J|A|C|S
ARTICLES
Published on Web 10/15/2004

Artificial Metalloenzymes: (Strept)avidin as Host for Enantioselective Hydrogenation by Achiral Biotinylated Rhodium−Diphosphine Complexes

Myriem Skander, Nicolas Humbert, Jérôme Collot, Julieta Gradinaru, Gérard Klein, Andreas Loosli, Jérôme Sauser,† Andrea Zocchi, François Gilardoni,‡ and Thomas R. Ward*

Contribution from the Institute of Chemistry, University of Neuchâtel, Avenue Bellevaux 51, CP2, CH-2007 Neuchâtel, Switzerland

Received April 22, 2004; E-mail: thomas.ward@unine.ch

Abstract: We report on the generation of artificial metalloenzymes based on the noncovalent incorporation of biotinylated rhodium−diphosphine complexes in (strept)avidin as host proteins. A chemogenetic optimization procedure allows one to optimize the enantioselectivity for the reduction of acetamidoacrylic acid (up to 96% ee (*R*) in streptavidin S112G and up to 80% ee (*S*) in WT avidin). The association constant between a prototypical cationic biotinylated rhodium−diphosphine catalyst precursor and the host proteins was determined at neutral pH: log $K_a = 7.7$ for avidin (pI = 10.4) and log $K_a = 7.1$ for streptavidin (pI = 6.4). It is shown that the optimal operating conditions for the enantioselective reduction are 5 bar at 30 °C with a 1% catalyst loading.

Introduction

Today, the preparation of enantiomerically pure compounds is one of the most actively pursued fields in synthetic chemistry. In the past 35 years, metal-catalyzed enantioselective transformations have enjoyed significant growth as it was recognized that these are among the most efficient ways to produce enantiomerically pure compounds.[1] This effort was rewarded by the 2001 Nobel Prize in Chemistry, which was shared by Knowles, Noyori, and Sharpless.[2−4]

Despite significant advances, it remains very difficult to predict the outcome of a metal-catalyzed enantioselective reaction. Indeed, the differences in energy involved in the transition state leading to both enantiomers of a desired product are too small to be reliably predicted or computed. As a consequence, the number of efficient enantioselective metal catalysts used for industrial applications remains limited.[5,6]

To get around the difficulty of predicting the enantioselectivity, combinatorial methodologies have successfully been applied to the discovery and to the development of new enantioselective catalysts.[7,8] These studies have highlighted the fact that many subtle experimental parameters (solvent, counterion, added salts, and so forth) often have a significant and unpredictable influence on the enantioselectivity of a reaction.[9] These weak contacts between a catalyst and its "nonbonded" environment are commonly referred to as the second coordination sphere.[10]

In recent years, enzymatic catalysis has emerged as an important alternative tool for the synthesis of enantiopure compounds.[11] Again, despite widespread academic and industrial research efforts, the number of industrial biocatalyst applications remains limited.[12]

Table 1 outlines a comparison of some of the most noteworthy features of homogeneous and of enzymatic catalysis. For example, homogeneous catalysts are more tolerant than enzymes toward variation in the size of a substrate; despite their difference in size, acetamidoacrylic and acetamidocinnamic acids are often both good substrates for homogeneous hydrogenation catalysts. Biocatalysts have evolved to target a single functionality even in the presence of other, perhaps more reactive, functionalities; enzymes do not require protective groups.

Recently, some of the inherent limitations of biocatalysts, including small substrate scope, operational stability, and availability of only one enantiomer of the product, have been overcome, thus expanding the scope of applications of bioca-

† Current address: Department of Chemistry, Oxford University, Mansfield Road, Oxford OX1 3TA, U.K.

‡ Current address: Inforsense Ltd., 48 Princes Gardens, SW7 2PE London, U.K.

(1) *Comprehensive Asymmetric Catalysis*; Jacobsen, E. N., Pfaltz, A., Yamamoto, H., Eds.; Springer: Berlin, 1999.
(2) Knowles, W. S. *Angew. Chem., Int. Ed.* **2002**, *41*, 1998−2007.
(3) Sharpless, K. B. *Angew. Chem., Int. Ed.* **2002**, *41*, 2024−2032.
(4) Noyori, R. *Angew. Chem., Int. Ed.* **2002**, *41*, 2008−2022.
(5) Blaser, H.-U.; Malan, C.; Pugin, B.; Spindler, F.; Steiner, H.; Studer, M. *Adv. Synth. Catal.* **2003**, *345*, 103−151.
(6) Blaser, H. U.; Spindler, F.; Studer, M. *Appl. Catal., A* **2001**, *221*, 119−143.
(7) Jandeleit, B.; Schaefer, D. J.; Powers, T. S.; Turner, H. W.; Weinberg, W. H. *Angew. Chem., Int. Ed.* **1999**, *38*, 2494−2532.
(8) Gennari, C.; Piarulli, U. *Chem. Rev.* **2003**, *103*, 3071−3100.
(9) Vogl, E. M.; Gröger, H.; Shibasaki, M. *Angew. Chem., Int. Ed.* **1999**, *38*, 1570−1577.
(10) Loeb, S. J. In *Comprehensive Supramolecular Chemistry*; Lehn, J.-M., Atwood, J. L., Davies, J. E. D., McNichol, D. D., Vögtle, F., Gokel, G. W., Eds.; Pergamon: Oxford, 1996; pp 733−753.
(11) Faber, K. *Biotransformations in Organic Chemistry*, 5th ed.; Springer: Berlin, 2004.
(12) Straathof, A. J. J.; Panke, S.; Schmid, A. *Curr. Opin. Biotechnol.* **2002**, *13*, 548−556.

Table 1. Comparing Homogeneous to Enzymatic Catalysis: Pros and Cons

	homogeneous catalysis	enzymatic catalysis
substrate scope	large	limited
enantiomers	both enantiomers accessible	single enantiomer
functional group tolerance	small	large
reaction repertoire	large	small
turnover numbers	small	large
solvent compatibility	large	small (aqueous)
optimization	chemical	genetic
second coordination sphere	poorly-defined	well-defined

talysis. In particular, directed evolution (i.e., genetic optimization) combined with high-throughput screening has significantly facilitated the optimization of customized enantioselective enzymes.[13−18]

Perhaps most relevant to this study, and in contrast to organometallic catalysts, is that enzymes exquisitely tailor both the first and the second coordination sphere of their active site to afford efficient and selective catalytic systems (Table 1).

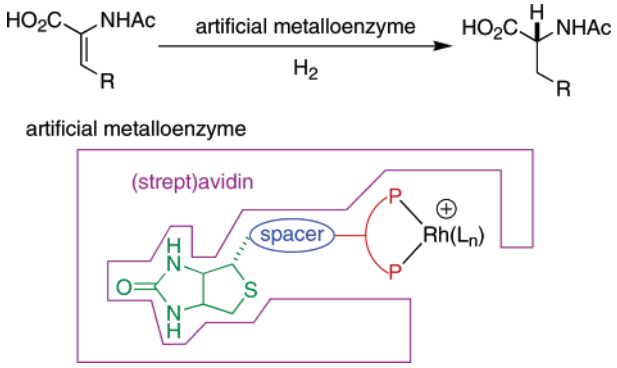
From these considerations, it appears that homogeneous and enzymatic catalyses are in many respects complementary. By anchoring an organometallic catalyst precursor into a host protein, we hoped to create artificial metalloenzymes with properties reminiscent of both kingdoms.

