ACS Catalysis

Letter

pubs.acs.org/acscatalysis

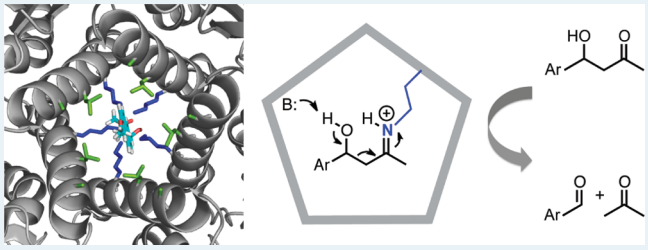
# Harnessing Protein Symmetry for Enzyme Design

Bigna Wörsdörfer, Lisa M. Henning, Richard Obexer, and Donald Hilvert*

Laboratory of Organic Chemistry, ETH Zürich, 8093 Zürich, Switzerland

Supporting Information

**ABSTRACT:** Cyclic protein oligomers are common in Nature. Here we show that the central pore of the pentameric ring-forming protein lumazine synthase from *Saccharomyces cerevisiae* (ScLS) can be rationally engineered to catalyze a retro-aldol reaction. The $C_5$-symmetry of the complex was exploited to equip the protein tunnel with a ring of five closely spaced lysines adjacent to an apolar site for substrate binding. The resulting system utilizes amine catalysis to promote the cleavage of (±)-methodol to 6-methoxy-2-naphthaldehyde and acetone with a >$10^3$-fold rate acceleration. The ease of organizing convergent functional groups within a protein pore may make the tunnels of many symmetric ring-shaped proteins useful starting points for creating designer enzymes.




**KEYWORDS:** *enzyme models, convergent design, amine catalysis, aldol reaction*

The majority of soluble and membrane-bound proteins and nearly all structural proteins found in living cells spontaneously form symmetric multimeric complexes.[1] The grooves, pockets, and pores generated at the symmetry interfaces of such supramolecular structures are potentially attractive sites for the generation of new function,[2] including catalysis. In a symmetric ring-shaped homooligomer the structural and functional effects of a tunnel mutation in one subunit result in multiple symmetry-related changes in close proximity within the complex, providing a facile means of arranging convergent functional groups. Although substrates and transition states usually lack symmetry, Nature shows that symmetric active sites are capable of efficiently transforming asymmetric molecules. Cleavage of diverse peptides at the $C_2$-symmetric active site of HIV protease is one prominent example.[3] Catalysis of dihydrofolate reduction in a 25 Å-long, $D_2$-symmetric tunnel passing through the middle of the homotetrameric R67 dihydrofolate reductase is another.[4,5]

*Saccharomyces cerevisiae* lumazine synthase (ScLS), which normally catalyzes the penultimate step in riboflavin biosynthesis, is a toroidal homopentamer.[6] The $C_5$-symmetric tunnel at the center of the pentameric ring is 12−18 Å from the natural active site (Figure 1A and 1B) and has no known function. Nevertheless, its dimensions—an average C$\alpha$ to C$\alpha$ diameter of ∼16 Å and a length of ∼25 Å—would be appropriate for binding small molecules and catalyzing simple reactions. Engineering binding or catalytic activity into such a site requires that the protein scaffold accommodate multiple clustered mutations. The ScLS tunnel has been shown to be extremely tolerant to mutation and to vastly different chemical environments.[7] Even simultaneous substitution of several residues lining the pore does not disrupt the protein's quaternary structure. Taking advantage of this property and simple mechanistic principles, we have equipped the ScLS

