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

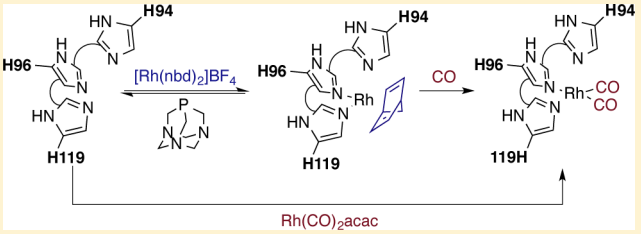
Generation, Characterization, and Tunable Reactivity of Organometallic Fragments Bound to a Protein Ligand

Hanna M. Key,[†,‡] Douglas S. Clark,[†,§] and John F. Hartwig*[,†,‡]

[†]Department of Chemistry, University of California, Berkeley, California 94720, United States,
[‡]Chemical Sciences Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, California 94720, United States,
[§]Department of Chemical and Biomolecular Engineering, Physical and Biological Sciences Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, California 94720, United States

Supporting Information

ABSTRACT: Organotransition metal complexes catalyze important synthetic transformations, and the development of these systems has rested on the detailed understanding of the structures and elementary reactions of discrete organometallic complexes bound to organic ligands. One strategy for the creation of new organometallic systems is to exploit the intricate and highly structured ligands found in natural metalloproteins. We report the preparation and characterization of discrete rhodium and iridium fragments bound site-specifically in a κ^2–fashion to the protein carbonic anhydrase as a ligand. The reactions of apo human carbonic anhydrase with $[Rh(nbd)_2]BF_4$ or $[M(CO)_2(acac)]$ (M=Rh, Ir) form proteins containing Rh or Ir with organometallic ligands. A colorimetric assay was developed to quantify rapidly the metal occupancy at the native metal-binding site, and ^{15}N-^{1}H NMR spectroscopy was used to establish the amino acids to which the metal is bound. IR spectroscopy and EXAFS revealed the presence and number of carbonyl ligands and the number total ligands, while UV-vis spectroscopy provided a signature to readily identify species that had been fully characterized. Exploiting these methods, we observed fundamental stoichiometric reactions of the artificial organometallic site of this protein, including reactions that simultaneously form and cleave metal-carbon bonds. The preparation and reactivity of these artificial organometallic proteins demonstrate the potential to study a new genre of organometallic complexes for which the rates and outcomes of organometallic reactions can be controlled by genetic manipulation of the protein scaffold.



INTRODUCTION

Transition metal complexes catalyze many important chemical reactions used to prepare molecules ranging from small commodity chemicals to complex, medicinally important natural products.[1] Fundamental studies on the synthesis and reactivity of organometallic complexes have created the framework on which new catalytic processes are developed. These studies have shown that the ligands in these catalysts modulate reactivity and selectivity via the electronic properties of the donor atom(s) and the particular steric environment that the ligand creates around the metal. Consequently, the expanding utility of this class of catalyst is inseparable from the development and iterative modification of organic ligands.

An alternative class of organometallic complexes to those containing ligands based on small molecules could exploit the intricate and highly structured coordination sphere of natural metalloproteins. Unlike many organic ligands, diverse proteins that bind metals can be prepared conveniently in a single step via recombinant expression, consistently purified in a single step by affinity chromatography, and rapidly derivatized by a variety of well-established mutagenesis techniques.[2] Due to these properties, protein-derived ligands could enable a strategy to control organometallic reactivity that complements that based on the more typical selection of small organic ligands. Ultimately, this strategy would include the efficient, directed evolution of a protein ligand most suited for the desired substrate and transformation.[3]

Organometallic complexes containing abiotic metals directly coordinated to proteins are a subset of the growing field of artificial metalloenzymes.[4,5] Most frequently, artificial metalloenzymes have been prepared by a strategy in which a known transition metal complex is anchored within a protein scaffold in a covalent or supramolecular fashion.[6−9] This strategy is distinguished by the fact that complete complexes of known catalytic activity are incorporated into the system; however, in this case the native binding site of the protein must be sufficiently large to accommodate an entire organometallic complex, and although the binding or attachment site of an anchoring group may be well-defined, the catalytic metal center does not necessarily fall within the native binding pocket of the protein host. Thus, the application of apo metalloproteins as ancillary ligands for organometallic complexes is a comple-

Received: April 29, 2015
Published: May 28, 2015

DOI: 10.1021/jacs.5b04431
J. Am. Chem. Soc. 2015, 137, 8261−8268

mentary tactic in which the artificial metal is positioned directly in the native binding pocket of the protein.

Although the number of examples in which metal complexes ligate a protein as the primary ancillary ligand and the resulting complex serves as a catalyst are limited, these reports nevertheless validate that protein-derived ligands have the potential to impose selectivity in catalytic reactions.[10−12] In one case, apo-ferritin has been shown to bind approximately 58 Rh fragments and the resulting protein to catalyze size-selective polymerization of phenyl acetylene.[10a] In another case, apo-carbonic anhydrase II was reconstituted with manganese to create an enantioselective artificial epoxidase.[11a] The same approach was then used to create carbonic anhydrase analogues containing rhodium in place of zinc. The resulting protein is reported to catalyze hydrogenation and regioselective hydroformylation reactions.[11b,c] Compared to ferritin, apo-carbonic anhydrase is simpler because the natural protein binds a single metal center; however, the binding location and coordination environment of the Rh center in the carbonic anhydrase analogue were not directly characterized in these studies.[11b,c] In fact, the complete, solution-state characterization of the metal center of a discrete, artificial organometallic protein has never been reported.

To create the foundation for further development of such systems, it is essential to synthesize discrete organometallic proteins and to fully characterize the number of metal fragments bound to the protein, the site to which the protein is bound, and the number and identity of the amino acids bound to the metal. If well-characterized complexes can be generated, the elementary reactivity of protein-ligated metal centers can then be observed, understood, and modulated to create the type of fundamental information that has fueled the development of small-molecule organometallic catalysts.