To be able to readily deconvolute the influence of the organometallic fragment from the influence of the protein, we focus on enantioselective catalysis. Incorporation of an achiral catalyst precursor in the host protein ensures that any level of enantioselection is induced by the second coordination sphere provided by the protein. Such an approach offers several appealing features: (i) the possibility of dissociating the activity (primarily dictated by the organometallic catalyst precursor) from the selectivity (governed by the host protein); (ii) the use of orthogonal diversity-generating procedures (molecular biology for the protein optimization as well as parallel synthesis for the organometallic fragment); and (iii) a novel approach to exploit weak interactions in enantioselective homogeneous catalysis.

Inspired by the seminal work of E. T. Kaiser,[19] several groups have recently developed methods to covalently modify proteins by incorporating transition-metal catalysts to yield hybrid catalytic systems with promising properties.[15,20−23]

The approach we focus on relies on a noncovalent incorporation (i.e., supramolecular) of the organometallic catalyst precursor in the protein. Since no chemical coupling step is required upon addition of the catalyst precursor to the protein, we reasoned that the integrity of the organometallic species would be warranted. To ensure the localization of the organometallic catalyst precursor within the protein, however, a very strong noncovalent host⊂guest (i.e., a protein inhibitor) system should be selected. The biotin−avidin system naturally comes to mind.[24] The principle of the biotin−avidin technology (often referred to as molecular velcro) relies on the extraordinary affinity of biotin for either avidin or streptavidin ($K_a \sim 1 \times 10^{14}$ M^{-1}).[25] Most importantly, it is generally accepted that derivatization of the valeric acid side chain of biotin does not affect significantly the strength of the biotin−avidin interaction as most stabilizing contacts are located on the bicyclic framework of (+)-biotin.[25] In the past 20 years, the biotin−avidin technology has found numerous applications in various fields of biotechnology, including ELISA, immunolabeling, affinity targeting, drug delivery, and so forth.[26] Whitesides was the first to suggest the use of the avidin as a host for enantioselective catalysis.[27,28] In the same spirit, Watanabe recently reported on the apomyoglobin as a host for Cr(III) Schiff base complexes in the oxidation of sulfides.[29]

Scheme 1. Artificial Metalloenzymes for Enantioselective Hydrogenation Reactions[a]



[a] The host protein (violet) displays high affinity for the anchor, biotin (green); introduction of a spacer (blue) and variation of the ligand scaffold (red) allow one to chemically optimize the enantioselectivity. Site-directed mutagenesis allows a genetic optimization of the host protein.

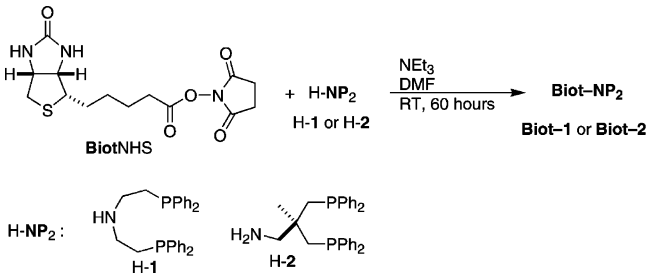
Herein, we report on the use of avidin and streptavidin (abbreviated hereafter (strept)avidin) as a host protein for biotinylated rhodium−diphosphine complexes. The resulting artificial metalloenzymes are tested in the enantioselective hydrogenation of acetamidoacrylic acid. Both chemical and genetic methodologies (i.e., chemogenetic optimization)[20] are used to optimize the enantioselectivity of the hybrid catalyst (Scheme 1).[30]

Results and Discussion

Ligand Synthesis. The approach used in this study relies on amide bond coupling reactions for the synthesis of the biotinylated ligands. Two different amino−diphosphine skeletons

(13) Powell, K. A.; Ramer, S. W.; del Cardayré, S. B.; Stemmer, W. P. C.; Tobin, M. B.; Longchamp, P. F.; Huisman, G. W. *Angew. Chem., Int. Ed.* **2001**, *40*, 3948−3959.
(14) Reetz, M. T. *Angew. Chem., Int. Ed.* **2002**, *41*, 1335−1338.
(15) Reetz, M. T. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *101*, 5716−5722.
(16) Fong, S.; Machajewski, T. D.; Mak, C. C.; Wong, C.-H. *Chem. Biol.* **2000**, *7*, 873−883.
(17) May, O.; Nguyen, P. T.; Arnold, F. H. *Nat. Biotechnol.* **2000**, *18*, 317−320.
(18) DeSantis, G.; Wong, K.; Farwell, B.; Chatman, K.; Zhu, Z.; Tomlinson, G.; Huang, H.; Tan, X.; Bibbs, L.; Chen, P.; Kretz, K.; Burk, M. J. *J. Am. Chem. Soc.* **2003**, *125*, 11476−11477.
(19) Kaiser, E. T.; Lawrence, D. S. *Science* **1984**, *226*, 505−511.
(20) Qi, D.; Tann, C.-M.; Haring, D.; Distefano, M. D. *Chem. Rev.* **2001**, *101*, 3081−3111.
(21) Reetz, M. T.; Rentzsch, M.; Pletsch, A.; Maywald, M. *Chimia* **2002**, *56*, 721−723.
(22) Nicholas, K. M.; Wentworth, P., Jr.; Harwig, C. W.; Wentworth, A. D.; Shafton, A.; Janda, K. D. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 2648−2653.
(23) Davis, B. G. *Curr. Opin. Biotechnol.* **2003**, *14*, 379−386.
(24) Kuntz, I. D.; Chen, K.; Sharp, K. A.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9997−10002.
(25) *Methods in Enzymology: Avidin−Biotin Technology*; Wilchek, M., Bayer, E. A., Eds.; Academic Press: San Diego, CA, 1990; Vol. 184.
(26) Wilchek, M.; Bayer, E. A. *Biomol. Eng.* **1999**, *16*, 1−4.
(27) Wilson, M. E.; Whitesides, G. M. *J. Am. Chem. Soc.* **1978**, *100*, 306−307.
(28) Lin, C.-C.; Lin, C.-W.; Chan, A. S. C. *Tetrahedron: Asymmetry* **1999**, *10*, 1887−1893.
(29) Ohashi, M.; Koshiyama, T.; Ueno, T.; Yanase, M.; Fujii, H.; Watanabe, Y. *Angew. Chem., Int. Ed.* **2003**, *42*, 1005−1008.
(30) Collot, J.; Gradinaru, J.; Humbert, N.; Skander, M.; Zocchi, A.; Ward, T. R. *J. Am. Chem. Soc.* **2003**, *125*, 9030−9031.

