Design and Evolution of Chimeric Streptavidin for Protein-Enabled Dual Gold Catalysis

Fadri Christoffel[‡,1,2], Nico V. Igareta[‡,1,2], Michela M. Pellizzoni[1,3], Laura Tiessler-Sala[4], Boris Lozhkin[1], Daniel C. Spiess[1,2], Agustí Lledós[4], Jean-Didier Maréchal[4*], Ryan L. Peterson[1,2,5*] Thomas R. Ward[1,2*]

[1]Department of Chemistry, University of Basel, Mattenstrasse 24a, BPR 1096, CH-4058, Basel, Switzerland
[2]National Center of Competence in Research, Molecular Systems Engineering, Basel, Switzerland
[3]Current address: Adolphe Merkle Institute, Chemin des Verdiers 4, Fribourg, Switzerland
[4]Unitat de Quimica Fisica, Departament de Quimica, Universitat Autonoma de Barcelona, Edifici C.n. 08193, Cerdanyola, Barcelona, Spain
[5]Current address: Texas State University, Department of Chemistry and Biochemistry 601 University Dr., San Marcos, TX 78666

‡ These authors contributed equally

Abstract

Artificial metalloenzymes (ArMs) result from anchoring an organometallic catalyst within an evolvable protein scaffold. Thanks to its dimer of dimers quaternary structure, streptavidin allows the precise positioning of two metal cofactors to activate a single substrate, thus expanding the reaction scope accessible to ArMs. To validate this concept, we report herein on our efforts to engineer and evolve an artificial hydroaminase based on dual-gold activation of alkynes. Guided by modelling, we designed a chimeric streptavidin equipped with a hydrophobic lid shielding its active site which enforces the advantageous positioning of two synergistic biotinylated gold cofactors. Three rounds of directed evolution using *E. coli* cell-free extracts led to the identification of mutants favouring either the anti-Markovnikov product (an indole carboxamide with 96% regioselectivity, 51 TONs) resulting from a dual gold σ,π-activation of an ethynylphenylurea substrate or the Markovnikov product (a phenyl-dihydroquinazolinone with 99% regioselectivity, 333 TONs) resulting from the π-activation of the alkyne by gold.

Introduction

Thanks to their unique affinity towards alkynes, allenes and alkenes, gold complexes have attracted significant attention for their catalytic potential.[1–3] In addition to the activation of unsaturated substrates via π-coordination, terminal alkynes undergo dual-gold activation via synergistic σ,π-coordination.[4–8] This mode of activation, which proceeds via a diaurated transition state, affords distinct products/regioisomers, significantly broadening the scope of gold-catalysed reactions. Such synergistic action of two metals in catalysis is reminiscent of polynuclear metalloenzymes, whereby (at least) two metals act in concert to catalyse challenging reactions.[9]

In the context of *in vivo* ligation and bioconjugation, alkynes occupy a place of choice as this functional group was shown to be bio-orthogonal, thus finding widespread use in click chemistry.[10,11] Although Cu and Ru are privileged catalysts in this context,[12,13] recent reports suggest that gold-complexes maintain catalytic activity in a cellular environment, albeit for a different type of reactivity.[14–18] To the best of our knowledge however, these biocompatible reactions rely on a π-activation of the alkyne, rather than the dual activation, so distinctive of gold-catalysis.

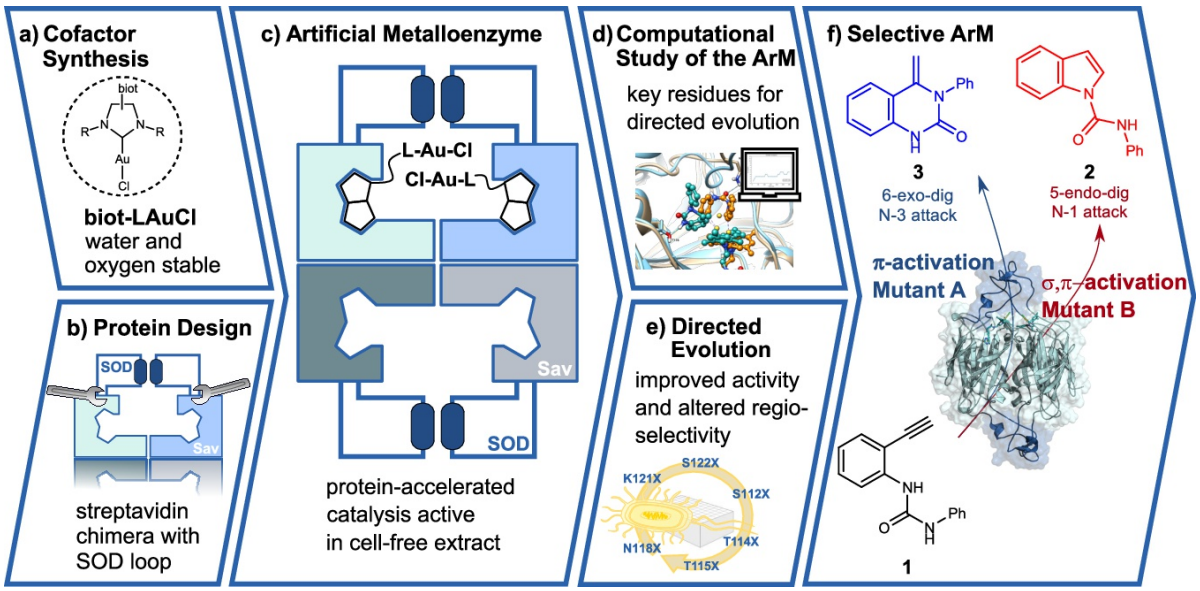


Figure 1 Engineering and evolving an artificial hydroaminase (HAMase) based on dual-gold activation of alkynes. Chemo-genetic optimization of the catalytic performance relies on combining: **a** biotinylated cofactor and **b** a tailored chimeric protein to assemble **c** an ArM equipped with two adjacent gold cofactors. Genetic optimization is guided by **d** modelling to identify **e** advantageous amino acids for directed evolution to favour **f** either σ,π-activation or π-activation of the alkyne to afford either indole **2** or quinazolinone **3** respectively.

With the aim of complementing natural enzymes,[19] artificial metalloenzymes (ArM hereafter) have experienced a renaissance in the past two decades.[20–24] For this purpose, an abiotic metal cofactor is compartmentalized within a protein scaffold which can be optimized by

genetic means. Thus far, more than 40 reactions can be catalysed by ArMs.[25] Current challenges in the field include; protein-accelerated catalysis, whereby a pre-catalyst is activated upon incorporation within the host protein,[26] dual catalysis[27,28] and compatibility of the ArM with a cytosolic environment.[29] Privileged scaffolds for ArMs include: carbonic anhydrase,[30] hemoproteins,[31,32] prolyl oligopeptidase,[33] lactococcal multiresistance regulator,[23] four helix-bundles,[34,35] nitrobindin,[36] human serum albumin,[37] and (strept)avidin.[20,38–40] The work presented herein capitalizes on the unique topology of Sav enabling the localization of two close-lying biotinylated probes within a hydrophobic environment. This enabled the engineering and evolution of a biocompatible artificial hydroaminase (HAMase hereafter) based on either single- or dual-gold activation of an alkyne, Figure 1.

Results

Design of the artificial hydroaminase

As reported by Asensio[5,41] and van der Vlugt[42], the regioselectivity for the hydroamination of ethynylurea **1** is by-and-large governed by the mode of activation of the alkyne by gold: the canonical π-activation favours the quinazolinone **3** (Markovnikov, 6-exo-dig addition product), while the dual σ,π-gold activation affords preferentially the indole **2** (anti-Markovnikov, 5-endo-dig addition product)[5,42,43] Upon π-coordination of the alkyne to gold, the pKa of the terminal C–H bond is lowered, thus favouring its deprotonation and coordination by a second gold to afford the σ,π-activation mode.[41] Accordingly, the spatial arrangement of the two gold species is critical in determining the regioselectivity of the reaction. We thus selected the gold-catalyzed cyclization of the ethynylurea **1** to engineer and evolve a dual-gold catalysed hydroaminase (HAMase) based on the biotin-streptavidin technology.

Thanks to its dimer of dimers quaternary structure, which places the valeric acid side chains of two proximal biotins 19.8 Å apart (pdb: 3ry2), we designed *N*-heterocyclic carbene ligands equipped with a biotin anchor introduced at various positions, Figure 2a.[44] We hypothesized that the relative position of two gold moieties within the biotin binding vestibule may influence the mode of alkyne-activation as reflected by the indole **2** vs the quinazolinone **3** ratio. Sterically-crowded imidazolium precursors were metalated using a one-step procedure,[45] and less-hindered carbenes were prepared through transmetalation of the silver-carbene complex, to afford the corresponding gold complexes: **biot-Au 1-5,** Supplementary Figure 1.[46] These are air- and water stable and can be stored for months as stock solutions in DMSO at 5 ºC.