The synthesis and characterization of such an organometallic protein is challenging. Preparing such complexes in high yield, determining which amino acid residues bind the metal, and monitoring the reactions of small molecules at the site created within a large protein require the judicious combination of methods used to characterize both organometallic complexes and proteins. Indeed, stoichiometric reactions of an abiological organometallic center bound to a protein ligand have not been observed directly.

We report the efficient synthesis and characterization of a series of discrete, well-defined organometallic complexes that ligate carbonic anhydrase through two of the three histidines of the native Zn-binding site. Methods to determine the extent and location of metal binding in a general fashion have been demonstrated, and the conversion of one organometallic protein to a second organometallic protein by the simultaneous cleavage and formation of metal−carbon bonds has been observed directly using these methods. Furthermore, we show that the structure of the protein surrounding the organometallic center controls the rates and products of these reactions in a manner that would be challenging to achieve using small-molecule ligands.

RESULTS AND DISCUSSION

Synthesis of Discrete Organometallic Proteins. To generate protein-derived ancillary ligands suitable for coordinating organometallic fragments in a defined manner, we followed the documented approach of removing native metals from natural metalloproteins (Figure 1).[11] The ligand substitution reaction of the resulting apo protein with an

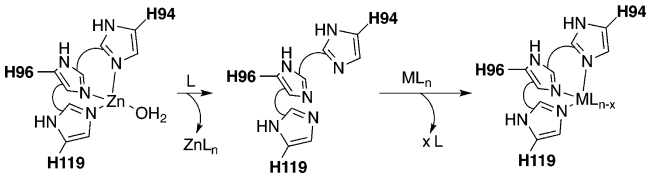


Figure 1. Standard method used to remove the native Zn fragment from CA.

organometallic fragment presents a series of challenges that are not encountered when conducting analogous preparations of small organometallic complexes. In particular, the separation of apo and metal-bound proteins and the separation of proteins in which the metal is bound to different sites are impractical. Therefore, the generation of discrete artificial organometallic proteins requires high-yielding and site-specific metalations. To develop such a reaction, efficient methods are needed to identify the metal precursor, reaction conditions, and protein mutants that form single complexes in high yield. X-ray crystallography, non-native mass spectrometry, and metal analysis are the predominant methods used to characterize artificial metalloenzymes,[4,5] but these methods, respectively, are too slow to evaluate the yields for formation of potential organometallic proteins and reactions of them, release the metal due to denaturation, and are silent concerning the binding site. Furthermore, X-ray crystallography cannot be used to characterize metalloprotein species that are unstable over time, such as the Rh proteins described in this study. Therefore, we sought alternative, solution-state methods that efficiently reveal the site-specificity of metal binding.

Initially, we explored methods to generate and characterize protein-ligated transition metal complexes using a mutant of human carbonic anhydrase II (CA). CA is an enzyme that binds zinc through three histidines,[13] and these strongly coordinating imidazole units could mimic classic polydentate ligands of organometallic complexes once Zn is removed to generate the apo-CA.[14] The mutant we studied (3*His-CA) lacks three surface histidine residues (Supporting Information Figure S1, Table S1) that could interfere with binding of the organometallic fragment to the original Zn binding site.[11b,c]

Although the combination of Rh precursors and apo-carbonic anhydrases has been previously reported,[11b,c] the metalation protocol that was employed is not suitable for efficient evaluation of a wide range of metal precursors. Previously, apo-carbonic anhydrase was dialyzed against a buffer solution that contained orders of magnitude excess rhodium compound. The elimination of excess Rh metal from the buffer and the de-coordination of Rh bound outside the Zn site were achieved by subsequent dialysis against Rh-free buffer. Beyond the impracticality of employing such a large excess of precious metal complexes, this method is limited to metal precursors with high water solubility and to those for which site-selective de-coordination of ligated metal fragments is possible.

To create an alternative, more practical, and general method to form organometallic proteins rapidly, we evaluated direct metalations involving the addition of nearly stoichiometric amounts of an organic stock solution of the appropriate rhodium precursor to a buffered solution of apo-CA.

To determine the selectivity of the noble metal fragments for coordination to the native Zn-binding site of CA, rather than the protein surface, we exploited the known hydrolysis of 4-nitrophenyl acetate catalyzed by cobalt-CA to produce 4-

DOI: 10.1021/jacs.5b04431
J. Am. Chem. Soc. 2015, 137, 8261−8268

nitrophenol (λ_{max} = 405 nm) (Figures S2 and S3).[15] Following the addition of a noble metal complex to the apo form of 3*His-CA (apo-3*His-CA), $CoCl_2$ was added to convert the remaining apo-CA to Co-CA. The amount of protein with noble metal coordinated at the Zn site would then correspond to the difference between the amount of apo-CA at the start of the reaction and the amount of Co-CA determined by the hydrolysis assay. This assay can be conducted with only 10 μg of protein in a 96-well format in 5 min, allowing the rapid assessment of different metal precursors, reaction stoichiometries, organic so-solvents, buffers, and other parameters. This method for assessing metal binding should be applicable to many other metalloprotein scaffolds for which activity or metal binding can be evaluated with a colorimetric assay.[16]

With this assay, we evaluated the binding of a series of 16-electron Rh(I) precursors for which small-molecule organometallic chemistry has been established and ligand substitution is often facile.[17] The addition of 1 equiv of $[Rh(cod)_2]BF_4$ (cod = cyclooctadiene) to apo-3*His-CA led to 45% Rh occupancy of the native Zn-binding site (Figures 2a, S4, and S5).

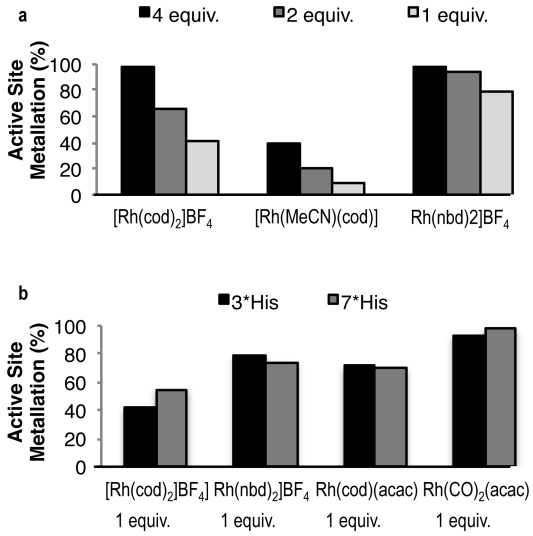


Figure 2. Preparation of organometallic proteins from carbonic anhydrase. (a) Comparison of the specificity of active-site metalation of Rh(I) precursors upon the addition of 1−4 equiv of Rh to apo-3*His-CA (67 μM) in 50 mM MES buffer, pH 5.3. Metalation of the active site was determined using the activity assay described in the text (and see Supporting Information Figure S2). (b) Comparison of the selectivity of metalation for the addition of 1 equiv of cationic and neutral rhodium sources.