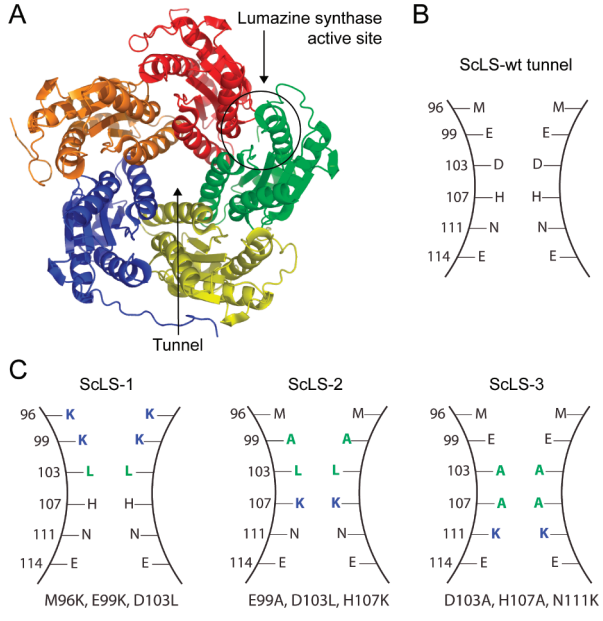



**Figure 1.** Equipping the ScLS tunnel with retro-aldolase activity. (A) Structure of the pentameric enzyme (PDB: 1hqk), showing the location of the natural active site and the central 5-fold symmetric pore. (B) Schematic representation of the residues lining the tunnel of ScLS-wt. (C) Scheme showing the residues lining the tunnel of the three retro-aldolase designs. The catalytic lysines (K) are shown in blue, the aliphatic alanine (A) and leucine (L) residues introduced to interact with a hydrophobic substrate are in green.

**Received:** January 30, 2012
**Revised:** March 22, 2012
**Published:** April 18, 2012

tunnel with a reactive lysine residue (Figure 1C) and significant retro-aldolase activity.

The amine-catalyzed retro-aldol reaction of methodol (**1**) to give naphthaldehyde (**3**) and acetone (Figure 2) represents an

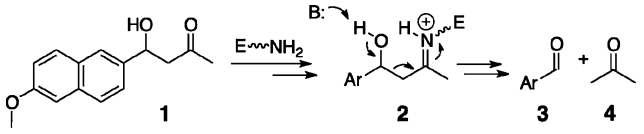



**Figure 2.** Amine-catalyzed retro-aldol reaction of methodol (**1**) proceeds via a Schiff base intermediate (**2**) to give 6-methoxy-2-naphthaldehyde (**3**) and acetone (**4**).

ideal model reaction since it is mechanistically well understood, the features required for catalysis are known, and a sensitive assay is available.[8] Natural class I aldolases exploit amine catalysis with the $\varepsilon$-amine group of a lysine at the active site serving as a critical functional group.[9−11] The catalytic cycle is initiated by nucleophilic attack of the unprotonated lysine on the carbonyl group of the substrate to give a Schiff base adduct (**2**). This iminium ion acts as an electron sink, facilitating cleavage of the adjacent C−C bond and generation of an aldehyde and an enamine. Protonation of the enamine and subsequent hydrolysis releases a ketone and regenerates the catalyst.

This mechanism has been mimicked by a variety of designed catalysts, including lysine-rich $\alpha$- and $\beta$-peptides,[12−15] small molecules,[16−18] catalytic antibodies,[19,20] and computationally designed enzymes.[21−23] Since imine formation is partially rate limiting in an aqueous environment, an important design feature for such catalysts is provision of a nucleophilic lysine with a depressed p$K_a$ to speed up this step. The p$K_a$ of the catalytic lysine can be lowered either by placing the side chain in a hydrophobic microenvironment or through Coulombic interactions with proximal positively charged residues, such as other lysines. The first strategy has been used successfully to create catalytic antibodies[19,20] and computationally designed enzymes[21,22] with significant aldolase activity. The second strategy has been used to design peptides[14,15] and cyclodextrin-based catalysts.[16,17]