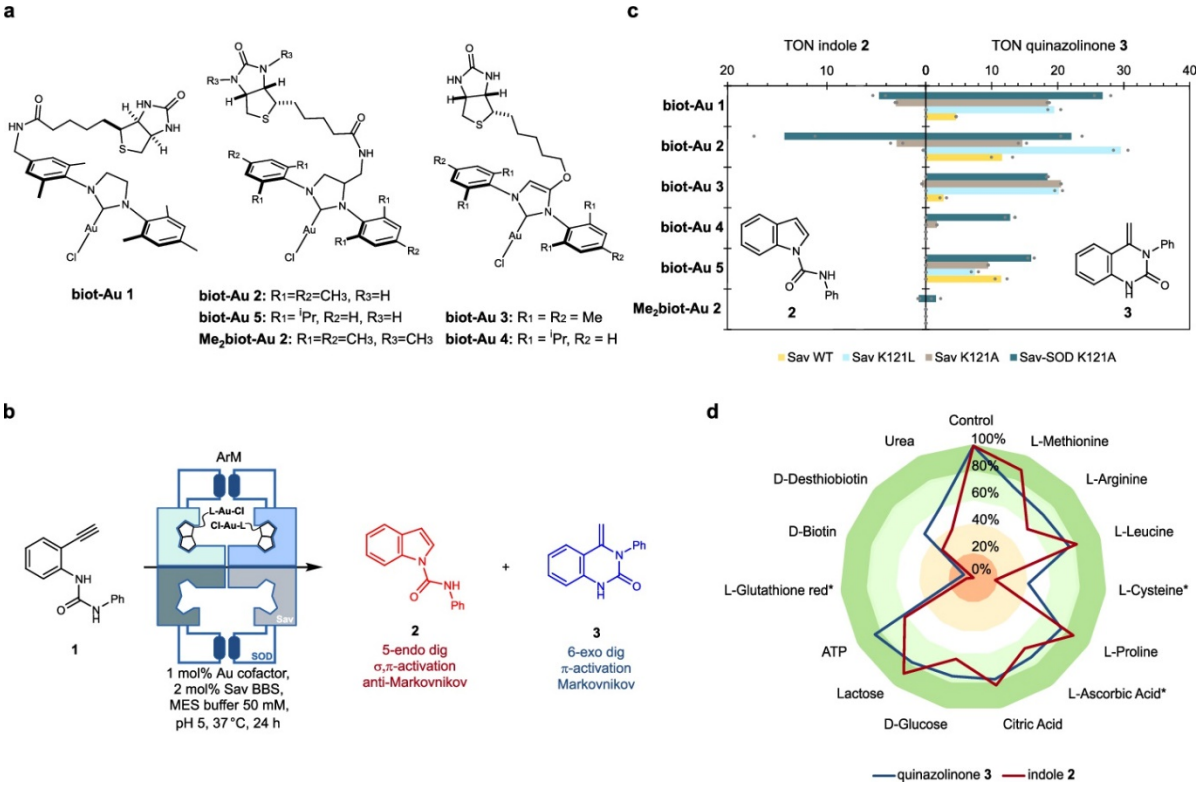


Figure 2 Chemo-genetic optimization of HAMase activity. a Biotinylated gold complexes **biot–Au 1-5** tested **b** in the presence of streptavidin isoforms **c** for the hydroamination of substrate **1** to afford indole **2** or quinazolinone **3**. Screening conditions using purified Sav samples: V_{tot} 200 μL (V_{DMSO} 15 μL), [Sub] 5 mM, [Au] 50 μM, [Sav] 100 μM, [MES-buffer] 50 mM, pH 5, 37 °C for 24 h. **d** Bio-additive-based screening of **biot-Au 2** · Sav-SOD K121A allows identifying detrimental cellular components. Reaction conditions: V_{tot} 200 μL (V_{DMSO} 15 μL), [Sub] 2.5 mM, [bio-additive] 2.5 mM, [**biot-Au 2**] 50 μM, [Sav] 100 μM, [MES-buffer] 50 mM, pH 5, 37 °C for 24 h. *cross-reaction with the substrate.

Engineering a chimeric protein with a hydrophobic biotin-binding vestibule

Initial studies using Au(I)-complexes bearing either commercially-available NHCs or **biot-Au 1-5** in buffered aqueous solutions afforded < 1 turnover (TON, using 7.5 % DMSO, [substrate **1**] = 5 mM and [Au] = 50 μM = 1 mol %, 24 hours at room temperature), Supplementary Table 1. Addition of wild-type streptavidin (Sav WT hereafter, 25 μM, corresponding to 100 μM biotin-binding sites, BBS) to a biotinylated cofactor **biot-Au 1-5** improved the catalytic activity leading to up to 12 TONs and affording the quinazolinone **3** exclusively, Figures 2b-2c and Supplementary Figure 2. This protein-acceleration phenomenon upon addition of Sav was not observed with cofactors devoid of biotin or with **Me$_2$biot-Au 2** which bears a dimethylated biotin anchor and thus a markedly decreased affinity towards streptavidin as highlighted by a HABA displacement assay, Supplementary

115 Figure 3. Fine-tuning the reaction conditions revealed that MES buffer at pH = 5 affords the
116 highest TONs, without significantly affecting the regioselectivity, Supplementary Figures 4-8.
117 Next, we selected **biot-Au 2** and screened it in the presence of a focused library of Sav
118 isoforms bearing mutations at S112 and/or K121. While the TON could be improved,
119 especially in the presence of small/hydrophobic residues at position K121, the critical **3**:**2**
120 ratio, diagnostic of the gold-activation mode, remained heavily biased in favour of the 6-exo-
121 dig product **3**, Supplementary Figure 9 and Supplementary Table 2. The addition of various
122 bio-additives in fiftyfold excess vs. catalyst was tolerated in most cases. Reduced glutathione
123 and cysteine poisoned the catalytic system, fortunately only at higher concentrations
124 (> 20x[**biot-Au 2**]). Strikingly, biotin inhibits the ArM already at low concentrations (> 4x[**biot-**
125 **Au 2**]). These results highlight the promising bio-robustness of the ArM, suggesting it may be
126 used in a cellular medium, Supplementary Table 3. With directed evolution in mind, this is a
127 highly desirable feature as it allows screening cell-free extracts (cfe), without the need to
128 purify the streptavidin mutants.[47]

129 Inspection of the > 30 X-ray structures of ArMs based on the biotin-streptavidin technology
130 reveal a tendency for biotinylated metal cofactors to be poorly localized within the biotin-
131 binding vestibule.[38] We attribute this to its shallow topology, thus allowing a biotinylated
132 cofactor to adopt multiple poses, resulting in reduced occupancy. Multiscale-modeling
133 strategies on some of these systems illustrate that the pronounced flexibility of the cofactor
134 may compromise the regiospecificity of the reaction.[48] With the aim of shielding the biotin-
135 binding vestibule, consisting of two eight-stranded β-barrels facing each other, we surveyed
136 the literature for naturally-occurring dimerization domains present in β-sheet-rich proteins.
137 We identified a potential candidate fitting this criterion: the superoxide dismutase C (sodC)
138 from *M. tuberculosis* (pdb: 1pzs[49]), which includes a ~ 30 amino acid dimerization domain.
139 This lid forms an interface that spans across the ~ 29 Å of the two Greek Key β-barrel
140 subunits of sodC. We thus set out to engineer a chimera consisting of the dimerization domain
141 of sodC inserted in the 3-4 loop of the Sav to yield a Sav-SOD, Figure 3a. To our delight,
142 Sav-SOD could be expressed in the soluble fraction in high yield (typically > 100 mg/L) in
143 shake flasks using *E. coli* BL21 DE3, Supplementary Figures 10-11.

144 To scrutinize the effect on the perturbation resulting from the addition of the sodC
145 dimerization domain, we performed Isothermal Titration Calorimetry measurements (ITC)
146 with biotin. Nanomolar binding affinity for biotin is retained over a wide range of temperatures
147 (10 to 40 ºC), Supplementary Figure 12 and Supplementary Table 4-5. At 25 ºC, the K_d is
148 4.2 nM with standard binding enthalpy ΔH^o = – 20.90 kcal/mol and binding enthalpy $T\Delta S^o$ =

149 – 9.49 kcal/mol. The standard free energy ΔG^o = – 11.4 kcal/mol at 25 ºC.[50] These
150 parameters suggest that the biotin-binding is primarily enthalpically-driven. Comparison of
151 the turn-on fluorescence upon incorporation of the biotinylated solvatochromic fluorescent
152 reporter 4-*N,N*-dimethylamino-1,8-naphthalimide (**biot-4DMN**) in Sav and Sav-SOD K121A
153 reveals a > 2.5 and 20 fold increase in fluorescence compared to the free **biot-4DMN**,
154 accompanied by a blue shift (λ_{em} = 532 and 526 nm vs. 556 nm), Supplementary Figure 13.
155 Such increase in fluorescence accompanied by an ipsochromic shift, have been attributed to
156 increased hydrophobicity.[50] Furthermore, thermal shift assay highlights an increased thermal
157 stability of the apo chimeric protein compared to apo Sav WT, Supplementary Figure 14. We
158 thus surmise that the SOD-lid stabilizes the protein and significantly contributes to provide a
159 hydrophobic and shielded environment for organometallic catalysis.

160 **Computational studies of the HAMase**

161 To gain structural and mechanistic insight on the influence of the protein scaffold, we refined
162 an integrative computational procedure that we have developed to model various ArMs,
163 Figure 3b.[51] We analysed the formation of the 6-exo-dig product **3** and 5-endo-dig product **2**
164 catalyzed by **biot-Au 2** · Sav WT, **biot-Au 2** · Sav K121A and **biot-Au 2** · Sav-SOD K121A.
165 For this purpose, we applied the following workflow: i) DFT calculations on the theozyme, the
166 core catalytic centre (urea **1** coordinated to one or two biotinylated gold catalysts) for both π-
167 and dual σ,π-activation modes. DFT calculations (B3LYP-D3 functional) were performed
168 using continuum water solvent conditions, Supplementary Table 6 and Supplementary
169 Figures 15-18.[52,53] ii) With no X-ray structure of the Sav-SOD available, structural modelling
170 (with Modeller[54]) was carried out on this system followed by classical Molecular Dynamics
171 (MD) simulations (AMBER Force field[55]) up to convergence, Supplementary Figure 19. Then,
172 after inclusion of **biot-Au 2** into the protein vestibule by protein-ligand docking, MD on the
173 three Sav scaffolds **biot-Au 2** · Sav-WT, **biot-Au 2** · Sav K121A and **biot-Au 2** · Sav-SOD
174 K121A were performed to determine the conformational space available for substrate
175 binding, Supplementary Figure 20. iii) Incorporation of the transition state structures for the
176 6-exo-dig and 5-endo-dig pathways in water (see i)) by protein-ligand docking approaches
177 (using GaudiMM[51] and Gold5.8.1[56]) into the most representative structures of the MD
178 simulations for the three Sav scaffolds (we term these "pseudo-transition states"),
179 Supplementary Figure 21, and iv) further refinement by MD of the best results obtained for
180 systems with reasonable predicted affinity (in iii)), Supplementary Figures 22-24. The final
181 simulations were analyzed focusing on: a) the complementarity of transition state structures
182 for the 6-exo-dig and 5-endo-dig within the Sav vestibule, and b) the number of gold cofactors

183 **biot-Au 2** involved (i.e. one or two), Supplementary Figures 25-26, Supplementary Table 7-
184 8. The influence of the host protein on the transition state structures helped identify amino
185 acid residues to randomize during directed evolution.

