Chimeric streptavidins as host proteins for artificial metalloenzymes

Chimeric Streptavidins as Host Proteins for Artificial Metalloenzymes

Michela M. Pellizzoni,[1†] Fabian Schwizer,[1] Christopher W. Wood,[2] Valerio Sabatino,[1] Yoann Cotelle,[3‡] Stefan Matile,[3] Derek N. Woolfson,[2] Thomas R. Ward[1,*]*

[1]University of Basel, Department of Chemistry, Mattenstrasse 24a, BPR 1096, CH 4002 Basel, Switzerland.

[2]School of Chemistry, University of Bristol, Bristol, BS8 1TS, United Kingdom.

[3]School of Chemistry and Biochemistry, National Centre of Competence in Research (NCCR) Chemical Biology, University of Geneva, Quai Ernest Ansermet 30, CH-1211 Geneva, Switzerland.

KEYWORDS Artificial metalloenzyme, chimeric protein, protein engineering, protein design, transfer hydrogenation, ring closing metathesis, anion π catalysis.

ABSTRACT The streptavidin scaffold was expanded with well-structured naturally occurring motifs. These chimeric scaffolds were tested as host for biotinylated catalysts as artificial metalloenzymes (ArM) for asymmetric transfer hydrogenation, ring closing metathesis and anion π catalysis. The additional second coordination sphere elements were shown to significantly influence both the activity and the selectivity of the resulting hybrid catalysts. These findings

have identified propitious chimeric streptavidins for future directed evolution efforts of artificial metalloenzymes.

INTRODUCTION: Artificial metalloenzymes (ArMs hereafter), as introduced by Whitesides in 1978,[1] aim to combine the best features of both enzymes and homogeneous catalysts. ArMs aren made by anchoring an abiotic cofactor within a host protein.[2-7] This field has gained popularity thanks to the widespread use of molecular biology and recombinant protein production. Since our first report on ArMs based on the biotin-streptavidin technology in 2003,[8] we have systematically pursued a chemo-genetic optimization strategy to improve the performance of these hybrid catalysts incrementally.[9-10] Critically, this versatile strategy relies on our ability of produce streptavidin (Sav) libraries and to screen these with a variety of different biotinylated cofactors. Having identified an active ArM, twenty to fifty single, double and triple point Sav mutants are typically screened using standard directed evolution schemes. Despite the versatility of site-directed mutagenesis and directed evolution,[11] point mutations do not allow the major remodeling of the active site. While such optimization efforts have proven successful, resulting in up to hundred fold increase in rate,[12] more-diverse libraries may allow further optimization of ArMs displaying catalytic efficiencies approaching those of natural enzymes.[13-15] The structures of ArMs based on the biotin-streptavidin technology highlights the narrow dispersion of the position of metal within the biotin-binding vestibule. With few exceptions however, each ArM requires a different Sav mutant for improved performance.[16]

Inspection of the biotin-binding vestibule of Sav where the catalytic event occurs suggests that it is ideally suited to host both the cofactor and its substrate. However, the low plasticity of Sav, as revealed by superposition of X-ray structures of Sav and metal-loaded Sav, as well as its shallow “active site” exposes the metal to the reaction medium (Figure 1). We hypothesized that

the introduction of additional structural features may offer opportunities to protect the active site. With this goal in mind, we set out to engineer Sav by introducing additional structural motifs around the biotin-binding vestibule. The chimeras presented herein were designed to enable the exploration of diverse topologies around the active site by the introduction of additional second coordination-sphere elements.[17]

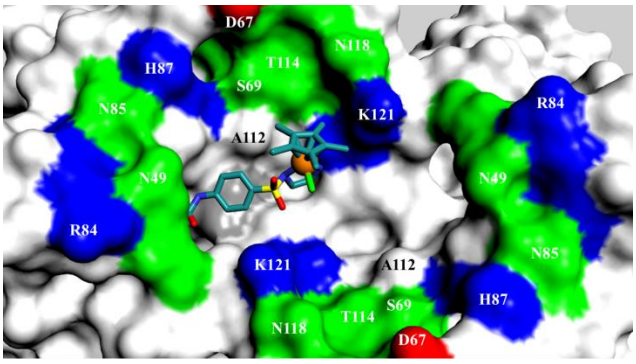


Figure 1. Close-up view of the X-ray structure of homotetrameric Sav S112A with one biotinylated cofactor [Cp*Ir(biot-*p*-L)Cl] bound (PDB ID 3PK2).[18] The surface representation displays a single cofactor (stick representation; Ir, orange sphere, Cl, green) and close-lying amino-acid residues (color code: red = acidic, blue = basic, green = polar, white = apolar). The shallowness of the biotin-binding vestibule (*i.e.*, the active site) suggests that engineering additional structural motifs may provide additional second coordination sphere elements around the biotinylated cofactor.

Thus, we set out to modify the ArMs genetically to introduce well-defined secondary structures around the active site. Thanks to the dimer-of-dimer nature of homotetrameric Sav, genetic modifications are reflected twice in each of the two active sites (Figure 2) and thus can be anticipated to have a significant shielding effect on the Sav-embedded cofactor.

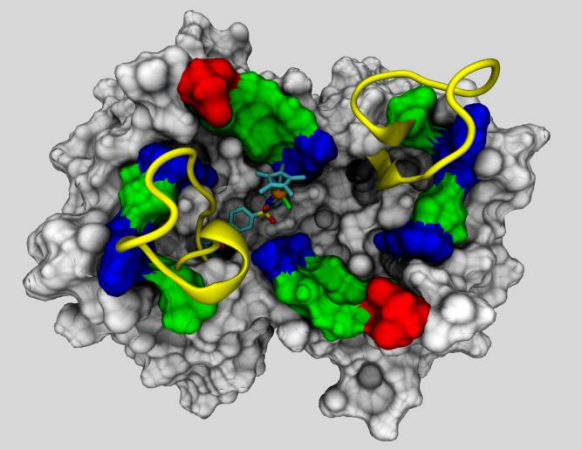


Figure 2. Cartoon representation of a homotetrameric Sav (surface representation) construct bearing an additional motif (entry 8 in Table 1) engineered in loop 3,4 (yellow cartoon representation). This model for the chimeric protein was generated from PDB ID 3PK2 using homology modeling and structure refinement with Yasara. (color code: red = acidic, blue = basic, green = polar, white = apolar)

RESULTS AND DISCUSSION

Protein design. Our initial efforts focused on inserting $(GGX)_n$ repeat sequences (where X = ..., and n = 1 – 8) at various loop positions that make up the biotin-binding vestibule. We selected the following positions for these insertions: G48–N49, T66–D67, R84–N85 and A117–N118 to introduce one or two $(GGX)_n$ repeat motifs (Table 1).

