an essential tool for insecticide design. The ligand-based approaches include quantitative structure-activity relationships (QSAR) and three-dimensional quantitative structure-activity relationships (3D-QSAR), including comparative molecular field analysis (CoMFA) and comparative molecular similarity indices analysis (CoMSIA). Both structure and ligand- based methods are a powerful tool that has been used nowadays in the field of computer-aided discovery of pesticides and insecticides to understand the structure-activity relationships of various insecticide compounds (20,35-39). They also provide a structural basis for the insecticide binding and resistance, which can guide the rational design of new and better insecticidal agents. More information on the methodology of computer-aided pesticide design can be found elsewhere (21-24,36-41). A short overview of those methods is described as following.

Homology modeling is a useful technique for protein structure prediction (42,43). Due to limited structural information of insect protein, this approach is often used as a primary tool for understanding the structural basis of insect enzyme. The predicted structure is based on sequence-structure alignment of a target sequence with a set of homologous sequence in which its 3D structure is known and will be used as template(s) (44). In insect, the templates are usually taken from available mammalian or other related species.

Molecular docking is a molecular modeling technique (45) that is basically used to describe the interactions between receptor and ligand when they are in a complex form. The strength or binding affinity between the receptor and the ligand is evaluated via the scoring functions (46). Several protein-ligand docking software with efficient searching algorithms are available, such as DOCK (47), GOLD (48) or AutoDoCK (49,50). The docking technique has now become a standard tool in the discovery of drug and insecticide. For insecticide design, this method is used to gain insight into the

Printed in the United States of America, 2014
ISBN: 978-1-63315-205-2

structural basis of insecticide binding and resistance as well as an identification of key amino acid residues responsible for the binding.

Molecular dynamics (MD) simulations provide more insights into the dynamics and structural properties of ligand-receptor complex in solution (51,52). It can be used together with molecular docking to evaluate the thermal and structural stability of the docked complex (53). It can also allow one to estimate the binding affinity of insecticide in target active site through the use of pairwise interaction energies approaches, such as MM-PBSA/MM-GBSA (54). The reliability of the MD simulation depends mainly on the quality of force fields. Many force fields have been developed for biomolecular simulations and implemented in various software packages, such as AMBER (55), CHARMM (36) and GROMACS (56). The advantage of MD simulation is that it could predict the ligand-receptor interactions extremely well within the protein receptors whose binding sites are highly flexible.

Quantitative Structure-Activity Relationship (QSAR) models are standard approaches for rational design and development of new insecticides since they are efficient in rapid prediction and virtual pre-screening of insecticidal activity (38, 57,58). There are many examples available in the literature in which QSAR models have been applied successfully for the screening of compounds for insecticidal activity (59-62). In principle, a good QSAR equation is required to correlate the biological activity (most often expressed by logarithms of equipotent molar activities) to a wide variety of parameters, including structural elements (Free Wilson analysis), physicochemical properties (Hansch analysis), or fields (3D-QSAR). The 3D-QSAR methods have been widely used since they require the 3D conformers of the molecule that were introduced by Cramer called Comparative Molecular Field Analysis (CoMFA) (58). Further developments on CoMFA lead to other alignment-dependent based methods such as Comparative Molecular Similarity Indices Analysis (CoMSIA) (57). The interpretation of these 3D-QSAR models are graphically described in terms of steric, electrostatic, hydrophobic, and hydrogen bond donor and acceptor molecular field interaction contributions. CoMFA and CoMSIA often provide a practical solution to an otherwise intractable problem of proper characterization of ligand–receptor interactions when a 3D structure of the receptor protein is absent. The QSAR models have also been developed together with the pharmacokinetic and pharmacodynamic (PBPK/PD) models and used as predictive models for human risk assessment to several insecticides such as pyrethroid (63), carbamate (64) and organophosphorous insecticides (68).

3. Recent Applications of Computational Methods in Insecticide Design

In order to demonstrate the impact of current computational methods on the in silico design of insecticides, four attractive targets specific to various insecticides including acetylcholinesterase, voltage-gated sodium channel, cytochrome P450

Printed in the United States of America, 2014
ISBN: 978-1-63315-205-2

monooxygenase and nicotinic acetylcholine receptors were described by some selected published works.

3.1 Acetylcholinesterase

Acetylcholinesterase (AChE) is a serine hydrolase vital for regulating the neurotransmitter acetylcholine in mammals, birds, and insects. This enzyme has a deep and narrow active-site gorge with a catalytic site at the bottom and a peripheral site at the entrance (see Fig. 2). AChE is a key enzyme of the cholinergic system because it regulates the level of acetylcholine and terminates nerve impulses by catalyzing the hydrolysis of acetylcholine, a neurotransmitter in many organisms including insects and humans. The significance of AChE in nerve system makes it an efficient target in developing pesticides. Its inhibition by interference of the building up of acetylcholine causes death. Organophosphate and carbamate compounds are current AChE insecticides that are believed to exert their activity by irreversible inhibition. Despite the proven efficacy of these insecticides, resistance-modified AChEs are still shown in many insect species (66-71), which lead to insensitivity of AChEs toward a variety of organophosphate and carbamate based compounds (72-76). Unfortunately, these insecticides, which contaminate air, water, soil, and food, are toxic not only to insects, but also to people and other animals. This is because the insecticides target a catalytic serine residue of AChE which can be found in both insects and mammals. Based on this

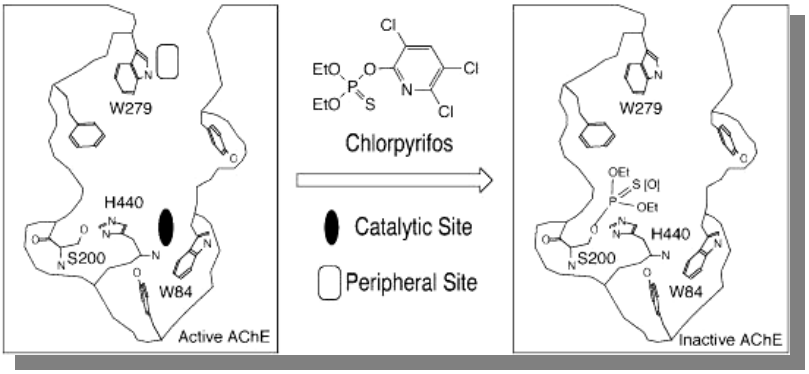


Fig. 2. Cartoon representation of the active-site gorge of AChE showing the locations of the peripheral and catalytic sites and conjugation of chlorpyrifos with the catalytic serine residue. Adapted from Pang et al. (80).

issue, many computational studies aim at designing and developing more specific and selective inhibitors/insecticidal agents (77) based on the SBD rational approach. For example, Doucet-Personeni et al. (78) described an SBD approach to design a novel

Printed in the United States of America, 2014
ISBN: 978-1-63315-205-2

reversible inhibitor with a view to developing insecticides with an improved selectivity for target species. They designed such inhibitor by combining the binding features of two previously known classes of AChE inhibitors i.e., tacrine and trifluoromethyl ketone as shown Fig. 3).

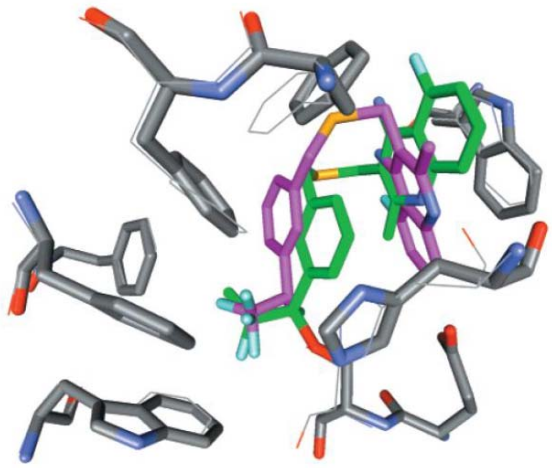


Fig. 3. A. close overlap between the experimentally observed (green carbon atoms) and predicted (purple carbon atoms) binding modes of the novel, reversible AChE inhibitor as designed by SBD approach. Adapted from Doucet-Personeni et al. (78).

Recently, Yuan-Ping Pang and his research groups (79,80) at the Mayo Clinic's Computer-Aided Molecular Design Laboratory in USA has discovered a novel and viable target site of AChEs through the use of computer-aided pesticide design, which leads to the new design and development of effective and environmentally safe insecticides. The new target site is a cysteine residue that presents at the active sites of greenbug and aphid AChEs but absents at those of mammalian AChEs and this residue could serve as a direct target for a new insecticide that would not affect humans and animals. In particular, Pang created 3D computer models based on genomic information of AChEs obtained from the greenbug (Schizaphis graminum) and the English grain aphid (Sitobion avenae) and discovered the pest-specific cysteine residue (C286) located at the opening of the AChE active site as shown in Fig. 4. He also carried his blueprint work with the Aphid AChEs (81) and the mosquito AChEs (82), providing a structural basis of insecticide binding that could be very useful for pest management control.