Inductively coupled plasma optical emission spectroscopy (ICP-OES) revealed an average of 0.94 Rh coordinated per CA; thus, about half of this rhodium complex binds to the Zn site, and half binds to other sites on the protein. The addition of superstoichiometric $[Rh(cod)_2]BF_4$ led to higher Rh occupancy of the native zinc binding site but resulted in concomitant binding of the excess Rh elsewhere on the protein (Figures 2a and S5).[18] Together, these methods revealed that $[Rh(COD)_2]BF_4$ undergoes ligand substitution with the apo-CA ligand, but that this reaction does not proceed with sufficient site specificity to form a discrete organometallic protein. Therefore, we sought alternative conditions to achieve more selective ligand substitution reactions that form Rh(bis-olefin)-CA complexes in high yield.

Although the buffer and the surface amino acid composition had minor effects on the selectivity of this reaction (Figure S6, S7), the ligands on rhodium did affect significantly the ligand substitution reaction that binds the metal precursor to the Zn site of the CA (Figure 2a). The reaction of $[Rh(cod)(MeCN)_2]BF_4$ with apo-3*His-CA resulted in low occupancy of the Zn site, likely because this cationic precursor with labile nitrile ligands binds predominantly to the protein surface. However, the reaction of 1 equiv of $[Rh(nbd)_2]BF_4$ with the same protein resulted in 79% occupation of the Zn site. Metalation of apo-3*His-CA with 1 equiv of the neutral complexes Rh(cod)(acac) (acac = acetylacetonate) and $Rh(CO)_2(acac)$, which contain dative ligands that do not dissociate readily, gave 71% and >90% yields of protein containing rhodium in the original Zn site, respectively (Figures 2b and S8). ICP-OES revealed that addition of 1 equiv of either $[Rh(nbd)_2]BF_4$ or $Rh(CO)_2(acac)$ to apo-CA resulted in the formation of a Rh-CA containing a Rh:protein ratio of approximately 1:1 (Figure S5). Thus, a vast majority of these two rhodium complexes bind to the original Zn site, with little bound elsewhere on the protein.

Following this comparison of the selectivity of Rh precursors for coordination to the Zn site, the resulting Rh-CA complexes were compared by UV−vis spectroscopy. The UV−vis spectrum of the Rh-CA prepared from $[Rh(nbd)_2]BF_4$ contained an absorption band at 410 nm, which is much different from the absorption of free $[Rh(nbd)_2]BF_4$ and distinguishable from the absorption of the Rh-CA generated from $[Rh(cod)_2]BF_4$ (Figures 3 and S4). The inequivalent

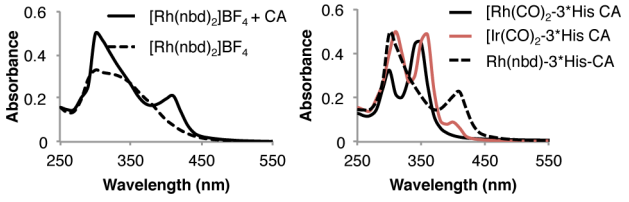


Figure 3. Characterization of [M]-CA products by UV−vis and $^{15}N-^{1}H$ HSQC spectroscopy. (a) UV−vis spectrum of the product formed from addition of 1 equiv of $[Rh(nbd)_2]BF_4$ to apo-CA, compared to the spectrum of $[Rh(nbd)_2]BF_4$ in the same buffer. (b) Comparison of the UV−vis spectra of Rh(nbd)-CA, $Rh(CO)_2$-CA, and $Ir(CO)_2$-CA.

spectra for the proteins generated from $[Rh(nbd)_2]BF_4$ and $[Rh(cod)_2]BF_4$ are likely due to slightly different diene ligands at the Rh center. The UV−vis spectrum of the Rh-CA protein generated from Rh(cod)(acac) matched that of the Rh-CA generated from $[Rh(cod)_2]BF_4$, consistent with generation of the same Rh(COD) fragment at the Zn site with these two precursors (Figure S4). In contrast to the low selectivity of metalation with the previously reported $[Rh(cod)_2]BF_4$ precursor, ligand substitution at the Zn site with Rh(cod)(acac) occurs with high selectivity. The UV−vis spectrum of the Rh-CA resulting from addition of $Rh(CO)_2(acac)$ is distinct from those of the other complexes (Figure 3).

Together, these operationally simple and broadly applicable methods to conduct and characterize ligand substitution reactions between apo metalloproteins and noble metal complexes enabled us to identify Rh precursors that react with apo-CA site-specifically and in high yield.

Characterization of the Protein−Metal Coordination Sphere. To characterize the coordination sphere of these

DOI: 10.1021/jacs.5b04431
J. Am. Chem. Soc. 2015, 137, 8261−8268

protein-ligated organometallic complexes fully and to enable monitoring of the reactivity of these complexes with organic substrates, we sought additional direct solution-state characterization of the protein-ligated metal center. Because the metal is likely to be bound by histidines, we used ^{15}N NMR spectroscopy to evaluate the identity and number of these amino acids bound to rhodium, much like one uses ^{31}P NMR spectroscopy to characterize phosphine-ligated metal complexes. Prior work has shown that ^{15}N–^{1}H HSQC spectroscopy can distinguish Zn-CA from apo-CA.[19]

A combination of ^{15}N–^{1}H HSQC NMR experiments revealed the coordination site of the rhodium bound to the CA mutants. One-bond HSQC spectra of ^{15}N-labeled CA required short experiment times with only 0.2 mM protein. However, this method detects only nitrogen atoms for which exchanges of the N–H protons are slower than the NMR time scale, and it does not detect the histidine nitrogens bound to rhodium (because of the absence of a proton bound to this nitrogen). Two-bond HSQC, which requires higher concentrations and longer experiment times, detects all histidine nitrogen atoms, with or without a bound proton. Thus, we used the one-bond HSQC experiment to monitor reactions and the two-bond HSQC experiment to fully characterize the products.