Scheme 2. Synthesis of Biotinylated Amidodiphosphine Ligands Devoid of Spacer, **Biot−1** and **Biot−2**

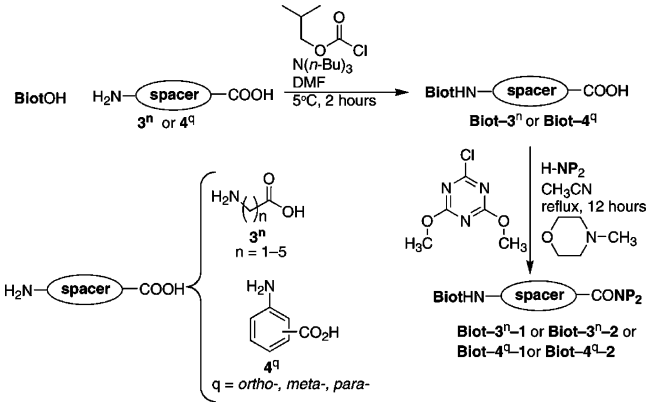


(**NP**$_2$) were tested in this study: the flexible ligand scaffold **1**, forming an eight-membered chelate ring, and the more-rigid ligand scaffold **2**, forming a six-membered chelate ring (Scheme 2). These contain an endocyclic secondary amine and an exocyclic primary amine, respectively. The biotinylated ligands, **Biot−1** and **Biot−2**, were synthesized using biotin *N*-hydroxysuccinimide (**Biot**NHS) in 67 and 69% yield, respectively (Scheme 2).

To probe the topography of the host protein, two different achiral amino acid spacer motifs were introduced between the biotin anchor and the diphosphine moieties: *n*-alkylamino acids, **3**^{n} (n = 1−5), as well as arylamino acids, **4**^{q} (q = *ortho*, *meta*, or *para*). We reasoned that the synthesis and isolation of biotinylated amino acids **Biot−3**^{n} and **Biot−4**^{q} would be most economical (since no racemization is possible even in the event of an intramolecular cyclization, typical of C → N peptide synthesis). Coupling of biotin to the unprotected amino acid was achieved using isobutylchloroformate as the activating agent.[31] The crude biotinylated amino acids were crystallized from EtOH/H_2O (1:1) upon acidification (53−81% yield). Coupling of **Biot−3**^{n} and **Biot−4**^{q} to the aminodiphosphine ligands, H−**1** and H−**2**, was best achieved using 2-chloro-4,6-dimethoxy-1,3,5-triazine (CDMT) as the activating agent in anhydrous acetonitrile.[32] Flash chromatography (SiO_2) under an inert atmosphere using $CHCl_3$/EtOH (7:1) allowed the isolation, in all but one case, of analytically pure biotinylated ligands **Biot−3**^{n}**−1** (n = 1−5, 29−64% yield), **Biot−4**^{q}**−1** (q = *meta* and *para*, 29 and 61% yield, respectively), **Biot−3**^{n}**−2** (n = 1−5, 20−50% yield), and **Biot−4**^{q}**−2** (q = *ortho*, *meta*, and *para*, 21−24% yield) (Scheme 3).

In the case of analytically pure biotinylated anthranilic acid, **Biot−4**^{ortho}, all attempts to couple aminodiphosphine H−**1** yielded **Biot−4**^{ortho}**−1**, but it was contaminated with **Biot−1**! This contamination escaped our scrutiny until recently.[30] Despite numerous attempts, we were unable to separate both phosphines, either by chromatography or by recrystallization. To circumvent this problem, a N → C coupling protocol, typical of peptide coupling methodologies, was followed. CDMT-assisted coupling of diphosphine H−**1** with anthranilic acid, **4**^{ortho}, affords **4**^{ortho}**−1**, which can be isolated in analytically pure form (49% yield) after flash chromatography (SiO_2, 7:1 $CHCl_3$/EtOH). Coupling with biotin is again achieved with CDMT to afford analytically pure **Biot−4**^{ortho}**−1** after flash chromatography (SiO_2, 7:1 $CHCl_3$/EtOH, 43% yield) (Scheme 4).

Scheme 3. Synthesis of Biotinylated Amidodiphosphine Ligands Incorporating Alkylamino Acid (**3**^{n}) or Arylamino Acid (**4**^{q}) Spacers Using a C → N Coupling Protocol



Protein Expression. For screening purposes, both avidin (Avi) and streptavidin (Sav) host proteins were tested. Because of the need for gram quantities of proteins and the prohibitive market price of these proteins, we set out to produce avidin and streptavidin by recombinant means.

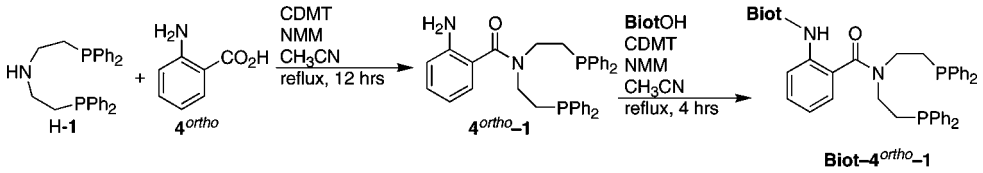
With directed-evolution experiments in mind, the expression of the host protein in the culture medium would considerably simplify the experimental procedure as the culture supernatant could be screened directly in catalysis. For this purpose, the methylotrophic yeast *Pichia pastoris*, strain GS115, in combination with the expression vector pPIC9K (Invitrogen), was selected for the extracellular production of avidin. Because we suspected that the isoelectric point of avidin may decrease the affinity of biotinylated cationic catalyst precursors, we expressed a recombinant glycosylated avidin (r-GAvi) with a lowered isoelectric point, pI = 5.4.[33] Compared to the wild-type avidin (WT Avi), r-GAvi contains the K3E, K9D, R122A, and R124A mutations,[34] as well as additional E-A-E at its N-terminus. Under biotin-depleted expression conditions using a minimal growth medium, 340 mg/L fully functional tetrameric avidin can be isolated and purified via affinity chromatography on an iminobiotin column.

Although reports for the extracellular expression of streptavidin (Sav) in *Bacillus subtilis* exist, high expression levels (up to 50 mg/L) are obtained only with biotin present in the growth medium,[35] thus requiring a denaturation−dialysis step before catalytic experiments are carried out with biotinylated catalyst precursors. Streptavidin can be conveniently expressed in the cytoplasm of *Escherichia coli*.[36−38] To favor the expression of soluble protein, we used the T7-tag, containing mature streptavidin plasmid, provided by Cantor and co-workers.[36] Under optimized conditions, the production of soluble streptavidin (WT Sav) reaches up to 210 mg/L.