Rowland and co-workers (83) have utilized computer modeling together with the biochemical experiment to investigate the inactivation of aphid AChE by two sulfhydryl reagents, 5,5'-dithiobis(2-nitrobenzoic acid) and N-ethylmaleimide. Several

Printed in the United States of America, 2014
ISBN: 978-1-63315-205-2

mutants were biochemically generated based on two key cysteines (Cys310 and Cys466, a cysteine homologous to Cys310) which were expected by the authors to be involved in inactivation by the sulfhydryl reagents. A computer model was used to explain different activities observed experimentally and revealed that the cysteine near the acyl pocket (Cys466) was more accessible. The authors also speculated that this inactivation might be due to the presence of Cys466 and proposed that this residue could be a target for a specific insecticide.

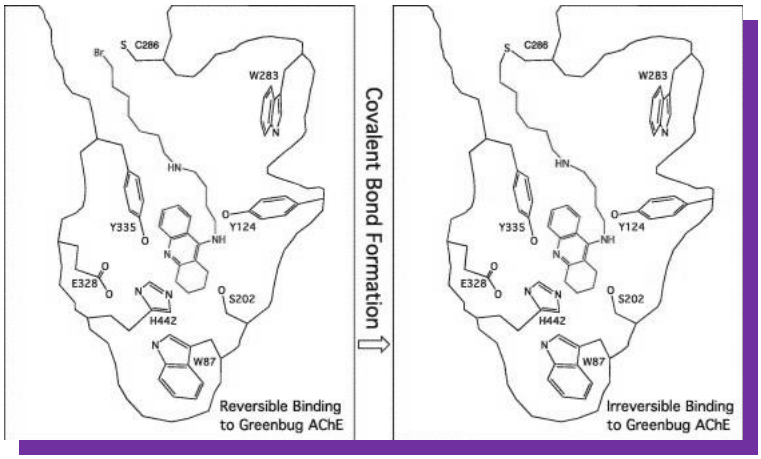


Fig. 4. Cartoon representation of the Schizaphis graminum AChE that is covalently bonded to an inhibitor upon binding to the active site. Adapted from Pang, (79).

Understanding the functional roles of AChE genes could be another important step for future insecticide discovery. While two AChE genes can be found in most insect species (84), their functional roles, on the other hand, are less known. A recent study by Lu and co-workers (84) has successfully discriminated a significant role of these two genes from the red flour beetle (*Tribolium castaneum*; TcAce1 and TcAce2) by completely deduced the genes and computationally predicted their 3D protein models by using homology modeling and MD simulations. The corresponding 3D protein models suggested that the TcAce1 protein is a robust acetylcholine (ACh) hydrolase and has susceptibility to sulfhydryl agents whereas the TcAce2 protein is not a catalytically efficient ACh hydrolase. These data provide researchers a better understand for the gene-function relationship of AChE enzyme and a promising strategy for future AChE-specific insecticide.

Printed in the United States of America, 2014
ISBN: 978-1-63315-205-2

3.2 Voltage-gated sodium channels

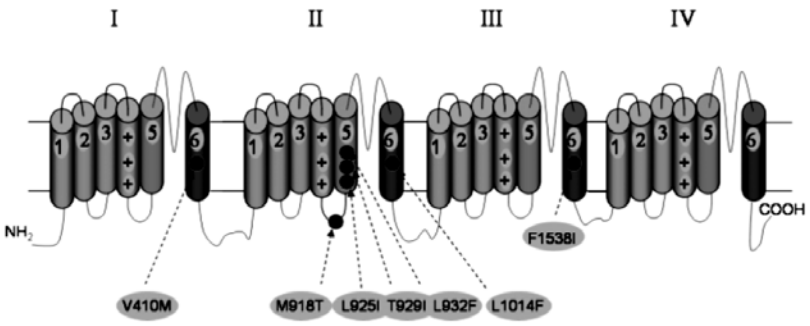


Fig. 5. A schematic representation of a Nav channel. The pore-forming α-subunit consists of a single polypeptide chain with four internally homologous domains (I–IV), each having six transmembrane helices S1–S6. The domains assemble to form a central aqueous pore, lined by the S5 and S6 helices and the S5–S6 linkers (P-loops). The S1–S4 helices of each domain assemble to form four physically discrete voltage sensors. The voltage dependence of channel activation is thought to derive from the movements of the four positively charged (+) S4 segments. The identity and location of some of the mutations associated with kdr resistance are shown (●), with residues numbered according to the sequence of the housefly (*Musca domestica*) Nav channel (EMBL accession: X96668). Adapted from Davies et al. (88).

Voltage-gated sodium (Nav) channels are pore-forming membrane proteins responsible for generation and propagation of action potentials in neurons and other excitable cells. They are well-known as the primary target of DDT and pyrethroid insecticides (5,85). These insecticides are so called “open channel blockers” affecting the Nav channel by slowing both activation and inactivation, which leads to hyperexcitability and death. Pyrethroid insecticides, which constitute a major class of insecticides used worldwide, bind to voltage-sensitive sodium channels and modify their gating kinetics, thereby disrupting nerve function (86). The pyrethroid binding site on the Nav channel has not been defined at the molecular level and remains a major unresolved issue in sodium channel pharmacology. To date, more than half a dozen mutations have been demonstrated to reduce the sensitivity of insect Nav channels to pyrethroids and a knockdown resistance (kdr) is concerned as a major mechanism of the resistance caused by point mutations in the Nav channel gene (6,87-90). Most resistance-conferring mutations were found in domain II and III of the transmembrane protein, particularly in the region comprised between transmembrane segments 4 and 6 (IIS4–IIS6 region) as shown in Fig. 5.

Since the existence of resistance to pyrethroid and DDT insecticides, the structural information of Nav channels are of particular importance for computer-aided insecticide design (91). Many efforts have been made to understand the binding

Printed in the United States of America, 2014
ISBN: 978-1-63315-205-2

structure of insecticide inhibitors and their corresponding mode of action through the use of computer modeling. Reilly and co-workers attempted to build a housefly Nav channel model in an open conformation (92) based on available X-ray structures of potassium channels (Kv). Their predicted model provides for the first time for the location of binding sites for several pyrethroid insecticides including DDT, highlighting the role of IIS4–S5 linker and the IIS5 and IIIS6 helices and the involvement of M918, T929 and F1538 in pyrethroid binding (Fig. 6).

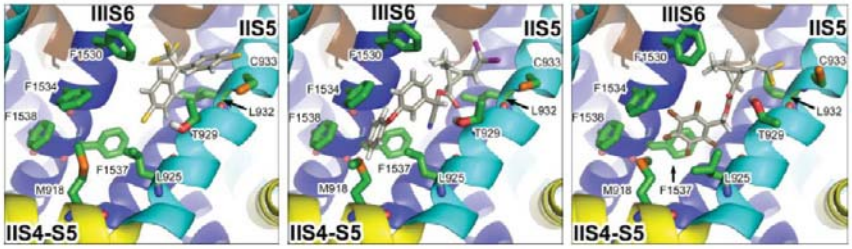


Fig. 6 Docking predictions for DDT (left-hand panel), deltamethrin (middle panel) and fenfluthrin (right-hand panel) with the voltage-gated sodium channel (insecticide structures shown in stick format). The IIS4–IIS6 regions are shown in cartoon (yellow, cyan, brown and blue respectively). Residues implicated in pyrethroid binding in various pest species (M918, L925, T929 and L932) are shown in green stick format. Adapted from Davies et al. (88).

The model also predicts that several additional residues in IIS5 and IIIS6 could be part of the pyrethroid binding site and might contribute to the different sensitivities between insect and mammalian Nav channels. Further analysis of twenty residues in IIIS6 of the cockroach Nav channel essential for the action of pyrethroid insecticides has been performed by using alanine-scanning mutagenesis and computer modeling (93). The experiment identified four residues, Ile1514, Gly1516, Phe1518 and Asn1522, involved the Nav channel insensitivity to pyrethroid insecticides. The Kv1.2-based homology model of the open Nav channel revealed that side chains of Ile1514, Phe1518 and Asn1522 are exposed towards helix IIS5 and linker IIS4-IIS5, which contain previously identified pyrethroid-interacting residues, whereas Ser1517 and Leu1521 face the inner pore where the batrachotoxin (BTX) receptor is located. Further work by this research group has also been done on the same target for other insecticides including BTX, deltamethrin, and BTG 502 (an alkylamide insecticide) and the computer model revealed the overlapping binding sites of those insecticides (94).