186 The DFT calculations revealed that the competition between π- and dual σ,π-activation modes
187 also operates in water, Supplementary Figure 15. The difference between the Gibbs energy
188 barriers for both pathways is about 2 kcal mol^{-1} (19.5 vs. 21.5 kcal mol^{-1} for 6-exo-dig and 5-
189 endo-dig mechanisms, respectively), suggesting that subtle changes in the first or second
190 coordination sphere of the metal may significantly affect the ratio between both products.
191 Calculations of the barriers in solvents of different dielectric constant suggest that the
192 regioselectivity is not significantly affected by the polarity of the medium, Supplementary
193 Figures 16-17.

194

195 We surmised that a shift in regioselectivity may be promoted by the protein environment. DFT
196 calculations carried out on small models were used to evaluate the impact on the ground-
197 and transition states structures of the gold complexes both in bulk water and in the confined,
198 hydrophobic environment provided by the biotin-binding vestibule. Scrutiny of the active site
199 led us to hypothesize that, following π-coordination, the alkyne's C-H may be deprotonated
200 by close-lying amino acids or a water molecule to afford a di-aurated σ,π–acetylide species:
201 (Au-C_σ = 2.01 Å) and (Au-C_π = 2.28, 2.33 Å) and a Au—Au distance ~3 Å, Supplementary
202 Figure 18.

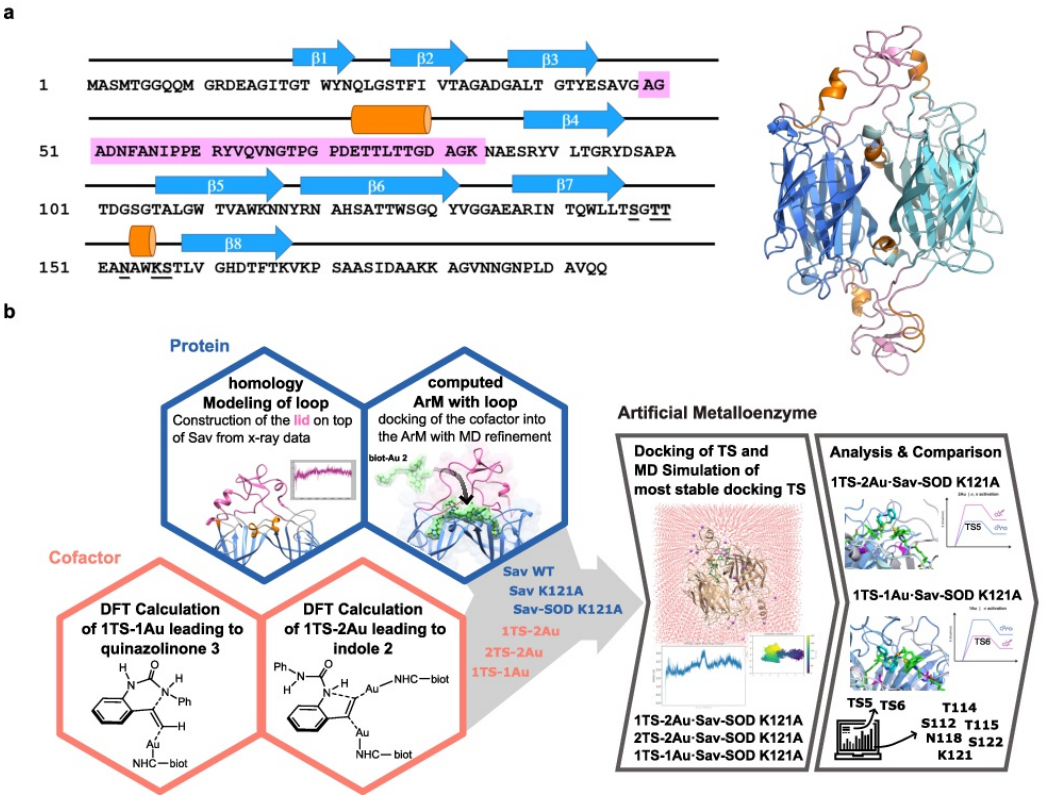


Figure 3 Design and structural characterization of the chimeric ArM: a Protein topology diagram of the sequence of one Sav-SOD monomer and the computed structure of chimeric Sav-SOD resulting from a 200 ns MD simulation; the SOD insert is highlighted in pink; residues subjected to saturation mutagenesis are underlined; blue arrows and orange cylinders represent β-sheets and α-helices respectively. **b** Workflow for the computational design and optimization of transition states in chimeric streptavidin. Identification of most promising amino acid residues to subject to mutagenesis.

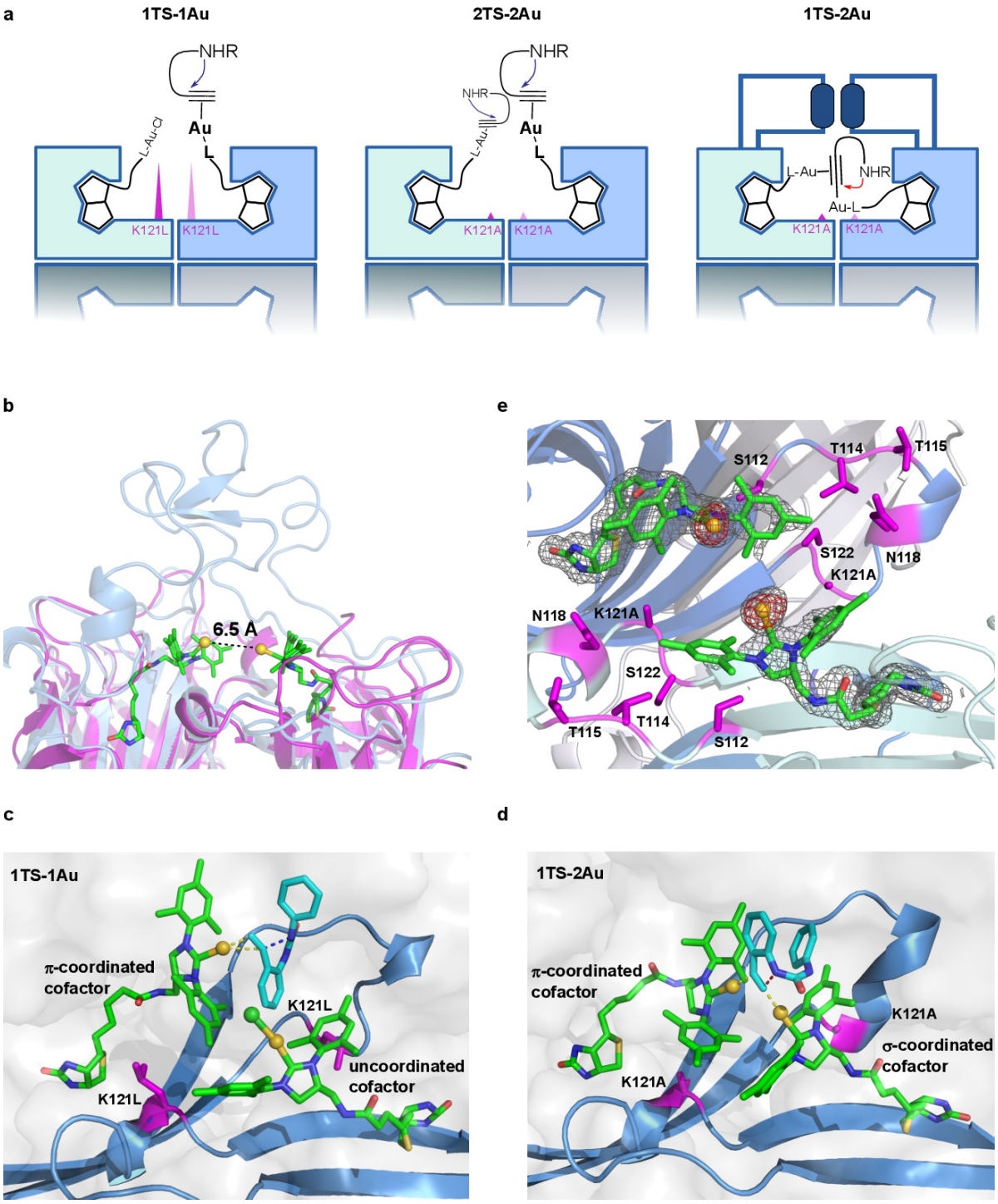


Figure 4 Analysis of the transition state structure and close-lying amino acid residues in chimeric Sav: a Schematic representation of postulated gold-catalyzed hydroamination reaction mechanisms within the biotin-binding vestibule. **b** X-ray crystal structure of **biot-Au 2** · Sav-SOD K121A (magenta) overlaid with the MD model (blue) (see Supplementary Figure 27 for details). **c,d** Computed transitions state for the **biot-Au 2**-catalyzed 6-exo-dig (**1TS-1Au**) and for the 5-endo-dig cyclization (**1TS-2Au**) docked within Sav-SOD K121A. **e** Close-up view of the X-ray structure of **biot-Au 2** · Sav-SOD K121A. Anomalous electron density (displayed as red mesh at 5σ) assigned to Au and modeled with a 50% occupancy; no electron density for Cl was detected. Electron density map (2Fo-Fc) for the **biot-Au 2** (displayed as grey mesh at 1σ). **biot-Au 2** (green stick), close-lying amino acid residues (magenta stick representation), Au as golden sphere, Cl as green sphere and the protein as cartoon representation.

222 Having identified transition state structures in water for the isolated cofactor (in the absence
223 of the protein scaffold), three different pseudo-transition states embedded in the protein were
224 evaluated: **1TS-1Au** and **2TS-2Au**-and **1TS-2Au**. The transition state **1TS-1Au** corresponds
225 to the π-activated TS which occupies half of the biotin-binding vestibule, with a second,
226 unligated **biot-Au 2** occupying the neighbouring biotin binding site. The transition state **2TS-**
227 **2Au** is similar to **1TS-1Au**, but includes two gold complexes, each activating an alkyne via π-
228 coordination. Finally, the transition state **1TS-2Au2** contains two gold complexes interacting
229 with a single alkyne substrate via σ,π-coordination, Figure 4a.