Table 1 Summary of the expressed Sav constructs bearing $(GGX)_n$ repeat sequences**.**

entry	Sav loop	loop sequence	N° of	Expression
1	3/4	G_{48}-GGSGGS-N_{49}	2	13.8
2	4/5	T_{66}-GGSGGS-D_{67}	2	16.3
3	5/6	R_{84}-GGSGGS-N_{85}	2	13.0
4	7/8	A_{117}-GGSGGS-N_{118}	2	not purified
5	3/4 + 4/5	G_{48}-GGSGGS-N_{49}	4	1.5
6	3/4 + 5/6	G_{48}-GGSGGS-N_{49}	4	0.9
7	3/4 + 7/8	G_{48}-GGSGGS-N_{49}	4	not purified
8	3/4	G_{48}-GGNGGNGGGGGVGGS-N_{49}	5	55.0
9	3/4	G_{48}-GGIGGSGGGGGHGGRGGGGGVGGS-N_{49}	8	not purified
10	3/4	G_{48}-GGNGGSGGGGGGGGSGGSGGS-N_{49}	7	not purified

11	3/4	G_{48}-GGRGGGGGHGGCGGVGGS-N_{49}	6	not purified

[a]Expressed as soluble and functional (i.e. binds biotin-4-fluorescein) tetrameric fraction. The expression conditions are detailed in the SI.

Having identified Sav positions tolerant to $(GGX)_n$ insertions, we set out to replace these GGX repeats, which we presumed would be unstructured, by naturally occurring motifs with a well-defined secondary structure. We pursued the construction of two different chimera streptavidin families: a) streptavidin containing an extended 2D topology structure (24 – 60 amino acid residues, 2D_Sav, hereafter, Table 2) and b) streptavidin containing shorter naturally occurring loops (5-12 amino acid residues, MP_Sav, hereafter, Table 3).

The first family with additional structured elements around the biotin-binding vestibule, 2D_Sav, was created by fusion with highly conserved and structured peptide sequences (see SI) that display close-lying *N* and *C* termini. For this purpose, the following sequences were selected: SH3 (Src Homology 3 domain, 60 aa) containing antiparallel β-sheets,[19] AR (Ankyrin, 30-34 aa),[20] HP (Villin Headpiece subdomain, 35 aa),[21] FPD (FoldIt Players Design, 24 aa)[22] containing an helix-turn-helix motif, and PPR (Penta-trico-Peptide Repeat, 35 aa) containing a coiled-coil domain[23] (Figure 3 and Table 2). Inspection of the X-ray structure of *mature* apo-Sav, which bears an N_{term}-T7-solubility tag at positions 1 – 12 and extends to position 159,[24] revealed that the C_{term} occupies the biotin-binding site.[25] Therefore, we introduced at the C_{term} either the SH3 or PPR consensus sequence to provide an additional well-structured secondary structure close to biotin-binding vestibule. The SH3, AR, HP and FDP motifs were introduced in four different loop regions of Sav taking into consideration their $N_{term} - C_{term}$ distance. Thus, nineteen different chimeric Sav genes were constructed (Figure 3 and Table 2).

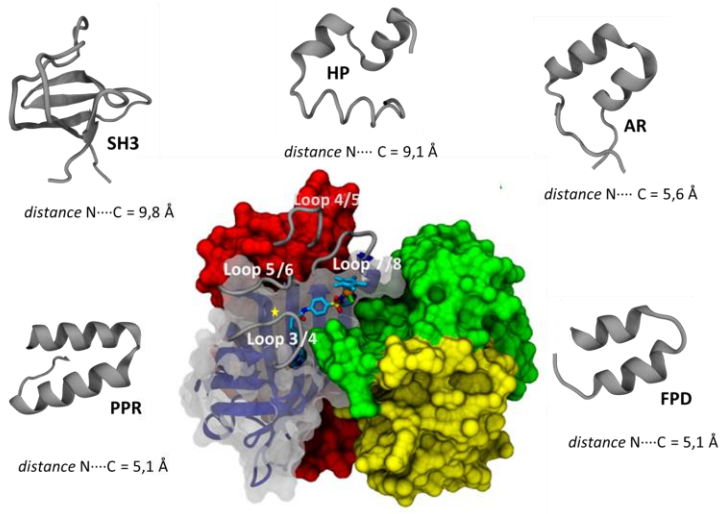


Figure 3 Surface display representation of the Sav tetramer. Monomer surfaces are highlighted in green, yellow, red and translucent grey. The Sav loops that were engineered to create chimera proteins are labelled: loop 3/4, residues 46 – 52; loop 4/5, residues 63 – 70; loop 5/6, residues 82 – 87; and loop 7/8, residues 113 – 117). The inserted structures with N_{term}-C_{term} distances (N · · · · ·C) are displayed as grey cartoon representation: SH3 (Src Homology 3 domain, ca. 60 aa, antiparallel β-sheets);[19] AR (Ankyrin, 30-34 aa, helix-turn-helix);[20] HP (Villin Headpiece subdomain, 35 aa, helix-turn-helix);[21] FPD (FoldIt Players Design, 24 aa, helix-turn-helix);[22] and PPR (Penta-trico-Peptide Repeat 35 aa, coiled coil).[23] The yellow the star highlights the position selected for the creation of Chimera MP_Sav family.

Table 2. Chimeric Proteins (2D_Sav's family) produced recombinantly in this study. Each protein is abbreviated with the acronym of the introduced motif; the number indicates the position of the insertion in the original Sav sequence.[a]

entry	abbreviation	MW (g/mol)[b]	solubility
1	PPR_159	20515.5	Yes
2	SH3_46-52	22877.2	After refolding
3	SH3_63-70	22819.2	No
4	SH3_82-87	22843.1	No
5	SH3_113-117	23016.4	After refolding
6	SH3_159	23347.7	Yes
7	AR_46-52	19457.4	After refolding
8	AR_64-70	19496.5	After refolding
9	AR_81-84	19650.5	No
10	AR_115-117	19798.8	After refolding
11	HP_46-52	20018.0	Yes
12	HP_63-70	19960.0	After refolding
13	HP_64-70	20057.2	Yes

14	HP_82-84	20325.4	Yes
15	HP_113-117	20157.3	No
16	FDP_46_52	18313.0	Yes
17	FDP_64-70	18352.1	After refolding
18	FDP_81-84	18506.2	Yes
19	FDP_115-117	18654.4	After refolding

[a] The expression and refolding conditions are detailed in the SI. [b] Molecular weight of the monomer.