3.3 Cytochrome P450 monooxygenases

Cytochrome P450 (CYP) monooxygenases are detoxifying enzymes important in pyrethroid metabolism. They have been implicated in insecticide resistance in many

Printed in the United States of America, 2014
ISBN: 978-1-63315-205-2

insects (95-97). Pyrethroids are synthetic substances based on compounds of the natural insecticide pyrethrum in Tansy flowers (Tanacetum). They have been successfully applied in fruit, vegetable and crop farming for decades. In insects, little information is available about the CYP structure. To date, crystal structures of insect CYPs have not been resolved and structural studies relying on in silico homology modeling approaches have been used to gain insight into the molecular basis of insecticide binding (53,98-104).

Chiu and co-workers (100) used molecular modeling, as supported by enzyme characterization, in deciphering the catalytic site geometry of CYP6Z1 and CYP6Z2 regarding to the mechanism of DDT metabolism. In particular, DDT and two pyrethroid insecticides (carbaryl and xanthotoxin) were docked into the binding pockets of two homology models of CYP6Z1and CYP6Z2. The binding predictions indicate that these two proteins have different catalytic sites with CYP6Z1 predicted to metabolize DDT and CYP6Z2 predicted not to bind this insecticide. These data have suggested that CYP6Z1 is considered as potential target for the design of inhibitors capable of reducing *An. gambiae* resistance to DDT and other insecticides.

The neonicotinoid imidacloprid is one of the most important insecticides worldwide. It is the newest and fastest-growing class among the major insecticides of chlorinated hydrocarbons, organophosphorus compounds, methylcarbamates, and pyrethroids. It is used extensively against the whitefly *Bemisia tabaci* (Hemiptera: Aleyrodidae), an insect pest of eminent importance globally, which was also the first pest to develop high levels of resistance against imidacloprid and other neonicotinoids in the field. Recent reports indicated that, in both the B and Q biotypes of B. tabaci, the resistant phenotype is associated with over-expression of the CYP6CM1 gene. According to this observation, Karunker and co-workers aimed at examining the potential target of the biotype Q variant of the CYP6CM1 enzyme (CYP6CM1vQ) for insecticide design (53). Using molecular docking and MD simulations, important interactions (see Fig. 7) between imidacloprid and the CYP6CM1vQ could be revealed. This predicted binding mode was found to be consistent with the biochemical experiment.

Jones and co-workers (102) predicted the CYP6G1, CYP12D1 and CYP6A2 homology models, which can provide a structural insight into insecticide resistance in flies overexpressing CYP enzymes with broad substrate specificities. Homology models are presented for CYP6G1, a CYP associated with resistance to DDT and neonicotinoids, and two other enzymes associated with insecticide resistance in *D. melanogaster*, CYP12D1 and CYP6A2. The models are based on a template of the X-ray structure of the phylogenetically related human CYP3A4, which is known for its broad substrate specificity. The model of CYP6G1 has a much smaller active site cavity than the template. Comparison of the DDT-CYP6G1 complex and a non-resistant CYP6A2 homology model implies that tight-fit recognition of this insecticide is important in CYP6G1. The active site can accommodate differently

Printed in the United States of America, 2014
ISBN: 978-1-63315-205-2

shaped substrates ranging from imidacloprid to malathion but not the pyrethroids permethrin and cyfluthrin.

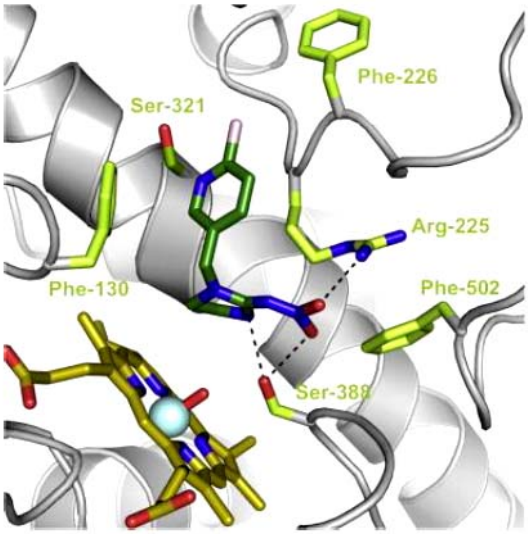


Fig. 7 A snapshot for the binding site model of the CYP6CM1vQ-imidacloprid complex taken from the MD simulations. Adaped from Karunker et al. (53).

Lertkiatmongkol and co-workers (103) have constructed homology models for CYP6AA3, CYP6P7, and CYP6P8 to understanding of molecular mechanisms underlying their binding sites toward insecticide substrates and inhibitors. They found differences in active site topologies among CYPAA3, CYP6P7, and CYP6P8 enzymes and suggested that the differences in metabolic activities among CYP enzymes in insects could be attributed to structural differences resulting in selectivity and different enzymatic activities against insecticides. Moreover, they also added a comment that information obtained using homology models has the potential to enhance the understanding of pyrethroid metabolism and detoxification mediated by CYP enzymes.

Stevenson and co-workers (104) have predicted the binding mode of deltamethrin and metabolites in CYP6M2 model. The availability of recombinant CYP6M2 and structural models provides an opportunity to identify the critical amino acid residues affecting substrate specificity and insecticide binding. For instance, their modeling studies indicate that two aromatic residues in particular, Phe 108 and Phe 121, may play a pi-stacking role in orientating deltamethrin in the active site. These residues are the subject of current mutagenesis studies.

Printed in the United States of America, 2014
ISBN: 978-1-63315-205-2

3.4 Nicotinic acetylcholine receptor

Nicotinic acetylcholine receptors (nAChRs) are major excitatory neurotransmitter receptors that are still one of the most attractive target sites for exploration in insecticide discovery (105). They are a target for insect-selective neonicotinoid insecticides (106). Because of the massive use of the neonicotinoid insecticides to control a variety of insect pest species (107), it is not surprising to observe evidence for the neonicotinoid resistance occurs by target site mutation (108). Current efforts (109-112) on computer-aided neonicotinoid design focus on understanding the source of diverse action and target-selectivity of the neonicotinoid insecticides, such as imidacoprid (113) and various computational approaches play essential roles in interpreting the biochemical results and in providing structural insight into the nature and the diversity of neonicotinoid-nAChR interactions, which will be useful for receptor structure-guided design of novel insecticidal compounds (114). An attempt to find an effective prediction pharmacophore model for neonicotinoid insecticides by using 3D-QSAR based on a series of nAChR agonists has been reported (115) and the obtained model can provide theoretical basis for designation and development of higher active insecticides.

Liu and co-workers (116) have used molecular modeling to understand selectivity of imidacloprid for fruit fly versus rat nAChRs. In particular, 3D models of fruit fly α1β2 and rat α4β2 nAChRs were generated by homology modeling, using the crystal structure of the acetylcholine-binding protein of Lymnaea stagnalis and the nAChR of Mus musculus as the templates, respectively. The conformational stability of the two models was studied by MD simulation and the quality of the models was confirmed. Especially, imidacloprid was docked into the putative binding site of the fruit fly α1β2 and rat α4β2 nAChRs by Surflex-docking. The calculated docking energies were in agreement with the experimental data and the putative binding sites were also consistent with the results from labeling and mutagenesis experiments. This research group has also applied the similar approach with γ-aminobutyric acid (GABA) receptors (117).

Li and co-workers (118) have successfully applied various computational approaches, together with available neonicotinoid database, to address the structural determinants for the novel potent inhibitors of Drosophila melanogaster nAChR (Dm-nAChR) and *Musca domestica* nAChR (Md-nAChR). In particular, 78 imidacloprid analogues of Dm-nAChR and Md-nAChR were collected and analyzed using 3D-QSAR models including COMFA and COMSIA. Homology modeling, MD simulation, and molecular docking were then used to probe the binding mode of these inhibitors of *Drosophila* and *Musca* containing the Y/S mutation associated with neonicotinoid resistance. The interaction of Dm-/Md-nAChR with neonicotinoid insecticides was explored and consistent with the 3D-QSAR analysis. Based on the computational analyses, the authors could be able to predict the essential features of

Printed in the United States of America, 2014
ISBN: 978-1-63315-205-2

imidacloprid-based nAChR inhibitors (Fig. 8), guiding the design of novel nAChR inhibitors with desired activities.