Site-Directed Mutagenesis. Site-directed mutagenesis was performed using the Quick-Change mutagenesis kit (Stratagene)

(31) Redeuilh, G.; Secco, C.; Baulieu, E.-E. *J. Biol. Chem.* **1985**, *260*, 3996−4002.
(32) Garrett, C. E.; Jiang, X.; Prasad, K.; Repic, O. *Tetrahedron Lett.* **2002**, *43*, 4161−4165.
(33) Zocchi, A.; Jobé, A. M.; Neuhaus, J.-M.; Ward, T. R. *Protein Expression Purif.* **2003**, *32*, 167−174.
(34) Marttila, A. T.; Airenne, K. J.; Laitinen, O. H.; Kulik, T.; Bayer, E. A.; Wilchek, M.; Kulomaa, M. S. *FEBS Lett.* **1998**, *441*, 313−317.
(35) Nagarajan, V.; Ramaley, R.; Albertson, H.; Chen, M. *Appl. Environ. Microbiol.* **1993**, *59*, 3894−3898.
(36) Sano, T.; Cantor, C. R. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 142−146.
(37) Gallizia, A.; de Lalla, C.; Nardone, E.; Santambrogio, P.; Brandazza, A.; Sidoli, A.; Arosio, P. *Protein Expression Purif.* **1998**, *14*, 192−196.
(38) Chilkoti, A.; Tan, P. H.; Stayton, P. S. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 1754−1758.

Scheme 4. Synthesis of **Biot−4^{ortho}−1** Using a N → C Coupling Protocol



on the pET11b-Sav plasmid.[36] Preliminary experiments suggested the presence of secondary structures within the gene in its single-stranded state, persistent even at a high temperature. Computational analysis confirmed secondary structures at 72 °C. DMSO (dimethyl sulfoxide) and glycerol are known to destabilize DNA secondary structures.[39] After screening different PCR conditions, we determined that site-directed mutagenesis was most efficient in the presence of 5% DMSO using an annealing temperature of 65 °C. These rather drastic conditions are made possible by the length of mutagenic primers (40 bp) and afford the full-length PCR product (i.e., 7000 bp).

Protein Activity. Following the protein expression and purification, we determined the activity of the protein using Gruber's protocol based on the fluorescence quenching of biotin-4-fluorescein upon binding.[47]

If, on the basis of this titration, the number of binding sites is $<3.8 \pm 0.2$, dialysis against Tris HCl (20 mM, pH 7.4) was repeated until this threshold was reached.

In the following discussion, the number of equivalents of biotinylated catalyst precursor was scaled to the actual number of binding sites. For example, for a protein with 3.7 binding sites, 3.0 equiv of biotinylated catalyst precursor corresponds, in reality, to $3.0/4.0 \times 3.7 = 2.78$ equiv.

Screening (Chemogenetic Optimization). The screening experiments were performed using the 18 ligands **Biot−1**, **Biot−2**, **Biot−3^n−1**, **Biot−3^n−2**, **Biot−4^q−1**, and **Biot−4^q−2** (n = 1−5; q = *ortho*, *meta*, and *para*), in combination with both wild-type proteins as well as five mutants proteins: WT Avi, r-GAvi, WT Sav, S112G, V47G, K80G, and P64G. (These latter four mutants are derived from WT Sav.)

By replacing an amino acid belonging to a loop of streptavidin by a glycine residue, we expected to perturb the loop structure and thus influence the enantioselectivity. Four mutations (Figure 1) were performed with the following goals. (i) The first goal is to alter the proximity of the metal site (S112 is located at one hinge of L7,8). This loop, which contains the hydrophobic lid (W120), is expected to be in close contact with the metal during catalysis. (ii) The next goal is to perturb the amino acids that are either participating in or neighboring the second coordination sphere of biotin[40] (V47 interacts with S45, which in turn forms a hydrogen bond with H−N of biotin, and K80 is adjacent to W79, which forms a hydrogen bond with T90 and in turn interacts with the thioether of biotin. (iii) The last goal is to perturb a loop very remote from the biotin binding site (replacing a proline residue (P64) may significantly affect the geometry of the distant loop).

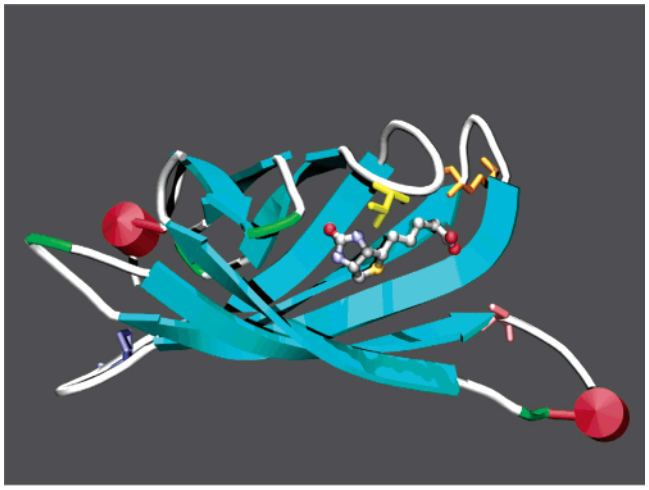


Figure 1. Schematic representation of a monomeric wild-type streptavidin with bound biotin (ball-and-stick), emphasizing the residues (stick) selected for site-directed mutagenesis: P64 (blue), V47 (yellow), K80 (orange), and S112 (pink).

For solubility reasons, acetamidoacrylic acid was preferred over acetamidocinnamic acid as a substrate for screening (5.65 mM substrate concentration). The experiments were performed with a 1 mol % [Rh(COD)(**Biot−ligand**)]$^+$ metal catalyst loading at room temperature and at 5 bar hydrogen pressure in a custom-designed glass autoclave holding up to 21 3-mL glass tubes (1.1 mL final volume, containing 0.1 mL of DMSO, used to solubilize the catalyst precursor).

Previous buffer screening experiments suggested that streptavidin performs best at pH 4.0, while avidin performs best at pH 7.0.[30] Screening was performed in acetate buffer (0.1 M, pH 4.0) for Sav and its mutants. For Avi and r-GAvi, experiments were carried out in MOPS buffer (0.1 M, pH 7.0), except with **Biot−1** in WT Avi and r-GAvi, where catalysis experiments were performed in phosphate buffer (0.07 M, pH 7.1).[41]

Table 2 and Figure 2 list the conversions and the enantiomeric excesses obtained for all ligand−protein combinations.[42]

To gain an unbiased view of the performance of the protein−ligand combinations, the matrix (and the corresponding transposed matrix) displayed in Table 2 was subjected to principal component analysis (PCA).[43] This statistical procedure transforms the set of correlated variables (proteins and ligands) into another set of uncorrelated (orthogonalized) variables (the principal components, PCs) that are ordered by decreasing variability (i.e., the first PC contains the most variability). These uncorrelated variables are linear combinations of the original

(39) Hung, T.; Mak, K.; Fong, K. *Nucl. Acids Res.* **1990**, *18*, 4953.
(40) Klumb, L. A.; Chu, V.; Stayton, P. S. *Biochemistry* **1998**, *37*, 7657−7663.
(41) All attempts to obtain reproducible and high activities and selectivities (ee > 20%) using **Biot−1** in Avi and r-GAvi in different buffers failed, suggesting that these particular ligand−protein combinations are very sensitive to pH and buffer composition.[27]
(42) The slight variations in enantiomeric excesses, compared to those of ref 30, are due to the systematic use of a glovebox to prepare catalytic runs. The enantiomeric excess reported for **Biot−4^{ortho}−1** in ref 30 was heavily biased by the previously unnoticed contamination of **Biot−1** in the system.
(43) Kachigan, S. K. *Multivariate Statistical Analysis: A Conceptual Introduction*, 2nd ed.; Radius Press: New York, 1991.
(44) Horsman, G. P.; Liu, A. M. F.; Henke, E.; Bornscheuer, U. T.; Kazlauskas, R. J. *Chem.−Eur. J.* **2003**, *9*, 1933−1939.
(45) Hofmann, K.; Titus, G.; Montibeller, J. A.; Finn, F. M. *Biochemistry* **1982**, *21*, 978−984.
(46) Green, N. M. *Methods Enzymol.* **1970**, *18*, 418−424.
(47) Kada, G.; Falk, H.; Gruber, H. J. *Biochim. Biophys. Acta* **1999**, *1427*, 33−43.