The tunnel of the ScLS pentameric ring is bounded by one five-turn $\alpha$-helix from each subunit (Figure 1A) and provides an excellent scaffold for placing multiple lysines in close proximity to one another. Introduction of one lysine per monomer would yield a ring of five closely spaced basic residues. Since simultaneous protonation of all five groups would be electrostatically unfavorable, this arrangement should favor deprotonation of one or more of the amines under physiological conditions, providing a catalytic group for reaction with a substrate molecule bound within the tunnel. To design ScLS tunnel variants with aldolase activity, residues that project into the tunnel at different depths were first identified in the crystal structure of ScLS. Lysines were introduced at these sites in PyMOL,[24] and the hydrophobic aldol substrate methodol was docked by hand within the tunnel to place its ketone near the lysine amines. Other tunnel residues were then replaced by apolar alanines or leucines to provide favorable binding interactions with the hydrophobic naphthyl group of the substrate. On the basis of these considerations, we chose the three designs, depicted in Figure 1C. The first features two rings of lysines at the N-terminal end of the tunnel and a deep binding pocket for the substrate. In the second design, the lysines are placed in the narrow middle region of the tunnel, and the hydrophobic substrate binding site is located toward the N-terminal entry port. In the third design, the lysines and hydrophobic residues are translated one helical turn toward the C-terminal end of the tunnel.

Individual mutations were introduced into the ScLS gene by site-directed mutagenesis. The three tunnel variants M96K/E99K/D103L (ScLS-1), E99A/D103L/H107K (ScLS-2), and D103A/H107A/N111K (ScLS-3) plus wild type ScLS (ScLS-wt) were produced in *Escherichia coli* and purified by $Ni^{2+}$-affinity chromatography using a C-terminal $His_6$-tag. All proteins were obtained in high yield (typically 60 mg to 100 mg per liter culture). Circular dichroism (CD) spectroscopy revealed no changes in secondary structure in the variants compared to wild type ScLS.

The activity of the designed LS tunnel variants for retro-aldol cleavage of racemic methodol was investigated by monitoring the formation of the fluorescent product 6-methoxy-2-naphthaldehyde (Figure 2). As test reaction conditions, 20 $\mu$M catalyst and 500 $\mu$M substrate were chosen. All variants showed detectable retro-aldol activity over the uncatalyzed reaction at pH 7.5 (Figure 3A). Even wild type ScLS exhibited nearly 4-fold higher initial rates than the uncatalyzed background reaction, probably due to lysines on the protein surface. One of the designed tunnel variants, ScLS-2, was particularly active. Methodol cleavage in the presence of this protein was 25 times faster than background and nearly seven times faster than wild type ScLS. The higher activity of ScLS-2 can be attributed to the mutations introduced into the tunnel, since they are the only changes relative to ScLS-wt. In the absence of crystal structures, we can only speculate as to why ScLS-2 is superior to the other two designs. However, the location of the catalytic lysines roughly in the middle of the pore, juxtaposed with a ring of leucines that snugly binds the substrate, may have been decisive. Modeling suggests that the engineered pocket in ScLS-3 is deeper but considerably wider than in ScLS-2 (Supporting Information, Figure S1), potentially reducing substrate affinity. In contrast, steric clashes between the substrate and the substantially narrower pore in ScLS-1 may diminish the efficacy of this design (Supporting Information, Figure S1). Moreover, although ScLS-1 has a larger number of lysines than the other two designs, the rotameric preferences of this amino acid would also tend to direct the side chains away from the engineered hydrophobic site toward solvent.

The properties of ScLS-2 were characterized in greater detail to provide insight into its mechanism of action. The tunnel mutations do not impair assembly of the pentameric ring as ScLS-2 elutes from a size exclusion chromatography column at a volume corresponding to the size of a pentamer. Because the proposed mechanism depends on a lysine residue with a perturbed p$K_a$, the reaction should be pH dependent. Consistent with this expectation, the reaction rate increases with increasing pH (Figure 3B). Not surprisingly, given the close proximity of the five lysines, the data are poorly fit by an equation for a catalyst possessing a single ionizable group. A much better fit is achieved if two ionizable groups with apparent p$K_a$ values of 7.7 and 9.4 are invoked (Figure 3B). Nevertheless, the rates at the lowest pH values are still higher than predicted by this model, so a significant population of deprotonated amine can apparently be generated even at pH 6. With increasing pH, additional lysines become deprotonated, increasing the local concentration of reactive amine at the active site. These residues might also mediate essential proton