The second family of chimera, MP_Sav, was generated by selecting natural loops with appropriate N_{term}–C_{term} distances for the substitution of Sav's residues between A46 and A50 (loop 3/4, Figure 3). To generate initial models of the chimeric proteins, a database of loop regions from X-ray protein crystal structures from the Protein Data Bank[REF] was created using ISAMBARD.[26] Loops were defined as any continuous region of backbone containing any mixture of random coil, hydrogen-bonded turn isolated beta-bridge or bend, as identified by DSSP.[27-28] Tools within ISAMBARD were then used to extract the backbone structure of the loop, which was stored in a database along with the PDB ID, chain and residue labels, surrounding regions of secondary structure, sequence, end-to-end distance (i.e. the distance between the C_α of the residue preceding the loop and the C_α of the residue immediately after the loop) and resolution of the X-ray structure. Redundant sequences were not removed from the database to allow any conformational diversity of the loops to be captured, with the rationale that different structures of the same, or homologous proteins, might contain very different loop conformations due to their highly flexible nature. Candidate loop designs were identified to span residue A46 and A50 of Sav (loop 3/4) based on an X-ray protein crystal structure (PDB ID 3PK2). Initially, these were filtered purely based on the end-to-end distance of the loop, requiring it to be within 0.5 Å of the A46 – A50 distance in 3PK2. Loops satisfying this criterion were fitted by aligning the backbone atoms of the first and the last residues of the loop with those

of residues A46 and A50, respectively. The root-mean square deviation (RMSD) of the distances between complementary backbone atom pairs was calculated to evaluate the quality of the fit (detailed in the SI). Models were sorted based on this quality and the candidate loops were evaluated manually considering the diversity of number and nature of amino acids, symmetry and proximity to the metal center. On this basis, twelve loops were selected for insertion into Sav between A46 and A50 (Table 3 and Figure 4). Subsequently, the family members of MP_Sav's were subjected to a first round of mutagenesis in which cationic arginine and/or lysine residues in the loop sequences, which are known to be detrimental to catalysis,[29] were mutated to either alanine or phenylalanine. Finally, a K121F mutation was made in the first and second generation of MP3 constructs because of its improved catalytic performance with various ArMs based on the biotin-streptavidin technology.[30]

Table 3. Selected chimera of MP_Sav family containing natural loops with commensurate N_{term}–C_{term} topologies inserted between A46 and A50 of Sav. ISAMBARD was used to create a database of loops regions from known X-ray crystal structures.[a]

entry	abbreviation	inserted sequence[a]	abbreviation	second generation MP_Sav[a]
1	WT	AVGNA[b]		
2	MP 1	GKTKG	MP 1 *K-F*	GATFG
3	MP 2	GRSRG		–
4	MP 3	GNLKYG[c]	MP 3 *K-A*	GNLAYG[c]
5	MP 4	GIDRNG		–
6	MP 5	GDMKPRG		–
7	MP 6	GHEKRDG	*MP 6 K-A_R-F*	GHEAFDG
8	MP 7	GKHNKPDDCG		–
9	MP 8	GRRQIGTRSG		–
10	MP 9	GEPFGGEKING	MP 9 *K-F*	GEPFGGEFING
11	MP 10	GGRVIPVKLGG		–
12	MP 11	GYLSSQNGQPG		–
13	MP 12	GTERPSKNSHPG	MP 12 *R-A_K-F*	GTEAPSFNSHPG

[a] The expression and refolding conditions are detailed in the SI. [b] Sequence deleted from the WT in the MP constructs. [c] Third generation of MP_Sav was generated combining this loop sequence and K121F

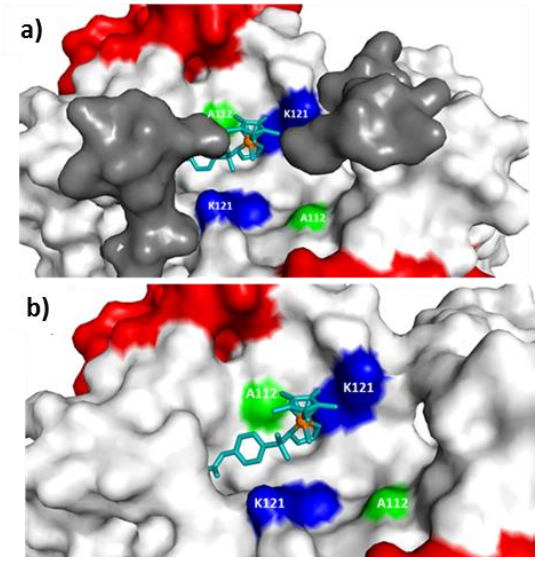


Figure 4. a) Model of MP_Sav member (MP_11) containing the loop GYLSSQNGQPG insertion (dark grey) between Sav position A46 and A50 generated from PDB file 3PK2. The insertion partially shields the biotin-binding vestibule. b) For comparison the surface representation of PDB ID 3PK2 is displayed in the same orientation. Color code: red, acidic residues; blue, basic; green, polar; and white, apolar.

Protein overexpression. BL21(DE3) *E. coli* cells and pET24 (+) plasmids was used for the overexpression of all chimeric streptavidins using an autoinduction medium.[31] After cell lysis, SDS gel analysis of the supernatant confirmed overexpression and the biotin-binding capacity of the proteins. For this purpose, biotin-4-fluorescein (B4F) was added to the Sav chimeras, and the presence of fluorescent bands in SDS PAGE confirmed correct folding and function of the Sav constructs. While the majority of the constructs with the shorter loops gave soluble and correctly folded proteins, the 2D_Sav family members had different folding behaviors depending on: i) the type of motif used; and ii) the site of insertion. Proteins containing PPR and SH3 at the *C* terminus were usually well folded, as were the Sav chimeras with either HP and FDP motifs engineered in either loop 3/4 or 5/6. As summarized in Table 2, some Sav constructs were

produced as inclusion bodies, however. Gratifyingly, these insoluble aggregates bound B4F, highlighting their biotin-binding activity. Screening various refolding buffers revealed that all insoluble chimeras could be refolded using a highly dilute MES buffer (see SI). Following this step when required, all constructs presented herein could be purified using a biotin-sepharose affinity column (Tables 2 and 3 and Figure 3). After purification, all the constructs were characterized by SDS PAGE and mass-spectroscopy (See SI).