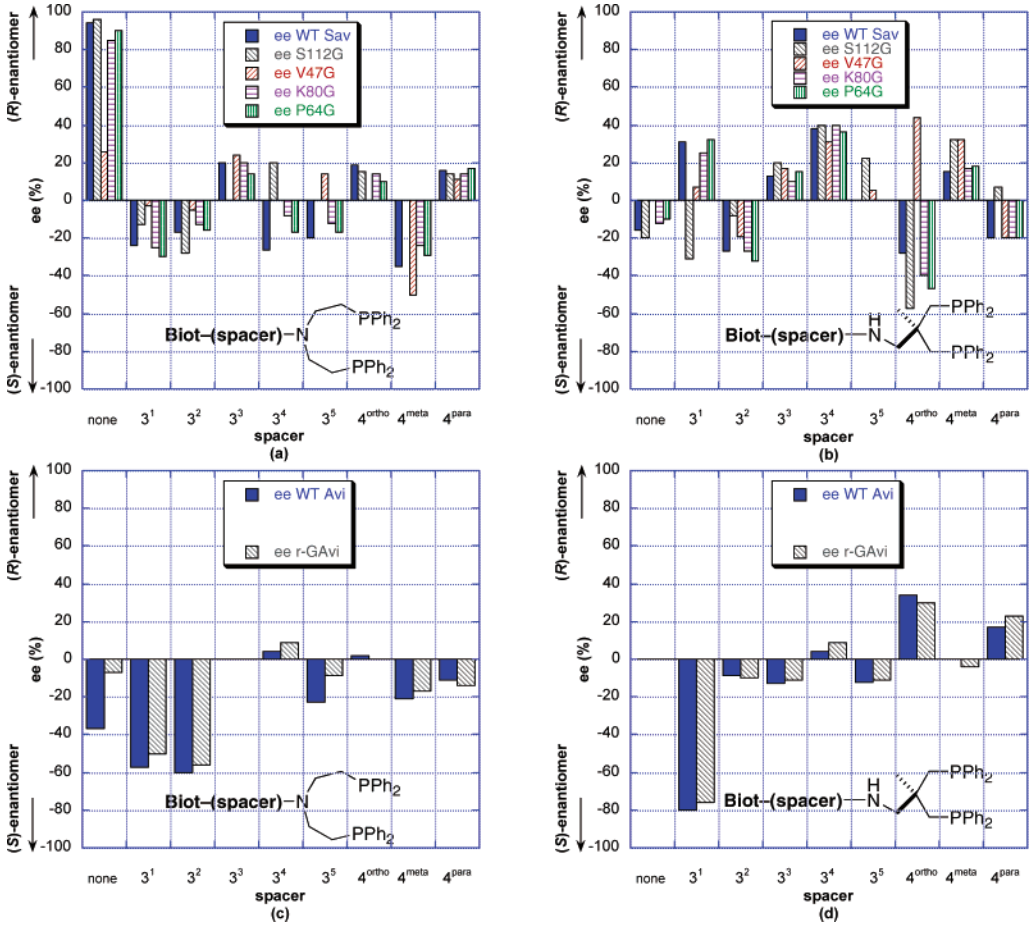


Figure 2. Graphic summary of the enantiomeric excess obtained for the enantioselective reduction of acetamidoacrylic acid with ligand scaffold **1** in the streptavidin family (a), with ligand scaffold **2** in the streptavidin family (b), with ligand scaffold **1** in the avidin family (c), and with ligand scaffold **2** in the avidin family (d).

Table 2. Enantiomeric Excesses (% ee) Obtained with the 18 Ligands in Combination with the 7 Proteins for the Reduction of Acetamidoacrylic Acid[a]

	WT Sav	S112G	V47G	K80G	P64G	WT Avi	r-GAvi
Biot−1	94	96	26	85	90	−39	−7
Biot−3^1−1	−24	−13	−3	−25	−30	−57	−50[b]
Biot−3^2−1	−17	−28	−5	−13	−16	−60	−56[c]
Biot−3^3−1	20	0	24	20	14	0	0[b]
Biot−3^4−1	−26	20	0	−8	−17	4	9[b]
Biot−3^5−1	−20	0	14	−12	−17	−23	−9[d]
Biot−4^{ortho}−1	19	15	0	14	10	2	0[e]
Biot−4^{meta}−1	−35	0	−50	−24	−29	−21	−17[c]
Biot−4^{para}−1	16	14	11	14	17	−11	−14[b]
Biot−2	−16[b]	−20	0	−12[f]	−10[d]	0[d]	0[e]
Biot−3^1−2	31	−31	7	25	32	−80	−76[b]
Biot−3^2−2	−27	−8	−19	−27	−32	−9	−10[d]
Biot−3^3−2	13	20	17	10	15	−13	−11
Biot−3^4−2	38	40	31	40	36	4	9
Biot−3^5−2	0	22	5	0	0	−12	−11[c]
Biot−4^{ortho}−2	−28	−57	44	−40	−47	34	30[g]
Biot−4^{meta}−2	15	32	32	17	18	0	−4[c]
Biot−4^{para}−2	−20[f]	7[d]	−20[d]	−20	−20[f]	17	23[c]

[a] Positive enantiomeric excess values are in favor of the (*R*) enantiomer, while negative enantiomeric excess values are in favor of the (*S*) enantiomer. Unless otherwise stated, all conversions were quantitative after 12 h at room temperature. [b] Conversion > 80%. [c] Conversion > 90%. [d] Conversion > 70%. [e] Conversion > 40%. [f] Conversion > 50%. [g] Conversion > 30%. [h] Conversion > 60%.

variables. The latter variables may be removed with minimal loss of information; in the present case, three PCs suffice to account for 96% of the variability in the data.

Hierarchical clustering was then applied on the scores of both the proteins and the ligands (using the matrix and the transposed matrix, respectively) in order to group similar entities. Hierarchical clustering may be represented by a two-dimensional diagram known as a dendrogram, which illustrates the fusions or divisions made at each successive stage of analysis (Figure 3). To assess the consistency of the statistical analysis performed on the original data sets, the experimental data set was split into 20 contiguous blocks that were randomly perturbed by the absolute experimental error estimated at 2%. The PCA performed on the modified data set, accounting for experimental uncertainties, clearly demonstrates that the experimental error does not affect the original underlying structures of the data sets.