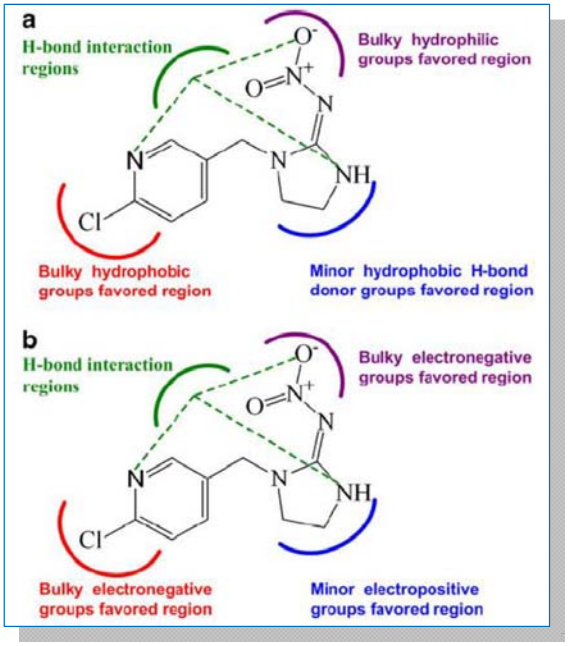


Fig. 8 Predicted binding feature of imidacloprid-based nAChR inhibitors toward Dm-nAChR (a) and Md-nAChR (b). Adapted from Li et al. (121).

Very recently, Cerón-Carrasco and co-workers (119) provide new insights for the key residues essential for the binding mode of imidacloprid by using ONIOM method (120). The calculations indicate the importance of Trp147 and Cys190–191, through weak CH···π interactions and both van der Waals and hydrogen-bond interactions, respectively. Furthermore, H-bonds between hydroxyl groups of both Ser189 and Tyr55 and the imidacloprid nitro group are pointed out. The participation of Ile118, whose main chain NH and carbonyl group are hydrogen-bonded with the imidacloprid pyridinic nitrogen through a water molecule, is characterized with the simulations indicating a significant contribution of this residue through van der Waals interactions. The contribution of a halogen-bond interaction between imidacloprid and acetylcholine binding protein, recently proposed in the literature, was confirmed.

Printed in the United States of America, 2014
ISBN: 978-1-63315-205-2

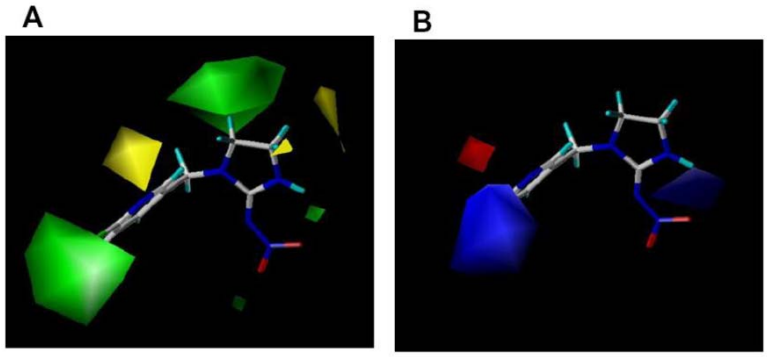


Fig. 9 Contour diagrams of steric (A) and electrostatic (B) fields with imidacloprid. In (A), the green and yellow areas indicate the steric-permissible and unfavorable regions for the receptor affinity, respectively. In (B), the red and blue areas indicate the regions where the more negative and positive electrostatic interactions with the receptor binding site increase the receptor affinity, respectively. Adapted from Nishiwaki et al. (121).

Recent studies have reported the combine use of simple synthesis and virtual pre-screening of series of neonicotinoid compounds of insect nAChR (121-123) and 3D-QSAR methods were used to guide the rational design of novel nAChR inhibitors. Nishiwaki and co-workers (124) present a series of imidacloprid derivatives with an alkylated imidazolidine ring via asymmetric synthesis and evaluate the insecticidal activity of those compounds against adult female housefly, Musca domestica nAChR. QSAR analysis of the receptor affinity demonstrated in Fig. 9 that the introduction of a substituent into the imidazolidine ring was fundamentally disadvantageous, but the introduction of a substituent at the R-5-position was permissible in the case of its small size. The binding model of the synthesized derivatives with the receptor supported the QSAR analysis, indicating the existence of space for a short alkyl group around the R-5-position in the ligand-binding site. Positive correlation was observed between the insecticidal activity and receptor affinity, suggesting that the receptor affinity was the primary factor in influencing the insecticidal activity. Other neonicotinoid derivatives have also been reported for virtual pre-screening and 3D-QSAR analysis (32). When combining new synthesis approaches with computer-guided neonicotinoid design, novel neonicotinoid insecticides with high affinities could be obtained, such as bis-neonicotinoid (125,126), where its unique binding mode and interaction is shown in Fig. 10.

5. Conclusion

Current applications of molecular modeling and computational methods on the rational design and development of insecticidal agents were emphasized in this review

Printed in the United States of America, 2014
ISBN: 978-1-63315-205-2

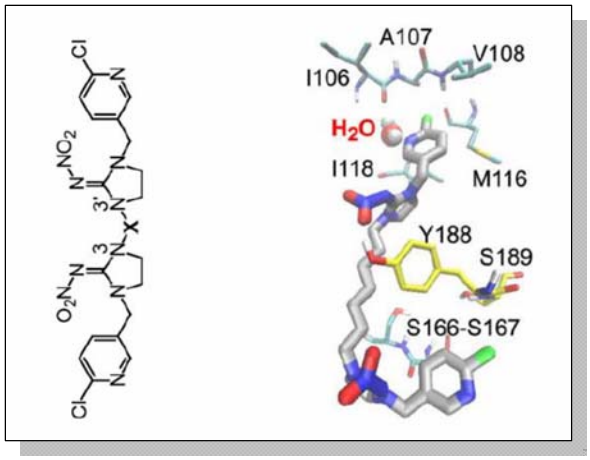


Fig. 10 A unique binding mode and interaction of novel bis-compound with acetylcholine-binding protein as obtained by molecular docking and MD simulation. Adapted from Ohno et al. (127).

by some recent published literatures. Computational methods such as docking and QSAR models are actively used nowadays and become a standard tool for insecticides design that can provide important insights into the nature of the compounds that act selectively and efficiently as potent inhibitors of the insecticide-targeting receptors, as exemplified by those of acetylcholinesterase, voltage-gated sodium channel, cytochrome P450 monooxygenase and nicotinic acetylcholine receptor. The advance of parallel synthesis methods and technology, the growing number of chemical and biological databases, and an explosion in currently available software tools are providing a much improved basis for the design of novel insecticides based on scaffold structure of known insecticides with desired activities and low toxicity.

6. Acknowledgement

This work is supported by grants from the University of Phayao (R020056216016) and the Thailand Research Fund (MRG5680143).

7. Reference

1. Brattsten, L.B., Holyoke, C.W.J., Leeper, J.R. and Raffa, K.F. (1986) Insecticide resistance: Challenge to pest management and basic research. *Science* 231: 1255-1260.
2. Hemingway, J., Field, L. and Vontas, J. (2002). An Overview of Insecticide Resistance. *Science* 298: 96-97.
3. Oakeshott, J.G., Home, I., Sutherland, T.D. and Russell, R.J. (2003) The genomics of insecticide resistance. *Genome Biol.* 4: 202.