Catalytic performance. Next, chimeric Sav were combined with various biotinylated cofactors and tested as artificial metalloenzymes in the following catalytic reactions: asymmetric transfer hydrogenation (ATH), ring-closing metathesis (RCM), and anion π-catalysis (ANPI). The respective biotinylated catalysts used are displayed in Figure 5.

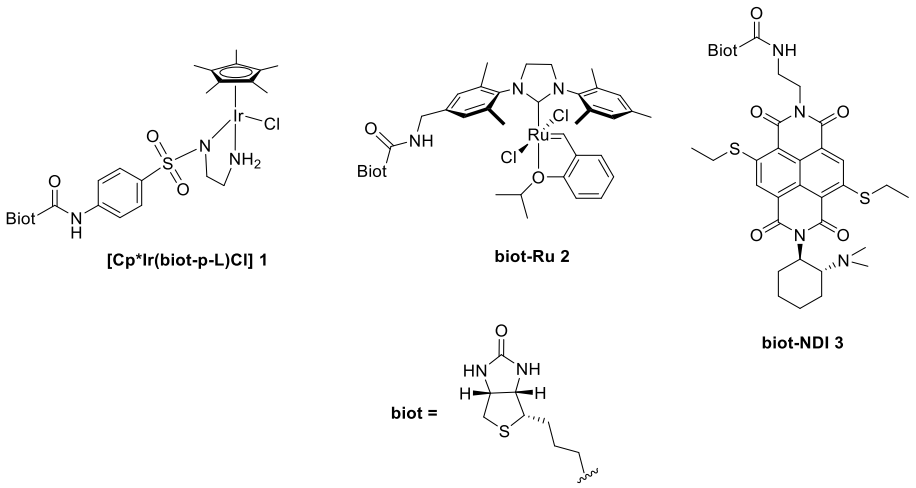
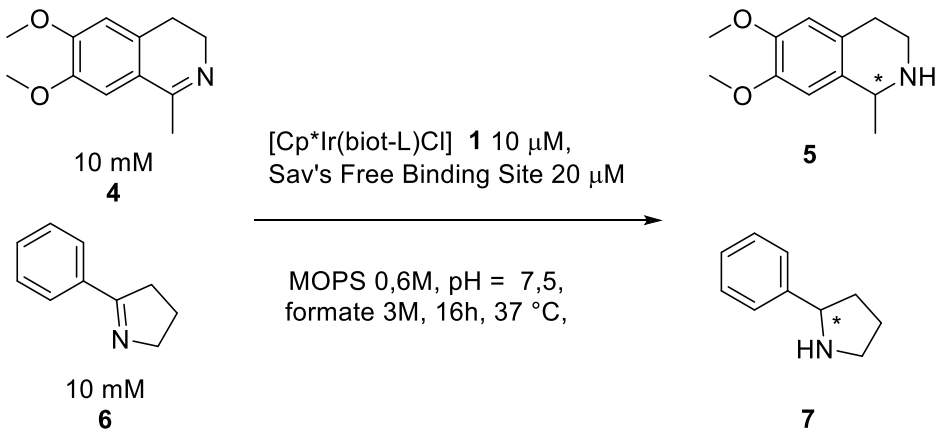


Figure 5. Biotinylated cofactors used to complete the holoenzymes for: i) asymmetric transfer hydrogenation reaction using [Cp*Ir(biot-L)Cl] **1**; ii) ring-closing metathesis using a Hoveyda–Grubbs second-generation catalyst (**biot-Ru 2** hereafter); and iii) anion π catalysis using naphthalenediimide (**biot-NDI 3** hereafter).

Previous work shows that the Ir-d^6 piano-stool [Cp*Ir(biot-L)Cl] **1** is the most active ATH catalysts for the reduction of prochiral imines in the presence of Sav.[32-33] Two different prochiral cyclic imines **4** and **6** were selected and the results of the asymmetric transfer hydrogenation, using [Cp*Ir(biot-L)Cl] **1,** are collected in Table 4. Compared to the free cofactor [Cp*Ir(biot-

L)Cl] **1** (Table 4 entries 1 and 2) and the corresponding benchmark ATHase [Cp*Ir(biot-L)Cl] · WT Sav (Table 4 entries 3 and 4), some of the engineered Sav constructs performed better, both in terms of TON (turnover number) and enantioselectivity. While the improvement of enantioselectivity was modest, the TON number improved two-fold for substrate **4** (Table 4 entries 5 and 6) and seven-fold for substrate **6** (Table 4 entries 7 and 8). Strikingly, the position of the insertions plays a critical role in the activity of the corresponding ATHase. Insertion in loop 3/4 (*i.e.*, positions 46 – 52) affords the most active ATHases, for both HP_46-52 and FDP_46-52 constructs, Table 4. For comparison, introduction of the HP in the loop 4/5 (*i.e.*, positions 64 – 70) had a negative impact on the ATHase activity (Table 4 entry 9). We hypothesize that the insertion between positions 46 – 52 projects the loop in the proximity of the Ir-center and thus has the highest (positive) impact on the ATHase's performance. ~~While the improvement in selectivity remains modest, the activity is significantly affected, resulting in up to a sevenfold increase in turnovers after sixteen hours.~~

Table 4. Selected ATHase results obtained with the 2D_Sav combined with [Cp*Ir(biot-L)Cl] **1**[a]



entry	2D_Sav	substrate	*e.e(R)*	TON	catalytic improvement vs WT Sav[b]

1	No protein	**4**	0	197	-
2	No protein	**6**	0	0	-
3	WT Sav	**4**	45	170	-
4	WT Sav	**6**	76	22	-
5	HP_46-52	**4**	59	223	1.3
6	FDP_46-52	**4**	51	340	2
7	HP_46-52	**6**	80	162	7.3
8	FDP_46-52	**6**	82	158	7.2
9	HP_64-70	**4**	22	14	0.1

[a] reaction conditions: substrates 10 mM, 16 h, 37 °C, [Cp*Ir(biot-L)Cl] **1** 10 µM, Sav's FBS 20 µM, MOPS 0,6 M, pH = 7,5, formate 3 M, V_{tot} = 200 µL.[b] Ratio between TON 2D_Sav and TON of WT Sav.

