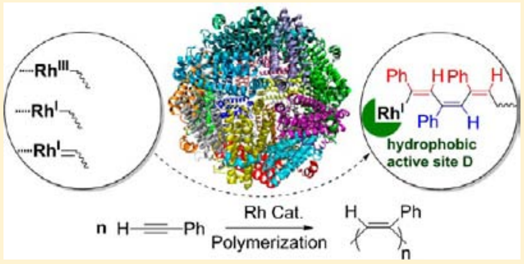
JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY
Article
pubs.acs.org/JACS

Catalytic Mechanism in Artificial Metalloenzyme: QM/MM Study of Phenylacetylene Polymerization by Rhodium Complex Encapsulated in *apo*-Ferritin

Zhuofeng Ke,[†] Satoshi Abe,[‡] Takafumi Ueno,[‡,∥] and Keiji Morokuma*[,†,§]

[†]Fukui Institute for Fundamental Chemistry, Kyoto University, Kyoto 606-8103, Japan
[‡]Department of Biomolecular Engineering, Tokyo Institute of Technology, 4259 Nagatsuta, Midori-ku, Yokohama 226-8501, Japan
[∥]Institute of Integrated Cell-Material Sciences (WPI-iCeMS), Kyoto University, Yoshida, Sakyo-ku, Kyoto 606-8501, Japan
[§]Cherry L. Emerson Center for Scientific Computation and Department of Chemistry, Emory University, Atlanta, Georgia 30322, United States

Supporting Information



ABSTRACT: Artificial metalloenzyme, composed of metal complex(es) and a host protein, is a promising way to mimic enzyme catalytic functions or develop novel enzyme-like catalysis. However, it is highly challenging to unveil the active site and exact reaction mechanism inside artificial metalloenzyme, which is the bottleneck in its rational design. We present a QM/MM study of the complicated reaction mechanism for the recently developed artificial metalloenzyme system, (**Rh(nbd)·*apo*-Fr**) (nbd = norbornadiene), which is composed of a rhodium complex $[Rh(nbd)Cl]_2$ and the recombinant horse L-chain *apo*-Ferritin. We found that binding sites suggested by the X-ray crystal structure, i.e., sites A, B, and C, are only precursors/intermediates, not true active sites for polymerization of phenylacetylene (PA). A new hydrophobic site, which we name D, is suggested to be the most plausible active site for polymerization. Active site D is generated after coordination of first monomer PA by extrusion of the $Rh^I(PA)$ complex to a hydrophobic pocket near site B. Polymerization occurs in site D via a Rh^I-insertion mechanism. A specific "hydrophobic region" composed by the hydrophobic active site D, the nonpolar 4-fold channel, and other hydrophobic residues nearby is found to facilitate accumulation, coordination, and insertion of PA for polymerization. Our results also demonstrate that the hydrophobic active site D can retain the native regio- and stereoselectivity of the Rh-catalyzed polymerization of PA without protein. This study highlights the importance of theoretical study in mechanistic elucidation and rational design of artificial metalloenzymes, indicating that even with X-ray crystal structures at hand we may still be far from fully understanding the active site and catalytic mechanism of artificial metalloenzymes.

1. INTRODUCTION

The efficiency and specificity of enzyme catalysis in nature inspire scientists to design and construct artificial enzymes to reproduce their catalytic functions or develop novel enzyme-like catalysis.[1−10] Metalloenzymes that account for nearly one-half of all enzymes in nature manipulate versatile biological processes.[4] One of the promising strategies is to incorporate a catalytic metal complex into a host protein to construct an artificial metalloenzyme to achieve novel catalytic features in the well-defined "womb" of protein.[1−6,11] This strategy is expected to combine the advantages of the metal complex in easy accessibility, broad substrate scope, and versatile transformation and the merits of the enzyme in high efficiency, specific selectivity, and green chemistry. Through genetic evolution of the protein component and chemical screening of the metal complex component, artificial metalloenzymes will provide a new horizon for application and development of catalysis.

Design of new artificial metalloenzymes is rather challenging. The interactions between the metal complex and the protein are complicated, and the structures of the metal complexes at binding sites can also highly vary. Understanding the active site and exact reaction mechanism inside the host protein is extremely important to development of artificial metalloenzymes.[7−9] However, characterization of the structural details of artificial metalloenzyme is very difficult. Even with some X-ray crystal structures of catalytic precursors or intermediates we are still far from understanding the catalysis process inside artificial metalloenzyme. In this paper, we will clarify, with QM/MM studies, an example of the intricate reaction mechanism of the recently developed artificial metalloenzyme system for polymerization.[12] Our results highlight the important role of theoretical study in mechanistic elucidation and rational design of artificial metalloenzymes.

Proteins can form different cages with size ranging from tens to hundreds of nanometers.[13,14] With this inspiration, a well-defined "womb" with spherical *apo*-Ferritin for polymerization

Received: June 5, 2012
Published: August 24, 2012

has been designed to constrain the molecular weight of the polymer. By encapsulating the rhodium catalyst $[Rh(nbd)Cl]_2$ (nbd = norbornadiene) into the inner cavity of the recombinant horse L-chain *apo*-Ferritin (**Rh(nbd)·*apo*-Fr**), poly(phenylacetylene) (PPA) with restricted molecular weight and a narrow molecular weight distribution has been obtained, see Figure 1.[12] Inside the *apo*-Ferritin, the size of the inner cavity of *apo*-Ferritin can well constrain the PPA molecule consisting of around 130 monomers. Although up to 72 molecules of the rhodium complexes have been found immobilized on the interior surface of *apo*-Ferritin, most of them cannot serve as active sites for polymerization because only several PPA chains were formed in the cage. X-ray characterization found that these ~60 immobilized rhodium complexes can be divided into three types of binding sites.[12]

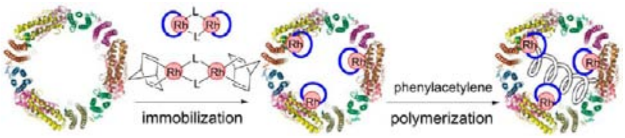


Figure 1. Polymerization of phenylacetylene by rhodium norbornadiene complexes encapsulated in the recombinant horse L-chain *apo*-Ferritin.

We recently studied computationally the detailed mechanism of polymerization of PA catalyzed by the homogeneous catalyst $[Rh(nbd)Cl]_2$ without protein environment.[15] There are complicated equilibriums of the reaction intermediates in solution. Rh^I, Rh^{III}, and Rh−carbene types of intermediates are all found to be thermodynamically and kinetically feasible at room temperature in solution.[15−18] We found that polymerization in solution prefers to undergo the Rh^I-insertion mechanism; the Rh^{III}-insertion and Rh−carbene metathesis mechanisms need to overcome higher reaction barriers. Furthermore, Rh-catalyzed polymerization was shown to prefer regio- and stereoselective 2,1-inserion of PA in solution, leading to the head-to-tail cis−trans PPA.[15]

Under the influence of the protein environment, how different is the polymerization mechanism inside the artificial metalloenzyme from that in homogeneous solution? It is not surprising that encapsulated rhodium complexes can form different types of active species in the artificial metalloenzyme system (**Rh(nbd)·*apo*-Fr**) due to the complicated equilibriums of rhodium complexes in solution. The variety of binding sites gives rise to a more complicated situation for the polymerization process inside the protein. Are these binding sites, presented by X-ray characterization, precursors, intermediates, or active sites for polymerization? Can **Rh(nbd)·*apo*-Fr** maintain the regio- and stereoselectivity features of the polymerization of PPA? With these questions in mind, QM/MM studies are carried out to provide deeper insights into the mechanism of polymerization in the artificial metalloenzyme **Rh(nbd)·*apo*-Fr** in order to advance rational design and development of artificial metalloenzymes.

2. COMPUTATIONAL MODELS AND DETAILS

2.1. Protein Setup and Preparatory MM Calculations. The initial structure of *apo*-Ferritin was taken from the previous study (PDB 2ZUR; resolution 1.80 Å).[12] Hydrogen atoms of all amino acids as well as missing atoms in Glu53, Glu63, Cys126, Glu130, and His173 were added to the 2ZUR structure with the VMD program.[19] There are three types of binding sites for **Rh(nbd)·*apo*-Fr**, as shown in Figure 2. Other heteroatoms that do not belong to *apo*-Ferritin in the PDB file as well as rhodium atoms were removed. Binding site C unveiled by X-ray diffraction involves an unusual cross-link between Cys48 with Rh(nbd). This cross-link was removed and the S−H bond of Cys48 recovered to build the structure for preliminary pure MM optimization. With the MM-optimized minimum, we subsequently reconstructed the binding sites for QM/MM studies based on the original X-ray structure. The protonation states and hydrogen-bonding networks were determined with the aid of PROPKA analysis.[20] The protonation states of histidine residues were further checked via visual inspection. The protonation states of histidine residues were determined as follows: HID49, HID114, HIE124, HID132, HID147, and HID173, where HID and HIE donate neutral histidines with protonated N_δ and N_ε atoms, respectively. From the monomer unit of the *apo*-Ferritin in PDB file 2ZUR, the entire structure of *apo*-Ferritin was constructed using crystallographic symmetry with Swiss PDB Viewer.[21] The prepared structure was then solvated in a cubic box of 146.3 Å and neutralized by addition of sodium counterions at the ionic concentration of 0.5 mol/L. The entire structure for 24 units contains 291 649 atoms including 4152 residues in 24 chains and 74 877 water molecules. To obtain a good initial geometry for the *apo*-Ferritin cage, optimizations were performed using the Amber suit of programs[22] (parm96.dat) on the obtained solvated 24 units *apo*-Ferritin structure with frozen backbone atoms, relaxed hydrogen atoms, and waters. The entire solvated protein structure was optimized subsequently without constraint.