Printed in the United States of America, 2014
ISBN: 978-1-63315-205-2

4. Georghiou, G.P. (1990) Managing resistance to agrochemicals. American Chemical Society, Washington DC.
5. Davies, T.G., Field, L.M., Usherwood, P.N. and Williamson, M.S. (2007) DDT, pyrethrins, pyrethroids and insect sodium channels. *IUBMB Life* 59: 151-162.
6. Dong, K. (2007) Insect sodium channels and insecticide resistance. *Invert. Neurosci* 7: 17-30.
7. Sattelle, D.B., Jones, A.K., Sattelle, B.M., Matsuda, K., Reenan, R. and Biggin, P.C. (2005) Edit, cut and paste in the nicotinic acetylcholine receptor gene family of Drosophila melanogaster. *BioEssays* 27: 366-376.
8. Jones, A.K., Raymond-Delpech, V., Thany, S.H., Gauthier, M. and Sattelle, D.B. (2006). The nicotinic acetylcholine receptor gene family of the honey bee, *Apis mellifera*. *Genome Res.* 16: 1422-1430.
9. Perry, T., McKenzie, J.A. and Batterham, P. (2007) A knockout strain of Drosophila melanogaster confers a high level of resistance to spinosad. *Insect Biochemistry and Molecular Biology* 37: 184-188.
10. Baxter, S.W., Chen, M., Dawson, A., Zhao, J.Z., Vogel, H., Shelton, A.M., Heckel, D.G. and Jiggins, C.D. (2010) Mis-spliced transcripts of nicotinic acetylcholine receptor alpha6 are associated with field evolved spinosad resistance in *Plutella xylostella* (L.). *PLoS Genet.* 6: e1000802.
11. Ffrench-Constant, R.H., Mortlock, D.P., Shaffer, C.D., MacIntyre, R.J. and Roush, R.T. (1991) Molecular cloning and transformation of cyclodiene resistance in Drosophila: an invertebrate gamma-aminobutyric acid subtype A receptor locus. *Proc. Natl. Acad. Sci. U S A* 88: 7209-7213.
12. Hosie, A.M., Baylis, H.A., Buckingham, S.D. and Sattelle, D.B. (1995) Actions of the insecticide fipronil, on dieldrin-sensitive and- resistant GABA receptors of Drosophila melanogaster. *Br. J. Pharmacol*. 115: 909-912.
13. Ranson, H., Claudianos, C., Ortelli, F., Abgrall, C., Hemingway, J., Sharakhova, M.V., Unger, M.F., Collins, F.H. and Feyereisen, R. (2002) Evolution of supergene families associated with insecticide resistance. *Science* 298: 179-181.
14. Scott, J.G. (1999) Cytochromes P450 and insecticide resistance. *Insect Biochem. Mol. Biol.* 29: 757-777.
15. Enayati, A.A., Ranson, H. and Hemingway, J. (2005) Insect glutathione transferases and insecticide resistance. *Insect Mol. Biol.* 14: 3-8.
16. Ahmad, M. (2007) Insecticide resistance mechanisms and their management in *Helicoverpa armigera* (HUBNER). *J. Agric. Res.* 45.
17. Joußen, N., Schuphan, I. and Schmidt, B. (2010) Metabolism of Methoxychlor by the P450-Monooxygenase CYP6G1 Involved in Insecticide Resistance of *Drosophila* melanogaster after Expression in Cell Cultures of Nicotiana tabacum. *Chem. Biodivers*. 7: 722-735.
18. Lindell, S.D., Pattenden, L.C. and Shannon, J. (2009) Combinatorial chemistry in the agrosciences. *Bioorg. Med. Chem.* 17: 4035-4046.
19. Tietjen, K., Drewes, M. and Stenzel, K. (2005) High Throughput Screening in Agrochemical Research. Comb. Chem. High Throughput Screen. 8: 589-594.
20. Walter, M.W. (2002). Structure-based design of agrochemicals. *Nat. Prod. Rep*. 19: 278-291.
21. Kapetanovic, I.M. (2008) Computer-aided drug discovery and development (CADDD): in silico-chemico-biological approach. *Chem. Biol. Interact.* 171: 165-176.
22. Zhang, S. (2011) Computer-aided drug discovery and development. Methods Mol Biol 716: 23-38.
23. Chu, H., Wang, J., Shen, H., Yang, Y., Zhu, W. and Li, G. (2012) Investigation of family 18 chitinases and inhibitors by computer-aided approaches. *Curr. Drug Targets* 13: 502-511.
24. Speck-Planche, A., Cordeiro, M.N., Guilarte-Montero, L. and Yera-Bueno, R. (2011) Current computational approaches towards the rational design of new insecticidal agents. *Curr. Comput. Aided Drug Des.* 7: 304-314.

Printed in the United States of America, 2014
ISBN: 978-1-63315-205-2

25. Xu, H. and Xiao, X. (2009) Natural products-based insecticidal agents 4. Semisynthesis and insecticidal activity of novel esters of 2-chloropodophyllotoxin against *Mythimna separata* Walker in vivo. *Bioorg. Med. Chem. Lett.* 19: 5415-5418.
26. Xu, H. and He, X.Q. (2010) Natural products-based insecticidal agents 6. Design, semisynthesis, and insecticidal activity of novel monomethyl phthalate derivatives of podophyllotoxin against *Mythimna separata* Walker in vivo. *Bioorg Med. Chem. Lett.* 20: 4503-4506.
27. Xu, H. and Wang, J.J. (2010) Natural products-based insecticidal agents 5. Design, semisynthesis and insecticidal activity of novel 4'-substituted benzenesulfonate derivatives of 4-deoxypodophyllotoxin against *Mythimna separata* Walker *in vivo*. *Bioorg. Med. Chem. Lett.* 20: 2500-2502.
28. Xu, H., Xiao, X. and Wang, Q.T. (2010) Natural products-based insecticidal agents 7. Semisynthesis and insecticidal activity of novel 4alpha-alkyloxy-2-chloropodophyllotoxin derivatives against *Mythimna separata* Walker in vivo. *Bioorg. Med. Chem. Lett.* 20: 5009-5012.
29. Xu, H. and Zhang, J.L. (2011) Natural products-based insecticidal agents 9. Design, semisynthesis and insecticidal activity of 28-acyloxy derivatives of toosendanin against *Mythimna separata* Walker in vivo. *Bioorg. Med. Chem. Lett.* 21: 1974-1977.
30. Xu, H. and Zhang, J.L. (2011) Natural products-based insecticidal agents 11. Synthesis and insecticidal activity of novel 4alpha-arylsulfonyloxybenzyloxy-2beta-chloropodophyllotoxin derivatives against *Mythimna separata* Walker *in vivo*. *Bioorg. Med. Chem. Lett.* 21: 5177-5180.
31. Dong, W., Xu, J., Xiong, L. and Li, Z. (2012) Synthesis, structure and insecticidal activities of some novel amides containing N-pyridylpyrazole moeities. *Molecules* 17: 10414-10428.
32. Su, W., Zhou, Y., Ma, Y.,Wang, L., Zhang, Z., Rui, C., Duan, H. and Qin, Z. (2012) N'-Nitro-2-hydrocarbylidenehydrazinecarboximidamides: design, synthesis, crystal structure, insecticidal activity, and structure-activity relationships. *J. Agric. Food. Chem.* 60: 5028-5034.
33. Scherkenbeck, J.r. (2009) Modern trends in agrochemistry. *Bioorg. Med. Chem.* 17: 4019.
34. Gong, J., Liu, X., Cao, X., Diao, Y., Gao, D., Li, H. and Qian, X. (2012) PTID: an integrated web resource and computational tool for agrochemical discovery. *Bioinformatics,* 29(2): 292-294.
35. Saini, V., Kumar, A. (2014) Computer Aided Pesticide Design: A Rational Tool for Supplementing DDT Lacunae. *Chemical Science Transactions* 3(2), 676-688.
36. Brooks, B.R., Brooks, C.L., Mackerell, A.D., Nilsson, L., Petrella, R.J., Roux, B., Won, Y., Archontis, G., Bartels, C., Boresch, S., Caflisch, A., Caves, L., Cui, Q., Dinner, A.R., Feig, M., Fischer, S., Gao, J., Hodoscek, M., Im, W., Kuczera, K., Lazaridis, T., Ma, J., Ovchinnikov, V., Paci, E., Pastor, R.W., Post, C.B., Pu, J.Z., Schaefer, M., Tidor, B., Venable, R.M., Woodcock, H.L., Wu, X., Yang, W., York, D.M. and Karplus, M. (2009) CHARMM: The biomolecular simulation program. *J. Comput. Chem.* 30: 1545-1614.
37. Benfenati, E. (2007) Quantitative Structure-Activity Relationships (QSAR) for Pesticide Regulatory Purposes Elsevier.
38. Naik, P.K., Sindhura, Singh, T. and Singh, H. (2009) Quantitative structure–activity relationship (QSAR) for insecticides: development of predictive in vivo insecticide activity models. *SAR QSAR Environ. Res.* 20: 551 – 566.
39. Verma, J., Khedkar, V.M. and Coutinho, E.C. (2010) 3D-QSAR in drug design - a review. *Curr. Top. Med. Chem.* 10: 95-115.
40. Comitani, F., Cohen, N., and Molteni, C. (2014) Insights into the binding of GABA to the insect RDL receptor from atomistic simulations: a comparison of models. *J.Comput Aided Mol. Des.* 28: 35-48.
41. Nakayama, A. (1990) Computer-Aided Molecular Design and Structure-Activity Relationships of Pesticides. *Journal of Pesticide Science*; ISSN:1348-589X; Vol. 15 (3): 481-494.
42. Marti-Renom, M.A., Stuart, A.C., Fiser, A., Sanchez, R., Melo, F. and Sali, A. (2000) Comparative protein structure modeling of genes and genomes. *Annu. Rev. Biophys. Biomol. Struct.* 29: 291-325.

