Enzyme Activity by Design: An Artificial Rhodium Hydroformylase for Linear Aldehydes

Amanda G. Jarvis,* Lorenz Obrecht, Peter J. Deuss, Wouter Laan, Emma K. Gibson, Peter P. Wells, and Paul C. J. Kamer*

Abstract: Artificial metalloenzymes (ArMs) are hybrid catalysts that offer a unique opportunity to combine the superior performance of natural protein structures with the unnatural reactivity of transition-metal catalytic centers. Therefore, they provide the prospect of highly selective and active catalytic chemical conversions for which natural enzymes are unavailable. Herein, we show how by rationally combining robust site-specific phosphine bioconjugation methods and a lipid-binding protein (SCP-2L), an artificial rhodium hydroformylase was developed that displays remarkable activities and selectivities for the biphasic production of long-chain linear aldehydes under benign aqueous conditions. Overall, this study demonstrates that judiciously chosen protein-binding scaffolds can be adapted to obtain metalloenzymes that provide the reactivity of the introduced metal center combined with specifically intended product selectivity.

The development of substrate- and product-specific catalytic processes that operate efficiently at mild reaction temperatures is a major challenge for the synthetic chemistry community. Enzymes are nature’s main catalysts and catalyze numerous chemical transformations, typically under benign conditions. However, many desired chemical reactions are not performed by nature, and therefore, there are no suitable natural enzymes. Artificial metalloenzymes (ArMs) provide a way to bridge that gap between natural and chemical synthesis, providing enzymes for unnatural reactions. Despite these successes, most ArMs do not meet the rates and performances achieved by natural enzymes. In addition, the molecular recognition and shape selectivity of proteins have typically not been exploited. The most successful approach to create ArMs has been the use of non-covalent anchoring strategies utilizing protein scaffolds with strong supramolecular recognition motifs such as avidin. In these ArMs, the binding site is used to carry the active metal center or bind metal-containing cofactors, restricting the possible applications of the binding properties of the protein. An alternative approach utilizes site-selective protein modification methods to incorporate transition metals into a wide range of protein scaffolds whilst leaving the protein’s innate binding capabilities largely intact. Any protein scaffold can be used, allowing the exploitation of the almost unlimited range of highly specific substrate-binding capabilities of proteins. Therefore, virtually any organometallic non-natural catalytic reaction can be merged with the sophisticated biological performance of enzymes, and this approach offers significant opportunities for the design of ArMs aiming at high selectivity by shape-selective product formation. Herein, we demonstrate the potential of such ArMs through the development of artificial rhodium enzymes derived from a protein scaffold that was selected for its apolar substrate-binding properties. These ArMs enable selective aldehyde formation in the biphasic hydroformylation of long-chain linear alkenes, a reaction for which no natural enzymes are known, and which is challenged in current industrial applications by the low solubility of the substrates in the aqueous phase.
We have previously reported a method that enables such an approach for the synthesis of ArMs containing metal-binding phosphine ligands (Scheme 1A). Rhodium–phosphine complexes are known to be highly active and robust hydroformylation catalysts, and thus our strategy provides the prospect of enzymatic hydroformylation reactions. Rhodium–protein hybrids tested to date in hydroformylation have utilized dative protein–rhodium interactions with limited success.

Although this has led to unprecedented linear selectivities for the hydroformylation of styrene, the exact nature of the active species, and thus the origin of the selectivity, is still unclear. Rhodium-catalyzed hydroformylation is used on a 80 000 ton scale to produce butyraldehyde from propene under biphasic conditions, allowing for the recovery of the expensive rhodium 3,3′,3′′-phosphanetriyltris(benzenesulfonic acid) trisodium salt catalyst (Rh/TPPTS). Long-chain aldehydes are desired by industry as they are important precursors for the production of detergents and plasticizers. This process is not feasible for long-chain alkenes (> 5 C atoms) owing to their low solubility in water.

Fatty acid transporter proteins contain apolar tunnels and clefts to bind their hydrophobic cargo. The steroid carrier protein type 2 like (SCP-2L) domain of the human multifunctional enzyme 2 (MFE-2) was identified as a suitable linear alken binding protein scaffold as it is known to bind a variety of linear aliphatic substrates and can be obtained in high yields. Our hypothesis was that the apolar tunnel present in SCP-2L (Figure 1B) would be able to facilitate the transport of alkenes to the aqueous environment and orient the starting alken along the tunnel towards the rhodium center, enabling the shape-selective production of the desired linear product (Figure 1A). To introduce the catalytic rhodium–phosphine complexes, two mutants containing unique cysteines at either end of the tunnel were prepared (SCP-2L V83C and SCP-2L A100C; Figure 1B). These two mutants, obtained in excellent yields, showed little structural permutations and similar aliphatic substrate binding capabilities as the wild-type (WT) protein (see the Supporting Information, Table S3). Both SCP-2L mutants were successfully modified with aldehyde phosphines P1–P3 through an acysteine modification strategy (Scheme 1A; for characterization data for SCP-2LV83C–P1–P3, see Ref. [5a]; for SCP-2L A100C-1–(P1–P3), see the Supporting Information).