The first dendogram clearly shows that the avidin and the streptavidin families are dissimilar. Among the streptavidins, WT Sav, K80G, and P64G form a cluster, which merges with V47G. Mutant S112G clearly stands out as a singular entity (Figure 3a). Hierarchical ordering of the ligands suggests that **Biot−1** forms a cluster with itself. The next four ligands, namely, **Biot−4^{ortho}−2**, **Biot−3^1−2**, **Biot−3^1−1**, and **Biot−3^2−1** (Figure 3b), form three subclusters laying in proximity.

In the avidin family, the best enantiomeric excesses are obtained with the glycine and the β-alanine spacers. In particular, **Biot−3^1−2** gives the highest enantiomeric excess in favor of the (*S*) enantiomer (80%). No significant variation with respect

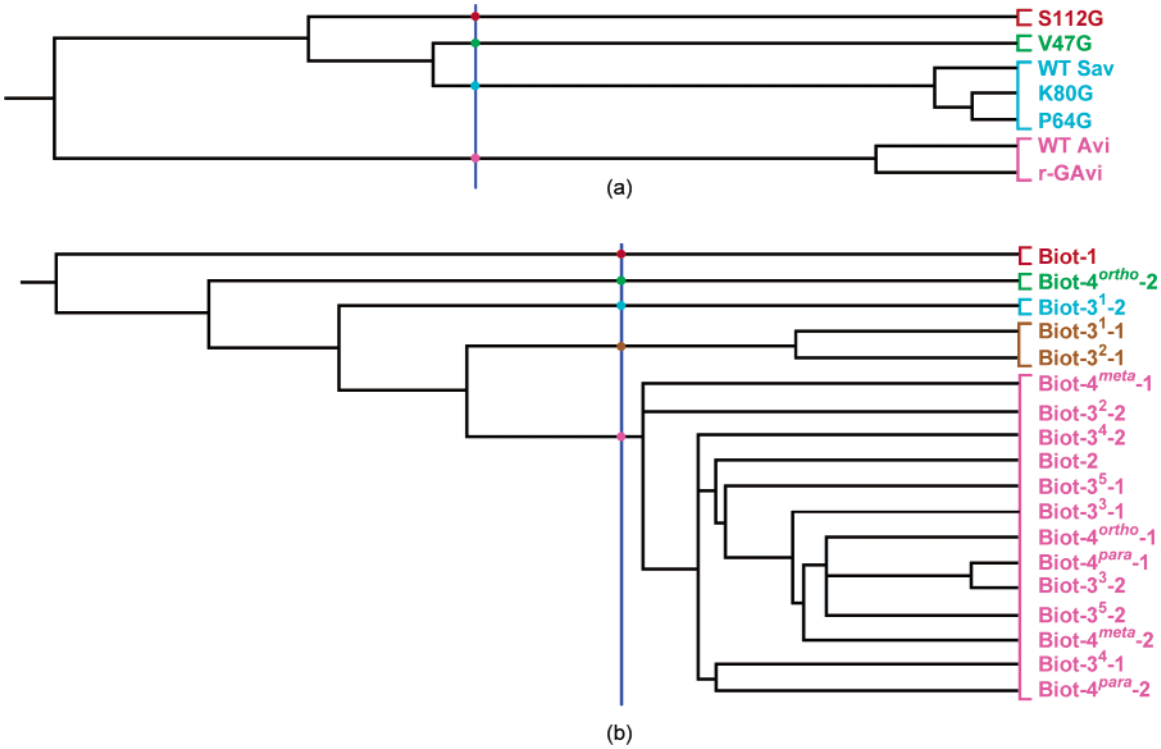


Figure 3. Dendrogram illustrating the fusions or divisions made at each successive stage of PCA. The analysis was performed on the score space obtained by a PCA. Hierarchical ordering of the proteins (a) and of the ligands (b). (IDs are listed along the *y*-axis; the *x*-axis measures intercluster distance.)

to WT Avi is observed with the r-GAvi (largest increase +6% for **Biot**−**4**^{para}−**2**, largest decrease −32% for **Biot**−**1**). Overall, both Avi and r-GAvi display a bias in favor of the (*S*) enantiomer (Figure 2c,d).

In the streptavidin family, neither K80G nor P64G provides much variability to the enantiomeric excess with respect to WT Sav (largest increase +19% for **Biot**−**4**^{ortho}−**2**, largest decrease −18% for **Biot**−**3**^{4}−**1**). In contrast, S112G and V47G show broad variability with respect to WT Sav. In particular, **Biot**−**4**^{ortho}−**2** exhibits the largest increase in enantiomeric excess (+29% in S112G, which corresponds to the highest enantiomeric excess (57%) in favor of the (*S*) enantiomer in the Sav family) and the largest decrease in enantiomeric excess (−72% in V47G) for all of the mutants tested. Along similar lines, **Biot**−**1** affords the highest enantiomeric excess in favor of the (*R*) enantiomer (96% in S112G) and the second largest decrease (−68% in V47G). These data suggest that **Biot**−**4**^{ortho}−**2** and **Biot**−**1** are particularly sensitive to the protein environment and may be good candidates for the genetic optimization of the selectivity. All of these observations are in full agreement with the dendograms displayed in Figure 3.

Although most conversions are quantitative within 12 h at room temperature, it appears that r-GAvi performs more poorly than all of the other proteins tested. In the streptavidin family, **Biot**−**2** and **Biot**−**4**^{para}−**2** tend to give conversions lower than those of all other spacer−ligand combinations (Table 2).

The data presented herein confirms (the common wisdom) that mutations close to the active site affect the selectivity more dramatically than do distant mutations.[44] Indeed, statistical analysis of the data clearly shows that mutations that either affect the biotin binding site via second coordination sphere interactions (V47G) or lie in proximity to the metal active site (S112G) provide more variability than either remote loop residues (P64 in Sav as well as K3, K9, R122, and R124 in Avi) or residues which are close but not directly involved in second coordination sphere interactions (K80G) with biotin. This latter mutation further suggests that removal of a positive charge in the vicinity of the catalytic moiety does not affect significantly the selectivity of the catalyst.

Stability Constant Determination. As suggested by Hofmann, the affinity of N^{ϵ}-iminobiotinyllysineamide for avidin is significantly lower than that of neutral iminobiotinylated amino acids. This unusual behavior was traced to the repulsion between the positively charged avidin host at neutral pH and the positively charged iminobiotinylated guest.[45] Considering the difference of the isoelectric points of avidin (pI = 10.4) and streptavidin (pI = 6.4), we speculated that the superior performance of streptavidin as the host protein may be caused by its higher affinity for the cationic biotinylated catalyst precursors. Since the first coordination sphere of the catalyst is achiral (with the exception of the chirality of the (+)-biotin bicyclic framework), any catalyst not incorporated within the host protein produces (nearly) racemic hydrogenation products.[30] We thus set out to determine the strength of the guest⊂host interaction of [Rh(COD)**Biot**−**1**]$^{+}$⊂(strept)avidin. In the field of study of biotin−avidin, it is widely accepted that the four biotin-binding events are noncooperative.[25,48] For the stability constant determination, we thus fit the data with a single binding constant.