dx.doi.org/10.1021/cs300076t | *ACS Catal.* 2012, 2, 982−985

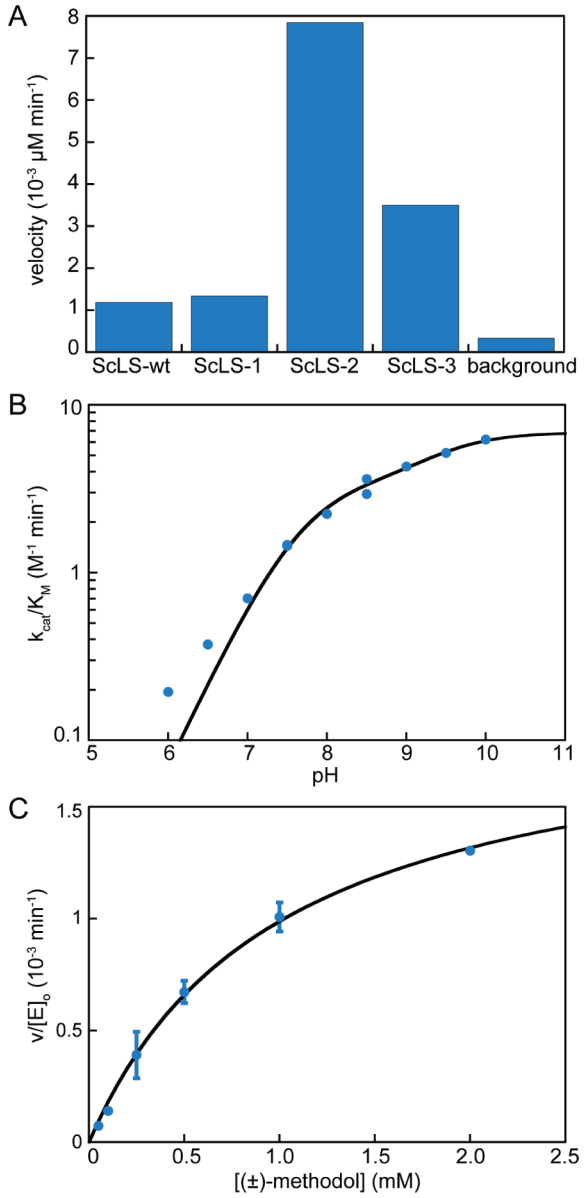



**Figure 3.** Kinetic characterization of ScLS-derived retro-aldolases. (A) Specific activities for the cleavage of **1** by ScLS-wt and the designs. (B) pH-rate profile for ScLS-2. Rates were determined at 50 μM substrate and 20 μM ScLS-2 pentamer and fitted to the equation $(k_{cat}/K_M)_{obs} = (k_{cat}/K_M)_{max}/(1 + 10^{pK_{a1}-pH}) + (k_{cat}/K_M)_{max}/(1 + 10^{pK_{a2}-pH})$. (C) Michaelis–Menten plot for ScLS-2. Reactions were performed in 50 mM sodium phosphate, pH 7.5 containing 150 mM NaCl and 5% MeCN, 20 μM ScLS pentamer and 50–2000 μM substrate at 30 °C.

transfers during the multistep reaction pathway (Figure 2). For comparison, the ε-amino group of lysine free in solution is essentially fully protonated under physiological conditions ($pK_a$ ~10.5) and hence much less reactive.[25] In natural enzymes such as the D-2-deoxyribose-5-phosphate aldolase (DERA)[10] and previously designed artificial aldolases the $pK_a$ value of the catalytic lysine ranges between 5.5 and 9.[14,15,20,25]