2.2. Binding Sites Construction. Because of the large number of atoms for the entire structure with 24 units, one monounit was abstracted to construct the binding sites of rhodium complex in *apo*-Ferritin to investigate the catalyzed polymerization mechanism inside **Rh(nbd)·*apo*-Fr**. Our QM/MM calculations were performed on the monomeric subunit because these active sites do not interfere with neighboring monomers, at least at the early stage of polymerization

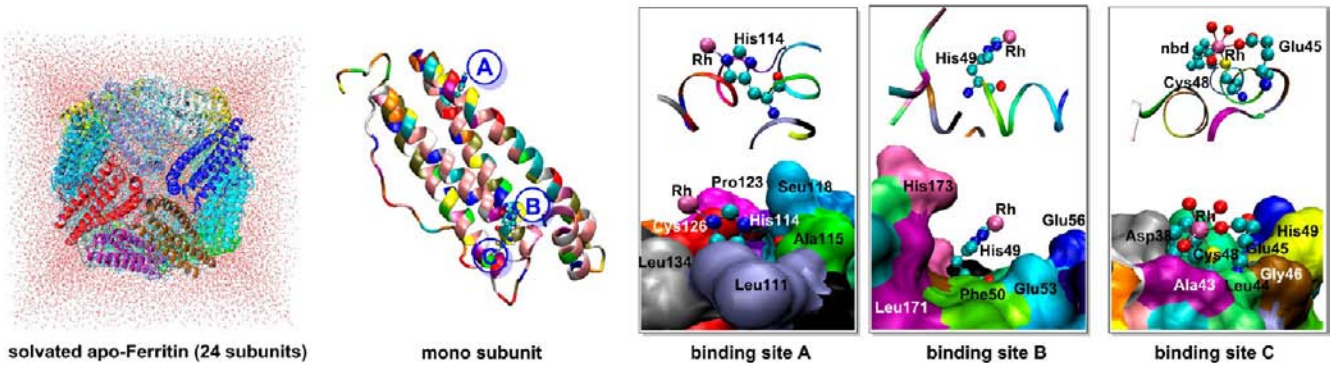


Figure 2. Solvated *apo*-Ferritin structures with 24 subunits, monounit structure with immobilized rhodium complexes, and structures of binding sites A, B, and C on each monounit (top, cartoon view; bottom, surface view).

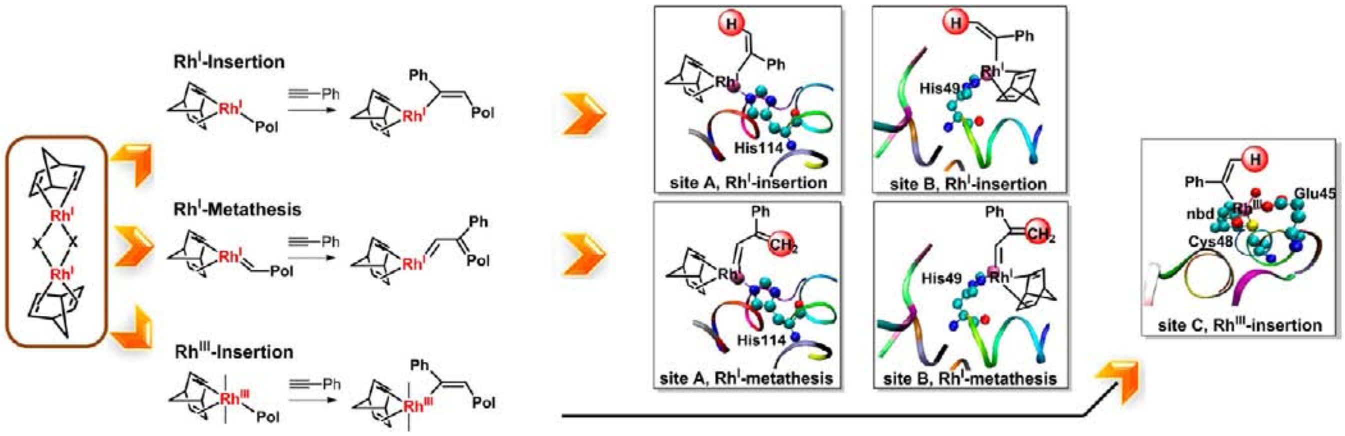


Figure 3. Plausible active sites A, B, and C inside **Rh(nbd)·*apo*-Fr** constructed based on the mechanisms of the previous theoretical study[15] of Rh–diene-catalyzed polymerization of PA: Rh$^{\text{I}}$ insertion, Rh$^{\text{I}}$ metathesis, or Rh$^{\text{III}}$ insertion. Pol stands for the polymer chain.

when the propagation chain is short. The 3D surface and ribbon views of the binding sites, A, B, and C sites, are shown in Figure 2. Binding site A is located at a shallow pocket composed of Leu111, Ala115, Seu118, Cys126, Glu130, and others. The rhodium metal center coordinates to His114 through a Rh–N$_\varepsilon$ dative interaction. Binding site B sits at a deep pocket surrounded by Phe50, Glu53, Glu56, Leu171, His173, and others. The rhodium metal center of binding site B binds to the N$_\varepsilon$ of His49. Binding site C is located at a pocket near site B, and its binding pocket is composed of Asp38, Ala43, Leu44, Glu45, Gly46, Cys48, and others. The rhodium metal center coordinates to both Cys48 and Glu45 through Rh–S and Rh–O bonds, respectively. For site C, nbd ligand and additional water ligands are also well characterized, showing that the rhodium center is six coordinated. In addition, there is an unusual cross-link between Cys48 and nbd.[12] For both binding sites A and B, other ligands of rhodium such as nbd are missing in the X-ray structure. Beside these binding sites, another binding site D is found to be crucial for polymerization of PA by our QM/MM studies (see Discussion later). This site D is just next to site B and is constituted of hydrocarbon side chains of Phe50, Lys143, and Leu171.

According to the mechanism obtained from our previously theoretical studies of polymerization of PA by the Rh-type catalyst without protein,[15] the catalyst promotes polymerization through Rh$^{\text{I}}$-insertion, Rh$^{\text{I}}$-metathesis, or Rh$^{\text{III}}$-insertion mechanism. Although Rh$^{\text{I}}$ insertion is preferred in solution, it is not clear how the protein environment of *apo*-Ferritin will influence the polymerization mechanism. The three binding sites discussed above have totally different microenvironments for polymerization. In particular, one notes that not all these binding sites can serve as active sites for polymerization; only a few PPA chains can be produced. Binding site C is a six-coordinate site. Therefore, if site C participates in the polymerization it should promote polymerization through Rh$^{\text{III}}$ insertion. In contrast, binding sites A and B, where the electron density of the other ligands is absent, are more likely to involve polymerization via Rh$^{\text{I}}$ insertion or Rh$^{\text{I}}$ metathesis with a lower coordination number on the rhodium center. With this rationalization in mind, we constructed plausible active sites for the present artificial metalloenzyme system **Rh(nbd)·*apo*-Fr**, as depicted on the right of Figure 3. Since the diene ligand has been documented to play a key role in the polymerization of alkynes,[23,24] the nbd ligand that is missing in the X-ray structure was retained at sites A and B. For Rh$^{\text{I}}$-insertion polymerization, a phenylethenyl ligand was added to represent the propagation chain, where the rest of the polymer chain was replaced with a hydrogen atom for computational convenience. For Rh$^{\text{I}}$-metathesis polymerization, a phenylethenyl carbene was added to represent the propagation chain, where the rest of the polymer chain was replaced by a methylene. We added the missing ligands based on the QM-optimized models in the gas phase. The positions of nbd and model chain were further adjusted for optimal space for propagation. In site C, the water molecule opposite to Cys48 is at the most open space for the propagation chain; therefore, it was replaced with a phenylethenyl ligand to construct a Rh$^{\text{III}}$ center for polymerization. The cross-link between nbd and Cys48 was also reconstructed based on the original X-ray structure for further QM/MM studies.

2.3. QM/MM Calculations. The multilayered quantum mechanical and molecular mechanical (QM/MM) method,[25−27] specifically, the two-layer ONIOM method,[28−36] was employed in this study as implemented in Gaussian09.[37] The QM part and MM parts were treated with the B3LYP method[38−40] and Amber force field,[41,42] respectively. The Stevens (SBK) triple-ξ valence basis set with the effective core potential (ECP) was used for rhodium and the 6-31G* basis set for the other elements in the QM part. The ONIOM total energy is obtained by an extrapolation approach

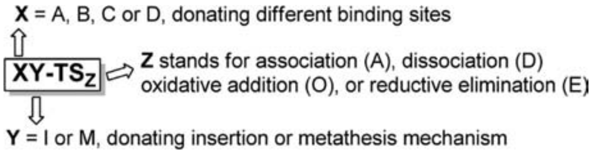
$$E_{\text{ONIOM}} = E_{\text{MM,real}} + E_{\text{QM,model}} - E_{\text{MM,model}} \tag{1}$$

where $E_{\text{MM,real}}$ is the energy of the real system calculated by MM and $E_{\text{QM,model}}$ and $E_{\text{MM,model}}$ are the energy of the model system calculated by QM and MM, respectively. The energy of the model system in the MM level will be canceled out via $E_{\text{MM,real}}$ minus $E_{\text{MM,model}}$. Therefore, the MM parameters for the rhodium complexes would not affect the accuracy of the ONIOM total energy, E_{ONIOM}. The ONIOM electronic embedding (EE) scheme was used to optimize all intermediates and transition states. Frequency calculations at the same level were further performed to verify key transition states and intermediates.