Printed in the United States of America, 2014
ISBN: 978-1-63315-205-2

43. Baker, D. and Sali, A. (2001) Protein structure prediction and structural genomics. *Science* 294: 93-96.
44. Jitonnom, J., Lomthaisong, K. and Lee, V.S. (2012) Computational design of peptide inhibitor based on modifications of proregion from *Plutella xylostella* midgut trypsin. *Chem. Biol. Drug Des.* 79: 583-593.
45. Lengauer, T. and Rarey, M. (1996) Computational methods for biomolecular docking. *Curr. Opin. Struct. Biol.* 6: 402-406.
46. Jain, A.N. (2006) Scoring functions for protein-ligand docking. *Curr. Protein Pept. Sci.* 7: 407–420.
47. Moustakas, D., Lang, P.T., Pegg, S., Pettersen, E., Kuntz, I., Brooijmans, N. and Rizzo, R. (2006) Development and validation of a modular, extensible docking program: DOCK 5. *J. Computer-Aided Molecular Design* 20: 601-619.
48. Verdonk, M.L., Cole, J.C., Hartshorn, M.J., Murray, C.W. and Taylor, R.D. (2003) Improved protein–ligand docking using GOLD. Proteins: Structure, Function, and *Bioinformatics* 52: 609-623.
49. Goodsell, D.S., Morris, G.M. and Olson, A.J. (1996) Automated Docking of Flexible Ligands: Applications of AutoDock. *J. Mol. Recognition* 9.
50. Morris, G.M., Huey, R. and Olson, A.J. (2002) Using AutoDock for Ligand-Receptor Docking. *Current Protocols in Bioinformatics*, John Wiley & Sons, Inc.
51. Jitonnom, J. and Mulholland, A. (2012) Insights into conformational changes of procarboxypeptidase A and B from simulations: a plausible explanation for different intrinsic activity. *Theor. Chem. Acc.* 131: 1-13.
52. Jitonnom, J. and Sontag, C. (2012) Comparative study on activation mechanism of carboxypeptidase A1, A2 and B: first insights from steered molecular dynamics simulations. *J. Mol. Graph. Model.* 38: 298-303.
53. Karunker, I., Morou, E., Nikou, D., Nauen, R., Sertchook, R., Stevenson, B.J., Paine, M.J., Morin, S. and Vontas, J. (2009) Structural model and functional characterization of the *Bemisia tabaci* CYP6CM1vQ, a cytochrome P450 associated with high levels of imidacloprid resistance. *Insect Biochem. Mol. Biol.* 39: 697-706.
54. Lee, V.S., Tue-ngeun, P., Nangola, S., Kitidee, K., Jitonnom, J., Nimmanpipug, P., Jiranusornkul, S. and Tayapiwatana, C. (2010) Pairwise decomposition of residue interaction energies of single chain Fv with HIV-1 p17 epitope variants. *Mol. Immunol.* 47: 982-990.
55. Salomon-Ferrer, R., Case, D.A. and Walker, R.C. (2012) An overview of the Amber biomolecular simulation package. *WIREs Comput. Mol. Sci.* 3: 198-210.
56. Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A.E. and Berendsen, H.J.C. (2005) GROMACS: Fast, flexible, and free. *J. Comput. Chem.* 26: 1701-1718.
57. Kubinyi, H., Folkers, G., Martin, Y. and Klebe, G. (2002) Comparative Molecular Similarity Indices Analysis: CoMSIA. 3D QSAR in Drug Design, Springer Netherlands. 3: 87-104.
58. Cramer, R.D. (2003) Topomer CoMFA: a design methodology for rapid lead optimization. *J. Med. Chem.* 46: 374-388.
59. Ford, M.G., Hoare, N.E., Hudson, B.D., Nevell, T.G. and Banting, L. (2002) QSAR studies of the pyrethroid insecticides: Part 3. A putative pharmacophore derived using methodology based on molecular dynamics and hierarchical cluster analysis. *J. Molecular Graphics and Modelling* 21: 29-36.
60. Stoddard, S.V., Yu, X., Potter, P.M. and Wadkins, R.M. (2004) In Silico Design and Evaluation of Carboxylesterase Inhibitors. *J. Pest. Sci.* 35: 240-249.
61. Bermudez-Saldana, J.M. and Cronin, M.T. (2006) Quantitative structure-activity relationships for the toxicity of organophosphorus and carbamate pesticides to the Rainbow trout *Onchorhyncus mykiss*. *Pest. Manag. Sci*. 62: 819-831.
62. Sparks, T.C., Crouse, G.D., Dripps, J.E., Anzeveno, P., Martynow, J., Deamicis, C.V. and Gifford, J. (2008) Neural network-based QSAR and insecticide discovery: spinetoram. *J. Comput. Aided Mol. Des.* 22: 393-401.

Printed in the United States of America, 2014
ISBN: 978-1-63315-205-2

63. Knaak, J.B., Dary, C.C., Zhang, X., Gerlach, R.W., Tornero-Velez, R., Chang, D.T., Goldsmith, R. and Blancato, J.N. (2012) Parameters for pyrethroid insecticide QSAR and PBPK/PD models for human risk assessment. *Rev. Environ. Contam. Toxicol.* 219: 1-114.
64. Knaak, J.B., Dary, C.C., Okino, M.S., Power, F.W., Zhang, X., Thompson, C.B., Tornero-Velez, R. and Blancato, J.N. (2008) Parameters for Carbamate Pesticide QSAR and PBPK/PD Models for Human Risk Assessment. *Rev. Environ. Contam. Toxicol.* 193: 53-212.
65. Knaak, J.B., Dary, C.C., Power, F., Thompson, C.B. and Blancato, J.N. (2004) Physicochemical and biological data for the development of predictive organophosphorus pesticide QSARs and PBPK/PD models for human risk assessment. *Crit. Rev. Toxicol.* 34: 143-207.
66. Mutero, A., Pralavorio, M., Bride, J.M. and Fournier, D. (1994) Resistance-associated point mutations in insecticide insensitive acetylcholinesterase. *Proc. Natl. Acad. Sci. USA* 91: 5922 - 5926.
67. Villatte, F., Ziliani, P., Marcel, V., Menozzi, P. and Fournier, D. (2000) A high number of mutations in insect acetylcholinesterase may provide insecticide resistance. *Pest Biochem. Physiol.* 67: 95 - 102.
68. Kozaki, T., Shono, T., Tomita, T. and Kono, Y. (2001) Fenitroxon insensitive acetylcholinesterase of the housefly, *Musca domestica* associated with point mutations. *Insect Biochem. Mol. Biol.* 31: 991 - 997.
69. Walsh, S.B., Dolden, T.A., Moores, G.D., Kristensen, M., Lewis, T., Devonshire, A.L. and Williamson, M.S. (2001) Identification and characterization of mutations in housefly (*Musca domestica*) acetylcholinesterase involved in insecticide resistance. *Biochem. J.* 359: 175 - 181.
70. Nabeshima, T., Kozaki, T., Tomita, T. and Kono, Y. (2003) An amino acid substitution on the second acetylcholinesterase in the pirimicarb-resistant strains of the peach potato aphid, *Myzus persicae*. *Biochem. Biophys. Res. Commun.* 307: 15 - 22.
71. Menozzi, P., Shi, M., Lougarre, A., Tang, Z. and Fournier, D. (2004) Mutations of acetylcholinesterase which confer insecticide resistance in Drosophila melanogaster populations. *BMC Evolutionary Biology* 4: 4.
72. Fournier, D. and Mutero, A. (1994) Modification of acetylcholinesterase as a mechanism of resistance to insecticides. *Comp. Biochem. Physiol.* 108C: 19 - 31.
73. Zhu, K.Y., Lee, S.H. and Clark, J.M. (1996) A point mutation of acetylcholinesterase associated with azinphosmethyl resistance and reduced fitness in Colorado potato beetle. *Pestic. Biochem. Physiol.* 55: 100 - 108.
74. Devonshire, A.L., Byrne, F.J., Moores, G.D. and Williamson, M.S. (1998) Biochemical and molecular characterisation of insecticide-insensitive acetylcholinesterases. In Structure and Function of Cholinesterases and Related Proteins: 491 - 496.
75. Vontas, J.G., Hejazi, M.J., Hawkes, N.J., Cosmidis, N., Loukas, M. and Hemingway, J. (2002) Resistance-associated point mutations of organophosphate insensitive acetylcholinesterase, in the olive fruit fly *Bactrocera oleae*. *Insect Mol. Biol.* 11: 329 - 336.
76. Hsu, J.C., Haymer, D.S., Wu, W.J. and Feng, H.T. (2006) Mutations in the acetylcholinesterase gene of Bactrocera dorsalis associated with resistance to organophosphorus insecticides. *Insect Biochem. Mol. Biol.* 36: 396-402.
77. Lang, G.-J., Yan Zhu, K. and Zhang, C.-X. (2012) Can Acetylcholinesterase Serve as a Target for Developing More Selective Insecticides? *Curr. Drug Targets* 13: 495-501.
78. Doucet-Personeni, C., Bentley, P.D., Fletcher, R.J., Kinkaid, A., Kryger, G., Pirard, B., Taylor, A., Taylor, R., Taylor, J., Viner, R., Silman, I., Sussman, J.L., Greenblatt, H.M. and Lewis, T. (2001) A structure-based design approach to the development of novel, reversible AChE inhibitors. *J. Med. Chem.* 44: 3203-3215.
79. Pang, Y.-P. (2007) Species marker for developing novel and safe pesticides. *Bioorg. Med. Chem. Lett.* 17: 197-199.