The catalytic efficiency of the active tunnel variant ScLS-2 was assessed by steady-state kinetic measurements at pH 7.5 and 30 °C. The enzyme-catalyzed retro-aldol reaction of racemic **1** follows Michaelis–Menten kinetics with a $k_{cat}$ of 2.0 × $10^{-3}$ $min^{-1}$ and a $K_M$ of 1.0 mM (Figure 3C). Comparison of the catalytic rate constant of ScLS-2 with the rate constant of the uncatalyzed reaction (6 × $10^{-7}$ $min^{-1}$) gives a rate acceleration ($k_{cat}/k_{uncat}$) of more than 3000. Moreover, the apparent second-order rate constant $k_{cat}/K_M$ (2 $M^{-1}$ $min^{-1}$) is 1.4 × $10^4$-fold larger than the second-order rate constant for the nonenzymic reaction promoted by butylamine, an analogue for the free lysine side chain. ScLS-2 also compares favorably with α-peptides that promote the cleavage of methodol. Thus, it has a >3-fold higher turnover number and a 3 to 6-fold larger $k_{cat}/K_M$ value than catalysts isolated from peptide phage libraries by reaction-based selection with 1,3-diketone inhibitors.[14] Its activity is also comparable to that of typical first-generation computationally designed retro-aldolases ($k_{cat}/K_M$ = 1 to 44 $M^{-1}$ $min^{-1}$).[21] However, ScLS-2 is 4 orders of magnitude less efficient than the catalytic antibody 38C2,[8] which itself is about 2 orders of magnitude less efficient than natural aldolase enzymes.[10]

The enantioselectivity of ScLS-2 was examined using the isolated (*R*) and (*S*) enantiomers of methodol. Comparison of their respective $k_{cat}/K_M$ values shows that the catalyst is *S*-selective, but the preference for retro-aldol cleavage of the (*S*) enantiomer is only a modest 1.4-fold. Other substrates and/or catalyst variants might conceivably exhibit higher levels of asymmetric induction.

Although ScLS-2 is much less active than the best aldolase antibodies and natural enzymes, its design, which focused on generation of a reactive lysine, is minimalistic. More effective catalysts of this multistep reaction typically possess additional functional groups to position substrate and orchestrate essential proton transfers, and introduction of such groups would be expected to enhance activity. Directed evolution has been effectively utilized to optimize many catalysts,[26−29] including computationally designed retro-aldolases,[22,23] and might conceivably be employed to boost the activity and enantioselectivity of the first-generation tunnel catalyst as well. Functional diversity in this system could be further increased by combining differently modified subunits to generate more complex heterooligomeric systems.

Our results demonstrate the feasibility of tailoring the central pore of a symmetric protein ring for a complex catalytic task. Given the ease of assembling convergent functional groups within the tunnel microenvironment, ring-forming proteins may be attractive starting points for accessing other enzyme activities.

## ■ ASSOCIATED CONTENT

### Supporting Information

Complete experimental procedures. This material is available free of charge via the Internet at http://pubs.acs.org.

## ■ AUTHOR INFORMATION


### Corresponding Author

*E-mail: hilvert@org.chem.ethz.ch.



### Funding

This work was supported by the Schweizerischer Nationalfond and the ETH Zürich.


### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We are grateful to Lars Giger and Sandro Tonazzi for the preparation and chiral resolution of methodol and to Florian

Seebeck and Kenneth J. Woycechowsky for stimulating discussions.