Considering the MP_Sav constructs, a general trend is observed in the reduction of substrate **4**: Whilst the first-round Lys/Arg-containing loop sequences had dramatically lower activities than and comparable selectivities to the benchmark ATHase [Cp*Ir(biot-L)Cl] **1** · WT Sav (Table 5 entries 3 *vs* entries 6 – 10), point mutations of Lys to Ala or Phe resulted in second-generation hybrid catalysts that rivaled or outperformed the WT ATHase (Table 5, entries 11 – 12). The [Cp*Ir(biot-L)Cl] **1** · Sav K121F ATHase displayed significantly improved TON although this was at the cost of *ee* ~~(13% ee (*S*) **5**, 970 TON and 45% ee (*R*) **5**, 170 TON for Sav K121F and WT, respectively)~~. Combining the beneficial K121F mutation and loops, with or without Lys residue, in all cases gives ATHase with lower TONs compared to Sav K121F (Table 5, entries 13 – 14). For substrate **6,** all second-generation MP_Sav performed better than the corresponding ATHase [Cp*Ir(biot-L)Cl] **1** · WT Sav (Table 5 entry 4 vs Table 5 entries 15 – 17). However [Cp*Ir(biot-L)Cl] **1** · Sav K121F followed by the chimera [Cp*Ir(biot-L)Cl] **1** · Sav HP_46-52 and [Cp*Ir(biot-L)Cl] **1** · Sav FDP_46-52 were the most active ATHases (Table 5 entry 5 and Table 4 entries 7 and 8). A summary of all ATHase experiments is presented in SI Tables S4 and S5.

Table 5. Selected ATHase results obtained for the MP_Sav introduced between position A46 and A50 combined with [Cp*Ir(biot-L)Cl] **1**.[a]

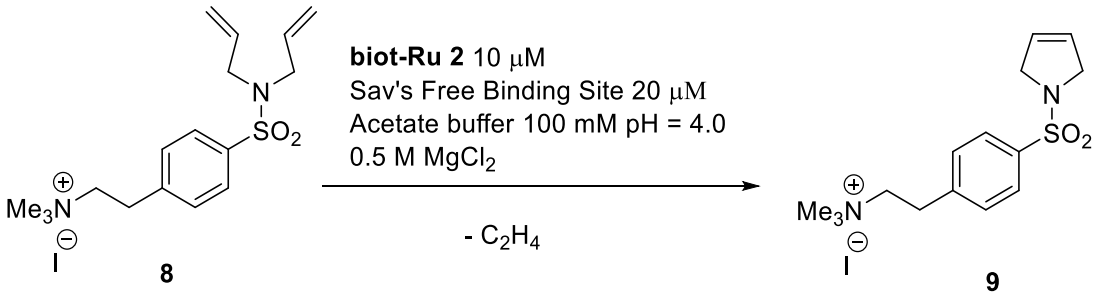
entry	MP_Sav[b,c]	substrate	*e.e*(*R*)	TON	catalytic improvement vs WT Sav[d]
1	No protein	**4**	0	197	-
2	No protein	**6**	0[d]	0	-
3	WT Sav	**4**	45	170	-
4	WT Sav	**6**	76	22	-
5	Sav K121F	**4**	-13 *(S)*	970	5.7
6	MP 3	**4**	53	6	0.04
7	MP 4	**4**	42	67	0.4
8	MP 9	**4**	53	82	0.48
9	MP 11	**4**	45	61	0.36
10	MP 12	**4**	62	15	0.08
11	MP 3_*K-A*	**4**	46	192	1.13
12	MP 9_*K-F*	**4**	49	198	1.16
13	MP 3_K121F	**4**	15	568	3.35
14	MP 3_*K-A*-K121F	**4**	14	462	2.7
15	MP 3_*K-A*	**6**	79	57	2.5
16	MP 6_*K-A*_*R-F*	**6**	77	100	4.55
17	MP 9_*K-F*	**6**	80	71	3.2

[a] reaction conditions: substrates 10 mM, 16 h, 37°C, [Cp*Ir(biot-L)Cl] **1** 10 µM, Sav's Free Binding Site (FBS) 20 µM, MOPS 0,6 M, pH = 7,5, formate 3 M, V_{tot} = 200 µL. [b] The numbering of the residues of Sav was kept identical to the WT numbering, eventhough the insertion may be placed before the position of the mutation. [c] Italicized one letter aminoacid abbreviations designate cationic aminoacids within the inserted loops that were mutated to either A or F to probe the effect of the charge on the catalytic performance (See Table 3 for details).[d] Ratio between TON MP_Sav and TON WT Sav.

Building on our previous work on olefin metathesis in aqueous phase,[34] we investigated the catalytic behavior of 2D_Sav and MP_Sav scaffolds in the presence of **biot-Ru 2**. The resulting artificial metathases were tested for the RCM of substrate **8**. Both the **biot-Ru 2** · Sav HP_64-70 and **biot-Ru 2** · Sav PPR_159 outperformed the bare catalyst (Table 6, entries 1 vs 4 and 5). As for the ATHases, introduction of a phenylalanine residue at position K121 improved catalytic

performance resulting in 105 TONs (Table 6, entry 3). The second- and the third-generation MP_Sav's variants originating from the Sav_MP3 sequence displayed improved activity (Table 6 entries 6-8). However, compared to **biot-Ru 2** ·Sav K121F, none of the newly designed chimeric Sav outperformed this single point mutant (Table 6 entries 4-8). A complete list of chimeric metathases is collected in SI Table S6.