The rhodium proteins (SCP-2LV83C/A100C–P1–P3–Rh) were obtained by the addition of Rh(acac)(CO)2. Other Rh precursors did not selectively bind to the phosphine (see Table S4 for MS and metal loading analysis). Their catalytic activity was investigated in the hydroformylation of 1-octene at 35°C and 80 bar synthesis gas (Figure 2, Scheme 1B, and Table S5). To minimize rhodium leaching and therefore false results from “free Rh” (leading to low selectivity (ca. 55% linearity) and high turnover numbers (TONs > 500; Table 1, entry 2)), a slight excess of protein (2 equiv) was used. Even though these reactions were performed at a relatively low temperature (typical industrial conditions are 125°C), significant hydroformylation activity was detected over 48 h for several of the rhodium–phosphine ArMs (Figure 2A). Reac-
phosphine selenide or ArM phosphine gold complexes showed that the rhodium was required for the hydroformylation reaction to occur (Table S5).

The structure of the phosphine cofactor was found to have a large effect on the TON, with the activity increasing by factors of 30 and 70, respectively, as the phosphine moved from the ortho (P1) to the meta position (P2) and from the ortho (P1) to the para position (P3) for SCP-2L V83C. The protein scaffold also influenced the reactivity, with a marked improvement in turnover found when using SCP-2L A100C–1–P3, achieving TONs of > 400, versus 75 for SCP-2L V83C–1–P3. The selectivity of the reaction was also found to depend on the combination of protein scaffold and phosphine cofactor applied. It ranged from 69 to 79% for the linear product (nonanal), matching typical selectivities for double-phosphine-ligated rhodium-catalyzed hydroformylation. Selectivities of 80% are rarely seen with monoligated P/Rh systems. Control reactions with P1–P3 or Rh(acac)(CO)–(PPh3) in neat alkene showed that with monophosphines in the organic phase of our reaction, selectivities of up to 74% can be achieved (Tables S8 and S9). Catalyst degradation and leaching of the phosphine and rhodium to the organic phase were not responsible for the observed selectivities as only minimal rhodium leaching and degradation were observed (Table S11 and Figures S12 and S13). It should be noted that our ArM system has a low P/Rh ratio, which, when using a benchmark biphasic Rh/TPPTS catalytic system, only gives activities and selectivities that correspond to rhodium leaching into the organic phase (Table 1, entries 2, 4, and 5). At the same concentration as in our ArM reactions, over 300 equivalents of the TPPTS ligand are needed to prevent metal leaching (Table 1, entries 4–6) and obtain high selectivities. The same effect of rhodium leaching was observed when the protein was simply mixed with Rh(acac)(CO)3 (Table 1, entry 3). This, alongside the differences between the two protein mutants, shows that it is the hybrid catalysts that are responsible for the hydroformylation results. SCP-2L A100C–1–P3 gave the best performance overall (79% nonanal, TON = 409; Table 1, entry 1). These conversions and selectivities are remarkable as a benchmark catalyst (Rh/TPPTS) gives negligible conversion approaching the detection limit (TON ≈ 1) when the TPPTS/Rh ratio is optimized to give similar selectivities to the metalloenzyme (TPPTS/Rh = 30:1 at ca. 10 times the Rh concentration of the ArM reaction gives 72% linearity; see Table S7).

Following the successful hydroformylation of 1-octene, these artificial metalloenzymes were tested in the hydroformylation of 1-decene, 1-dodecene, and 1-octadecene (Figure 2B). When TPPTS was used, a tenfold rate decrease was observed on the addition of two carbon atoms to the chain owing to the reduced water solubility of the alkene. Using the ArM, a less than fourfold

![Figure 2. Activity (colored bars) and selectivity (black squares) A) of the catalytic hydroformylation of 1-octene using different artificial metalloenzymes (the values for P1 have been magnified) and B) for the hydroformylation of alkenes with different chain lengths. Standard reaction conditions: 80 bar CO/H2 (1:1), 35°C, 625 rpm, 0.5 mL of catalyst solution and 0.5 mL of alkene containing 9% (v/v) n-heptane and 1% (v/v) diphenyl ether as internal standards. The Rh concentrations were determined by ICP-MS and used to calculate the TONs, recorded values between 20 and 100 nmol Rh. Conversions and linear selectivities were obtained by GC analysis with a minimum of three runs. Error bars show the standard deviation. For further details, see Section S4.](image)