In the QM/MM studies, residues and ligands that directly coordinate to rhodium, including nbd and model polymer chain, were treated with QM. The remaining part of the protein was treated with MM. The MM part was further divided into the optimized MM region and the frozen MM region. Residues with any atom within 12 Å of the binding sites were assigned to the optimized MM region. The van der Waals parameters for the rhodium center were taken from Dunbar's paper.[43] Since polymerization occurs on the interior surface of *apo*-Ferritin, transportation, coordination of PA monomer, and insertion/metathesis steps will cause significant rearrangement of the water molecules and the H-bonding network, disturbing the continuity of the potential energy surface. Therefore, the solvated water molecules were removed during QM/MM studies. However, we used several strategies to preserve the conformation of protein in the X-ray structure. (1) We used the MM minimized structure (minimized with solvated water) as the initial structure. MD has double-edged effects, which sometimes may lead to a less realistic equilibrium state. The characterized X-ray structure should be chosen as the starting point, especially when accurate MM parameters for Rh complexes are unfortunately not available yet. (2) We froze the outer part of the protein during QM/MM optimization to maintain the conformation of the protein to that of the MM minimum in solution. The rmsd values for all stationary points along the potential energy surface (PES) of each reaction pathway are shown in Figure S12, Supporting Information. The rmsd values are very small, and the conformation and side chains of the residues of the protein are well preserved. (3) Further QM/MM study on the key reaction pathways with a 5 Å water shell obtained very similar results, leading to the same discussion and conclusion, which strongly suggests that models with a solvated water shell and models with a frozen outer part of the protein are consistent with each other (Figure S11, Supporting Information). All of the above results guarantee that QM/MM-optimized protein structures in our manuscript preserved the conformation in the original X-ray structure.

Our previous studies found that the calculated gas-phase results are similar to the results in solution for alkyne polymerization.[15] All energies (with zero-point energy correction (ZPC) of the QM part) used in the discussion are relative to the reactant complex of the catalyst, including both the PA monomer and binding site. For further analysis of the energy contribution from the protein, the ONIOM mechanical embedding (ME) scheme was used for single-point calculations at the ONIOM-EE-optimized structures. Pure QM optimization at the B3LYP level on the model systems in the gas phase was also performed for comparison with ONIOM results. The nomenclature for all stationary points is illustrated in Chart 1.

Chart 1. Nomenclature



2.4. ONIOM Energy Component Analysis. We performed energy decomposition analysis to analyze the contributions of energy components (electrostatic, vdW, polarization, etc.) for energy changes between intermediates or transition states. To separate the polarization effect, the electrostatic components were obtained from ONIOM-ME single-point calculations on the ONIOM-EE-optimized structures via turning off the nonbonded Coulomb interaction terms in the MM potential function. The charges of atoms in the QM part were taken from the Merz−Kollman scheme of electrostatic potential fit, and the Amber charges were kept for the atoms in the MM part. The vdW components were obtained with ONIOM-ME single-point calculations on the ONIOM-EE-optimized structures via turning off the vdW interaction terms in the MM potential function. Polarization energies were obtained from the difference between ONIOM-EE and ONIOM-ME single-point energies on the ONIOM-EE-optimized geometries. QM deformation energies are the changes in intrinsic QM energies in QM/MM stationary points. With the above methods, the energy changes of each pair of stationary points can be divided into aforementioned energy components.

3. RESULTS AND DISCUSSION

3.1. Binding Sites A, B, and C. The ONIOM-optimized structures for binding sites A, B, and C are shown in Figure 4. The corresponding DFT-optimized structures for the QM models are shown in Figure S2, Supporting Information. In most cases the ONIOM-EE structural parameters are very similar to the DFT-optimized parameters of the QM model. For instance, for the binding sites for Rh^I insertion, i.e., **AI-1** and **BI-1**, the Rh−C^1 bond length between the Rh^I center and the phenylethenyl group is calculated to be 2.08 and 2.07 Å, respectively, in good agreement with similar X-ray structures reported in the literature.[44−46] The Rh=C^1 bonds in the carbene type of binding sites for Rh^I metathesis, **AM-1** and **BM-1**, are calculated to be 1.88 and 1.89 Å, respectively, shorter than the Rh−C^1 bond in **AI-1** and **BI-1**; the Rh=C^1 bonds here have typical double-bond character.[47,48] The Rh−N distances in all these binding site structures are in the range 2.10−2.17 Å, consistent with the Rh−N distance in binding site A of the 2ZUR X-ray structure (2.13 Å).

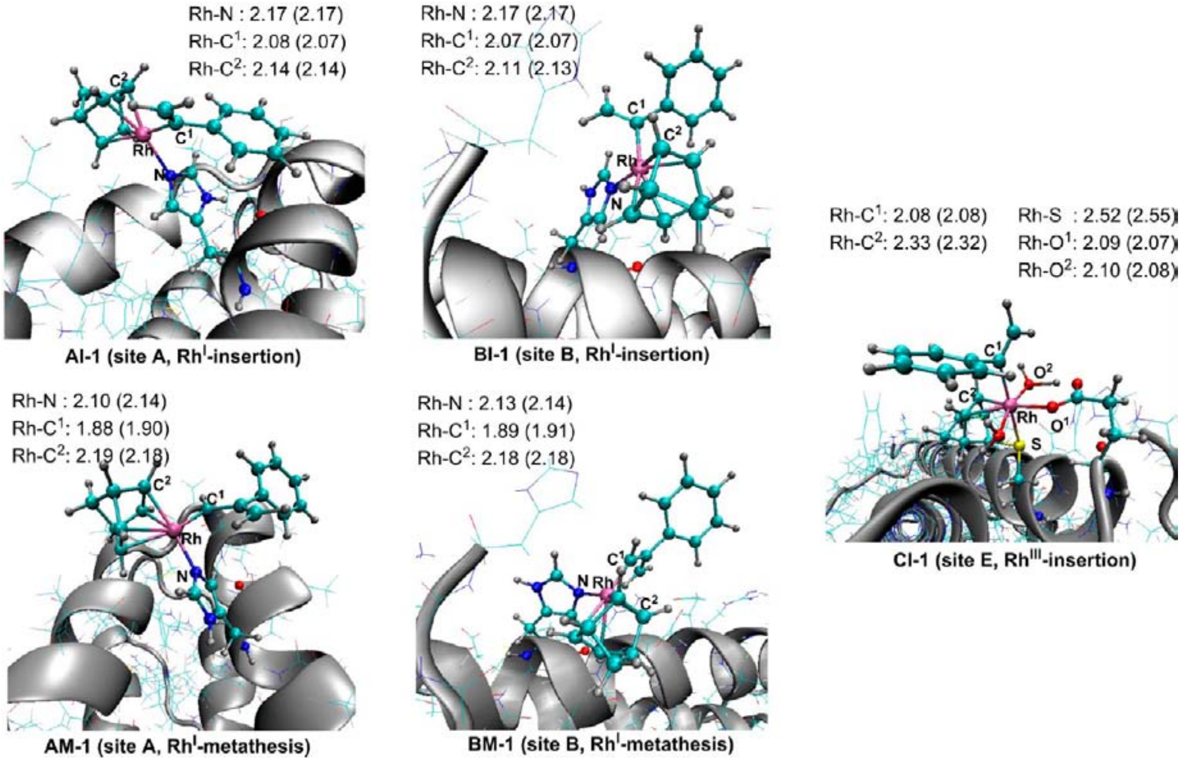


Figure 4. ONIOM-EE-optimized structures for binding sites A, B, and C for polymerization of PA inside **Rh(nbd)·*apo*-Fr**. Bond distances are given in Angstroms. Data in parentheses are obtained from DFT-optimized structures for the QM model system in the gas phase.

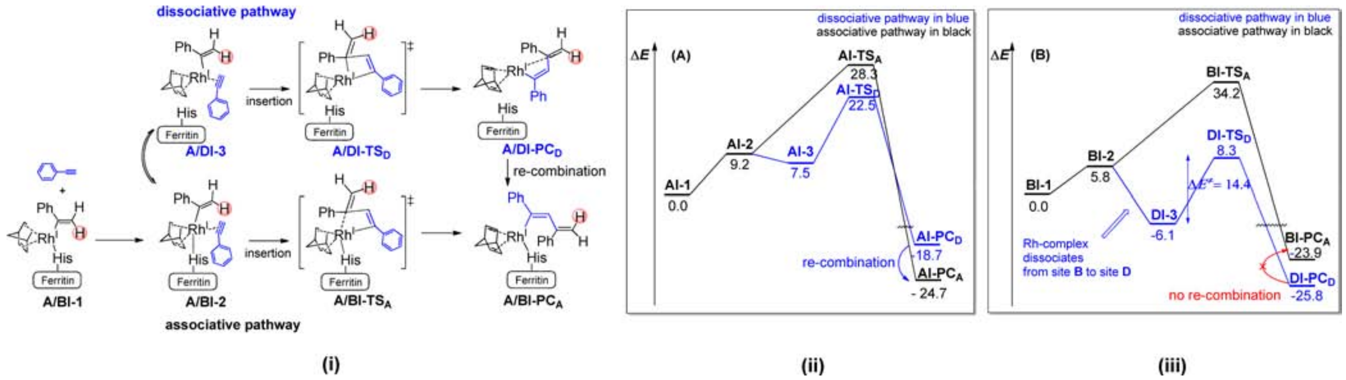


Figure 5. (i) Dissociative (in blue) and associative (in black) Rh^I-insertion mechanisms for polymerization of PA in binding sites A and B (and D) of **Rh(nbd)·*apo*-Fr**. (ii and iii) ONIOM-EE PES profiles with ZPC in kcal/mol.