Table 6. Selected results for RCM of diallyl substrate **8** using 2D_Sav and MP_Sav.[a]

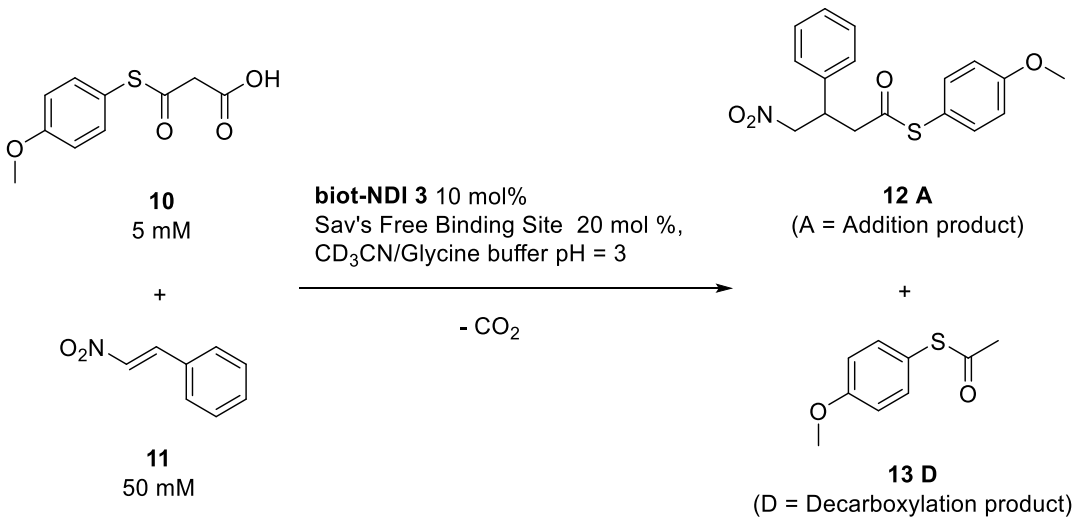


entry	Sav [b,c]	TON
1	No protein	38
2	WT Sav	87
3	K121F	105
4	HP_64-70	56
5	PPR_159	64
6	MP 3_*K-A*	61
7	MP 3_K121F	88
8	MP 3_*K-A*-K121F	98

[a] reaction conditions: substrate **8** 10 mM, **biot-Ru 2** 10 μM, 16 h, 37°C, Chim_Sav and MP_Sav 20 μM Free Binding Site, acetate buffer 100 mM, 0.5 M $MgCl_2$ pH = 4.0, V_{tot} = 200 μL. [b] The numbering of the residues of Sav were kept as the WT numbering, eventhough the insertion may be placed before the position of the mutation. [c] Italicized residues designate cationic aminoacids in the inserted loops that were mutated to either A or F to probe the effect of the charge on the catalytic performance (See Table 3 for details)

The Michael addition of malonic acid half thioester **10** (MAHT) was investigated next. This reaction is of interest as it mimics the first step of the polyketide biosynthesis. Importantly, anion-π interactions were shown to have a significant impact on the chemoselectivty of this reaction by selectively producing the disfavored addition product **12 A**. Without an electron-deficient π-system as the catalyst, the decarboxylation product **13 D** is generated almost exclusively.[35-36] Naphthalenediimides (NDIs) are π-acidic surfaces with proven efficiencies as primary component in anion-π catalysts. By linking a biotin to an NDI π-surface and incorporation within Sav anion-π enzymes result. These artificial enzymes produce the addition product **12 A** with an exquisite chemoselectivity (A/D ratio > 30) and enantioselectivity (*ee* = 95% for **biot-NDI 3**· Sav S112Y)[37] through synergism between NDIs and Sav mutants. However, due the inhibition by anions, the use of a more confined cavity may be of interest. With this goal in mind, the catalysis was tested with the previously established conditions.[37] Most of the chimeric Sav's produced the addition product **12 A** with excellent chemoselectivity (**12 A**/ **13 D** > 30) and moderate yields (up to 61 Yield%, Table 7 entry 2). Three of the chimeric Sav's were inactive. We posit that the lack of reaction may be attributed to an important shielding above the NDI surface embedded in either MP 12 Sav, Sav FDP_46-52 and Sav HP_46-52. Only **biot-NDI 3** · Sav PPR_159, containing the coiled-coil domain at the C_{term} afforded **12 A** with an enantioselectivity (Table 7, entry 3), thus competing with the most efficient anion-π organocatalysts achieved to date.[36, 38] A summary of all π experiments is presented in SI, Table S7.

Table 7. Selected results obtained for anion-π catalysis using **biot-NDI 3** · 2D_Sav or **biot-NDI 3** · MP_Sav.[a]



entry	Sav[b,c]	yield %	12 A/13 D	*ee* (%)
1	WT Sav	60	>30	41
2	MP 12_*R-A_K-F*	61	>30	0
3	PPR_159	57	>30	22
4	MP 3_*K-A*_K121F	55	>30	0
5	MP 3_K121F	51	>30	0
6	MP 9	23	>30	0
7	MP 9_*K-F*	45	>30	0
8	MP 12	traces	nd	0
9	FDP_46-52	0	nd	nd
10	HP_46-52	0	nd	nd

[a]Reaction conditions: **biot NDI 3** 10 mol%, Sav's FBS 20 mol %, CD_3CN/Glycine buffer pH = 3, substrate **10** 5 mM, substrate **11** 50 mM. [b] The numbering of the residues of Sav were kept as that of WT Sav, eventhough the insertion may be placed before the position of the mutation. [c] Italicized residues designate cationic aminoacids in the inserted loops that were mutated to either A or F to probe the effect of the charge on the catalytic performance (See Table 3 for details).

OUTLOOK

The catalytic performance of ArMs derived from the chimeric Savs described herein suggest that these straightforward modifications to the host can have significant impacts on catalytic

activity. Thanks to the remarkable quaternary stability of the Sav scaffold, introduction of additional structural motifs lead, in the vast majority of cases, to soluble and functional chimeric Savs. For the ATHase, the resulting activity was up to sevenfold higher than the parent WT Sav. Thus, this strategy of embellishing loops proximal to the active site offers a versatile means to complement directed evolution efforts to optimize the performance of the ArMs based on the biotin-streptavidin technology. The newly introduced structural motifs might be further engineered to form a protective lid over the active site, resulting in ArMs able to perform catalytic reactions in cell free extracts or in the cytoplasm.

EXPERIMENTAL SECTION

General procedure for the asymmetric transfer hydrogenation. 10 µl of proteins stock solution in Milli-Q H2O (200 µM free binding site) was added to 185 µl of reaction buffer followed by the addition of 5 µl of the biotinylated metal complex [Cp*Ir(biot-L)Cl] **1** from a stock solution (0.4 mM in DMSO). The solution was mixed for 20 min at 37 °C and 800 rpm in a thermo-mixer for precomplexation. Finally, 5 µl of substrate **4** or **6** stock solution (400 mM in DMSO) was added and the mixture was stirred at 37 °C for 16 h. Subsequently, 20% NaOH solution was added to the reaction mixture, followed by the addition of CH_2Cl_2. After mixing, the organic phase was separated and dried with anhydrous Na_2SO_4. Solids were separated through centrifugation and the supernatant was analyzed by using HPLC or GC.