The site at which the rhodium fragment binds was revealed by the differences in chemical shifts between the ^{15}N–^{1}H HSQC spectra of 3*His-apo-CA and the analogous spectra of the product from addition of the Rh precursors to this protein. The one-bond HSQC spectrum of apo-3*His-CA contains resonances for four of the nine histidine residues (Figure 4). By

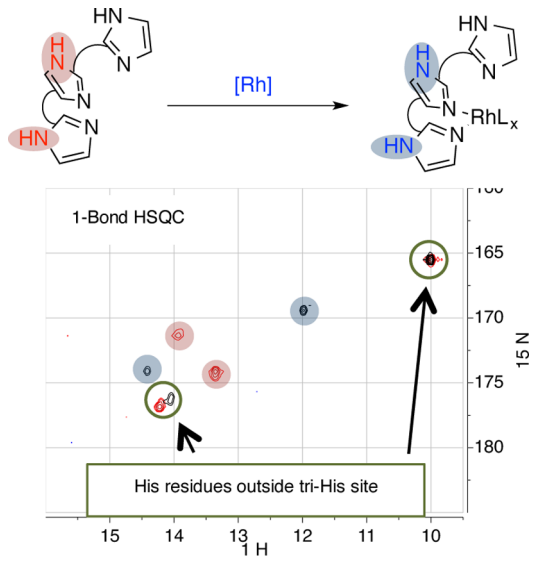


Figure 4. Comparison of the one-bond ^{15}N–^{1}H HSQC spectra of apo-3*His-CA (red) and the Rh-CA formed from the addition of [Rh(nbd)$_2$]BF$_4$ to 3*His-apo-CA (blue). Only N–H nitrogens that do not rapidly exchange (highlighted) are observed in this experiment.

comparing this spectrum to that of 3*His-Zn-CA, we determined that two of the four observed signals correspond to histidines in the tri-His site that binds zinc, and two correspond to histidines outside the tri-His site. Thus, if rhodium binds to the tri-His site, the chemical shift of the two tri-His signals in apo-3*His-CA should change upon addition of the rhodium complex, while those of the two signals from histidines outside the tri-His site should not change. Indeed, addition of 1 equiv of [Rh(nbd)$_2$]BF$_4$ to apo-3*His-CA caused only the chemical shifts of the signals corresponding to histidines in the tri-His site to change (Figures 4 and S9). This result provides strong evidence that the Rh fragment binds to the tri-His site of CA.

The two Rh-CA proteins generated by addition of [Rh(nbd)$_2$]BF$_4$ and Rh(CO)$_2$(acac) to apo-3*His-CA were also distinguishable by one-bond HSQC spectroscopy. The addition of 1 equiv of Rh(CO)$_2$(acac) to apo-3*His-CA generated a Rh-3*His-CA protein with a ^{15}N–^{1}H HSQC spectrum that is distinct from that of the Rh-3*His-CA generated from [Rh(nbd)$_2$]BF$_4$, as well as that of apo-3*His-CA (Figure 5). These data corroborate our conclusion from UV–vis spectroscopy that these two Rh precursors generate different Rh-CA proteins. Furthermore, the ability to distinguish among different Rh-CA's rapidly in solution-state samples using a protein concentration of 0.2 mM makes this method directly applicable to the monitoring of catalytic reactions in situ, because catalysts are commonly used at this concentration or higher.

The identity of the Rh-bound nitrogens and the coordination number of the Rh in the organometallic protein were revealed by two-bond ^{15}N–^{1}H HSQC spectroscopy. To reduce the total number of histidine signals in the two-bond HSQC spectrum, we acquired HSQC data of the mutant 6*His-CA (that lacks six surface histidine residues) instead of 3*His-CA that lacks three (3*His-CA, Table S1). The two-bond ^{15}N–^{1}H HSQC spectrum of the product from addition of 1 equiv of [Rh(nbd)$_2$]BF$_4$ to apo-6*His-CA contains ^{15}N NMR signals corresponding to the four nitrogens in the two histidines bound to rhodium (Figure 6). The ^{15}N NMR chemical shifts corresponding to two of the nitrogens lie near the region in which Zn-bound nitrogens resonate in the ^{15}N NMR spectrum of Zn-CA, and the signals corresponding to the other two nitrogen atoms lie in the regions in which unligated histidines resonate. The spectrum of the product from addition of Rh(CO)$_2$(acac) to apo-6*His also contained signals corresponding to two metalated nitrogens (Figure S10). From these data, we conclude that the rhodium is bound by two histidine residues of a κ^2-CA ligand in both Rh-CA's.

To determine which two His residues ligate Rh, we prepared Rh-CA complexes from the three mutants, 3*His-H94A-CA, 3*His-H96A-CA, and 3*His-H119A-CA, each of which lacks one His residue in the metal-binding site. The UV–vis spectra of the proteins formed from adding [Rh(nbd)$_2$]BF$_4$ or Rh(CO)$_2$(acac) to the three mutants are similar to each other and to those obtained from the apo-3*His protein (Figures S11 and S12). Thus, any pair of histidine residues from the original Zn-binding site is capable of coordinating the Rh fragment as a κ^2-CA ligand. However, the resonances in the ^{15}N–^{1}H HSQC spectra of the two proteins, Rh-3*His-H94A and Rh-3*His-CA, formed from addition of Rh(CO)$_2$(acac) are identical to each other, while the resonances of the analogous proteins, Rh-3*His-H96A-CA and Rh-3*His-H119A-CA, are different from those of Rh-3*His-CA (Figures 5b and S13). These data show that H96 and H119 are the two histidine residues that coordinate Rh within the native tri-histidine site.