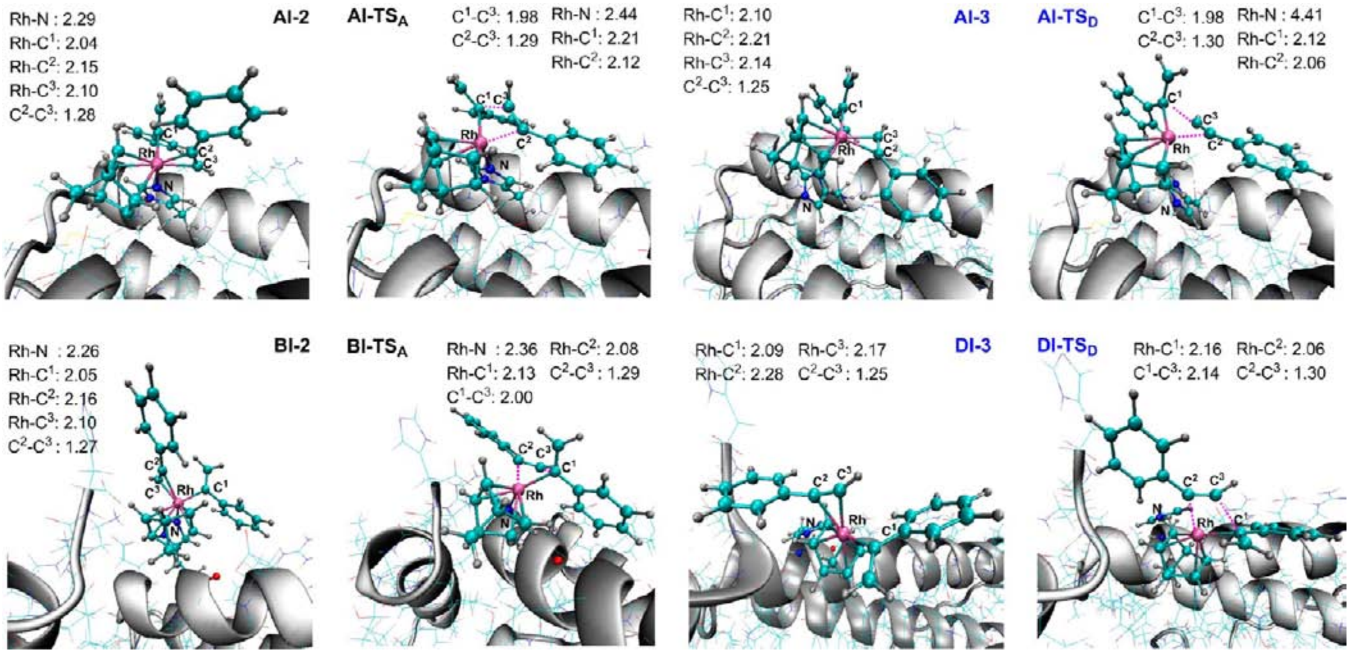


Figure 6. Key intermediates and transition states for polymerization of PA via the Rh^I-insertion mechanism in binding sites A and B (and D) of **Rh(nbd)·*apo*-Fr**. Bond distances are given in Angstroms.

Introduction of a phenylethenyl as a model propagation chain gives **CI-1** for Rh^{III} binding site C, as shown in Figure 4. The calculated Rh−S is 2.52 Å, longer than the Rh−S bond length in the 2ZUR X-ray structure (2.25 Å). Elongation of the Rh−S bond mainly arises from the trans effect of the phenylethenyl group.[49] We further evaluate site C (Figure S3, Supporting Information) with three water molecules as ligands. As expected, the Rh−S bond length becomes shorter, 2.32 Å, in better agreement with the bond length in the 2ZUR X-ray structure.

We also evaluated the possibility of a Rh^I center for binding site C. As expected, the Rh^I center is less likely to adopt a six-coordinated geometry due to the filled d_{z^2} orbital. Upon optimization from an initial six-coordinated structure, the Rh^I center spontaneously changed into a four-coordinated structure, with two water ligands detached from the metal center (Figure S4, Supporting Information), which is inconsistent with the X-ray structures. Reduction of the Rh^{III} center to a Rh^I center requires a reductive elimination step, which is not feasible in site C, where the coordinated ligands are water, a neutral sulfur atom, and a double bond.

3.2. Rh^I Insertion in Sites A and B (and D).

Starting from binding sites A and B, at first we investigated polymerization of PA via Rh^I insertion. There is one histidine residue acting as ligand of the rhodium center in binding sites A and B. Rh^I insertion can occur through two possible pathways (Figure 5i). One is the associative pathway in which histidine stays associated with the metal center during coordination and insertion of monomer PA. The other one is the dissociative pathway, where the incoming monomer PA requires detachment of histidine from the metal center, and insertion occurs from a four-coordinated intermediate. Whether the metal center recombines or not with the histidine residue after insertion would depend on the binding energy of the active center and the histidine ligand.

The ONIOM-EE PES profiles for Rh^I insertion at site A are shown in Figure 5ii. Along the associative pathway, coordination of reactant PA forms a quasi-square-pyramidal intermediate **AI-2**, uphill by 9.2 kcal/mol. A quasi-square-pyramidal Rh^I intermediate is expected to be less stable than a square-planar Rh^I structure, because the nonbonding d_{z^2} orbital of the square-planar Rh^I center becomes antibonding in the five-coordinate **AI-2**. The deformation energy is estimated to be 3.1 kcal/mol, comparing the QM part of **AI-2** to the optimized QM model in the gas phase. Additional destabilization comes from the electrostatic and polarization effect between the coming monomer and residues around site A. Therefore the associative Rh^{I}-insertion mechanism is not likely to be favorable. As can be seen in Figure 5ii, the five-coordinated transition state **AI-TS$_A$** is higher in energy by 5.8 kcal/mol than the dissociative transition state **AI-TS$_D$**. The reaction barrier for associative Rh^{I} insertion is 28.3 kcal/mol relative to **AI-1**. In the dissociative pathway of site A, coordination of PA requires detachment of histidine from rhodium in **AI-2** to form a relatively stable intermediate **AI-3**. Dissociative Rh^{I} insertion via **AI-TS$_D$** from **AI-3** has a reaction barrier of 22.5 kcal/mol, relative to **AI-1**. In **AI-TS$_D$** (Figure 6), the Rh−C^1 and Rh−C^2 bond distance are 2.12 and 2.06 Å, respectively, shorter than those in the associative **AI-TS$_A$**. The lower transition state energy of **AI-TS$_D$**, compared with **AI-TS$_A$**, is attributed to further formation of the Rh−C^2 bond and less cleavage of the Rh−C^1 bond because of a smaller repulsion from the protein environment. Rh^{I}-insertion polymerization in site B presents an interestingly different story. The shape of site B is quite different from site A, as shown in Figure 2. While site A is shallow, site B is deep, surrounded by residues Glu53, Glu56, Leu171, Lys172, and His173. Along the associative pathway, coordination of PA produces a reactant complex **BI-2**, which is uphill by 5.8 kcal/mol relative to **BI-1**. However, the associative transition state **BI-TS$_A$** cannot be located despite extensive search and probably does not exist due to severe constraint in the deep site B. During insertion, the triple bond of PA is required to rotate onto the plane of nbd, Rh, and phenylethenyl, which results in severe repulsion between the monomer and residues Lys172 and His173, as can be seen in **BI-TS$_A$** of Figure 6. Therefore, the complex is automatically extruded from the deep site B accompanied with cleavage of the Rh−N bond while trying to locate the associative transition state. To evaluate the barrier for associative insertion, an ONIOM-EE relaxed scan was performed, as shown in Figure S5, Supporting Information. The scanned surface suggests that the approximate reaction barrier for the associative Rh^{I} insertion is as high as 34.2 kcal/mol.

The dissociative Rh^{I}-insertion pathway in site B requires detachment of histidine upon PA coordination. Different from dissociative Rh^{I} insertion in site A, the released Rh(PA) complex is spontaneously adsorbed by a hydrophobic pocket close to site B. This new binding pocket for the rhodium active species will be called site D, which is composed of hydrocarbon side chains of Phe50, Lys143, and Leu171. Noteworthy, this hydrophobic site D plays a key role in the polymerization of PA in **Rh(nbd)·*apo*-Fr**. Further detailed discussions will be given in section 3.5. The structures of site B before and after dissociation are shown in Figure S13, Supporting Information. All positions of residues are well preserved. The movement of the complex from site B to neighboring site D is only ~3 Å (nbd as reference). The formed intermediate **DI-3** in site D is lower in energy than **BI-1** by 6.1 kcal/mol and than **BI-2** by 11.9 kcal/mol. Another isomer of **DI-3** was also found with a $C^2{\equiv}C^3$ triple bond perpendicular to the plane of nbd, Rh, and phenylethenyl and is higher in energy than **DI-3** by only 0.3 kcal/mol. **DI-3** then undergoes Rh^{I} insertion through the dissociative transition state **DI-TS$_D$**. The reaction barrier of this insertion step is only 14.4 kcal/mol relative to **DI-3**. Therefore, the dissociative Rh^{I} insertion from site B (in site D) is much more favored than the associative Rh^{I} insertion. This is also more favorable when compared with polymerization of PA in binding site A, where the dissociative Rh^{I} insertion needs to overcome a barrier of 22.5 kcal/mol.

3.3. Rh^{III} Insertion in Site C. Site C has a six-coordinated Rh^{III} center as revealed by the X-ray crystal structure, which is quite different from sites A and B. Starting from **CI-1**, Rh^{III}-insertion polymerization requires replacement of one of the solvent water molecules by monomer PA, Figure 7. The water molecule located at the open position has more space to accept the monomer and thus is replaced with the coming PA, leading to intermediate **CI-2**, Figure 8. **CI-2** is higher in energy than **CI-1** by 5.5 kcal/mol. The binding between the $C^2{\equiv}C^3$ triple bond of PA and the Rh^{III} center is relatively weak. The Rh−C^3 bond length is 2.08 Å, and the Rh−C^2 distance is as long as 3.07 Å, due to the crowded environment of the six-coordinated Rh^{III} center. The Rh^{III}-insertion transition state **CI-TS** is calculated to be very high in energy (44.7 kcal/mol). This agrees well with our previous suggestion that the Rh^{III} center is not as reactive as the Rh^{I} center for triple-bond insertion.[15] The weak activation of the PA triple bond in **CI-2** and the structural constraints of the crowded six-coordinated Rh^{III} center both contribute to the high barrier for Rh^{III} insertion in site C. It can be anticipated that with a bulkier polymer chain instead of the model the Rh^{III}-insertion polymerization pathway will become even more difficult.