General procedure for ring closing metathesis. In a 1.2 mL glass vial, 10 µl of protein stock solution (400 µM) and 5 µl **biot Ru 2** (400 µM) were added to 85 µl of reaction buffer and incubated at 37°C for 20 min. After incubation, 100 µl of substrate 20 mM were added to the reaction mixture and the reaction was stirred at 37°C for 16 hours at 1000 rpm. After reaction,

methanol and the internal standard were added to the reaction mixture and the whole volume transferred to Eppendorf tubes for centrifugation. The supernatant was then transferred into HPLC vials containing MQ water and the sample analyzed by UPLC-MS for the quantification of product **9**.

General procedure for anion π catalysis. Stock solutions of substrates **10** (40 mM), **11** (400 mM) and **biot NDI 3** (2 mM) were prepared in CD_3CN. Solutions of substrates **10** should be freshly prepared as decarboxylation to afford **13 D** build-up on aging. Solutions were prepared by mixing successively streptavidin WT or mutants, biotinylated ligand **3,** substrates **10**, **11** and stirred at 20 °C. After 24h, the mixture was extracted with $CDCl_3$, dried over Na_2SO_4, and filtered and analyzed by ^{1}H-NMR .

ASSOCIATED CONTENT

Supporting Information. The following files are available free of charge.

Supplementary natural loop design, cloning procedure, protein expression and purification, analytical methods (PDF)

AUTHOR INFORMATION

Corresponding Authors

*E-mail: thomas.ward@unibas.ch

D.N.Woolfson@bristol.ac.uk

Present Address

† Adolphe Merkle Institute, University of Fribourg, Chemin des Verdiers 4, 1700 Fribourg, Switzerland.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT Generous funding for this research was provided by the Swiss National Science Foundation (grant 200020 162348 to TRW) as well as the NCCR Molecular Systems Engineering (to SM and TRW). Precious metal salts were kindly provided by Umicore. CWW and DNW are supported by a European Research Council Advanced Grant (340764). DNW holds a Royal Society Wolfson Research Merit Award.

REFERENCES

1. Wilson, M. E.; Whitesides, G. M., Conversion of a protein to a homogeneous asymmetric hydrogenation catalyst by site-specific modification with a diphosphinerhodium(I) moiety. *J. Am. Chem. Soc.* **1978,** *100* (1), 306-307.
2. Heinisch, T.; Pellizzoni, M.; Durrenberger, M.; Tinberg, C. E.; Kohler, V.; Klehr, J.; Haussinger, D.; Baker, D.; Ward, T. R., Improving the Catalytic Performance of an Artificial Metalloenzyme by Computational Design. *J. Am. Chem. Soc.* **2015,** *137* (32), 10414-10419.
3. Yang, H.; Srivastava, P.; Zhang, C.; Lewis, J. C., A General Method for Artificial Metalloenzyme Formation through Strain-Promoted Azide-Alkyne Cycloaddition. *ChemBioChem* **2014,** *15* (2), 223-227.
4. Song, W. J.; Tezcan, F. A., A designed supramolecular protein assembly with in vivo enzymatic activity. *Science* **2014,** *346* (6216), 1525-1528.
5. Bos, J.; Browne, W. R.; Driessen, A. J. M.; Roelfes, G., Supramolecular Assembly of Artificial Metalloenzymes Based on the Dimeric Protein LmrR as Promiscuous Scaffold. *J. Am. Chem. Soc.* **2015,** *137*, 9796-9799.
6. Srivastava, P.; Yang, H.; Ellis-Guardiola, K.; Lewis, J. C., Engineering a dirhodium artificial metalloenzyme for selective olefin cyclopropanation. *Nat. Commun.* **2015,** *6*, 7789.
7. Bos, J.; Roelfes, G., Artificial metalloenzymes for enantioselective catalysis. *Curr. Opin. Chem. Biol.* **2014,** *19*, 135-143.
8. Collot, J.; Gradinaru, J.; Humbert, N.; Skander, M.; Zocchi, A.; Ward, T. R., Artificial Metalloenzymes for Enantioselective Catalysis Based on Biotin-Avidin. *J. Am. Chem. Soc.* **2003,** *125* (30), 9030-9031.