The host protein was initially loaded with a large excess of 2-(4′-hydroxyazobenzene)benzoic acid (HABA), ensuring nearly complete saturation of the biotin binding sites (eq 1). Addition of [Rh(COD)(**Biot**−**1**)]$^{+}$ aliquots caused a decrease of both

(48) Jones, M. L.; Kurzban, G. P. *Biochemistry* **1995**, *34*, 11750−11756.

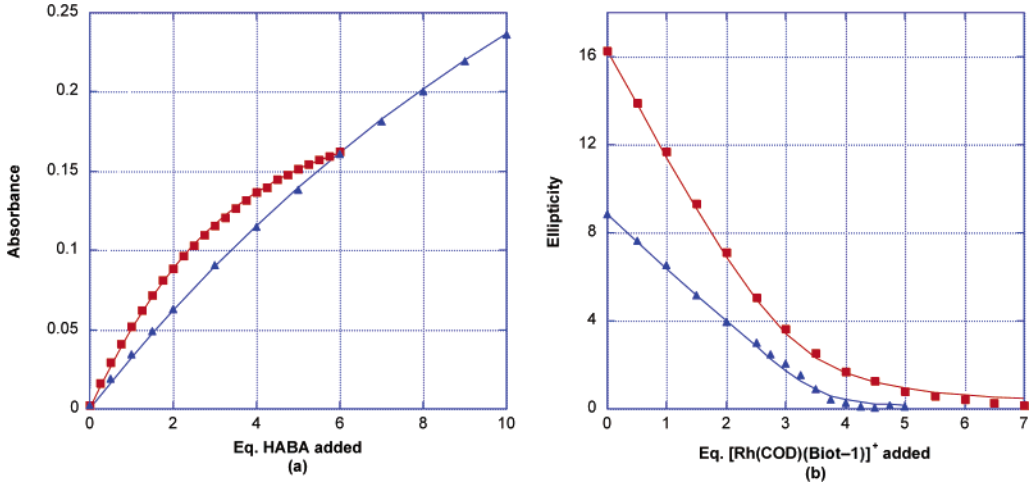


Figure 4. Two-step procedure for the determination of the overall stability constant for [Rh(COD)(**Biot**−**1**)]$^+$⊂(strept)avidin. The appearance of an absorption band (λ_{max} = 506 nm) upon addition of HABA aliquots to (strept)avidin allows the determination of the K_{HABA} for HABA⊂(strept)avidin (a). Decrease of the CD signal (λ_{max} = 506 nm) of HABA⊂(strept)avidin for the displacement equilibrium K_{exch} between HABA and [Rh(COD)(**Biot**−**1**)]$^+$ (b). Blue triangles are the measured points for WT Sav; the blue line is the fitted curve for WT Sav. Red squares are the measured points for WT Avi; the red line is the fitted curve for WT Avi.

visible and CD signals (λ_{max} = 506 nm) of HABA⊂(strept)-avidin, allowing the determination of the stability constant for the displacement equilibrium in eq 2.

$$\text{HABA} + \text{(strept)avidin} \rightleftarrows \text{HABA}\subset\text{(strept)avidin} \quad (1)$$

$$\text{HABA}\subset\text{(strept)avidin} + [\text{Rh(COD)(}\mathbf{Biot}-\mathbf{1}\text{)}]^+ \rightleftarrows [\text{Rh(COD)(}\mathbf{Biot}-\mathbf{1}\text{)}]^+\subset\text{(strept)avidin} + \text{HABA} \quad (2)$$

The stability constant, K_{HABA} (eq 3), for WT Avi (log K_{HABA} = 5.5 (R^2 = 0.996))[46] and for WT Sav (log K_{HABA} = 3.5 (R^2 = 0.997))[46] was determined by following the appearance of a visible band centered at 506 nm that was caused by the HABA⊂(strept)avidin interaction. With these values at hand, and assuming a single binding constant (i.e., noncooperative binding events), we determined the stability constant, $K_{[\text{Rh(COD)}\mathbf{Biot}-\mathbf{1}]^+\subset\text{(strept)avidin}}$, according to eq 5.

$$K_{HABA} = \frac{[\text{HABA}\subset\text{(strept)avidin}]}{[\text{HABA}] \times [\text{(strept)avidin}]} \quad (3)$$

$$K_{exch} = \frac{[\text{Rh(COD)(}\mathbf{Biot}-\mathbf{1}\text{)}\subset\text{(strept)avidin}] \times [\text{HABA}]}{[\text{HABA}\subset\text{(strept)avidin}] \times [\text{Rh(COD)(}\mathbf{Biot}-\mathbf{1}\text{)}]} \quad (4)$$

$$K_{[\text{Rh(COD)(}\mathbf{Biot}-\mathbf{1}\text{)}]^+\subset\text{(strept)avidin}} = \frac{[\{\text{Rh(COD)(}\mathbf{Biot}-\mathbf{1}\text{)}\}\subset\text{(strept)avidin}]}{[\text{Rh(COD)(}\mathbf{Biot}-\mathbf{1}\text{)}] \times [\text{(strept)avidin}]} = K_{HABA} \times K_{exch} \quad (5)$$

Since the biotinylated catalyst is not soluble in water, methanol was used to solubilize the stock solution of [Rh(COD)-(**Biot**−**1**)]$^+$. Using the known total concentrations, $[\text{HABA}]_{total}$, $[\text{(strept)avidin}]_{total}$, and $[\text{Rh(COD)(}\mathbf{Biot}-\mathbf{1}\text{)}]^+_{total}$, and rearranging eq 4, one can derive a single quadratic equation to estimate the concentration of the absorbing species, [HABA⊂(strept)-avidin] (eq 6):

$$(1 - K_{exch}) \times [\text{HABA}\subset\text{(strept)avidin}]^2 + (K_{exch} \times \{[\text{(strept)avidin}]_{total} - [\text{Rh(COD)(}\mathbf{Biot}-\mathbf{1}\text{)}]\} - [\text{HABA}]_{total} - [\text{(strept)avidin}]_{total}) \times [\text{HABA}\subset\text{(strept)avidin}] + [\text{HABA}]_{total} \times [\text{(strept)avidin}]_{total} = 0 \quad (6)$$

For a given value of the fitting parameter, K_{exch}, the corresponding theoretical titration curve is calculated, and a least-squares minimization is performed until the sum of all data points $(\text{signal}_{calculated} - \text{signal}_{measured})^2$ reaches a minimum. The calculated and the measured curves are depicted in Figure 4.

From these stability constant determinations (pH 7.0, phosphate buffer, I = 0.3 M), we conclude that the affinity of [Rh-(COD)(**Biot**−**1**)]$^+$ for the considered host protein is independent of the isoelectric point of the protein and is essentially constant: log $K_{[\text{Rh(COD)(}\mathbf{Biot}-\mathbf{1}\text{)}]^+\subset\text{Avi}}$ = 7.7 and log K_{exch} = 2.2 (R^2 = 0.998), and log $K_{[\text{Rh(COD)(}\mathbf{Biot}-\mathbf{1}\text{)}]^+\subset\text{Sav}}$ = 7.2 and log K_{exch} = 3.7 (R^2 = 0.997). The quality of the fits using single binding constants suggests that there is no significant cooperativity upon binding.