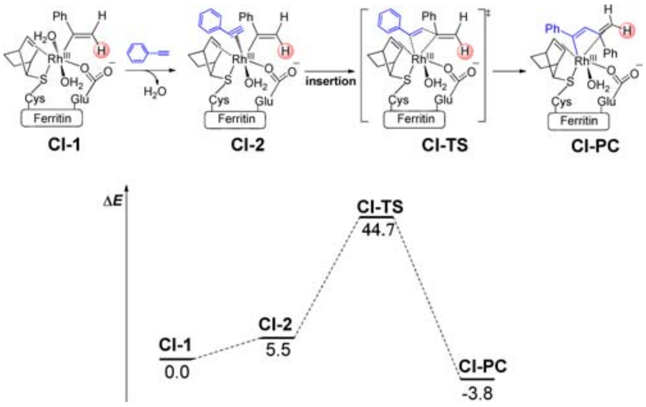


Figure 7. ONIOM-EE PES profiles (with ZPC in kcal/mol) for Rh^{III}-insertion mechanisms for polymerization of PA in binding site C of **Rh(nbd)·*apo*-Fr**.

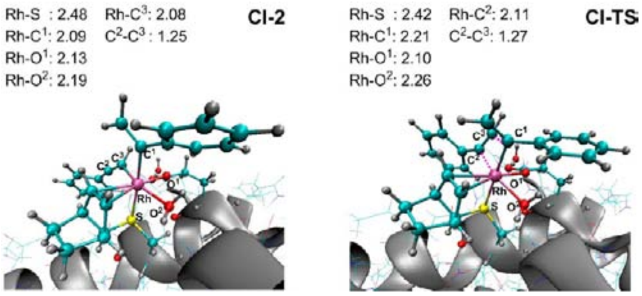


Figure 8. Key intermediates and transition states for polymerization of PA via the Rh^{III}-insertion mechanism in binding site C of **Rh(nbd)·*apo*-Fr**. Bond distances are given in Angstroms.

3.4. Rh^{I} Metathesis in Sites A and B. After evaluating the insertion mechanism, we further investigated the feasibility of the Rh^I-metathesis polymerization mechanism in sites A and B, which is promoted by the Rh^I carbene species **AM-1** and **BM-1** (Figure 4). Along the Rh^I-metathesis pathway, the reaction between the Rh═C double bond and the triple bond of PA forms a metallocyclobutene intermediate, followed by a ring opening to finish the propagation, as shown in Figure 9. A dissociative pathway for the Rh^I metathesis is difficult due to the strong interaction between the positive Rh^I(nbd)−carbene center and the residue histidine (Figure S6, Supporting Information). Replacement of histidine with PA to form a four-coordinated Rh^I(PA)−carbene complex is highly endothermic (16.0 kcal/mol). The metathesis barrier is evaluated to be 21.5 kcal/mol. Therefore, the dissociative Rh^I-metathesis mechanism is not further considered.

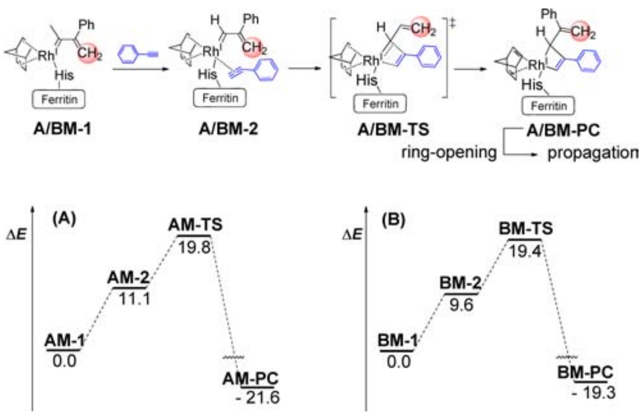


Figure 9. ONIOM-EE PES profiles (with ZPC in kcal/mol) for Rh^I-metathesis mechanisms for polymerization of PA in binding sites A and B of **Rh(nbd)·*apo*-Fr**.

In site A, **AM-1** coordinates with PA to form intermediate **AM-2**, uphill by 11.1 kcal/mol. PA coordinates to the rhodium center with only one carbon atom (Rh−C^3 = 2.35 Å), Figure 10. The filled d_{z^2} orbital of rhodium tends to interact with the π^* orbital of the PA triple bond. The subsequent metathesis step via transition state **AM-TS** has a barrier of 19.8 kcal/mol relative to **AM-1**. In site B, we obtained similar results. Coordination of PA to **BM-1** forms **BM-2**, which is higher in energy by 9.6 kcal/mol. The metathesis transition state **BM-TS** has an activation energy of 19.4 kcal/mol relative to **BM-1**. Different from the insertion mechanism, where the deep narrow site B is less favorable for the associative insertion than the shallow site A, the Rh^I-metathesis polymerization has similar PES profiles for both sites A and B. Rh^I insertion requires significant rotation of monomer PA in site B (around 90° from **BI-2** to **BI-TS$_A$**). In contrast, the orientation of PA changes much less during metathesis. PA only rotates about 34° from **BM-2** to **BM-TS**. Therefore, the Rh^I-metathesis reactions in sites A and B have similar reaction barriers (∼20 kcal/mol) as geometric constraints raised by PA rotations that are quite similar for both sites. The distance C^1−C^2 is 2.38 Å in **AM-TS** and 2.59 Å in **BM-TS**. These parameter deviations arise from the polarization effect between the QM part and the MM part in sites A and B, especially between the phenyl group of PA and the hydrocarbon side chains of surrounded residues; with the ONIOM-ME method, the optimized transition structures for **AM-TS** and **BM-TS** actually have a quite similar C^1−C^2 distance (2.45 and 2.47 Å, respectively). Rh−N is 2.28 Å in **AM-TS** and 2.14 Å in **BM-TS**. Compared with **AM-TS**, shortening of the Rh−N bond compensates for elongation of the C^1−C^2 bond in **BM-TS**, leading to similar reaction barriers (∼20 kcal/mol).

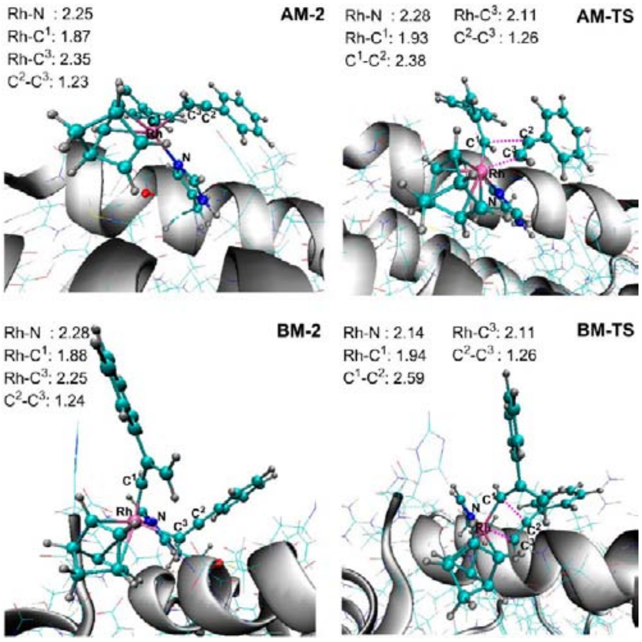
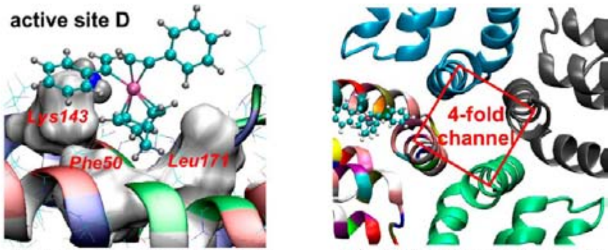


Figure 10. Key intermediates and transition states for polymerization of PA via the Rh^I-insertion mechanism in binding sites A and B of **Rh(nbd)·*apo*-Fr**. Bond distances are given in Angstroms.

3.5. Polymerization Behavior in Rh(nbd)·*apo*-Fr.

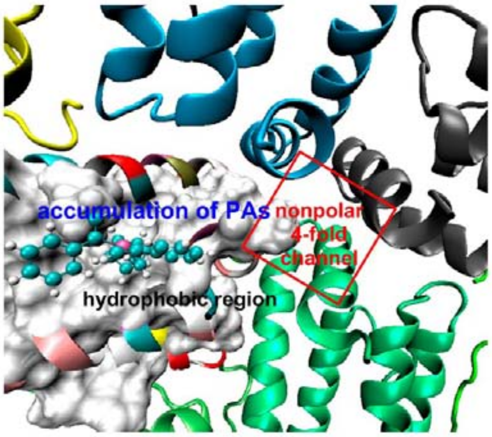
3.5.1. Hydrophobic Environment and Plausible Active Site D. Analysis of all polymerization mechanisms discussed above indicates that the Rh^I-insertion pathway in the hydrophobic pocket (site D) dissociated from site B is the most feasible mechanism for polymerization of PA in **Rh(nbd)·*apo*-Fr**, with a transition state energy of 8.3 kcal/mol. The insertion barrier in site D is calculated to be 14.4 kcal/mol. In comparison, Rh^I insertion in site A needs to overcome a reaction barrier as high as 22.5 kcal/mol. Carbene metathesis in sites A and B also has a reaction barrier of around 20 kcal/mol. Polymerization in site C via Rh^{III} insertion is even less possible, with a reaction barrier of 34.2 kcal/mol. Therefore, our QM/MM studies suggest that polymerization of PA in **Rh(nbd)·*apo*-Fr** is catalyzed by the Rh^I species docked to a hydrophobic site D, which may be the most plausible active site for polymerization.