9. Pordea, A.; Ward, T. R., Chemogenetic protein engineering: an efficient tool for the optimization of artificial metalloenzymes. *Chem. Commun. (Cambridge, U. K.)* **2008,** (36), 4239-4249.
10. Kohler, V.; Wilson, Y. M.; Durrenberger, M.; Ghislieri, D.; Churakova, E.; Quinto, T.; Knorr, L.; Haussinger, D.; Hollmann, F.; Turner, N. J., Synthetic Cascades Are Enabled by Combining Biocatalysts with Artificial Metalloenzymes. *Nat. Chem.* **2013,** *5* (2), 93-99.
11. Renata, H.; Wang, Z. J.; Arnold, F. H., Expanding the Enzyme Universe: Accessing Non-Natural Reactions by Mechanism-Guided Directed Evolution. *Angew. Chem., Int. Ed.* **2015,** *54* (11), 3351-3367.
12. Hyster, T. K.; Knorr, L.; Ward, T. R.; Rovis, T., Biotinylated Rh(III) complexes in engineered streptavidin for accelerated asymmetric C-H activation. *Science* **2012,** *338* (6106), 500-503.
13. Matsuo, T.; Fukumoto, K.; Watanabe, T.; Hayashi, T., Precise Design of Artificial Cofactors for Enhancing Peroxidase Activity of Myoglobin: Myoglobin Mutant H64D Reconstituted with a "Single-Winged Cofactor" Is Equivalent to Native Horseradish Peroxidase in Oxidation Activity. *Chem. Asian J.* **2011,** *6* (9), 2491-2499.
14. Dydio, P.; Key, H. M.; Nazarenko, A.; Rha, J. Y.-E.; Seyedkazemi, V.; Clark, D. S.; Hartwig, J. F., An artificial metalloenzyme with the kinetics of native enzymes. *Science* **2016,** *354* (6308), 102-106.
15. Obexer, R.; Godina, A.; Griffiths, A. D.; Garrabou, X.; Mittl, P. R. E.; Baker, D.; Hilvert, D., Emergence of a catalytic tetrad during evolution of a highly active artificial aldolase. *Nature chemistry* **2017,** *9*, 50-56.
16. Ward, T. R., Artificial Metalloenzymes Based on the Biotin-Avidin Technology: Enantioselective Catalysis and Beyond. *Acc. Chem. Res.* **2011,** *44* (1), 47-57.
17. Romero, P. A.; Arnold, F. H., Exploring protein fitness landscapes by directed evolution. *Nat. Rev. Mol. Cell Biol.* **2009,** *10* (12), 866-876.
18. Schwizer, F.; Kohler, V.; Durrenberger, M.; Knorr, L.; Ward, T. R., Genetic Optimization of the Catalytic Efficiency of Artificial Imine Reductases Based on Biotin-Streptavidin Technology. *ACS Catal.* **2013,** *3* (8), 1752-1755.
19. Kurochkina, N.; Guha, U., SH3 domains: modules of protein–protein interactions. *Biophys Rev* **2013,** *5*.
20. Kramer, M. A.; Wetzel, S. K.; Plückthun, A.; Mittl, P. R. E.; Grütter, M. G., Structural Determinants for Improved Stability of Designed Ankyrin Repeat Proteins with a Redesigned C-Capping Module. *Journal of Molecular Biology* **2010,** *404*, 381–391.
21. Chiu, T. K.; Kubelka, J.; Herbst-Irmer, R.; Eaton, W. A.; Hofrichter, J.; Davies, D. R., High-resolution x-ray crystal structures of the villin headpiece subdomain, an ultrafast folding protein. *PNAS* **2005,** *102* (21), 7517–7522.
22. Eiben, C. B.; Siegel, J. B.; Bale, J. B.; Shen, B. W.; Foldit Players; Stoddard, B. L.; Popovic, Z.; Baker, D., Increased Diels-Alderase activity through backbone remodeling guided by Foldit players. *Nature biotechnology* **2012,** *30* (2), 190-194.
23. Coquille, S.; Filipovska, A.; Chia, T.; Rajappa, L.; Lingford, J. P.; Razif, M. F. M.; Thore, S.; Rackham, O., An artificial PPR scaffold for programmable RNA recognition. *Nat Commun* **2014,** *5*, 5729.
24. Sano, T.; Cantor, C. R., Expression of a cloned streptavidin gene in Escherichia coli. *Proc. Natl. Acad. Sci. USA* **1990,** *87*, 142-146.

25. Le Trong, I.; Humbert, N.; Ward, T. R.; Stenkamp, R. E., Crystallographic Analysis of a Full-length Streptavidin with Its C-terminal Polypeptide Bound in the Biotin Binding Site. *J. Mol. Biol.* **2006,** *356*, 738–745.
26. Wood, C. W.; Heal, J. W.; Thomson, A. R.; Bartlett, G. J.; Ibarra, A. A.; Brady, R. L.; Sessions, R. B.; Woolfson, D. N., ISAMBARD: an open-source computational environment for biomolecular analysis, modelling and design. *Bioinformatics* **2017**, 1–8.
27. Touw, W. G.; Baakman, C.; Black, J.; te Beek, T. A. H.; Krieger, E.; Joosten, R. P.; Vriend, G., A series of PDB-related databanks for everyday needs. *Nucleic Acids Research* **2015,** *43*, D364–D368.
28. Kabsch, W.; Sander, C., Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983,** *22* (12), 2577–2637
29. Schwizer, F.; Okamoto, Y.; Heinisch, T.; Gu, Y.; Pellizzoni, M. M.; Lebrun, V.; Reuter, R.; Köhler, V.; Lewis, J. C.; Ward, T. R., Artificial Metalloenzymes: Reaction Scope and Optimization Strategies. *Chem. Rev.* **2017**.
30. Letondor, C.; Pordea, A.; Humbert, N.; Ivanova, A.; Mazurek, S.; Novic, M.; Ward, T. R., Artificial Transfer Hydrogenases Based on the Biotin-(Strept)avidin Technology: Fine Tuning the Selectivity by Saturation Mutagenesis of the Host Protein. *J. Am. Chem. Soc.* **2006,** *128* (25), 8320-8328.
31. Studier, F. W., Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification* **2005,** *41*, 207–234.
32. Durrenberger, M.; Heinisch, T.; Wilson, Y. M.; Rossel, T.; Nogueira, E.; Knorr, L.; Zimbron, M. J.; Pierron, J.; Mutschler, A.; Kersten, K.; Schirmer, T.; Ward, T. R., Artificial Transfer Hydrogenases for the Enantioselective Reduction of Cyclic Imines. *Angew. Chem., Int. Ed.* **2011,** *50* (13), 3026-3029.
33. Pellizzoni, M. M.; Gandolfi, R.; Fusè, M.; Facchetti, G.; Contini, A.; Isabella, R., Evaluation of Chemical Diversity of Biotinylated Chiral 1,3-diamines as a Catalytic Moiety in Artificial Imine Reductase. *ChemCatChem* **2016**, 1665–1670.
34. Kajetanowicz, A.; Chatterjee, A.; Reuter, R.; Ward, T. R., Biotinylated Metathesis Catalysts: Synthesis and Performance in Ring Closing Metathesis. *Catal. Lett.* **2014,** *144* (3), 373-379.
35. Cotelle, Y.; Benz, S.; Avestro, A.-J.; Ward, T. R.; Sakai, N.; Matile, S., Anion-π Catalysis of Enolate Chemistry: Rigidified Leonard Turns as a General Motif to Run Reactions on Aromatic Surfaces. *Angew. Chem., Int. Ed.* **2016,** *55* (13), 4275-4279.
36. J. López-Andarias; A.Frontera; Matile, S., Anion-π Catalysis on Fullerenes. . *J. Am. Chem. Soc.* **2017,** *Article ASAP*.
37. Cotelle, Y.; Lebrun, V.; Sakai, N.; Ward, T. R.; Matile, S., Anion-π Enzymes. *ACS Cent. Sci.* **2016,** *2* (6), 388-393.
38. C. Wang; Matile, S., Anion-π Catalysts with Axial Chirality. *Chem. Eur. J.* **2017,** *23*, 11955-11960.