From these data, we conclude that at a given pH, the affinity of a cationic catalyst precursor for a protein does not depend on the isoelectric point of the host protein. Thus, both avidin and streptavidin should be suitable host proteins. The difference in performance is not caused by a difference in affinity, but most likely, it is caused by a difference in their respective binding pocket topography.

Pressure and Temperature. The performance of [Rh(COD)-(**Biot**−**1**)]$^+$⊂(strept)avidin was evaluated as a function of pressure and temperature (Figure 5a,b, respectively). The temperature study reveals that the artificial metalloenzyme performs sluggishly up to 20 °C. Hydrogenation of acetamidoacrylic acid at 5 bar during 12 h at 0 and 10 °C affords acetamidoalanine with a 33% conversion (80% ee (*R*)) and a 73% conversion (85% ee (*R*)), respectively. The optimal temperature for hydrogenation is 30 °C, affording a quantitative conversion after 12 h and 94% ee (*R*). Above this temperature, the enantioselectivity decreases (Figure 5b). Hydrogenation at 1 and 5 bar both gives the same enantioselectivity (94% ee (*R*),

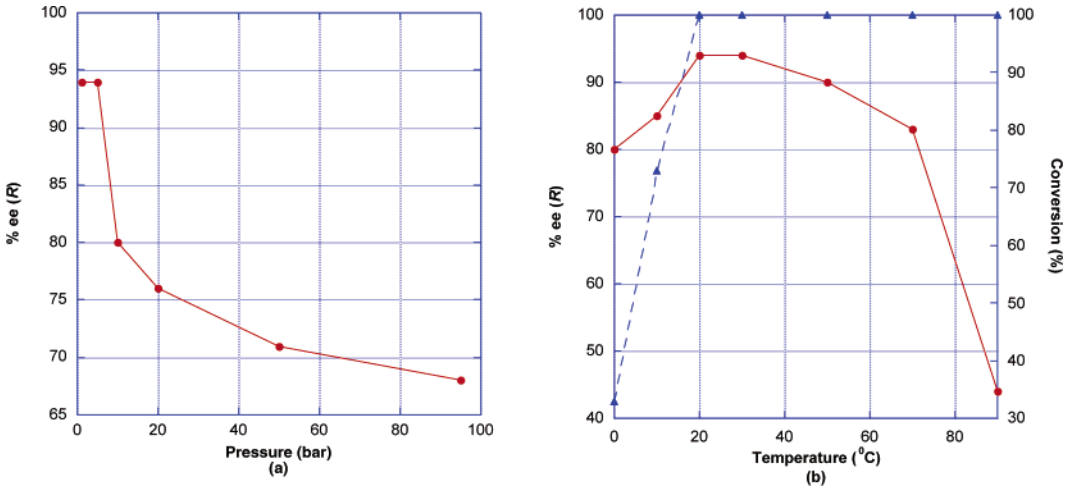


Figure 5. Activity and selectivity of [Rh(COD)(**Biot**−**1**)]$^+$⊂WT Sav as a function of pressure (a) and temperature (b). The red circles represent the enantioselectivity, and the triangles represent the conversion. All pressure experiments were performed at 30 °C and were quantitative after 12 h.

quantitative conversion at 30 °C after 12 h). Thereafter, the enantiomeric excess rapidly decreases as a function of pressure (Figure 5a). It appears that the magnitude of the variations of both the activity and the selectivity is more pronounced than that for most enantioselective hydrogenation catalysts (although exceptions exist),[49] suggesting that the protein conformation is affected both by temperature and by pressure.

The ideal operating conditions are 5 bar hydrogen pressure and 30 °C, thus ensuring high enantioselectivities and rapid conversions.

Outlook

The work presented herein demonstrates the versatility of a combined chemogenetic optimization procedure for the creation of artificial metalloenzymes based on the biotin−avidin technology.

Screening of seven host proteins (two wild-type and five mutant proteins), with 18 biotin−spacer−ligand scaffolds, reveals the best protein−ligand combinations produce (*R*)-acetamidoalanine (96% ee in S112G with **Biot**−**1**) or (*S*)-acetamidoalanine (80% ee in WT Avi with **Biot**−**3^1**−**2**; 57% ee in S112G with **Biot**−**4^{ortho}**−**2**).

Statistical analysis of the data suggests future optimization avenues. (i) Streptavidin and its mutants display more variability than avidin and its mutants. (ii) Position S112 (and to a lesser extent V47) in streptavidin appears to be particularly sensitive to mutation. (iii) Biotinylated ligands, **Biot**−**1**, **Biot**−**3^1**−**2**, **Biot**−**3^1**−**1**, **Biot**−**3^2**−**1**, and **Biot**−**4^{ortho}**−**2**, display the most variability. All of the other ligands tested behave similarly and display less variability.

More detailed studies with [Rh(COD)(**Biot**−**1**)]$^+$ in both WT Avi and WT Sav uncover the following features. (i) The affinity of the cationic catalyst precursor for both host proteins is essentially identical, despite their marked difference in pI. Since the repulsion energy is inversely proportional to the distance between the two point charges, remote positive charges on the host protein's side chains do not influence the affinity of the cationic catalyst precursor. (ii) The marked pressure and temperature dependence of the activity and the selectivity of the artificial metalloenzymes were revealed.

From the considerations outlined in Table 1, we conclude that such artificial metalloenzymes display features which are reminiscent both of homogeneous catalysts and of enzymes. In particular, we have shown that the second coordination sphere can be exploited either by mutagenesis or by chemical means (by the introduction of a spacer between the biotin anchor and the metal) to produce more active and selective hydrogenation catalysts.

Targeted mutations (e.g., introduction of a metal-coordinating amino acid) within the second coordination sphere should make these systems more enzyme-like as the active site will be tailored for the needs of a particular catalytic cycle. For this purpose, we favor an optimization procedure inspired by statistical analysis and molecular modeling over a high-throughput screening using directed evolution. Current efforts in this direction are in progress within our group.

Acknowledgment. This work was funded by the Swiss National Science Foundation, CERC3, and the Canton of Neuchâtel. We thank Belovo Egg Science and Technology for a generous gift of egg white avidin. We thank C. R. Cantor for the streptavidin gene, J.-M. Neuhaus, P. Schürmann (University of Neuchâtel), and P. Arosio (University of Milano) for their help in setting up the streptavidin and avidin productions, and Z. Lei (University of Berne) for the help with the mutagenesis experiments.

Supporting Information Available: Complete experimental procedures, stability constant determination, characterization of all intermediates and ligands, and hydrogenation protocol. This material is available free of charge via the Internet at http://pubs.acs.org.

JA0476718

(49) Landis, C. R.; Halpern, J. *J. Am. Chem. Soc.* **1987**, *109*, 1746−1754.