This hydrophobic pocket of site D is composed of hydrophobic type of residues, Phe50 and Leu170, and the hydrocarbon side chain of Lys143 (Figure 11i). Coordination of PA extrudes the Rh(PA) complex from site B to site D. The formed **DI-3** is lower in energy by −11.9 kcal/mol than **BI-2**. Energy analysis of the contributions of the protein environments shows that the electrostatic effect contributes the most (−25.0 kcal/mol) and the vdW effect contributes the second most (−7.0 kcal/mol) (Figure 12). Polarization has some negative effect (11.5 kcal/mol) when the Rh(PA) complex moves from site B to site D. The electrostatic effect is more or less overestimated with the ONIOM-EE method. However, the total contribution of the electrostatic and polar effects is quite reasonable and predicted to be −13.5 kcal/mol. The driving forces for Rh(PA) complex to move from site B to site D upon coordination with monomer are as follows. (1) Electrostatic stabilization of the Rh(PA) complex having nonpolar ligands. Further analysis of the electrostatic contributions from each



(i) Active site D at hydrophobic pocket composed of Phe50, Lys143 and Leu171.

(ii) Active site D near the hydrophobic 4-fold channel of apo-Ferritin.



(iii) Accumulation of PA monomers to the hydrophobic region composed by the 4-fold channel, the hydrophobic site D and other hydrophobic residues nearby. Hydrophobic residues are highlighted with grey surface in one subunit.

Figure 11. Structure and likely function of the plausible active site D in the hydrophobic pocket near site B and the hydrophobic region for polymerization in **Rh(nbd)·*apo*-Fr**.

Figure 12. Total energy changes and analysis of their major components: energy changes (ΔE) for extrusion of Rh(PA) complex from site B (**BI-2**) to site D (**DI-3**); reaction barrier ($\Delta E^{\ddagger}$) of Rh$^{\text{I}}$ insertion at **DI-TS$_{\text{D}}$** in the active site D (relative to **DI-3**); selectivity ($\Delta\Delta E^{\ddagger}$) between 2,1-insertion and 1,2-insertion in site D. QM is the deformation energy of the QM part. vdW, elect, and polar are the van der Waals, electrostatic, and polarization energies of the protein environment, respectively.

residue in the protein (Figure S14, Supporting Information) shows that the electrostatic stabilization mainly comes from Glu53 (−28.0 kcal/mol), Lys143(−43.0 kcal/mol), and Lys172 (−13.2 kcal/mol). In contrast, electrostatic destabilization results from Arg52 (30.6 kcal/mol) and His173 (26.0 kcal/mol). (2) Enhanced vdW interactions between the hydrocarbon side chains of the hydrophobic pocket and the nonpolar ligands of the Rh(PA) complex. The vdW contributions of Phe50, Lys143, and Leu170 are estimated to be −1.3, −1.9, and −3.6 kcal/mol, respectively. (3) Release of the steric constraints in site B. Site B has a deep narrow shape. After coordination of PA, the bulky Rh(PA) complex is uncomfortable (as can be seen from the highly unstable BI-TS$_{\text{A}}$). Extrusion of the Rh(PA) complex benefits from release of the steric constraints. Overall, around −11.9 kcal/mol stabilization energy can be obtained for transfer of the Rh(PA) complex from site B to site D. After coordination and insertion of the first PA, the active center remains in this hydrophobic site D to continue the propagation, as shown in the PES profiles in Figure 5iii.

The hydrophobic site D is located near the 4-fold channel of *apo*-Ferritin (Figure 11ii). This 4-fold channel is mainly lined with the nonpolar Leu residues, which constitutes an extraordinary hydrophobic environment in the interior surface of *apo*-Ferritin and is considered to favor accumulation of nonpolar PA substrates. In contrast to the hydrophobic 4-fold channel, the 3-fold channel in Ferritin is a polar channel, where the hydrophilic environment is suggested to facilitate penetration and transmission of metal ions.[50−52] Interestingly, the hydrophobic site D near the 4-fold channel revealed by this QM/MM study rationalizes nicely the polymerization behavior in **Rh(nbd)·*apo*-Fr**. As can be seen, the hydrophobic site D, the nonpolar 4-fold channel, and other nonpolar side chains of residues nearby (Figure 11iii) all together form a region that may be called a "hydrophobic region". This hydrophobic region is very convenient for accumulation, coordination, and insertion of PA during polymerization. These results are in good agreement with the experimental findings that PA derivatives bearing negatively charged carboxylic or phosphonic acid substituents could not be polymerized in **Rh(nbd)·*apo*-Fr**,[12] due to their difficulties in penetration and accumulation around the hydrophobic region. It is worth noting that extrusion of the metal center from His49 in site B to be adsorbed in the hydrophobic site D needs to overcome some barrier, which is expected to be higher than 5.8 kcal/mol (the relative energy of **BI-2**). Only some of the metal centers in site B would be extruded to act as active sites for polymerization, depending on the concentration of the monomer. This again well explains the previous experimental observations that only 2−8 polymer chains can be produced inside *apo*-Ferritin, depending on the PA concentration, with molecular weight restricted by the size of the inner space of *apo*-Ferritin.[12]

We further calculated the energy of the Rh(nbd) complex in site D without the phenylacetylene substrate. Results show that without substrate, the structure of the Rh(nbd) complex in site D is higher in energy than in site B (BI-1) by 6.1 kcal/mol, indicating that the Rh complex prefers to stay in site B before addition of the substrate, which is exactly what was observed in the X-ray structure. Therefore, QM/MM studies and experimental observation together strongly suggest that a hydrophobic site D near site B, to which the dissociated rhodium center is attracted, rather than sites A, B, and C themselves initially presented by X-ray crystal structure, is the most plausible active site for polymerization of PA inside **Rh(nbd)·*apo*-Fr**. It should be noted that, as shown in Figure 11iii, there is a big hydrophobic area there for polymerization. We cannot totally exclude the possibility that other part of this hydrophobic area may also be able to host the dissociated Rh complex, leading to a similar mechanistic story. Site D is proposed to be the most possible docking site for Rh$^{\text{I}}$ complex, because it is closest to site B. Furthermore, the calculated binding energy of Rh complexes in sites B and D, as discussed above, strongly supports that the Rh complex prefers to dissociate from site B and dock to site D after coordination of monomer PA. Most importantly, our conclusion is in good agreement with the experimental observations that only a few active sites can dissociate from site B to the hydrophobic region to produce only several PPA chains. We carefully checked the electron density of the X-ray structure of **Rh(nbd)·*apo*-Fr** before and after polymerization; however, no direct information can be found due to the low density of the active sites inside **Rh(nbd)·*apo*-Fr**. This is reasonable because there are in total around 60 immobilized rhodium complexes inside the Ferritin, and very few of them can be dissociated from site B and serve as active sites.

It is also interesting to find out how the hydrophobic site D affects the polymerization reactivity inside the artificial enzyme, **Rh(nbd)·*apo*-Fr**. From DFT optimization of the QM model in the gas phase the intrinsic reaction barrier of $\mathrm{Rh^I}$ insertion with respect to the Rh(PA) complex **QI-3** is 6.8 kcal/mol (Figure S7 and S8, Supporting Information). $\mathrm{Rh^I}$ insertion in site D of **Rh(nbd)·*apo*-Fr** is calculated to be 14.4 kcal/mol, higher than the intrinsic barrier of $\mathrm{Rh^I}$ insertion in the gas phase. Actually, from **DI-3** to **DI-TS$_\mathrm{D}$**, the deformation energy of the QM part in our QM/MM calculation is 7.6 kcal/mol (Figure 12), which is very similar to the intrinsic barrier of the QM model in the gas phase. The other major energy components for the insertion barrier in site D are electronic and polarization effects (2.3 and 4.6 kcal/mol, respectively). Therefore, the overall barrier becomes higher than the intrinsic insertion barrier without the protein environment. On the other hand, the vdW interaction changes much less from reaction complex **DI-3** to transition state **DI-TS$_\mathrm{D}$**, contributing only 0.6 kcal/mol to the insertion barrier.

3.5.2. Termination Process. To provide further insight into the polymerization behavior and further evaluate the feasibility of the proposed active site in the hydrophobic pocket, ONIOM-EE studies were also carried out to investigate the termination process in active site D. There are two possible ways for active site D to release the polymer chain.[15] One is direct hydrogen transfer from the coordinated PA to the propagation chain (Figure 13). The other is oxidative addition of PA to form a $\mathrm{Rh^{III}}$−hydride intermediate, which subsequently undergoes reductive elimination between the hydride and the propagation chain to release the polymer. ONIOM-EE PES profiles of the termination process in site D are shown in Figure 13, with ONIOM-EE-optimized structures depicted in Figure S9, Supporting Information. β-H elimination is a commonly observed termination process for olefin polymerization. In the present case, the β-hydrogen highlighted in red in **DI-3** in Figure 13 is a model atom to represent the polymer chain, which should be repeated phenylethenyl units. However, the only β-hydrogen in the propagation chain is trans to the metal center, which is difficult to be abstracted due to the nonrotatable double bond. Therefore, β-H elimination is a less plausible termination process for polymerization of PA.

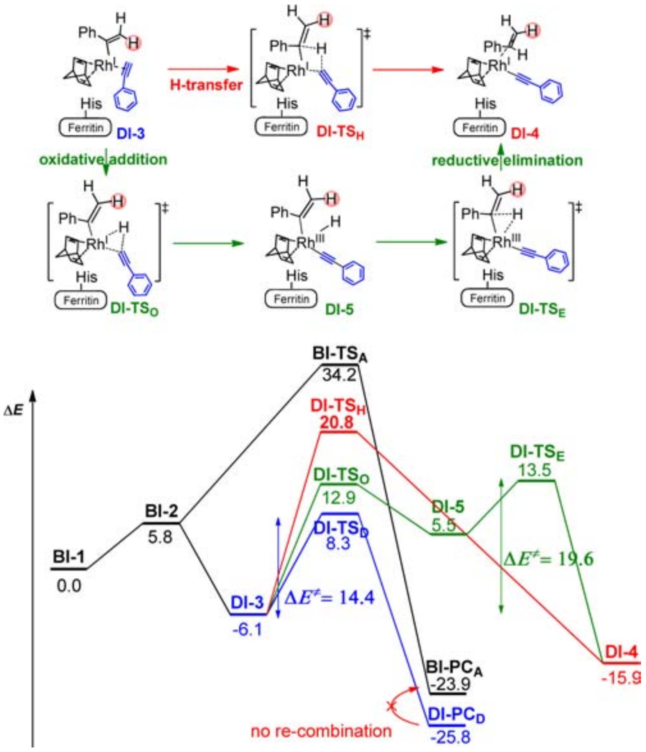


Figure 13. ONIOM-EE PES profiles (with ZPC in kcal/mol) of the termination process for polymerization in active site D.