Printed in the United States of America, 2014
ISBN: 978-1-63315-205-2

80. Pang, Y.P., Brimijoin, S., Ragsdale, D.W., Zhu, K.Y. and Suranyi, R. (2012) Novel and viable acetylcholinesterase target site for developing effective and environmentally safe insecticides. *Curr. Drug Targets* 13: 471-482.
81. Pang, Y.-P., Singh, S.K., Gao, Y., Lassiter, T.L., Mishra, R.K., Zhu, K.Y. and Brimijoin, S. (2009) Selective and Irreversible Inhibitors of Aphid Acetylcholinesterases: Steps Toward Human-Safe Insecticides. *PLOS One* 4: e4349.
82. Pang, Y.P., Ekstrom, F., Polsinelli, G.A., Gao, Y., Rana, S., Hua, D.H., Andersson, B., Andersson, P.O., Peng, L., Singh, S.K., Mishra, R.K., Zhu, K.Y., Fallon, A.M., Ragsdale, D.W. and Brimijoin, S. (2009) Selective and irreversible inhibitors of mosquito acetylcholinesterases for controlling malaria and other mosquito-borne diseases. *PLOS One* 4: e6851.
83. Rowland, M., Tsigelny, I., Wolfe, M. and Pezzementi, L. (2008) Inactivation of an invertebrate acetylcholinesterase by sulfhydryl reagents: a reconsideration of the implications for insecticide design. *Chem. Biol. Interact.* 175: 73-75.
84. Lu, Y., Pang, Y.-P., Park, Y., Gao, X., Yao, J., Zhang, X. and Zhu, K.Y. (2012) Genome Organization, Phylogenies, Expression Patterns, and Three-Dimensional Protein Models of Two Acetylcholinesterase Genes from the Red Flour Beetle. *PLOS One* 7: e32288.
85. Zlotkin, E. (1999) The insect voltage-gated sodium channel as target of insecticides. *Annu. Rev. Entomol.* 44: 429-455.
86. Soderlund, D.M., Clark, J.M., Sheets, L.P., Mullin, L.S., Piccirillo, V.J., Sargent, D., Stevens, J.T. and Weiner, M.L. (2002) Mechanisms of pyrethroid neurotoxicity: implications for cumulative risk assessment. *Toxicology* 171: 3-59.
87. Usherwood, P.N., Davies, T.G., Mellor, I.R., O'Reilly, A.O., Peng, F., Vais, H., Khambay, B.P., Field, L.M. and Williamson, M.S. (2007) Mutations in DIIS5 and the DIIS4-S5 linker of Drosophila melanogaster sodium channel define binding domains for pyrethroids and DDT. *FEBS Lett.* 581: 5485-5492.
88. Davies, T.E., O'Reilly, A.O., Field, L.M., Wallace, B. and Williamson, M.S. (2008) Knockdown resistance to DDT and pyrethroids: from target-site mutations to molecular modelling. *Pest Manag. Sci.* 64: 1126-1130.
89. Burton, M.J., Mellor, I.R., Duce, I.R., Davies, T.G.E., Field, L.M. and Williamson, M.S. (2011) Differential resistance of insect sodium channels with kdr mutations to deltamethrin, permethrin and DDT. *Insect Biochem. Mol. Biol.* 41: 723-732.
90. Karatolos, N., Gorman, K., Williamson, M.S. and Denholm, I. (2012) Mutations in the sodium channel associated with pyrethroid resistance in the greenhouse whitefly, *Trialeurodes vaporariorum*. *Pest Manag. Sci.* 68: 834-838.
91. Bloomquist, J.R. (1996) Ion Channels as Targets for Insecticides. *Annual Review of Entomology* 41: 163-190.
92. O'Reilly, A.O., Khambay, B.P., Williamson, M.S., Field, L.M., Wallace, B.A. and Davies, T.G. (2006) Modelling insecticide-binding sites in the voltage-gated sodium channel. *Biochem. J.* 396: 255-263.
93. Du, Y., Lee, J.E., Nomura, Y., Zhang, T., Zhorov, B.S. and Dong, K. (2009) Identification of a cluster of residues in transmembrane segment 6 of domain III of the cockroach sodium channel essential for the action of pyrethroid insecticides. *Biochem. J.* 419: 377-385.
94. Du, Y., Garden, D., Khambay, B., Zhorov, B.S. and Dong, K. (2011) Batrachotoxin, pyrethroids, and BTG 502 share overlapping binding sites on insect sodium channels. *Mol. Pharmacol.* 80: 426-433.
95. Feyereisen, R. (1999) Insect P450 enzymes. *Annu. Rev. Entomol.* 44: 507-533.
96. Joußen, N., Agnolet, S., Lorenz, S., Schöne, S.E., Ellinger, R., Schneider, B. and Heckel, D.G. (2012) Resistance of Australian *Helicoverpa armigera* to fenvalerate is due to the chimeric P450 enzyme CYP337B3. *Proc. Natl. Acad. Sci. USA* 109: 15206-15211.

Printed in the United States of America, 2014
ISBN: 978-1-63315-205-2

97. David, J.-P., Ismail, H.M., Chandor-Proust, A. and Paine, M.J.I. (2013) Role of cytochrome P450s in insecticide resistance: impact on the control of -borne diseases and use of insecticides on Earth. *Phil. Trans. R. Soc.* B 368.

98. Baudry, J., Li, W., Pan, L., Berenbaum, M.R. and Schuler, M.A. (2003) Molecular docking of substrates and inhibitors in the catalytic site of CYP6B1, an insect cytochrome P450 monooxygenase. *Protein Engineering* 16: 577-587.

99. Rupasinghe, S.G., Wen, Z., Chiu, T.L. and Schuler, M.A. (2007) Helicoverpa zea CYP6B8 and CYP321A1: different molecular solutions to the problem of metabolizing plant toxins and insecticides. *Protein Eng. Des. Sel.* 20: 615-624.

100. Chiu, T.-L., Wen, Z., Rupasinghe, S.G. and Schuler, M.A. (2008) Comparative molecular modeling of *Anopheles gambiae* CYP6Z1, a mosquito P450 capable of metabolizing DDT. *Proc. Natl. Acad. Sci. USA* 105: 8855-8860.

101. McLaughlin, L.A., Niazi, U., Bibby, J., David, J.P., Vontas, J., Hemingway, J., Ranson, H., Sutcliffe, M.J. and Paine, M.J.I. (2008) Characterization of inhibitors and substrates of *Anopheles gambiae* CYP6Z2. *Insect Mol. Biol.* 17: 125-135.

102. Jones, R.T., Bakker, S.E., Stone, D., Shuttleworth, S.N., Boundy, S., McCart, C., Daborn, P.J., ffrench-Constant, R.H. and van den Elsen, J.M. (2010) Homology modelling of Drosophila cytochrome P450 enzymes associated with insecticide resistance. *Pest Manag. Sci.* 66: 1106-1115.

103. Lertkiatmongkol, P., Jenwitheesuk, E. and Rongnoparut, P. (2011) Homology modeling of mosquito cytochrome P450 enzymes involved in pyrethroid metabolism: insights into differences in substrate selectivity. *BMC Res Notes* 4: 321.