To reveal the small-molecule ligands bound to rhodium in the Rh-3*His-CA proteins, we used native nanoESI-MS. In the mass spectrum of the Rh-CA generated from [Rh(nbd)$_2$]BF$_4$, a mass corresponding to Rh(nbd)$^+$ is observed, indicating that the Rh site contains one diene and one κ^2-CA ligand (Figure S14). In the mass spectrum of the Rh-CA generated from Rh(CO)$_2$(acac), masses corresponding to both Rh(CO)-CA

DOI: 10.1021/jacs.5b04431
J. Am. Chem. Soc. 2015, 137, 8261–8268

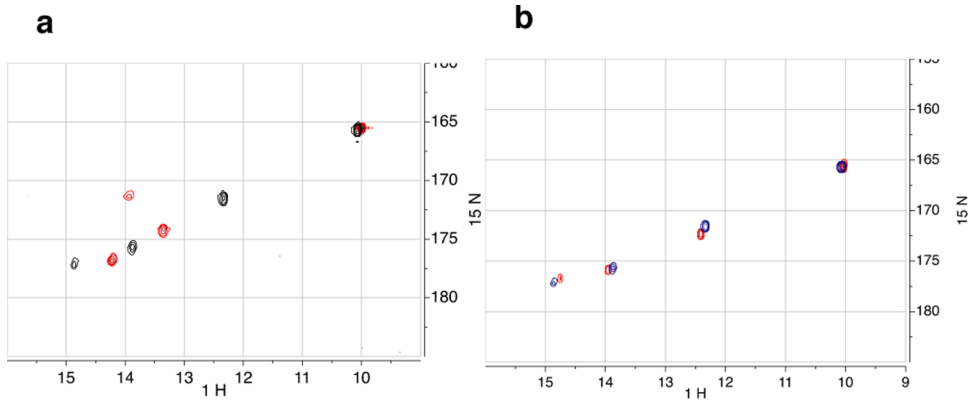


Figure 5. Characterization of the product of the reaction of apo-3*His-CA with $Rh(CO)_2(acac)$ by one-bond $^{15}N-^{1}H$ HSQC. (a) One-bond $^{15}N-^{1}H$ HSQC spectrum of apo-3*His-CA (red) and the spectra of the Rh-CA formed from the addition of $Rh(CO)_2(acac)$ to 3*His-apo-CA (blue). (b) One-bond $^{15}N-^{1}H$ HSQC spectra of $Rh(CO)_2$-3*His-CA (red) and $Rh(CO)_2$-H94A-3*His-CA (blue).

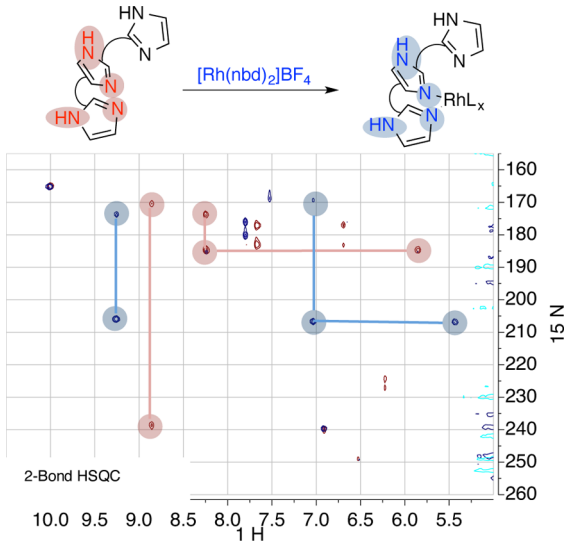


Figure 6. Comparison of the two-bond $^{15}N-^{1}H$ HSQC spectra of apo-6*His-CA (red) and the Rh-CA formed from the addition of $[Rh(nbd)_2]BF_4$ to 6*His-apo-CA (blue). As a result of metalation, the signals from the histidine residues of the tri-His site in apo-6*His-CA are different from those in the Rh-CA (highlighted).

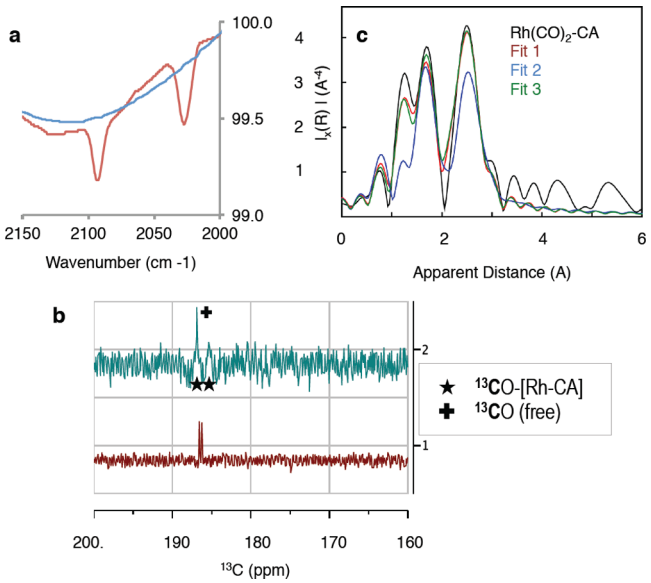


Figure 7. Characterization of $Rh(CO)_2$-CA. (a) Comparison of the IR spectra of $Rh(CO)_2$-CA (red) and apo-CA (blue). (b) ^{13}C NMR spectra of $Rh(^{13}CO)_2$-CA (top) and $Rh(^{13}CO)_2(acac)$ (bottom). (c) Comparison of the EXAFS of $Rh(CO)_2$-CA to those of the three model complexes, $Rh(His)_2(CO)_2$ (Fit 1), $Rh(His)_3(CO)_1$ (Fit 2), and $Rh(His)_3(CO)_2$ (Fit 3).

and $Rh(CO)_2$-CA were observed. These observed ions could result from a mixture of products or the dissociation of one CO ligand during the mass spectrometry (Figure S14).