230 Molecular modelling of **biot-Au 2** · Sav WT provides the best fit for the pseudo-transition
231 state **1TS-1Au** when inserted into the catalytic vestibule of Sav, Supplementary Table 6 and
232 Supplementary Figure 21. Worse complementarities were computed for **2TS-2Au** and severe
233 clashes with the amino acids on position 121 were predicted for **1TS-2Au**. All these pseudo-
234 transition state structures are rather solvent-exposed, with minimal impact of the second
235 coordination sphere, resulting in similar energy barriers and similar predicted regioselectivity
236 compared to the free cofactor. Both the free cofactor and **biot-Au 2** · Sav WT are thus
237 predicted to favour π–activation to afford quinazolinone **3**. Gratifyingly, *in vitro* experiments
238 support the model, revealing that **biot-Au 2** · Sav WT affords exclusively quinazolinone **3**,
239 Figure 2c.

240 In the case of **biot-Au 2** · Sav K121A, similar or worse fitting scores compared to Sav WT
241 are computed for all pseudo-transition states, except for **1TS-2Au**. In fact, the score
242 associated per cofactor reveals that the binding affinity of **1TS-2Au** is close to the value
243 obtained for **1TS-1Au** (that corresponds to a single transition state geometry with no
244 geometric constraint from the adjacent Sav monomer), Supplementary Table 6. These results
245 suggest that a single-point mutation K121A improves the docking score of **1TS-2Au**
246 compared to **1TS-1Au**. This reflects a modest shift in favour of σ,π-alkyne coordination which
247 should allow for the formation of the anti-Markovnikov product indole **2**. These findings are
248 supported by experiments: the regioselectivity (**3:2**) varies from 100:0 **(**for **biot-**
249 **Au 2** · Sav WT) to 83:17 for **biot-Au 2** · Sav K121A, Figure 2c and Table 1.

250 X-ray quality crystals of **biot-Au 2** · Sav-SOD (including various mutants) were obtained
251 through co-crystallization, Supplementary Figure 27 and Supplementary Table 9. Although a
252 resolution down to 1.8 Å was achieved, the SOD-lid could not be fully resolved due to its high
253 flexibility, pdb: 7ALX. We thus set out to model the structure of Sav-SOD K121A, starting
254 with homology modeling, followed by a long classical MD (200 ns), Figure 4b and

255 Supplementary Figure 19.[57] The resulting models are stable with the SOD-lid presenting the
256 highest degree of flexibility (RMSF 2.63). Next, we docked **biot-Au 2** into the system. Good
257 complementarities were obtained by collective motion of the entire SOD-lid, Supplementary
258 Figures 20-21. This hydrophobic lid contributes to shield both cofactors from the solvent. A
259 second MD (300 ns) placed both **biot-Au 2** cofactors sufficiently close to synergistically
260 engage in σ,π-activation of a terminal alkyne, Supplementary Figures 22-24. To our delight,
261 **biot-Au 2** · Sav-SOD K121A indeed displayed the highest 5-endo-dig regioselectivity (i.e.
262 62:38 for **3:2**) and TONs of up to 40, Figure 2c and Table 1. We thus selected **biot-**
263 **Au 2** · Sav-SOD K121A for directed evolution.

264 Docking of the three pseudo-transition state models into Sav-SOD K121A reveal that the best
265 complementarity is obtained for **1TS-1Au** and **1TS-2Au** (especially for TS5), Figures 4c-d.
266 Scoring values for **2TS-2Au** were extremely low as there is limited space for such a large
267 pseudo-transition state in the Sav-SOD's vestibule, Supplementary Table 6. This also
268 suggests that the possibility of another substrate approaching **1TS-2Au** to form **2TS-2Au** is
269 unlikely in chimeric Sav-SOD. Depending on the transition state, the substrate(s) occupies
270 different positions within the active site, Figure 4c vs 4d. The most notable difference is a
271 180-degree rotation of the substrate between **1TS-1Au** (6-exo-dig) compared to **1TS-2Au** (5-
272 endo-dig). Further residue- contribution analysis of **1TS-2Au** and **1TS-1Au** (to afford 5-endo-
273 dig and 6-exo-dig products respectively) in the active ArM were performed. Calculations were
274 carried out using Cytoscape[58] as implemented in UCSF Chimera[59] (for a qualitative analysis
275 of main interactions along the MDs) followed by MMGBSA (to extract indicative energetic
276 values), Supplementary Figures 25-26 and Supplementary Tables 7-8. Direct interactions of
277 Au with close-lying amino acids are very weak (e.g. purely VdW contacts). Both pseudo-
278 transition state structures reveal common interactions between the coordinated substrate **1**
279 and the residues SOD-N8, SOD-I9, SOD-A3 and Sav-N118. For **1TS-1Au**, additional
280 contacts were identified: from Sav-S112 to S122, especially T114. This increased number of
281 contacts is traced back to the high level of flexibility of the **1TS-1Au** versus the **1TS-2Au**. As
282 the SOD lid is highly flexible (and disordered in the X-ray structure), we selected close-lying
283 residues belonging to Sav rather than the SOD-lid. Accordingly, the following amino acids
284 were selected for the directed evolution campaign: S112, T114, T115, N118, K121 and S122,
285 Figure 4e.

286

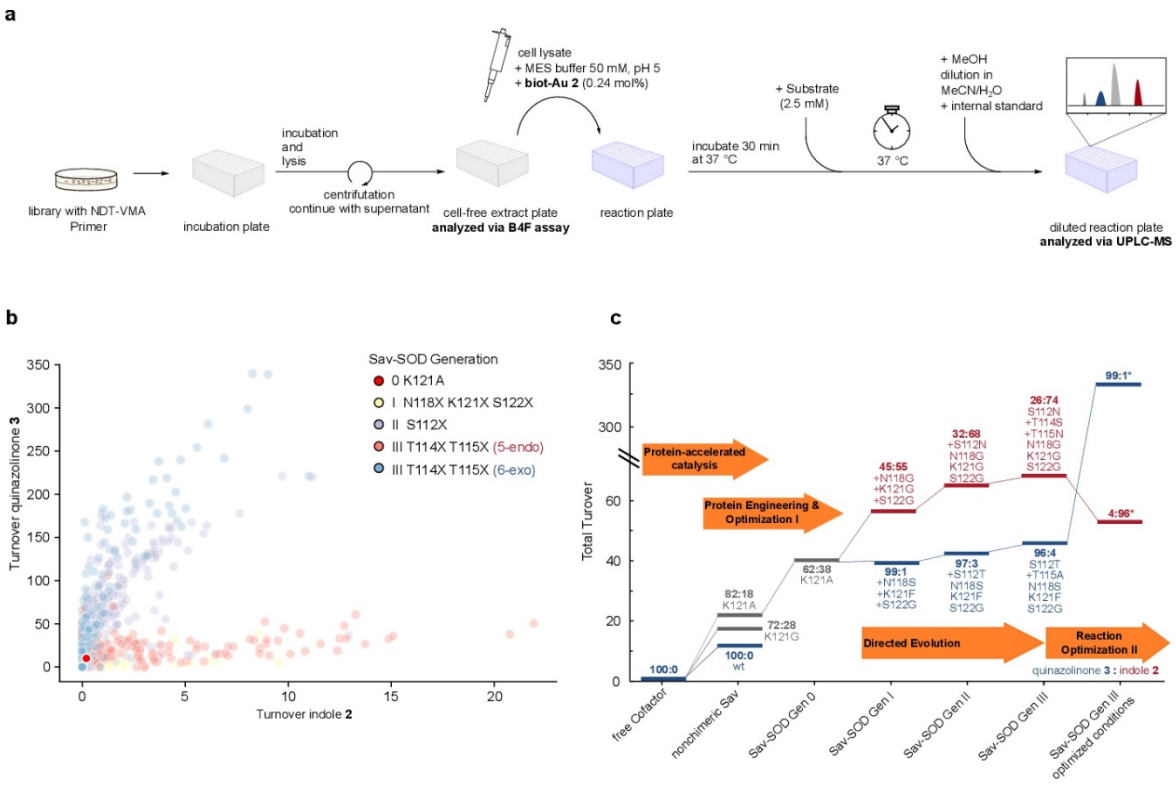


Figure 5 Directed Evolution of an HAMase Based on Sav-SOD. a Streamlined protocol for the genetic optimization of **biot-Au 2** · Sav-SOD in cfe using a 96-well plate assay. **b** Scatterplot of 2500 cfe experiments displaying the two evolutionary trajectories Sav-SOD GGG (red) and Sav-SOD SFG (blue), for amino acid positions selected for mutagenesis, see panel **c**, ultimately leading to the identification of **biot-Au 2** · Sav-SOD T-A-SFG and **biot-Au 2** · Sav-SOD N-SN-GGG. Reaction conditions for 96-well cfe screening: V_{tot} 400 μL (V_{DMSO} 20 μL), [Sub] 2.5 mM, [Cofactor] 6.25 μM, MES-buffer:lysis-buffer (1:1), pH ~5.5, sealed at 37 °C for 4 days. **c** Evolution trajectory of HAMase using purified Sav isoforms of the evolution of **biot-Au 2** · Sav-SOD. Reaction conditions for catalysis with purified protein samples: V_{tot} 200 μL (V_{DMSO} 15 μL), [Sub] 5 mM, [Cofactor] 50 μM, [Sav] 100 μM, [MES-buffer] 50 mM, pH 5, 37 °C for 24 h except optimized reaction conditions* for 6-exo-dig: V_{tot} 400 μL (365 μL of MES-buffer, 50 mM, pH 5) [Sub] 2.5 mM, [**biot-Au 2]** 2.5 μM**,** [Sav] 5 μM, 39 °C for 72 h and for 5-endo-dig: V_{tot} 100 μL (V_{MES} 45 μL,), [Sub] 15 mM, [Diamide] 15 mM, [**biot-Au 2**] 100 μM, [Sav] 200 μM, 37 °C for 72 h.