The ONIOM-EE profiles indicate that the H-transfer transition state **DI-TS$_\mathrm{H}$** has a reaction barrier of 26.9 kcal/mol relative to **DI-3**, which is much higher than the dissociative $\mathrm{Rh^I}$-insertion barrier (**DI-TS$_\mathrm{D}$**, $\Delta E^\ddagger$ = 14.4 kcal/mol). Oxidative addition of PA via **DI-TS$_\mathrm{O}$** is easier with a reaction barrier of 19.0 kcal/mol. Subsequent reductive elimination from **DI-5** to **DI-4** via **DI-TS$_\mathrm{E}$** is calculated to have a barrier of 19.6 kcal/mol, relative to **DI-3**. The polymer chain in **DI-4** then can be released by exchange with solvent or monomer. This oxidative-addition/reductive-elimination pathway still cannot effectively compete with the insertion process, making the growth process more favorable to the termination process. Looking back to the $\mathrm{Rh^I}$ insertion in site A ($\Delta E^\ddagger$ = 22.5 kcal/mol), the $\mathrm{Rh^I}$ metathesis in sites A and B ($\Delta E^\ddagger$ = 19.8 and 19.4 kcal/mol, respectively), and the $\mathrm{Rh^{III}}$ insertion in site C ($\Delta E^\ddagger$ = 34.2 kcal/mol), one can see that they will all suffer from the competition of the likely termination processes and cannot lead to successful polymerization. Consequently, the hydrophobic site D is further rationalized to be a plausible active site by evaluation of the termination process.

The size of the inner cavity of *apo*-Ferritin is assumed to play a key role in constraining the MW of the PPA molecule. Each layer of a helical screw PPA polymer is around 4 A containing ∼6 monomers.[15] Roughly speaking, there would be around 120 monomers, which is in good agreement with our findings that PPA has around 130 monomer each polymer. We expect that polymerization would terminate when the polymer fills the size of the cavity, which makes further propagation difficult. Unfortunately, the whole system with 24-meric protein and big polymer chains is too large for QM/MM calculations.

3.5.3. Overall Mechanism. The overall mechanism of the polymerization of PA in the artificial metalloenzyme system, **Rh(nbd)·*apo*-Fr**, can be proposed and summarized in Figure 14. After coordination of monomer PA to the $\mathrm{Rh^I}$ complex bonded to site B, the true active site for polymerization is generated by extrusion of the $\mathrm{Rh^I}$(PA) complex to a hydrophobic pocket nearby. Site D, which is located near the nonpolar 4-fold channel of *apo*-Ferritin, is suggested to be the most plausible active for polymerization as it is the closest hydrophobic pocket to site B that can hold the dissociated Rh complexes. The PA monomers, accumulated in the hydrophobic region, are coordinated with the metal center and then incorporated into the polymer chain via $\mathrm{Rh^I}$ insertion. After insertion, the $\mathrm{Rh^I}$ complex does not recombine to His49 of site B. The entire polymerization process takes place in the hydrophobic region on the interior surface of *apo*-Ferritin near

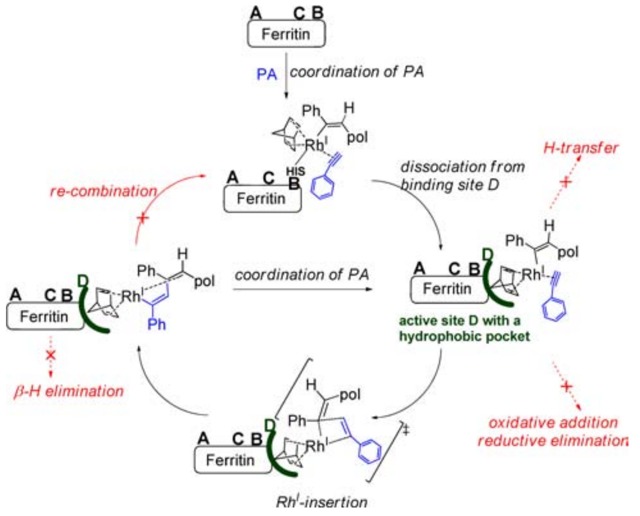


Figure 14. Proposed overall mechanism of polymerization of PA in the artificial metalloenzyme, **Rh(nbd)·*apo*-Fr**.

the 4-fold channel (Figure 11iii). Rh$^{\rm I}$ insertion is more favorable than the termination processes, e.g., H transfer, oxidative-addition/reductive-elimination, or β-H elimination, and successively propagates to yield PPA with restricted molecular weight in the constrained inner space of the *apo*-Ferritin.

3.5.4. Selectivity. Regio- and stereoselectivity is essential for synthesis of stereoregular PPA polymer. Rh complexes are reported to be able to generate highly stereoregular cis−trans PPA in solution, and our previous theoretical study confirmed this selectivity.[15] Inside **Rh(nbd)·*apo*-Fr**, the hydrophobic site D is found to play a key role in polymerization. Can polymerization of PA in **Rh(nbd)·*apo*-Fr** maintain the selectivity under the influence of the residues around this hydrophobic site? With this question in mind, further ONIOM-EE studies were performed to compare two different insertion pathways, 2,1-insertion and 1,2-insertion.

As shown in Figure 15, **DI-3** undergoes 2,1-insertion via **DI-TS$_{\rm D}$** to form a head-to-tail product **DI-PC$_{\rm D}$**. On the other hand, rotation of PA leads to isomer **DI-3$_{\rm iso}$**, which undergoes 1,2-insertion via transition state **DI-TS$_{\rm Diso}$** to form head-to-head product **DI-PC$_{\rm iso}$**. DFT optimization of the model system in the gas phase without *apo*-Ferritin found that 2,1-insertion is preferred over 1,2-insertion by 3.6 kcal/mol (Figure S10, Supporting Information). This selectivity resulted from the unique conjugative transition state for the 2,1-insertion.[15] Essentially the same feature can also be seen in **DI-TS$_{\rm D}$** (Figure 6) and **DI-TS$_{\rm Diso}$** (Figure 16i) structures, where the conjugation system of PA is essentially conserved in **DI-TS$_{\rm D}$** but broken in **DI-TS$_{\rm Diso}$**. The ONIOM-EE calculation predicts a preference of 2,1-insertion over 1,2-insertion by 2.3 kcal/mol (Figure 16ii)), which is slightly lower that the results of DFT optimization of the model system in the gas phase (3.6 kcal/mol) due to the influence of the protein environment. Further energy analysis of QM/MM results indicates that the contribution of the QM deformation is 3.1 kcal/mol, nearly the same as the DFT gas-phase results of the model system. The vdW, electrostatic, and polarization effects from the protein environment to the selectivity are relatively small. They are calculated to be −2.5, 1.3, and −0.6 kcal/mol, respectively. The overall influence of the *apo*-Ferritin environment retains well the selectivity during insertion. It can be expected that with a propagation chain bulkier than the model chain the selectivity would increase due to the steric effect.[15] Consequently, in the artificial metalloenzyme system, **Rh(nbd)·*apo*-Fr**, the micro-environment around active site D can retain the specific selectivity of polymerization, leading to head-to-tail cis−trans PPA, which was observed in the experimental studies.

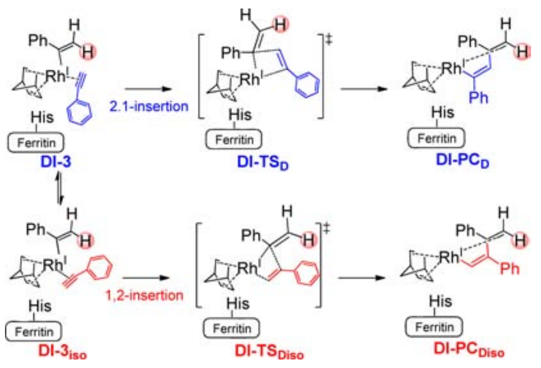


Figure 15. 2,1-Insertion and 1,2-insertion for polymerization of PA in active site D of **Rh(nbd)·*apo*-Fr**.

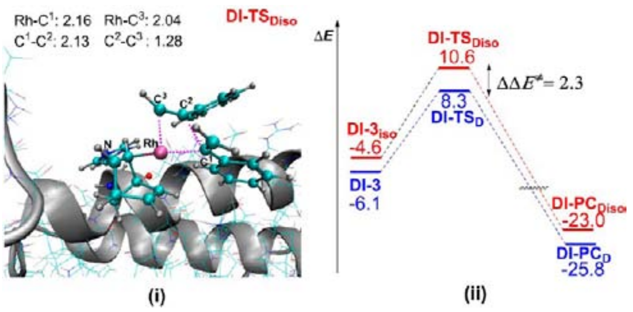


Figure 16. (i) Transition state structure for 1,2-insertion, and (ii) PES profiles (with ZPC in kcal/mol) for 2,1-insertion and 1,2-insertion for polymerization of PA in the hydrophobic active site D of **Rh(nbd)·*apo*-Fr**.

4. CONCLUSIONS

ONIOM-EE studies were performed to clarify the intricate polymerization mechanism of the recently developed artificial metalloenzyme system, **Rh(nbd)·*apo*-Fr**. We evaluated various possible mechanisms for all binding sites suggested by the X-ray crystal structure, i.e., sites A, B, and C. Our QM/MM studies found that sites A, B, and C themselves are only precursors/intermediates, not true active sites for polymerization of PA inside **Rh(nbd)·*apo*-Fr**. A new site D is suggested to be the most plausible active site for polymerization. This active site D is generated by extrusion of the Rh$^{\rm I}$(PA) complex to a hydrophobic pocket near site B after coordination of first monomer PA. Rh$^{\rm I}$ insertion is the most plausible polymerization mechanism. The metal center does not recombine to His49 of site B even after olefin insertion, remaining in the hydrophobic pocket for further propagation.