104. Stevenson, B.J., Bibby, J., Pignatelli, P., Muangnoicharoen, S., O'Neill, P.M., Lian, L.-Y., Müller, P., Nikou, D., Steven, A., Hemingway, J., Sutcliffe, M.J. and Paine, M.J.I. (2011) Cytochrome P450 6M2 from the malaria vector *Anopheles gambiae* metabolizes pyrethroids: Sequential metabolism of deltamethrin revealed. *Insect Mol. Biol.* 41: 492-502.

105. Jeschke, P., Nauen, R. and Beck, M.E. (2013) Nicotinic acetylcholine receptor agonists: a milestone for modern crop protection. *Angew. Chem. Int. Ed. Engl.* 52: 9464-9485.

106. Jeschke, P. and Nauen, R. (2008) Neonicotinoids-from zero to hero in insecticide chemistry. *Pest Manag. Sci*. 64: 1084-1098.

107. Jeschke, P., Nauen, R., Schindler, M. and Elbert, A. (2011) Overview of the status and global strategy for neonicotinoids. *J. Agric. Food. Chem.* 59: 2897-2908.

108. Liu, Z., Williamson, M.S., Lansdell, S.J., Han, Z., Denholm, I. and Millar, N.S. (2006) A nicotinic acetylcholine receptor mutation (Y151S) causes reduced agonist potency to a range of neonicotinoid insecticides. *J. Neurochem*. 99: 1273-1281.

109. Matsuda, K., Shimomura, M., Ihara, M., Akamatsu, M. and Sattelle, D.B. (2005) Neonicotinoids show selective and diverse actions on their nicotinic receptor targets: electrophysiology, molecular biology, and receptor modeling studies. *Biosci. Biotechnol. Biochem.* 69: 1442-1452.

110. Matsuda, K., Kanaoka, S., Akamatsu, M. and Sattelle, D.B. (2009) Diverse actions and target-site selectivity of neonicotinoids: structural insights. *Mol. Pharmacol*. 76: 1-10.

111. Tomizawa, M. and Casida, J.E. (2009) Molecular recognition of neonicotinoid insecticides: the determinants of life or death. *Acc. Chem. Res.* 42: 260-269.

112. Li, J., Shao, Y., Ding, Z., Bao, H., Liu, Z., Han, Z. and Millar, N.S. (2010) Native subunit composition of two insect nicotinic receptor subtypes with differing affinities for the insecticide imidacloprid. *Insect Biochemistry and Molecular Biology* 40: 17-22.

113. Kagabu, S. (2010) Discovery of imidacloprid and further developments from strategic molecular designs. *J. Agric. Food. Chem.* 59: 2887-2896.

Printed in the United States of America, 2014
ISBN: 978-1-63315-205-2

114. Tomizawa, M., Kagabu, S. and Casida, J.E. (2011) Receptor structure-guided neonicotinoid design. *J. Agric. Food. Chem.* 59: 2918-2922.
115. Li, J., Ju, X.L. and Jiang, F.C. (2008). Pharmacophore model for neonicotinoid insecticides. *Chinese Chem. Lett.* 19: 619-622.
116. Liu, G.-Y., Ju, X.-L. and Cheng, J. (2010) Selectivity of Imidacloprid for fruit fly versus rat nicotinic acetylcholine receptors by molecular modeling. *J. Mol. Model* 16: 993-1002.
117. Cheng, J., Ju, X.-L., Chen, X.-Y. and Liu, G.-Y. (2009) Homology modeling of human $\alpha 1\beta 2\gamma 2$ and house fly $\beta 3$ GABA receptor channels and Surflex-docking of fipronil. *J. Mol. Model* 15: 1145-1153.
118. Li, Q., Kong, X., Xiao, Z., Zhang, L., Wang, F., Zhang, H., Li, Y. and Wang, Y. (2012. Structural determinants of imidacloprid-based nicotinic acetylcholine receptor inhibitors identified using 3D-QSAR, docking and molecular dynamics. *J. Mol. Model* 18: 2279-2289.
119. Cerón-Carrasco, J.P., Jacquemin, D., Graton, J., Thany, S. and Le Questel, J.-Y. (2013) New Insights on the Molecular Recognition of Imidacloprid with *Aplysia californica* AChBP: A Computational Study. *J. Phys. Chem.* B 117: 3944-3953.
120. Chung, L.W., Hirao, H., Li, X. and Morokuma, K. (2013) The ONIOM method: its foundation and applications to metalloenzymes and photobiology. *WIREs Comput. Mol. Sci.* 2: 327-350.
121. Nishiwaki, H., Nakagawa, Y., Ueno, T., Kagabu, S. and Nishimura, K. (2001) Insecticidal and binding activities of N3-substituted imidacloprid derivatives against the housefly *Musca domestica* and the α-bungarotoxin binding sites of nicotinic acetylcholine receptors. *Pest Management Science* 57: 810-814.
122. Nishimura, K., Kiriyama, K. and Kagabu, S. (2006) Quantitative structure– activity relationships of imidacloprid and its analogs with substituents at the C5 position on the pyridine ring in the neuroblocking activity. *J. Pesticide Science* 31: 110-115.
123. Kagabu, S., Nishiwaki, H., Sato, K., Hibi, M., Yamaoka, N. and Nakagawa, Y. (2002) Nicotinic acetylcholine receptor binding of imidacloprid-related diaza compounds with various ring sizes and their insecticidal activity against *Musca domestica. Pest Management Science* 58: 483-490.
124. Nishiwaki, H., Kuriyama, M., Nagaoka, H., Kato, A., Akamatsu, M., Yamauchi, S. and Shuto, Y. (2012) Synthesis of imidacloprid derivatives with a chiral alkylated imidazolidine ring and evaluation of their insecticidal activity and affinity to the nicotinic acetylcholine receptor. *Bioorg. Med. Chem. Lett.* 20: 6305-6312.
125. Sun, C., Jin, J., Zhu, J., Wang, H., Yang, D. and Xing, J. (2010) Discovery of bis-aromatic ring neonicotinoid analogues fixed as cis-configuration: Synthesis, insecticidal activities, and molecular docking studies. *Bioorganic & Medicinal Chemistry Letters* 20: 3301-3305.
126. Sun, C.W., Fang, T., Wang, J., Hao, Z.B. and Nan, S.B. (2012) Synthesis, insecticidal activity, crystal structure, and molecular docking studies of nitenpyram analogues with an omega-hydroxyalkyl ester arm anchored on the tetrahydropyrimidine ring. *J. Agric. Food Chem.* 60: 9553-9561.
127. Ohno, I., Tomizawaa, M., Durkin, K.A., Casida, J.E. and Kagabu, S. (2009) Bis-neonicotinoid insecticides: Observed and predicted binding interactions with the nicotinic receptor. *Bioorg. Med. Chem. Lett.* 19: 3449-3452.

Article History: Received 15th June, 2013; Revised 20th August 2013 and Accepted 15th March 2014 and Published on 30th October 2014.

Printed in the United States of America, 2014
ISBN: 978-1-63315-205-2

Butterfly

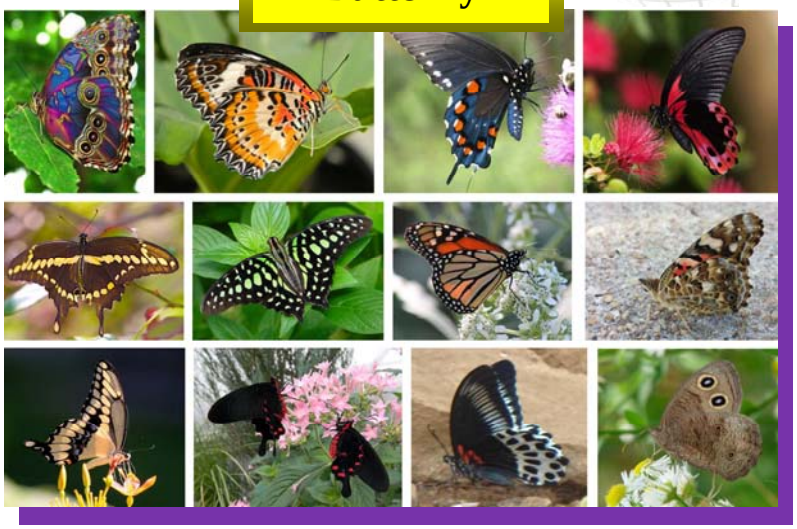


Photo Credit: Jill Sterit Murphy, Dennis Hoffman, Dave Clermont, Juan Carlos GM., Marrisia Ruth, Charleen EM., Chandrasekar, R.

Printed in the United States of America, 2014
ISBN: 978-1-63315-205-2