Directed evolution of the HAMase based on Sav-SOD

Having engineered an evolvable hydrophobic environment lining the biotin-binding vestibule and identified promising residues, we set out to optimize the HAMase by directed evolution. Building on the computational insight, we selected **biot-Au 2** · Sav-SOD K121A as starting point for the iterative saturation mutagenesis. The free cofactor **biot-Au 2** did not display significant HAMase activity in the absence of Sav, either using MES buffer or cfe, Table 1 and Supplementary Table 10. We hypothesize that as **biot-Au 2** is insoluble in the aqueous cfe, the gold is shielded from poisoning by the cellular debris (in particular soluble thiols). Upon compartmentalization within Sav, a soluble and active ArM results, whereby the protein (partially) shields the cofactor from these detrimental metabolites, thus restoring catalytic activity, as previously observed in related studies.[37,60] We were delighted to observe HAMase activity upon addition of **biot-Au 2** to *E. coli* cfe (BL21 DE3) containing Sav and Sav-SOD,

Supplementary Table 10. This strategy allows bypassing the laborious protein purification step and complements our previous high-throughput screening platforms based on periplasmic and surface-display.[25,61,62]

Table 1 Selected Results of evolved ArMs using purified HAMases[a]. Best hits for either regioisomer are highlighted in bold[b,c] and were tested on preparative scale[d,e].

Entry	Sav-Mutants	Total TON[a]	Selectivity (**3:2**)
1	no streptavidin	0.3 ± 0.1	100:0
2	*Sav K121A*	*21.2 ± 0.9*	*82:18*
3	Sav-SOD K121A	39.3 ± 1.9	62:38
4	Sav-SOD N118G K121G S122G	57.5 ± 4.7	45:55
5	Sav-SOD S112N N118G K121G S122G	64.6 ± 2.1	32:68
6	**Sav-SOD S112N T114S T115N N118G K121G S122G**	**51.2 ± 3.5[b]** **39[d]**	**4:96[b]** **8:92[d]**
7	Sav-SOD N118S K121F S122G	38.6 ± 1.0	99:1
8	Sav-SOD S112T N118S K121F S122G	42.0 ± 1.1	97:3
9	**Sav-SOD S112T T115A N118S K121F S122G**	**333 ± 57[c]** **94[e]**	**99:1[c]** **>99:1[e]**

[a]The analytical experiments were carried out in quadruplicates. The combined turnover for both products in relation to [**biot-Au 2**] is displayed as mean ± SD (n=4). The quinazolinone:indole ratio (**3:2**) with standard reaction conditions: V_{tot} 200 μL (V_{DMSO} 15 μL), [Sub] 5 mM, [**biot-Au 2**] 50 μM, [Sav] 100 μM, [MES-buffer] 50 mM, pH 5, 37 °C for 24 h.; [b]Optimized reaction conditions 5-endo-dig with the following changes: V_{tot} 100 μL (V_{MES} 45 μL, V_{DMSO} 20 μL), [Sub] 15 mM, [Diamide] 15 mM, [**biot-Au 2**] 100 μM, [Sav] 200 μM, 37 °C for 72 h; [c]Optimized reaction conditions 6-exo-dig with the following changes: V_{tot} 400 μL (365 μL of MES-buffer 50 mM, pH 5) [Sub] 2.5 mM, [**biot-Au 2]** 2.5 μM**,** [Sav] 5 μM, 39 °C for 72 h; [d]Preparative reaction (0.2 mmol): 72% yield of indole **2** and 6% yield of quinazolinone **3**; [e]Preparative reaction (0.1 mmol): 47% yield of quinazolinone **3** and no isolable trace of indole **2**.

First, positions N118, K121 and S122 were mutated simultaneously using a library consisting of NDT codons at these three positions, Figure 5a, Supplementary Table 11 and Supplementary Figure 28-29.[63]. For this purpose, thirteen 96-well plates, representing 24% coverage, were grown. Each well was subjected to a biotin-4-fluorescein (B4F) binding assay to identify biotin-binding Sav-SOD-mutants that were then evaluated in catalysis. Twelve hits including Sav-SOD N118G K121G S122G (GGG hereafter) and Sav-SOD N118S K121F S122G (SFG hereafter) were quantified by UPLC-MS and displayed either the highest TON or the highest regioselectivity. They were sequenced, expressed, purified by affinity chromatography and tested *in vitro*. Next, the twelve hits were subjected to another round of directed evolution targeting position S112 with 18-possible amino acid combinations using VMA and NDT codons, Supplementary Table 11-12. The best six hits were expressed, purified and evaluated. Mutants Sav-SOD S112N N118G K121G S122G (N-GGG) and Sav-SOD S112T N118S K121F S122G (T-SFG) were identified as most promising for 5-endo-dig

331 product **2** and 6-exo-dig product **3** respectively. These two quadruple mutants were subjected
332 to another round of mutagenesis, targeting simultaneously positions T114 and T115.

333 From this screening campaign, the following trends emerge; large amino acids at position
334 K121 favour the 6-exo-dig product **3**, suggesting that the clashes observed via computational
335 studies between K121 and the σ,π-gold intermediate are also valid for other mutants.
336 Reactions conditions used for cfe experiments tend to favour the formation of 6-exo-dig
337 product **3**. After two rounds of directed evolution starting from Sav-SOD K121A, clear
338 evolutionary trajectories favouring either regioisomer emerge (the scaffold GGG favours the
339 indole **2** and SFG favours the quinazolinone **3** pathways respectively), Figure 5b and
340 Supplementary Figure 28. To validate these trends, the most promising Sav chimeras were
341 expressed in shake flasks and purified. Significant improvement of activity and selectivity
342 over various generations were observed. Ultimately, we identified two chimeras Sav-SOD
343 S112N T114S T115N N118G K121G S122G (N-SN-GGG) and Sav-SOD S112T T115A
344 N118S K121F S122G (T-A-SFG) for the production of either regioisomer: **biot-Au 2** · Sav-
345 SOD T-A-SFG and Sav-SOD N-SN-GGG afforded 333 and 51 TONs with a ratio **3:2** of 99:1
346 and 4:96 respectively, Figure 5c, Table 1 and Supplementary Table 13. To further validate
347 these results, the two best chimeras were scaled-up (> 500 mg) and the two HAMases were
348 tested on preparative scale (0.1 mmol and 0.2 mmol respectively): **biot-Au 2** · Sav-SOD T-
349 A-SFG afforded quinazolinone **3** with 47% isolated yield and no isolable amount of indole **2**
350 (> 99:1 regioselectivity) while **biot-Au 2** · Sav-SOD N-SN-GGG afforded quinazolinone **3**
351 and indole **2** in 6% and 72% isolated yield respectively (8:92 regioselectivity), Supplementary
352 Figures 30-33.

353 **Dual gold mechanism and substrate scope**

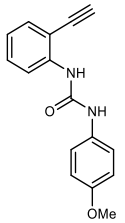
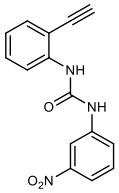
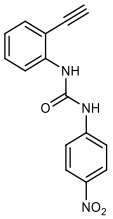
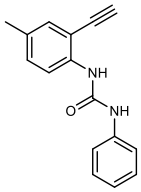
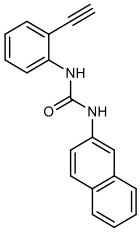
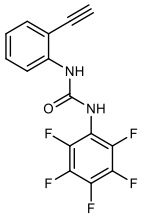
354 To confirm that the 5-endo-dig product **2** isolated using the evolved HAMase Sav-SOD N-SN-
355 GGG and Sav-SOD T-A-SFG indeed results from a dual-gold mechanism –rather than a single-
356 gold mechanism– we carried out preparative-scale reactions in D_2O. The dual activation
357 mechanism proceeds via a di-aurated intermediate, which then undergoes deutero-deauration to
358 regenerate the catalyst and release the (di-deuterated) product **2**-d_2, Supplementary Figure 34.
359 Gratifyingly, both ^{1}H-NMR and HR-MS studies confirm that the isolated indole **2**-d_2, is indeed
360 equally di-deuterated at its C2 and C3 positions (> 98% by ^{1}H-NMR), Supplementary Figures 34-
361 35. We thus conclude that the formation of the indole **2** through either evolved HAMase indeed
362 proceeds via a dual-gold mechanism.

363 Finally, six structurally-related ethynylureas **1a-1f** were tested for their regioselective
364 cyclization in the presence of the two evolved chimeric HAMases, Table 2 and

365 Supplementary Table 14. In some cases, the limited solubility and water-stability of these
366 aromatic ureas proved challenging. Nevertheless, both chimeras afforded preferentially
367 either the quinazolinone- (Sav-SOD T-A-SFG) or the indole products (Sav-SOD N-SN-GGG).

368

369 **Table 2** Substrate scope using the fourth generation of **biot-Au 2** · Sav-SOD mutants obtained with the optimized
370 reaction conditions for either 5-endo-dig[a] or 6-exo-dig[b] products.

Substrate / Mutant	 1a	 1b	 1c	 1d	 1e	 1f
Sav-SOD S112N T114S T115N N118G K121G S122G[a]	45.9 (7:93)	35.9 (19:81)	22.0 (15:85)	15.3 (1:99)	5.2 (4:96)	6.9 (50:50)
Sav-SOD S112T T115A N118S K121F S122G[b]	104.7 (96:4)	35.3 (97:3)	41.7 (97:3)	26.0 (97:3)	16.7 (85:15)[c]	6.0 (96:4)[c]

The combined turnover for both products is displayed in relation to [**biot-Au 2**] as well as the quinazolinone:indole ratio (**3:2**). [a]Optimized reaction conditions 5-endo-dig: V_{tot} 100 μL (V_{MES} 45 μL, V_{DMSO} 20 μL), [Sub] 15 mM, [Diamide] 15 mM, [**biot-Au 2**] 100 μM, [Sav] 200 μM, [MES-buffer] 50 mM, pH 5, 37 °C for 24 h; [b]Optimized reaction conditions 6-exo-dig: V_{tot} 400 μL (V_{MES} 365 μL, V_{DMSO} 12 μL) [Sub] 2.5 mM, **[biot-Au 2]** 6.25 μM**,** [Sav] 12.5 μM, [MES-buffer] 50 mM, pH 5, 39 °C for 48 h; [c] with following changes: **[biot-Au 2]** 2.5 μM**,** [Sav] 5 μM, for 24 h. For a summary of results using the unoptimized reaction conditions used in Table 1, see Supplementary Table 14.