**Table 1: Control reactions for the aqueous hydroformylation of 1-octene.**

<table>
<thead>
<tr>
<th>Entry</th>
<th>Catalyst</th>
<th>TON</th>
<th>Linear aldehyde [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ArM: SCP-2L A100C–1–P3–Rh^[a]</td>
<td>408.7 (57.79)</td>
<td>78.8 (4.86)</td>
</tr>
<tr>
<td>2</td>
<td>Rh(acac)(CO)^[b]</td>
<td>529.7 (53.30)</td>
<td>55.3 (0.67)</td>
</tr>
<tr>
<td>3</td>
<td>protein scaffold A100C treated with Rh^[c]</td>
<td>123.5 (38.19)</td>
<td>57.8 (0.07)</td>
</tr>
<tr>
<td>4</td>
<td>Rh/TPPTS (1:12)^[d]</td>
<td>2245 (674)</td>
<td>58.9 (0.41)</td>
</tr>
<tr>
<td>5</td>
<td>Rh/TPPTS (1:20)^[e]</td>
<td>700 (190)</td>
<td>56.5 (0.08)</td>
</tr>
<tr>
<td>6</td>
<td>Rh/TPPTS (1:300)^[f]</td>
<td>5.4 (3.25)</td>
<td>65.9 (8.55)</td>
</tr>
<tr>
<td>7</td>
<td>Rh/TPPTS (1:300) and SCP-2L A100C^[f]</td>
<td>9.4 (1.01)</td>
<td>60.7 (1.48)</td>
</tr>
<tr>
<td>8</td>
<td>Rh/TPPTS (1:300) and WT SCP-2L^[f]</td>
<td>5.1 (2.48)</td>
<td>67.0 (6.64)</td>
</tr>
</tbody>
</table>

Standard conditions: 80 bar syngas (1:1), 35°C, 625 rpm, 0.5 mL of catalyst solution and 0.5 mL of 1-octene containing 9% (v/v) n-heptane and 1% (v/v) diphenyl ether. The Rh concentration was determined by ICP-MS for entry 1. Conversions and linear selectivities were obtained by GC analysis of a minimum of three runs. Standard deviations given in parentheses. [a] P/Rh = 1.3, 23 nmol Rh. [b] 41.25 h, total volume: 0.4 mL 1-octene, no water, 150 nmol Rh. [c] Treated with Rh and washed in the same manner as the metalloproteins. [d] 10.0 nmol Rh. [e] 30 nmol Rh. [f] Two equivalents of the protein relative to Rh.
decrease in activity was observed upon going from 1-octene to 1-decane, and an only tenfold decrease when going to octadecene. Under selectivity-optimized conditions (high ligand concentrations to give adequate linear selectivity while preventing Rh leaching), Rh/TPPTS displayed no significant activity (Figure 2A, Table 1, entry 6, and Table S7). Control experiments using the Rh/TPPTS system in the presence of the protein scaffold showed no significant difference in turnover for 1-octene (Table 1, entry 6 vs. 7 and 8), providing evidence that the increase in activity for the ArMs cannot solely be explained by the protein acting as a phase-transfer reagent. We therefore attribute the higher-than-expected activity to the presence of the lipid-binding tunnel in the protein scaffold in direct proximity of the Rh center.

Overall, the selectivities for the linear hydroformylation products were remarkably high for monophosphine-ligated rhodium in water, indicating that the protein scaffold counterbalances the lack of phosphine ligands. In addition, both the phosphine cofactor and the protein mutant affect the activity of the reaction. To better understand the observed selectivity of our hydroformylase, we investigated the local environment of the Rh center in the protein scaffold. Both the X-ray absorption near-edge structure (XANES) and extended X-ray absorption fine structure (EXAFS) of SCP-2L A100C–P3-Rh at the Rh K edge were assessed. Comparing the XANES of SCP-2L A100C–P3–Rh to model Rh complexes (Figure S10) suggested the loss of carbonyl functionalities from the Rh(acac)(CO)3 precursor. This was further supported by the lack of CO stretches in the IR spectrum and fitting the EXAFS data of SCP-2L A100C–P3–Rh (Table S1).

The fitting model applied used characteristic scattering paths from both acac and PPh3Ar ligands, with a derived Rh–P coordination number of two. EXAFS is unable to distinguish between scattering neighbors of like atomic number, especially where $Z = \pm 1$. The prospect of two phosphorus atoms coordinated to the rhodium center appeared unlikely owing to the steric congestion resulting from placing two protein scaffolds around the metal. Thus we postulated the possibility of one of the observed Rh–P neighbors arising from a Rh–S interaction (Figure 3A); the protein scaffold contains functionalized sulfur in the form of methionine, at the N-terminus and on the flexible \( \beta \)-helices, and the introduced cysteine (Figure 3B).