## ■ ABBREVIATIONS

ScLS, *Saccharomyces cerevisiae* lumazine synthase

## ■ REFERENCES

(1) Goodsell, D. S.; Olson, A. J. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 105−153.
(2) Monod, J.; Wyman, J.; Changeux, J. P. *J. Mol. Biol.* **1965**, *12*, 88−118.
(3) Prabu-Jeyabalan, M.; Nalivaika, E.; Schiffer, C. A. *Structure* **2002**, *10*, 369−381.
(4) Narayana, N.; Matthews, D. A.; Howell, E. E.; Xuong, N. H. *Nat. Struct. Biol.* **1995**, *2*, 1018−1025.
(5) Schmitzer, A. R.; Lepine, F.; Pelletier, J. N. *Protein Eng. Des. Sel.* **2004**, *17*, 809−819.
(6) Meining, W.; Mörtl, S.; Fischer, M.; Cushman, M.; Bacher, A.; Ladenstein, R. *J. Mol. Biol.* **2000**, *299*, 181−197.
(7) Woycechowsky, K. J.; Seebeck, F. P.; Hilvert, D. *Protein Sci.* **2006**, *15*, 1106−1114.
(8) List, B.; Barbas, C. F., III; Lerner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 15351−15355.
(9) Gefflaut, T.; Blonski, C.; Perie, J.; Willson, M. *Prog. Biophys. Mol. Biol.* **1995**, *63*, 301−340.
(10) Heine, A.; DeSantis, G.; Luz, J. G.; Mitchell, M.; Wong, C. H.; Wilson, I. A. *Science* **2001**, *294*, 369−374.
(11) Fullerton, S. W. B.; Griffiths, J. S.; Merkel, A. B.; Cheriyan, M.; Wymer, N. J.; Hutchins, M. J.; Fierke, C. A.; Toone, E. J.; Naismith, J. H. *Bioorg. Med. Chem.* **2006**, *14*, 3002−3010.
(12) Johnsson, K.; Allemann, R. K.; Widmer, H.; Benner, S. A. *Nature* **1993**, *365*, 530−532.
(13) Weston, C. J.; Cureton, C. H.; Calvert, M. J.; Smart, O. S.; Allemann, R. K. *ChemBioChem* **2004**, *5*, 1075−1080.
(14) Tanaka, F.; Fuller, R.; Barbas, C. F., III *Biochemistry* **2005**, *44*, 7583−7592.
(15) Müller, M. M.; Windsor, M. A.; Pomerantz, W. C.; Gellman, S. H.; Hilvert, D. *Angew. Chem., Int. Ed.* **2009**, *48*, 922−925.
(16) Tagaki, W.; Yamamoto, H. *Tetrahedron Lett.* **1991**, *32*, 1207−1208.
(17) Yuan, D. Q.; Dong, S. D.; Breslow, R. *Tetrahedron Lett.* **1998**, *39*, 7673−7676.
(18) Mukherjee, S.; Yang, J. W.; Hoffmann, S.; List, B. *Chem. Rev.* **2007**, *107*, 5471−5569.
(19) Wagner, J.; Lerner, R. A.; Barbas, C. F., III *Science* **1995**, *270*, 1797−1800.
(20) Barbas, C. F., III; Heine, A.; Zhong, G. F.; Hoffmann, T.; Gramatikova, S.; Björnestedt, R.; List, B.; Anderson, J.; Stura, E. A.; Wilson, I. A.; Lerner, R. A. *Science* **1997**, *278*, 2085−2092.
(21) Jiang, L.; Althoff, E. A.; Clemente, F. R.; Doyle, L.; Röthlisberger, D.; Zanghellini, A.; Gallaher, J. L.; Betker, J. L.; Tanaka, F.; Barbas, C. F., III; Hilvert, D.; Houk, K. N.; Stoddard, B. L.; Baker, D. *Science* **2008**, *319*, 1387−1391.
(22) Wang, L.; Althoff, E. A.; Bolduc, J.; Jiang, L.; Moody, J.; Lassila, J. K.; Giger, L.; Hilvert, D.; Stoddard, B.; Baker, D. *J. Mol. Biol.* **2012**, *415*, 615−625.
(23) Althoff, E. A.; Wang, L.; Jiang, L.; Giger, L.; Lassila, J. K.; Wang, Z.; Smith, M.; Hari, S.; Kast, P.; Herschlag, D.; Hilvert, D.; Baker, D. *Protein Sci.* **2012**, *21*, 717−726.
(24) *The PyMOL Molecular Graphics System*; Schrödinger, LLC: Portland, OR.
(25) Lassila, J. K.; Baker, D.; Herschlag, D. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 4937−4942.
(26) Romero, P. A.; Arnold, F. H. *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 866−876.
(27) Turner, N. J. *Nat. Chem. Biol.* **2009**, *5*, 567−573.
(28) Jäckel, C.; Hilvert, D. *Curr. Opin. Biotechnol.* **2010**, *21*, 753−759.
(29) Reetz, M. T. *Angew Chem., Int. Ed.* **2011**, *50*, 138−174.