371

372 Conclusions

373 Natural metalloenzymes often rely on dual-catalysis to functionalize challenging substrates.
374 Capitalizing on the unique topology of the biotin-binding vestibule of streptavidin, we
375 designed an artificial hydroaminase that proceeds via a σ,π-activation of a terminal alkyne by
376 two biotinylated gold cofactors. To ensure the positioning of the two gold moieties,
377 streptavidin's biotin-binding vestibule was equipped with a hydrophobic lid, borrowed from
378 superoxide dismutase c.[49] *In silico* modelling of the resulting chimeric HAMase provided
379 insight into the two mechanistic manifolds, and revealed close-lying amino-acid residues to
380 target by directed evolution, to favour the preferential formation of the anti-Markovnikov
381 product indole **2** over the Markovnikov product quinazolinone **3**. These two products result
382 from the dual-gold σ,π-activation and the single gold π-activation reaction manifolds
383 respectively. Thus far, optimization of ArMs'-performance was mostly focused on optimizing
384 enantioselectivity.[22,25] In addition to displaying remarkable levels of enantiocontrol, enzymes

385 excel at imposing catalyst control to address regioselectivity challenges. Herein, we
386 combined protein engineering and directed evolution to fine-tune the second coordination
387 sphere around the abiotic cofactor to control the regioselectivity of the hydroamination
388 reaction. Despite its pronounced thiophylicity, the biotinylated cofactor **biot-Au 2** could be
389 used in the presence of *E. coli* cell free extracts, thus significantly simplifying the directed
390 evolution campaign. We tentatively assign this feature to the cofactor's insolubility in the
391 reaction medium. Upon solubilization resulting from binding to Sav, the cofactor is partially
392 shielded from thiols thus affording biocompatible, active and selective artificial
393 hydroaminases. Current efforts are aimed at integrating this versatile dual-activation reaction
394 *in vivo* to complement the natural metabolism.

395 Methods

396 General information

397 Chemicals were purchased from Sigma Aldrich, Acros Organics, Alfa Aesar or Fluorochem
398 and used without further purification. Dry solvents were directly purchased from Acros
399 Organics and used without further purification. Water used for molecular biology and in the
400 catalytic reactions was purified by Milli-Q Advantage system. Degassed solvents were
401 prepared via three freeze-pump-thaw cycles. All catalytic reactions were carried out with non-
402 degassed solvents under air. Temperature was maintained using Thermowatch-controlled
403 heating blocks. ^{1}H NMR (500 MHz) and ^{13}C NMR (126 MHz) spectra were recorded at room
404 temperature on a Bruker 500 MHz or 600 MHz spectrometer. GC-MS analysis was performed
405 on a Shimadzu GCMS-QP2010S equipped with Agilent HP1-1MS (length: 30 m; Diameter:
406 0.25 mm; Film: 0.25 μM). GC column flow 2.05 mL/min Helium. High-resolution mass
407 spectrometry (HR-MS) was performed on a Bruker maXis II QTOF ESI mass spectrometer
408 coupled to a Shimadzu LC. Fluorescence assays were recorded on a Tecan fluorimeter
409 Infinite M1000Pro. UPLC experiments were performed on an Acquity UPLC-H Class Bio from
410 Waters equipped with a PDA set to 254 nm and a SQ detector 2 with the following column:
411 ACQUITY UPLC, HSS T3 1.8 µm, 2.1 × 100 mm. Molecular biology reagents were
412 purchased from New England Biolabs (NEB), Integrated DNA Technologies (IDT), and
413 Macherey-Nagel and were used as described in the accompanying protocols. Mutations were
414 verified by Sanger sequencing performed by Microsynth (Balgach, Switzerland).

415 Substrate & product synthesis

416 The urea bearing substrates **1** and **1a-1f** were prepared according to a modified procedure
417 of Medio-Simon and co-workers.[5] The quinazolininone products **3** and **3a-3f** were prepared
418 according to a modified procedure of Asensio and co-workers.[41] The indole products **2** and
419 **2a-2f** were synthetized via procedure B1 or B2 (see supplementary information for detailed
420 procedure).

421 Synthesis of biotin analogues and complexes

422 Biotin mesylate was synthetized as follows. A mixture of Biotinol (1 eq), dry CH_2Cl_2 (30 mL),
423 pyridine (10 mL) and diisopropylethylamine (2.5 eq) was cooled to 0 °C, followed by a slow
424 addition of mesyl chloride (3 eq). The solution was stirred 3 hours at 0 °C and quenched. The
425 solution was extracted, dried and concentrated *in vacuo*. The product was precipitated from
426 CH_2Cl_2 by the addition of diethyl ether to yield the product. *N,N*-dimethyl biotin
427 pentafluorophenol was prepared according to a modified procedure of Ward and co-
428 workers.[46] The synthesis of **4DMN** dye was achieved with slight modification of previously
429 published protocols.[64] A detailed description of the synthesis for the gold complexes **biot-Au**
430 **1-5** can be found in the supplementary information.

431 Reaction screening

432 Initial screening experiments were carried out using the following conditions: MES-buffer
433 (50 mM, pH< 5), purified protein (100 μM BBS), cofactor (50 μM in DMSO) and substrate **1**
434 (5 mM in DMSO). After 24 hours of shaking, the reactions were quenched by the addition of
435 methanol (200 μL), followed by centrifugation (14'000 rpm, 10 °C, 10 min). In UPLC vials, the
436 reaction mixture (20 μL) was diluted in UPLC media (980 μL, 1:1 MeCN:MQ-Water containing
437 20.4 μM of phthalane as internal standard) before UPLC-MS analysis. Control experiments
438 were carried out with the same concentrations, see supplementary for detailed information.

439 Protein design and characterization

440 The Sav-SOD chimera gene was synthesized and cloned in to pET28a-(+) vector. Protein
441 production and purification was carried out as previously described. Biotin binding affinities
442 were determined using a Microcal VP-ITC as previously reported by Stayton⁶⁵
443 and Coworkers.[65] For the fluorescent-probe binding assay sodium phosphate buffer (40 mM,
444 pH 7.0) employing Sav K121A or Sav-SOD K121A (1.25 mM) with **biot-4DMN** (40 μM) was
445 used. The excitation wavelength (420 nm) and the emission spectra were collected every
446 1 nm from 450 – 850 nm. Protein melting curves were determined using the protein thermal

447 shift reagent kit (Applied Biosystems) as directed using a StepOne Real Time PCR system
448 and buffer described for ITC analysis. A total sample volume of 20 μL/well was used with a
449 final concentration of 5.25 mg/mL of protein and +/– (0.25 mg/mL) biotin. For HR-MS analysis
450 the proteins were dissolved in acidic Mili-Q water (0.1% formic acid, pH 2.5) with a final
451 concentration of 0.2 mg/mL and clarified by centrifugation. A HPLC (Shimadzu, equipped
452 with a Jupiter© 5 μm C4 300 Å)-ESI QTof (Bruker Daltonics, ESI MaxisII QTof MS) system
453 was used to record the data, see Supplementary Figure 36 and Supplementary Table 15.

454 QM calculations

455 DFT calculations on TSs and cofactor **biot-Au 2** were performed with Gaussian09 program
456 with B3LYP[66] functional including Grimme's dispersion D3 for geometry optimization and
457 frequency calculations. Calculations were carried out in water solvent (SMD continuum
458 model) with ε= 78.35, except when testing the effect of the protein environment (ε =2, 4, 7,
459 20 and 35). The basis set 6-31G(d+p) was used for non-metallic atoms and SDD for Au
460 (including f polarization function). Energies in water were refined using Def2QZVP for Au and
461 Def2TZVP for non-metallic atoms.

462

463 Protein set up and SOD construct

464 For Sav-WT and Sav-K121A systems, calculations were carried out using the x-ray crystal
465 available (pdb code: 3RY2). The systems were prepared by removing waters, ions and small
466 ligands (except biotin). Duplicate conformers of amino acids were removed and hydrogen
467 atoms were added using Chimera UCSF. The Lys121 was mutated to Ala using the Dunbrack
468 rotamer library.[59] Because the X-ray structure of Sav-SOD-K121A was not fully resolved, a
469 homology model was built for Sav-SOD-K121A using Modeller9.21[54] using loop modelling for
470 the missing region (Ala37-Lys71 from PDB:1PZS). The best model was selected according
471 to the SCOPE score and submitted to a MD simulation of 200ns with Sav region constrained
472 to allow SOD stabilization and accommodation into Sav.

473 Protein-ligand docking

474 The DFT optimized structures of the cofactor **biot-Au 2** and the different transition state
475 structures were incorporated into the binding site of the three protein (i.e. Sav-WT, Sav-
476 K121A, Sav-SOD-K121A) by protein-ligand dockings. This was performed on the most

477 populated clusters from previous MD simulations for Sav-WT, Sav-K121A, Sav-SOD-K121A
478 and using a covalent protocol with a fixed biotin. Two software were used. **GaudiMM**[67] and
479 **GOLD5.8.1**.[55] All results were visualized and analysed using GaudiView.[51]

480 Molecular Dynamics

481 All MD simulations were prepared with the xleap from AMBER18.[55] The systems were
482 solvated in a cubic box with a neutral charge (neutralization with Na^+ and Cl^-). The
483 AMBER14SB force field was used for proteins, GAFF for non-standard residues, ions94.lif
484 for ions and TIP3P for water. The parameters for **biot-Au 2** and the corresponding transition
485 states were calculated using MCPB.py and charges with RESP. The force constants and
486 equilibrium parameters between metal and residues were obtained through the Seminario
487 method. Simulations were run with the OpenMM7.1 toolkit with the OMMprotocol. The
488 Langevin integrator was used with a time step of 1fs with periodic conditions. The simulation
489 was performed with the SHAKE approximation and using a barostat coupled to a bath of
490 1.01325 bar, a cut-off 1nm for non-bonded interactions (short-range electrostatic and Van
491 der Waals interactions), and the PME method for long-range electrostatic interactions. Initial
492 energy minimizations were performed (400000 steps) followed by equilibration steps to heat
493 the system from 100K to 300K. Finally, production runs of 300ns were carried out which
494 allow to reach convergence.