To probe the possibility of a nearby methionine coordinating to the Rh center, the selenomethionine derivatives of the protein scaffold were expressed and purified. Rhodium complexation of the phosphine-modified SeMet protein with Rh(acac)(CO)3 gave similar results as for SCP-2L A100C–P3, showing the RhCO adducts in the mass spectra, and also resulted in a small shift in the \(^7\)Se NMR spectrum upon Rh addition (Figure S11). There was a profound change in the Rh K edge EXAFS data of SeMet-A100C–P3–Rh compared to SCP-2L A100C–P3–Rh (Figure 3C). The differences observed can be rationalized by coordination to a neighbor of higher atomic number, with a larger backscattering amplitude; between the analogous systems, Rh–S interactions have been replaced by Rh–Se interactions. The EXAFS analysis supports our hypothesis that there is monophosphine coordination, coupled with further interaction with sulfur from a methionine residue. Moreover, through observation of this Rh–S/Se interaction, we have direct evidence for the coordination of the protein scaffold to the Rh center.

We were intrigued if the methionine coordination would have an effect on the catalytic performance of the ArM, or if the methionine would just decoordinate under the reaction conditions to give the same active catalyst, and thus the same catalytic results in all cases. To probe this further, four mutants of SCP-2L A100C were prepared in which each Met residue was replaced by alanine, and the proteins were then modified as above. The results of the initial catalytic experiments are ambiguous; the activity of SCP-2L A100C M105A–P3–Rh was on average higher than that of A100C or the other mutants (TON = 112(± 33) and TON = 60–86 for the others, see Table S10), which could indicate that M105 is indeed involved in rhodium coordination. However, this can also be due to other factors, such as structural changes or decreased protein stability. More detailed studies are required before
firm conclusions can be drawn on the role of the methionine residues in the hydroformylation by these ArMs.

In summary, we have shown that rhodium phosphate modified SCP-2Ls are linear-selective catalysts in the hydroformylation of long-chain alkenes. SCP-2L A100C–P3–Rh showed a rate enhancement of at least $10^4$ compared to the traditional Rh/TPPTS system in the biphasic hydroformylation of 1-octene and 1-decene. This demonstrates that a protein chosen for its specific binding properties can be converted into an enzyme in which these properties are used to transmit product selectivity. Combining this technology with the recent advances in chemical biology will allow us to rapidly engineer highly selective catalysts that operate under benign conditions. Moving forward, we believe that this approach has the potential to be used for a whole range of reactions that traditionally use phosphines as ligands and to convert these into biocatalytic processes. In the long term, as chemoenzymatic optimization is used to improve activity, this could open the door to a new era of biocatalytic chemical production.

Acknowledgements

This work was supported by the EU through NEST Adventure STREP Project Artzymes (FP6-2003-NEST-B3 15471), Marie Curie Excellence Grants. Artzyme Catalysis (MEXT-2004-014320), and through a Marie Curie Individual Fellowship project ArtOxiZymes to A.G.J. (H2020-MSCA-IF-2014-657755). We thank the EPSRC for funding through the EPSRC critical mass grant “Clean catalysts for sustainable development” (EP/J018139/1) and Sasol (CASE studentship to P.J.D.). The UK Catalysis Hub is kindly thanked for resources and support provided via our membership of the UK Catalysis Hub Consortium, which is funded by the EPSRC (EP/K014706/2, EP/K014668/1, EP/K014854/1, EP/K014714/1, and EP/M013219/1). We acknowledge Diamond Light Source for beamtime through award SP10306, and thank the beamline scientists Dr Ian Giano and Dr. Giannantionio Cibin. We thank Dr. Sally Shirran and Dr. Catherine Botting for advice on mass spectrometry, and Dr. Thomas Lebl for help with $^{79}$Se NMR spectroscopy. We also thank Dr. T. Glumoff (University of Oulu, Finland) for providing plasmids. A.G.J. thanks Dr. Irshad Baig for assistance with WT SCP-2L, protein production. P.J.D. thanks Yiteng Zhang for assistance.

Conflict of interest

The authors declare no conflict of interest.

Keywords: artificial metalloenzymes · catalyst design · hydroformylation · phosphines · rhodium

How to cite: Angew. Chem. Int. Ed. 2017, 56, 13596–13600

Angew. Chem. 2017, 129, 13784–13788


Manuscript received: June 6, 2017
Accepted manuscript online: August 25, 2017
Version of record online: September 13, 2017