To clarify the number of CO ligands in the $Rh(CO)_x$-CA protein, IR, EXAFS, and ^{13}C NMR spectroscopy were performed. The IR spectrum contains two CO stretching frequencies, consistent with the symmetric and antisymmetric stretches of two CO ligands in equivalent or nearly equivalent environments (Figure 7a), and the ^{13}C NMR spectrum contains two ^{13}CO signals (Figure 7b). The EXAFS spectrum of the sample matched the predicted spectrum of Rh-(histidine)$_2(CO)_2$ and poorly matched the predicted spectrum of Rh(histidine)$_3$(CO) (Figure 7c, Table S2). Together, these data indicate that the structure of the product of the reaction of apo-CA with $Rh(CO)_2(acac)$ is $Rh(CO)_2(\kappa^2$-CA), in which the two CO ligands lie in similar but chemically inequivalent environments.

To assess the generality of these synthetic and analytical methods, we studied the reactions of apo-CA with $[Ir(cod)_2]$-BF_4 and $Ir(CO)_2(acac)$. As observed for the Rh analogue, the reaction of $[Ir(cod)_2]BF_4$ with apo-CA led to low occupancy of the metal-binding site, whereas the reaction of apo-CA with $Ir(CO)_2(acac)$ led to >95% yield of the protein containing Ir at the original Zn site (Figure S15). UV–vis spectroscopy (Figure 3), $^{15}N-^{1}H$ HSQC spectroscopy (Figure S16), and nano-ESI-MS (Figure S14) were fully consistent with a protein containing an $Ir(\kappa^2$-CA)$(CO)_2$ site. Thus, this set of characterizations to determine the coordination mode of an organometallic fragment bound to a protein ligand can be broadly applied to protein scaffolds expected to coordinate metals using at least one histidine residue.

Characterization of Organometallic Reactions. The methods used to characterize the structures of the artificial organometallic proteins in the solution phase were selected to be equally suited to monitor reactions of the protein-ligated organometallic complexes. Specifically, we investigated organometallic ligand substitutions of Rh(nbd)-CA that would cleave and form metal–carbon bonds in the same reaction (Figure

DOI: 10.1021/jacs.5b04431
J. Am. Chem. Soc. 2015, 137, 8261–8268

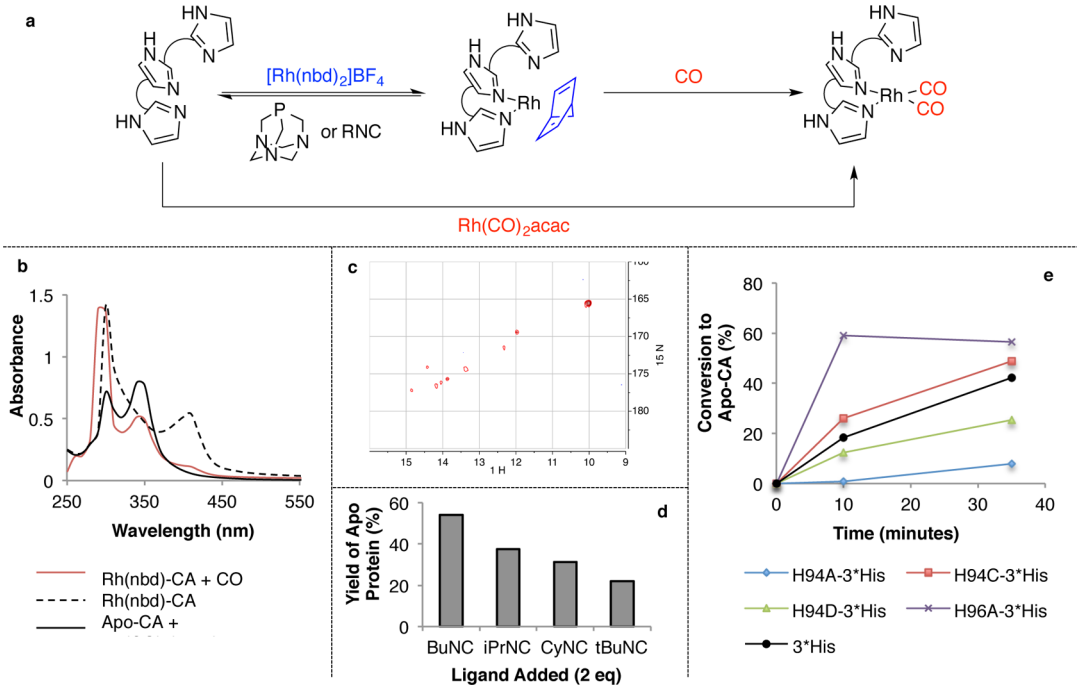


Figure 8. Ligand substitution reactions of Rh(nbd)-CA. (a) Summary of ligand substitution reactions achieved with Rh(nbd)-CA. Depiction of imidazole groups represents their relative orientation as histidine residues within CA. (b) Characterization of the ligand substitution reaction of CO (1 atm) with Rh(nbd)-CA. (c) One-bond $^{15}N-^{1}H$ HSQC spectrum of the reaction mixture formed by adding CO to a solution of Rh(nbd)-CA. (d) Analysis of the size selectivity of the substitution of the protein ligand from Rh(nbd)-CA with isocyanides. (e) Time course for the formation of apo-CA in the ligand substitution reaction of the phosphine PTA with several mutants of Rh(nbd)-CA.

8a). The reaction of Rh(nbd)-CA with CO (1 atm) resulted in significant conversion of the starting material, as observed by UV–vis spectroscopy (Figures 8b and S17). The decay of the absorption corresponding to Rh(nbd)-3*His-CA was accompanied by the appearance of an absorption at 350 nm, which matches the λ_{max} of $Rh(CO)_2$-CA. Based on the A_{350} value, $Rh(CO)_2$-CA formed in 53% yield, while 15% of the Rh(nbd)-CA starting material remained. Further analysis of the reaction composition by $^{15}N-^{1}H$ HSQC spectroscopy also was consistent with the presence of $Rh(CO)_2$-CA and Rh(nbd)-CA; the remaining protein in solution was shown by this method to be apo-CA, presumably from substitution of the κ^2-CA ligand (Figure 8c). To ensure full conversion, the reaction was conducted with a higher pressure of CO; indeed, the reaction of Rh(nbd)-3*His-CA with 100 psi of CO formed $Rh(CO)_2$-CA in 80% yield, as determined by UV–vis spectroscopy (Figure S17). This reaction is the first documented stoichiometric reaction to cleave and form metal–carbon bonds within an artificial protein.