495 Clustering, RMSD, all-to-all RMSD, PCA analysis were analysed with MDtraj and cpptraj.
496 Residue contributions analysis was performed using StructureViz[58] from the MD analysed in
497 UCSF Chimera and the MMGBSA method implemented in the module MMPBSA.py.[68]

498 Directed evolution and screening procedure

499 Mutant plasmid libraries were transformed into BL21 (DE3) cells and selected on LB-Agar
500 plates containing kanamycin (50 mg/L). Individual colonies were used and protein production
501 was carried out in 96-deep well plates as previously described.[61] Cell-free extracts were used
502 for reaction screening and hit identification. Hit selection was based on the measured activity
503 and selectivity in relation to the B4F assay as the expression levels varied. For the first round
504 of directed evolution, mutagenesis on positions N118, K121 and S122 was carried out,
505 followed by position S112 and finally T114 and T115. Site directed mutagenesis was
506 achieved using the primers listed in Supplementary Table 10 followed by BsaI/DpnI digestion
507 and ligated using T4 ligase.

508 ## ArM crystallography and native MS characterization

509 For co-crystallization experiments 10 μL **biot-Au 2** (20 mM in DMSO) was added to a protein
510 solution (2 mg/mL in 1 mL 20 mM Tris-HCl, pH 7.4). After 18 h of soaking at rt, the solution
511 was subjected to concentration and buffer exchange (water) using a Nanosep Centrifugal
512 Device with an Omega Membrane (3K, Pall). Crystals grew within 21 days and were flash
513 frozen. X-ray diffraction patterns were collected at the Swiss Light Source beam line PXIII
514 (100 K, wavelength of 0.999 Å). For detailed information see Supplementary Table 9. The
515 finalized structure was uploaded to the pdb database and can be found under the code 7ALX.
516 Sav and Sav-SOD based ArMs were further characterized by native MS, see Supplementary
517 Figure 37-40 and Supplementary Table 16-17.

518

519 ## Data availability

520 Data relating to the materials and methods, detailed substrate and cofactor synthesis,
521 optimization studies, catalytic experiments, protein expression, MD & DFT calculations,
522 selected UPLC-MS chromatograms, HRMS spectra and NMR studies are available in the
523 Supplementary Information. Crystallographic data for **biot-Au 2** · Sav-SOD K121A is
524 available free of charge from the PDB under reference number 7ALX. All other data are
525 available from the authors upon request.

526

527 ## References

528 1. McLean, E. B. & Lee, A. L. Golden potential. *Nat. Chem.* **11**, 760–761 (2019).

529 2. Dorel, R. & Echavarren, A. M. Gold(I)-Catalyzed Activation of Alkynes for the
530 Construction of Molecular Complexity. *Chem. Rev.* **115**, 9028–9072 (2015).

531 3. Pflästerer, D. & Hashmi, A. S. K. Gold catalysis in total synthesis - Recent
532 achievements. *Chemical Society Reviews* vol. 45 1331–1367 (2016).

533 4. Stephen, A., Hashmi, K., Braun, I., Rudolph, M. & Rominger, F. The role of gold
534 acetylides as a selectivity trigger and the importance of gem-diaurated species
535 in the gold-catalyzed hydroarylating-aromatization of arene-diynes.

536 *Organometallics* **31**, 644–661 (2012).

537 5. Gimeno, A., Medio-Simón, M., De Arellano, C. R., Asensio, G. & Cuenca, A. B.
538 NHC-stabilized gold(I) complexes: Suitable catalysts for 6-exo-dig
539 heterocyclization of 1-(o-Ethynylaryl)ureas. *Org. Lett.* **12**, 1900–1903 (2010).

540 6. Ye, L., Wang, Y., Aue, D. H. & Zhang, L. Experimental and computational
541 evidence for gold vinylidenes: Generation from terminal alkynes via a
542 bifurcation pathway and facile C-H insertions. *J. Am. Chem. Soc.* **134**, 31–34
543 (2012).

544 7. Zhao, X., Rudolph, M. & Hashmi, A. S. K. Dual gold catalysis-an update. *Chem.*
545 *Commun.* **55**, 12127–12135 (2019).

546 8. Wang, W. *et al.* Dinuclear gold catalysis. *Chem. Soc. Rev.* **50**, 1874–1912
547 (2021).

548 9. Bertini, I., Gray, H. B., Valentine, J. S. & Stiefel, E. I. *Biological Inorganic*
549 *Chemistry : Structure and Reactivity*. (CA: University Science Books, 2007).

550 10. Jewett, J. C. & Bertozzi, C. R. Cu-free click cycloaddition reactions in chemical
551 biology. *Chem. Soc. Rev.* **39**, 1272–1279 (2010).

552 11. Devaraj, N. K. The Future of Bioorthogonal Chemistry. *ACS Cent. Sci.* **4**, 952–
553 959 (2018).

554 12. Pickens, C. J., Johnson, S. N., Pressnall, M. M., Leon, M. A. & Berkland, C. J.
555 Practical Considerations, Challenges, and Limitations of Bioconjugation via
556 Azide-Alkyne Cycloaddition. *Bioconjug. Chem.* **29**, 686–701 (2018).

557 13. Boren, B. C. *et al.* Ruthenium-catalyzed azide-alkyne cycloaddition: Scope and
558 mechanism. *J. Am. Chem. Soc.* **130**, 8923–8930 (2008).

559 14. Vidal, C., Tomás-Gamasa, M., Destito, P., López, F. & Mascareñas, J. L.
560 Concurrent and orthogonal gold(I) and ruthenium(II) catalysis inside living cells.
561 *Nat. Commun.* **9**, 1–9 (2018).

562 15. Tsubokura, K. *et al.* In Vivo Gold Complex Catalysis within Live Mice. *Angew.*
563 *Chemie - Int. Ed.* **56**, 3579–3584 (2017).

564 16. Pérez-López, A. M. *et al.* Gold-Triggered Uncaging Chemistry in Living
565 Systems. *Angew. Chem. Int. Ed.* **56**, 12548–12552 (2017).

566 17. Jung Jou, M. *et al.* Highly selective fluorescent probe for Au(III) based on
567 cyclization of propargylamide. *Chem. Commun.* **46**, 7218–7220 (2009).

568 18. Wang, J. B., Wu, Q. Q., Min, Y. Z., Liu, Y. Z. & Song, Q. H. A novel fluorescent
569 probe for Au(III)/Au(I) ions based on an intramolecular hydroamination of a
570 Bodipy derivative and its application to bioimaging. *Chem. Commun.* **48**, 744–
571 746 (2012).

572 19. Chen, K. & Arnold, F. H. Engineering new catalytic activities in enzymes. *Nature*
573 *Catalysis* vol. 3 203–213 (2020).

574 20. Wilson, M. E. & Whitesides, G. M. Conversion of a Protein to a Homogeneous
575 Asymmetric Hydrogenation Catalyst by Site-Specific Modification with a
576 Diphosphinerhodium(I) Moiety. *J. Am. Chem. Soc.* **100**, 306–307 (1978).

577 21. Studer, S. *et al.* Evolution of a highly active and enantiospecific metalloenzyme
578 from short peptides. *Science* **362**, 1285–1288 (2018).

579 22. Gu, Y., Natoli, S. N., Liu, Z., Clark, D. S. & Hartwig, J. F. Site-Selective
580 Functionalization of (sp3)C–H Bonds Catalyzed by Artificial Metalloenzymes
581 Containing an Iridium-Porphyrin Cofactor. *Angew. Chemie - Int. Ed.* **58**, 13954–
582 13960 (2019).

583 23. Roelfes, G. LmrR : A Privileged Scaffold for Artificial Metalloenzymes. *Acc.*
584 *Chem. Res.* **52**, 545–556 (2019).

585 24. Song, W. J. & Tezcan, F. A. A designed supramolecular protein assembly with
586 in vivo enzymatic activity. *Science* **346**, 1525–1528 (2014).

587 25. Schwizer, F. *et al.* Artificial Metalloenzymes: Reaction Scope and Optimization

588 Strategies. *Chem. Rev.* **118**, 142–231 (2018).

589 26. Hyster, T. K., Knörr, L., Ward, T. R. & Rovis, T. Biotinylated Rh(III) complexes
590 in engineered streptavidin for accelerated asymmetric C-H activation. *Science*
591 **338**, 500–503 (2012).

592 27. Zhou, Z. & Roelfes, G. Synergistic catalysis in an artificial enzyme by
593 simultaneous action of two abiological catalytic sites. *Nat. Catal.* **3**, 289–294
594 (2020).

595 28. Alonso, S. *et al.* Genetically engineered proteins with two active sites for
596 enhanced biocatalysis and synergistic chemo- and biocatalysis. *Nat. Catal.* **3**,
597 319–328 (2020).

598 29. Martínez-Calvo, M. *et al.* Intracellular Deprotection Reactions Mediated by
599 Palladium Complexes Equipped with Designed Phosphine Ligands. *ACS Catal.*
600 **8**, 6055–6061 (2018).

601 30. Monnard, F. W., Nogueira, E. S., Heinisch, T., Schirmer, T. & Ward, T. R.
602 Human carbonic anhydrase II as host protein for the creation of artificial
603 metalloenzymes: The asymmetric transfer hydrogenation of imines. *Chem. Sci.*
604 **4**, 3269–3274 (2013).

605 31. Oohora, K., Onoda, A. & Hayashi, T. Hemoproteins Reconstituted with Artificial
606 Metal Complexes as Biohybrid Catalysts. *Acc. Chem. Res.* **52**, 945–954 (2019).

607 32. Mirts, E. N., Petrik, I. D., Hosseinzadeh, P., Nilges, M. J. & Lu, Y. A designed
608 heme-[4Fe-4S] metalloenzyme catalyzes sulfite reduction like the native
609 enzyme. *Science* **361**, 1098–1101 (2018).