The hydrophobic pocket of site D, which is composed of hydrophobic types of residues, Phe50 and Leu170, and the hydrocarbon side chain of Lys143, plays a key role for the polymerization behavior inside **Rh(nbd)·*apo*-Fr**. The hydrophobic active site D, nonpolar 4-fold channel, and other hydrophobic residues nearby constitute a specific "hydrophobic region" for accumulation, coordination, and insertion of PA during polymerization. Notably, this hydrophobic active site D in the artificial metalloenzyme, **Rh(nbd)·*apo*-Fr**, well retains the regio- and stereoselectivity feature of Rh-catalyzed polymerization of PA found without the protein.

Our QM/MM study demonstrates an example of a complicated reaction mechanism of artificial metalloenzymes. Binding sites from X-ray crystal structures do not always guarantee the true active site for catalytic reaction of artificial metalloenzyme, especially when binding interactions of the metal complexes and protein can highly vary. Theoretical studies are essential in providing deeper insights to the active site and the catalytic behavior of artificial metalloenzymes. The binding and active sites are not always the same, which inspires us to believe that reactivity should be tuned based on the true active site and binding sites, on the other hand, may serve as a useful tool in improving the loading amount of catalyst. Furthermore, the hydrophobic or hydrophilic region of the protein may have a significant influence on catalysis design.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

Complete refs 22 and 37; figures containing detailed information about other important QM/MM or QM structures, other PES profiles, and Cartesian coordinates of QM atoms. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

morokuma@emory.edu

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Z.K. acknowledges the Fukui Institute Fellowship. The authors thank Drs. Miho Hatanaka, Hajime Hirao, and Fengyi Liu for helpful discussions. This work is in part supported by the Japan Science and Technology Agency (JST) with a Core Research for Evolutional Science and Technology (CREST) grant in the Area of High Performance Computing for Multiscale and Multiphysics Phenomena. Computer resources allocated at Research Center of Computer Science (RCCS) at the Institute for Molecular Science (IMS) are also acknowledged.

■ REFERENCES

(1) Ward, T. R. *Acc. Chem. Res.* **2011**, *44*, 47−57.

(2) Hillard, E. A.; Jaouen, G. *Organometallics* **2011**, *30*, 20−27.

(3) (a) Heinisch, T.; Ward, T. R. *Curr. Opin. Chem. Biol.* **2010**, *14*, 184−199. (b) Deuss, P. J.; den Heeten, R.; Laan, W.; Kamer, P. C. J. *Chem.—Eur. J.* **2011**, *17*, 4680−4698.

(4) Lu, Y.; Yeung, N.; Sieracki, N.; Marshall, N. M. *Nature* **2009**, *460*, 855−862.

(5) Abe, S.; Ueno, T.; Watanabe, Y. *Top. Organomet. Chem.* **2009**, *25*, 25−43.

(6) Reetz, M. T. *Top. Organomet. Chem.* **2009**, *25*, 63−92.

(7) Yeung, N.; Lin, Y. W.; Gao, Y. G.; Zhao, X.; Russell, B. S.; Lei, L. Y.; Miner, K. D.; Robinson, H.; Lu, Y. *Nature* **2009**, *462*, 1079−U144.

(8) Rothlisberger, D.; Khersonsky, O.; Wollacott, A. M.; Jiang, L.; DeChancie, J.; Betker, J.; Gallaher, J. L.; Althoff, E. A.; Zanghellini, A.; Dym, O.; Albeck, S.; Houk, K. N.; Tawfik, D. S.; Baker, D. *Nature* **2008**, *453*, 190−U4.

(9) Jiang, L.; Althoff, E. A.; Clemente, F. R.; Doyle, L.; Rothlisberger, D.; Zanghellini, A.; Gallaher, J. L.; Betker, J. L.; Tanaka, F.; Barbas, C. F.; Hilvert, D.; Houk, K. N.; Stoddard, B. L.; Baker, D. *Science* **2008**, *319*, 1387−1391.

(10) Qi, D. F.; Tann, C. M.; Haring, D.; Distefano, M. D. *Chem. Rev.* **2001**, *101*, 3081−3111.

(11) Ringenberg, M. R.; Ward, T. R. *Chem Commun* **2011**, *47*, 8470−8476.

(12) Abe, S.; Hirata, K.; Ueno, T.; Morino, K.; Shimizu, N.; Yamamoto, M.; Takata, M.; Yashima, E.; Watanabe, Y. *J. Am. Chem. Soc.* **2009**, *131*, 6958−6960.

(13) Meldrum, F. C.; Wade, V. J.; Nimmo, D. L.; Heywood, B. R.; Mann, S. *Nature* **1991**, *349*, 684−687.

(14) Liu, X. F.; Theil, E. C. *Acc. Chem. Res.* **2005**, *38*, 167−175.

(15) Ke, Z.; Abe, S.; Ueno, T.; Morokuma, K. *J. Am. Chem. Soc.* **2011**, *133*, 7926−7941.

(16) Lynam, J. M. *Chem.—Eur. J.* **2010**, *16*, 8238−8247.

(17) Vastine, B. A.; Hall, M. B. *Organometallics* **2008**, *27*, 4325−4333.

(18) Cowley, M. J.; Lynam, J. M.; Slattery, J. M. *Dalton Trans.* **2008**, 4552−4554.

(19) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graph.* **1996**, *14*, 33−&.

(20) Li, H.; Robertson, A. D.; Jensen, J. H. *Proteins: Struct., Funct., Bioinf.* **2005**, *61*, 704−721.

(21) Guex, N.; Peitsch, M. C. *Electrophoresis* **1997**, *18*, 2714−2723.

(22) Case, D. A.; et al. *Amber 9*; University of California: San Francisco, 2006.

(23) Liu, J. Z.; Lam, J. W. Y.; Tang, B. Z. *Chem. Rev.* **2009**, *109*, 5799−5867.

(24) Mayershofer, M. G.; Nuyken, O. *J. Polym. Sci., Part A: Polym. Chem.* **2005**, *43*, 5723−5747.

(25) Lin, H.; Truhlar, D. G. *Theor. Chem. Acc.* **2007**, *117*, 185−199.

(26) Senn, H. M.; Thiel, W. *Angew. Chem., Int. Ed.* **2009**, *48*, 1198−1229.

(27) Menikarachchi, L. C.; Gascon, J. A. *Curr. Top. Med. Chem.* **2010**, *10*, 46−54.

(28) Maseras, F.; Morokuma, K. *J. Comput. Chem.* **1995**, *16*, 1170−1179.

(29) Svensson, M.; Humbel, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 3654−3661.

(30) Svensson, M.; Humbel, S.; Froese, R. D. J.; Matsubara, T.; Sieber, S.; Morokuma, K. *J. Phys. Chem.* **1996**, *100*, 19357−19363.

(31) Matsubara, T.; Sieber, S.; Morokuma, K. *Int. J. Quantum Chem.* **1996**, *60*, 1101−1109.

(32) Humbel, S.; Sieber, S.; Morokuma, K. *J. Chem. Phys.* **1996**, *105*, 1959−1967.

(33) Dapprich, S.; Komáromi, I.; Byun, S.; Morokuma, K.; Frisch, M. J. *J. Mol. Struct. (THEOCHEM)* **1999**, *461*, 1−21.

(34) Vreven, T.; Morokuma, K.; Farkas, O.; Schlegel, H. B.; Frisch, M. J. *J. Comput. Chem.* **2003**, *24*, 760−769.

(35) Vreven, T.; Byun, K. S.; Komáromi, I.; Dapprich, S., Jr.; J., A. M.; Morokuma, K.; Frisch, M. J. *J. Chem. Theory Comput.* **2006**, *2*, 815−826.

(36) Lundberg, M.; Kawatsu, T.; Vreven, T.; Frisch, M. J.; Morokuma, K. *J. Chem. Theory Comput.* **2009**, *5*, 222−234.

(37) Frisch, M. J.; et al. *Gaussian 09*, Revision A.02; Gaussian, Inc.: Wallingford, CT, 2009.

(38) Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200−1211.

(39) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.

(40) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.

(41) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.

(42) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.

(43) Dunbar, K. R.; Kang, M.; Chifotides, H. T. *Biochemistry* **2008**, *47*, 2265−2276.

(44) Itazaki, M.; Yoda, C.; Nishihara, Y.; Osakada, K. *Organometallics* **2004**, *23*, 5402.

(45) Laubender, M.; Werner, H. *Angew. Chem., Int. Ed. Engl.* **1998**, *37*, 150.
(46) Werner, H.; Wiedemann, R.; Steinert, P.; Wolf, J. *Chem.—Eur. J.* **1997**, *3*, 127.
(47) Pechmann, T.; Brandt, C. D.; Werner, H. *Organometallics* **2003**, *22*, 3004.
(48) Werner, H.; Schwab, P.; Bleuel, E.; Mahr, N.; Steinert, P.; Wolf, J. *Chem.—Eur. J.* **1997**, *3*, 1375.
(49) Garcia, M. P.; Jimenez, M. V.; Lahoz, F. J.; Lopez, J. A.; Oro, L. A. *J. Chem. Soc., Dalton Trans.* **1998**, 4211.
(50) Yang, X. K.; Chasteen, N. D. *Biophys. J.* **1996**, *71*, 1587−1595.
(51) Thei, E. C. *Curr. Opin. Chem. Biol.* **2011**, *15*, 304−311.
(52) Bou-Abdallah, F.; Zhao, G.; Biasiotto, G.; Poli, M.; Arosio, P.; Chasteen, N. D. *J. Am. Chem. Soc.* **2008**, *130*, 17801−17811.