Protein-Controlled Reactivity. With our established methods to observe directly ligand substitution reactions occurring at the metal center, we investigated the potential of the protein scaffold to affect the rates and selectivities of ligand substitutions at the Rh-CA sites. The reactions of phosphines and isocyanides with Rh(nbd)-3*His-CA displace the κ^2-CA ligand from the Rh center (Figure 8a), and the protein scaffold can be modified to modulate both the rate and selectivity of these reactions occurring at the protein-ligated Rh. The extent of these reactions was monitored by UV–vis spectroscopy and the NPA hydrolysis activity assay (vide supra). The substitution of the κ^2-CA ligand from the metal center by isocyanides to form apo-CA depended on the steric properties of the added ligand. The reaction of Rh(nbd)-3*His-CA with the unencumbered *n*BuNC progressed to the highest conversion, while reaction with the most hindered *t*BuNC proceeded to the lowest conversion (Figure 8d).

Addition of the water-soluble phosphine 1,3,5-triaza-7-phosphatricyclo[3.3.1.13.7]decane (PTA) to Rh(nbd)-3*His-CA also resulted in complete conversion of the starting material. This reaction yielded $[Rh(nbd)(PTA)_2]^+$ and apo-CA, as determined by UV–vis (Figure S18), ^{31}P NMR spectroscopy (Figure S19), and $^{15}N-^{1}H$ HSQC (Figure S20) spectroscopy. Addition of excess $ZnSO_4$ to the resulting apo-CA generated Zn-3*His-CA, indicating that the protein structure remained intact throughout the introduction and elimination of rhodium (Figure S20). Reactions of PTA with Rh(nbd)-3*His-CA containing different mutations near the metal center indicated that the progress of this substitution reaction can be controlled by the protein structure. The proteins Rh-3*His-H94A and Rh-3*His-H94D were less reactive toward this ligand substitution than were Rh-3*His and Rh-3*His-H94C, while the Rh-3*His-H96A mutant was more reactive (Figures 8e and S21). The dependence of the reactivity of Rh(nbd)-3*His-CA on the size of the reagent and the identity of the mutant illustrates the potential to assess directly the effects of mutations and protein structure on an elementary reaction of an organometallic protein.

Previous publications state that Rh-CA's prepared from $[Rh(COD)_2]BF_4$ and $Rh(CO)_2(acac)$ are active catalysts for hydrogenation and hydroformylation of stilbenes and styrenes.[11b,c] However, we found that the discrete organometallic protein complexes Rh(cod)-CA, Rh(nbd)-CA, and $Rh(CO)_2$-CA do not catalyze the hydrogenation or hydroformylation of a range of potential substrates for these reactions, including terminal, 1,1-disubstituted, and internal alkenes and vinylarenes. These findings suggest that the active catalyst of the

DOI: 10.1021/jacs.5b04431
J. Am. Chem. Soc. 2015, 137, 8261–8268

previously reported systems was not a Rh center ligated at the native Zn site of carbonic anhydrase. Instead, it is more likely that these reactions are catalyzed by a dissociated Rh fragment or a fragment associated with a different site on the protein. Most important for the current work, our efforts to synthesize and characterize discrete organometallic proteins by the methods described here permit critical assessment of catalysis by specific complexes, much like one tests the competence of a small organometallic complex to be an intermediate in a catalytic process.

CONCLUSION

Although the catalytic reactivity of artificial organometallic proteins has been previously documented, the potential to understand and improve this class of catalysts has been limited by the absence of methods to synthesize and characterize a range of discrete catalyst systems and to evaluate their fundamental reactivity. The synthesis, characterization, and reactivity of the artificial organometallic proteins documented in the current work demonstrate the potential to gain information on the elementary reactions of discrete organometallic protein complexes.

The expedient solution-state characterization methods presented in this work are crucial for determining the site and extent of metal binding and for evaluating the organometallic reactivity of the Rh-CA's. The efficient use of colorimetric activity assays to indirectly quantify the specificity of metal binding should be broadly applicable for the rapid assessment of a wide range of combinations of protein scaffold libraries and transition metal precursors. After a selective coordination process is identified, the binding mode of the protein ligand can be fully characterized by heteronuclear NMR spectroscopy, and the small-molecule ligands at the metal site can be identified by the combination of ESI-MS, IR, and NMR spectroscopy. Even the elementary reactions of these sites can be studied rapidly, setting the stage to reveal the ability of these macromolecular ligands to control the reactivity of the organometallic site in these artificial proteins and to create the potential to achieve selectivities beyond those observed with traditional organometallic complexes containing small organic ligands.

ASSOCIATED CONTENT

Supporting Information

Detailed experimental methods, additional figures, and complete characterizations of the prepared Rh-CA's. The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.5b04431.

AUTHOR INFORMATION

Corresponding Author

*jhartwig@berkeley.edu

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported by the Director, Office of Science, U.S. Department of Energy, under Contract No. DE-AC02-05CH11231 and by the NSF (graduate research fellowship). We thank the QB3MacroLab facility (subcloning, shared equipment), Dr. Jeffery Pelton and the QB3 Central California 900 MHz NMR facility (useful discussion and instrumentation, supported by NIH grant gm 68933), Dr. Tony Iavarone and the QB3Mass Spectrometry Facility (nano-ESI data collection, supported by NIH grant 1S10RR022393-01), Dr. Junko Yano for EXAFS data collection and analysis, and Prof. Carol Fierke (U. Michigan) for the wild-type carbonic anhydrase gene. Dedicated to Prof. Stephen J. Lippard on the occasion of his 75th birthday.

REFERENCES

(1) Hartwig, J. F. *Organotransition metal chemistry*; 2010, University Science Books, Mill Valley, CA.

(2) Andreini, C.; Bertini, I.; Cavallaro, G.; Holliday, G. L.; Thornton, J. M. *J. Biol. Inorg. Chem.* **2008**, *8*, 1205−1218.