610 33. Lewis, J. C. Beyond the Second Coordination Sphere: Engineering Dirhodium
611 Artificial Metalloenzymes To Enable Protein Control of Transition Metal
612 Catalysis. *Acc. Chem. Res.* **52**, 576–584 (2019).

613 34. Lombardi, A., Pirro, F., Maglio, O., Chino, M. & DeGrado, W. F. De novo design
614 of four-helix bundle metalloproteins: One scaffold, diverse reactivities. *Acc.*

615 *Chem. Res.* **52**, 1148–1159 (2019).

616 35. Chino, M. *et al.* Artificial Diiron Enzymes with a de Novo Designed Four-Helix
617 Bundle Structure. *European Journal of Inorganic Chemistry* vol. 2015 3371–
618 3390 (2015).

619 36. Grimm, A. R. *et al.* A Whole Cell E. coli Display Platform for Artificial
620 Metalloenzymes: Poly(phenylacetylene) Production with a Rhodium-
621 Nitrobindin Metalloprotein. *ACS Catal.* **8**, 2611–2614 (2018).

622 37. Eda, S. *et al.* Biocompatibility and therapeutic potential of glycosylated albumin
623 artificial metalloenzymes. *Nat. Catal.* **2**, 780–792 (2019).

624 38. Liang, A. D., Serrano-Plana, J., Peterson, R. L. & Ward, T. R. Artificial
625 Metalloenzymes Based on the Biotin-Streptavidin Technology: Enzymatic
626 Cascades and Directed Evolution. *Acc. Chem. Res.* **52**, 585–595 (2019).

627 39. Heinisch, T. & Ward, T. R. Artificial Metalloenzymes Based on the Biotin-
628 Streptavidin Technology: Challenges and Opportunities. *Acc. Chem. Res.* **49**,
629 1711–1721 (2016).

630 40. Vornholt, T. *et al.* Systematic engineering of artificial metalloenzymes for new-
631 to-nature reactions. *Sci. Adv.* **7**, 1–12 (2021).

632 41. Gimeno, A. *et al.* Competitive gold-activation modes in terminal alkynes: An
633 experimental and mechanistic study. *Chem. - A Eur. J.* **20**, 683–688 (2014).

634 42. Vreeken, V. *et al.* Well-Defined Dinuclear Gold Complexes for Preorganization-
635 Induced Selective Dual Gold Catalysis. *Angew. Chem. Int. Ed.* **55**, 10042–
636 10046 (2016).

637 43. Ye, D. *et al.* Gold-catalyzed intramolecular hydroamination of terminal alkynes
638 in aqueous media: Efficient and regioselective synthesis of indole-1-
639 carboxamides. *Green Chem.* **11**, 1201–1208 (2009).

640 44. Breker, V., Sak, H., Baracchi-Krause, G. & Krause, N. Synthesis and properties

641 of a biotin-tagged NHC-gold complex. *Tetrahedron Lett.* **56**, 3390–3392 (2015).

642 45. Collado, A., Gómez-Suárez, A., Martin, A. R., Slawin, A. M. Z. & Nolan, S. P.
643 Straightforward synthesis of [Au(NHC)X] (NHC = N-heterocyclic carbene, X =
644 Cl, Br, I) complexes. *Chem. Commun.* **49**, 5541–5543 (2013).

645 46. Kajetanowicz, A., Chatterjee, A., Reuter, R. & Ward, T. R. Biotinylated
646 metathesis catalysts: Synthesis and performance in ring closing metathesis.
647 *Catal. Letters* **144**, 373–379 (2014).

648 47. Anhäuser, L., Teders, M., Rentmeister, A. & Glorius, F. Bio-additive-based
649 screening: toward evaluation of the biocompatibility of chemical reactions. *Nat.*
650 *Protoc.* **14**, 2599–2626 (2019).

651 48. Muñoz Robles, V., Vidossich, P., Lledós, A., Ward, T. R. & Maréchal, J. D.
652 Computational insights on an artificial imine reductase based on the biotin-
653 streptavidin technology. *ACS Catal.* **4**, 833–842 (2014).

654 49. Spagnolo, L. *et al.* Unique features of the sodC-encoded superoxide dismutase
655 from Mycobacterium tuberculosis, a fully functional copper-containing enzyme
656 lacking zinc in the active site. *J. Biol. Chem.* **279**, 33447–33455 (2004).

657 50. Loving, G. & Imperiali, B. Thiol-reactive derivatives of the solvatochromic 4-
658 N,N-dimethylamino-1,8- naphthalimide fluorophore: A highly sensitive toolset
659 for the detection of biomolecular interactions. *Bioconjug. Chem.* **20**, 2133–2141
660 (2009).

661 51. Alonso-Cotchico, L., RodrÍguez-Guerra, J., Lledós, A. & Maréchal, J. D.
662 Molecular Modeling for Artificial Metalloenzyme Design and Optimization. *Acc.*
663 *Chem. Res.* **53**, 896–905 (2020).

664 52. Marenich, A. V., Cramer, C. J. & Truhlar, D. G. Universal solvation model based
665 on solute electron density and on a continuum model of the solvent defined by
666 the bulk dielectric constant and atomic surface tensions. *J. Phys. Chem. B* **113**,
667 6378–6396 (2009).

53. Kiss, P. T. & Baranyai, A. A systematic development of a polarizable potential of water. *J. Chem. Phys.* **138**, 204507 (2013).

54. Webb, B. & Sali, A. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinforma.* **2016**, 5.6.1-5.6.37 (2016).

55. Case, D. A. *et al.* Amber 2018. *Univ. California, San Fr. 2018* 1–923 (2018).

56. Gareth Jones, Peter Willett, Robert C. Glen , Andrew R. Leach, R. T. Development and Validation of a Genetic Algorithm for Flexible Docking Gareth. *J. Mol. Biol.* **267**, 727–748 (1997).

57. Nanda, V. & Koder, R. L. Designing artificial enzymes by intuition and computation. *Nat. Chem.* **2**, 15–24 (2010).

58. Morris, J. H., Huang, C. C., Babbitt, P. C. & Ferrin, T. E. StructureViz: Linking Cytoscape and UCSF Chimera. *Bioinformatics* **23**, 2345–2347 (2007).

59. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

60. Wilson, Y. M., Dürrenberger, M., Nogueira, E. S. & Ward, T. R. Neutralizing the detrimental effect of glutathione on precious metal catalysts. *J. Am. Chem. Soc.* **136**, 8928–8932 (2014).

61. Hestericová, M. *et al.* Directed Evolution of an Artificial Imine Reductase. *Angew. Chem. Int. Ed.* **57**, 1863–1868 (2018).

62. Jeschek, M. *et al.* Directed evolution of artificial metalloenzymes for in vivo metathesis. *Nature* **537**, 661–665 (2016).

63. Qu, G., Li, A., Acevedo-Rocha, C. G., Sun, Z. & Reetz, M. T. The Crucial Role of Methodology Development in Directed Evolution of Selective Enzymes. *Angew. Chem. Int. Ed.* vol. 59 13204–13231 (2020).

64. Wen, J. *et al.* Naphthalimide-rhodamine based fluorescent probe for ratiometric sensing of cellular pH. *Chinese Chem. Lett.* **28**, 2005–2008 (2017).

65. Chu, V., Stayton, P. S., Freitag, S., Le Trong, I. & Stenkamp, R. E. Thermodynamic and structural consequences of flexible loop deletion by circular permutation in the streptavidin-biotin system. *Protein Sci.* **7**, 848–859 (1998).

66. Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **98**, 5648–5652 (1993).

67. Rodríguez-Guerra Pedregal, J., Sciortino, G., Guasp, J., Municoy, M. & Maréchal, J.-D. GaudiMM: A modular multi-objective platform for molecular modeling. *J. Comput. Chem.* **38**, 2118–2126 (2017).

68. Miller, B. R. *et al.* MMPBSA.py: An efficient program for end-state free energy calculations. *J. Chem. Theory Comput.* **8**, 3314–3321 (2012).

Acknowledgements

TRW thanks the ERC advanced grant (the DrEAM, grant agreement 694424), the Swiss National Science Foundation (Grant SNF Grant 200020_182046) and the NCCR Molecular Systems Engineering for generous support. We thank the Analytical Team of the Chemistry Department of the University of Basel, in particular Michael Pfeffer and Sylvie Mittelheisser for HR MS analysis and Prof. Dr. Daniel Häussinger for assistance with the two-dimensional NMR experiments. We thank Juliane Klehr and Amanda Santos Kron for their assistance with protein expression and protein purification as well as Dr. Johannes Georg Rebelein for assistance with the protein crystallography. LTS, AL and JDM thank the Spanish MINECO (grant CTQ2017-87889-P) and the Generalitat de Catalunya (2017SGR1323) for the financial support. LTS thanks the Spanish Ministerio de Ciencia, Innovación y Universidades (grant FPU18/05895) for the financial support. We thank Dr. Giuseppe Sciortino and José Emilio Sánchez Aparicio for assistance with the molecular modelling set up and analysis.

Correction to: Nature Catalysis https://doi.org/10.1038/s41929-021-00651-9, published online 2 August 2021.

In the version of this article initially published, there was an omission in the Acknowledgements section. The text has now been amended to include the following: “B.L.

723 thanks the European Union's Horizon 2020 research and innovation program under the Marie
724 Skłodowska-Curie grant agreement no. 765497 (THERACAT) for generous support."

725 The change has been made in the HTML and PDF versions of the article.

726

727 Author Contributions

728 TRW, RLP and FC conceived and designed the study. FC, MMP and BL contributed to the
729 synthesis of the substrates, products and complexes. NVI, DCS, RLP and FC contributed to
730 mutagenesis, protein expression, protein purification and protein characterization. NVI
731 performed the crystallization, X-ray structure determinations and the native MS experiments.
732 FC performed the catalytic, preparative, and deuterium-labelling experiments, designed the
733 screening protocol and recorded the data. TRW, FC and NVI analysed the data. JDM, AL,
734 and LTS contributed to the molecular modelling experiments. TRW, FC and NVI wrote the
735 manuscript, which was further supplemented through contributions from RLP, NVI and JDM.
736 All authors have given approval to the final version of the manuscript.

737 Competing Interests

738 The authors declare no competing interests.

739