(3) Reetz, M. T.; Peyralans, J. J.-P.; Maichele, A.; Fu, Y.; Maywald, M. *Chem. Commun.* **2006**, 4318−4320.

(4) Wilson, M. E.; Whitesides, G. M. *J. Am. Chem. Soc.* **1978**, *100*, 306−307.

(5) (a) Rosati, F.; Roelfes, G. *ChemCatChem* **2010**, *2*, 916−927. (b) Lewis, J. C. *ACS Catal.* **2013**, *3*, 2954−2975.

(6) Carey, J. R.; Ma, S. K.; Pfister, T. D.; Garner, D. K.; Kim, H. K.; Abramite, J. A.; Wang, Z.; Guo, Z.; Lu, Y. *J. Am. Chem. Soc.* **2004**, *126*, 10812−10813.

(7) Dürrenberger, M.; Heinisch, T.; Wilson, Y. M.; Rossel, T.; Nogueira, E.; Knörr, L.; Mutschler, A.; Kersten, K.; Zimbron, M. J.; Pierron, J.; Schirmer, T.; Ward, T. R. *Angew. Chem., Int. Ed.* **2011**, *50*, 3026−3029.

(8) Ohashi, M.; Koshiyama[1], T.; Ueno, T.; Yanase, M.; Fujii, H.; Watanabe, Y. *Angew. Chem., Int. Ed.* **2003**, *42*, 1005−1008.

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

(10) Examples of prior work based on ferritin: (a) 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. (b) Abe, S.; Niemeyer, J.; Abe, M.; Takezawa, Y.; Ueno, T.; Hikage, T.; Erker, G.; Watanabe, Y. *J. Am. Chem. Soc.* **2008**, *130*, 10512−10514. (c) Abe, S.; Hikage, T.; Watanabe, Y.; Kitagawa, S.; Ueno, T. *Inorg. Chem.* **2010**, *49*, 6967−6973. (d) Ueno, T.; Abe, M.; Hirata, K.; Abe, S.; Suzuki, M.; Shimizu, N.; Yamamoto, M.; Takata, M.; Watanabe, Y. *J. Am. Chem. Soc.* **2009**, *131*, 5094−5100.

(11) Prior work based on carbonic anhydrases: (a) Okrasa, K.; Kazlauskas, R. J. *Chem.—Eur. J.* **2006**, *12*, 1587−1596. (b) Jing, Q.; Kazlauskas, R. J. *ChemCatChem* **2010**, *2*, 953−957. (c) Jing, Q.; Okrasa, K.; Kazlauskas, R. J. *Chem.—Eur. J.* **2009**, *15*, 1370−1376. (d) Fernández Gacio, A.; Codina, A.; Fastrez, J.; Riant, O.; Soumillion, P. *ChemBioChem* **2006**, *7*, 1013−1016.

(12) Podtetenieff, J.; Taglieber, A.; Bill, E.; Reijerse, E. J.; Reetz, M. T. *Angew. Chem., Int. Ed.* **2010**, *49*, 5151−5155.

(13) Krishnamurthy, V. M.; Kaufman, G. K.; Urbach, G. K.; Gitlin, I.; Gudiksen, K. L.; Weibel, D. B.; Whitesides, G. M. *Chem. Rev.* **2008**, *108*, 946−1051.

(14) Togni, A.; Venanzi, L. M. *Angew. Chem., Int. Ed.* **1994**, *33*, 497−526.

(15) (a) Thorslund, A.; Lindskog, S. *Eur. J. Biochem.* **1967**, *3*, 117−123. (b) Hakansson, K.; Wehnert, A.; Liljas, A. *Acta Crystallogr., Sect. D* **1994**, *50*, 93−100.

(16) (a) Winn-Deen, E. S.; David, H.; Sigler, G.; Chavez, R. *Clin. Chem.* **1988**, *34*, 2005−2008. (b) Plocke, D. J.; Levinthal, C.; Vallee, B. L. *Biochemistry* **1962**, *1*, 373−378. (c) Wang, J.; Sheppard, G. S.; Lou, P.; Kawai, M.; Park, C.; Egan, D. A.; Schneider, A.; Bouska, J.; Lesniewski, R.; Henkin, J. *Biochemistry* **2003**, *42*, 5035−5042. (d) Houchin, O. B. *Clin. Chem.* **1958**, *4*, 519−523.

(17) (a) Ojima, I., Vu, T., Bonafoux, D. Organometallic complexes of rhodium. In *Houben-Weyl, Methods of Molecular Transformations*; Trost, B. M., Lautens, M., Eds.; Thieme: Stuttgurt, 2001; Vol. *1*, pp 531−616. (b) Fagnou, K.; Lautens, M. *Chem. Rev.* **2003**, *103*, 169−196.

(18) Although prior work (ref 11c) reported that dialysis can be used to eliminate excess rhodium from the protein surface (as determined by ICP-MS), we are unable to reproduce these results, in part because

DOI: 10.1021/jacs.5b04431
J. Am. Chem. Soc. 2015, 137, 8261−8268

this publication does not describe the method for sample preparation. We found from a series of ICP-OES measurements that a high concentration of nitric acid and sufficient sample heating is necessary to detect all the Rh in protein-containing samples. Milder methods (without heating or with lower acid concentrations) reported for metal analysis of metalloproteins failed to detect all Rh in control samples (ref 20).

(19) Shimahara, H.; Yoshida, T.; Shibata, Y.; Shimizu, M.; Kyogoku, Y.; Sakiyama, F.; Nakazawa, T.; Tate, S.-I.; Ohki, S.-Y.; Kato, T.; Moriyama, H.; Kishida, K.-I.; Tano, Y.; Ohkubo, T.; Kobayashi, Y. *J. Biol. Chem.* **2007**, *282*, 9646−9656.

(20) Robinson, N.; Waldron, K.; Tottey, S.; Bessant, C. *Nature Protoc. Exch.* **2008**, DOI: 10.1038/nprot.2008.236.

DOI: 10.1021/jacs.5b04431
J. Am. Chem. Soc. 2015, 137, 8261−8268