JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

# Fast Knoevenagel Condensations Catalyzed by an Artificial Schiff-Base-Forming Enzyme

Xavier Garrabou, Basile I. M. Wicky, and Donald Hilvert*

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

Supporting Information

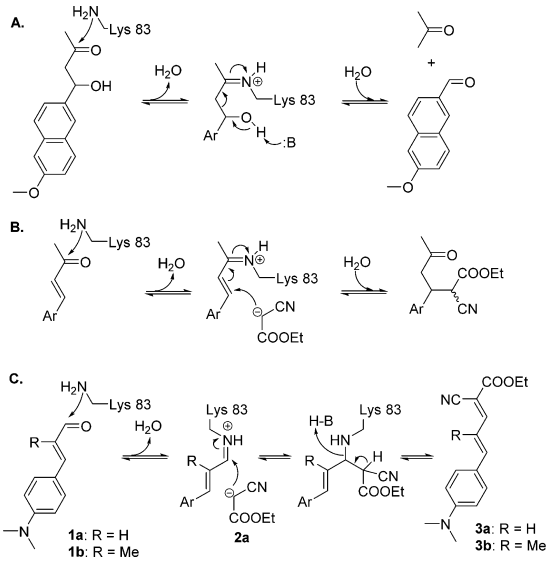
**ABSTRACT:** The simple catalytic motifs utilized by enzymes created by computational design and directed evolution constitute a potentially valuable source of chemical promiscuity. Here we show that the artificial retro-aldolase RA95.5-8 is able to use a reactive lysine in a hydrophobic pocket to accelerate promiscuous Knoevenagel condensations of electron-rich aldehydes and activated methylene donors. Optimization of this activity by directed evolution afforded an efficient enzyme variant with a catalytic proficiency of $5 \times 10^{11}$ $M^{-1}$ and a $>10^8$-fold catalytic advantage over simple primary and secondary amines. Divergent evolution of de novo enzymes in this way could be a promising strategy for creating tailored biocatalysts for many synthetically useful reactions.

Enzymes are increasingly finding application in the chemical industry for their ability to catalyze diverse reactions with exacting selectivity and high turnovers under benign reaction conditions. If natural enzymes are not suitable for a specific transformation of interest, substrate scope and stereochemical preferences can often be tailored through protein engineering.[1] Nevertheless, even minor modifications can disrupt the complex arrays of functional groups that characterize natural active sites, so more drastic reprogramming is often challenging. For this reason, artificial enzymes created by computational design and directed evolution represent a potentially useful source of new activities.[2,3] Their simple catalytic motifs, embedded in evolvable scaffolds, can facilitate rapid functional diversification.

A family of artificial aldolases is illustrative of this approach.[4] These enzymes were computationally designed to accelerate the aldol cleavage of methodol via covalent amine catalysis by a nucleophilic lysine in an apolar pocket (Scheme 1A). Although their starting activities are typically modest, substantial improvements can be achieved by directed evolution.[5−7] In the case of the retro-aldolase RA95, for instance, multiple rounds of mutagenesis and screening increased catalytic efficiency more than 4000-fold.[8] In addition to retro-aldol activity, the optimized variant RA95.5-8 was found to promote Michael reactions by activation of unsaturated ketones as transient iminium adducts with the catalytic lysine (Scheme 1B).[9] Here, the promiscuous activity of this enzyme is shown to include Knoevenagel condensations, a cornerstone transformation in amine catalysis.[10]

A set of aldehydes was tested as potential acceptor substrates against typical nucleophiles for Knoevenagel reactions (Schemes S1 and S2). As previously reported,[11] ethyl 2-cyanoacetate (**2a**) readily reacts with a range of aliphatic and aromatic aldehydes in the absence of catalyst. In contrast, aldehydes **1a** and **1b** (Scheme 1C), which are deactivated by the presence of a strong electron-donating group, are poor reaction partners for **2a**. In the presence of RA95.5-8, however, **1a** and **1b** condense rapidly with **2a** to give the corresponding Knoevenagel products **3a**/**b**. Precipitation of **3a**/**b** in aqueous medium ensures that the enzymatic condensations proceed in high yield (isolated: 93% and 83%, respectively). Product formation was not observed with less acidic Michael donors (**2c**−**e** in Scheme S2).

**Scheme 1. Promiscuous Reactions Catalyzed by RA95.5-8**




We optimized RA95.5-8 for the Knoevenagel reaction of **1a** and **2a** by directed evolution. First, 16 active-site residues (Figure S1) were independently randomized, and variants with enhanced activity were identified using a microtiter plate assay to monitor aldehyde depletion spectrophotometrically at 350 nm (see Supporting Information (SI)). Five mutations that conferred enhanced activity (L9V, F89W, N135G, K210L, and S233G) were shuffled. A triple mutant (L9V/N135G/K210L RA95.5-8, or KN.1) was identified that exhibited a 7.5-fold activity enhancement over the starting catalyst in the plate assay. Further activity improvements relied on screening larger

**Received:** January 23, 2016
**Published:** May 19, 2016





DOI: 10.1021/jacs.6b00816
J. Am. Chem. Soc. 2016, 138, 6972−6974

libraries created by error-prone PCR and gene shuffling. A medium-throughput assay was devised to screen 10 000–25 000 variants on agar plates. Large Petri plates containing ~2000 colonies were sprayed with a buffered solution of **1a** (1 mM) and **2a** (2.5–40 mM), and after 5–20 min incubation, colonies expressing active enzymes were detected by the insoluble red Knoevenagel product **3a** that formed. This assay allowed fast, qualitative identification of the most active variants in the library (see Figure 1A), which could be subsequently produced and analyzed in individual photometric assays. The donor concentration in the assay was decreased after each round of mutagenesis and screening to detect increasingly active variants.

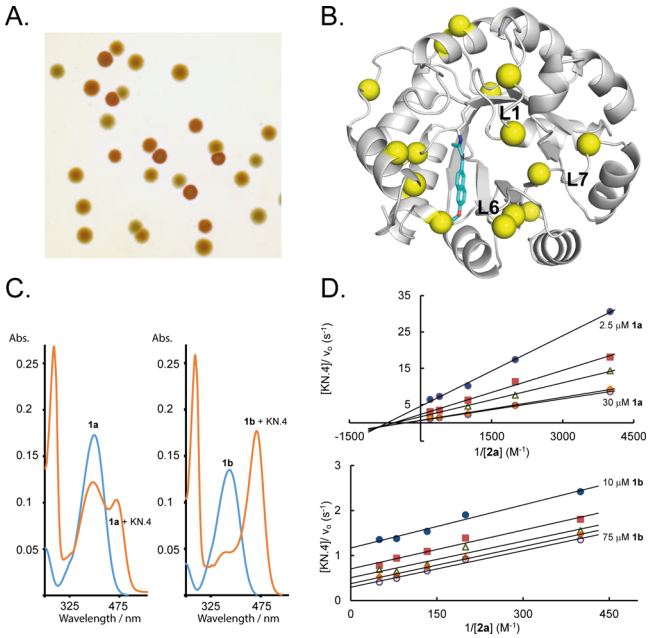



**Figure 1.** Evolution and characterization of the Knoevenagel catalyst KN.4. (A) Agar plate assay for the Knoevenagel reaction of **1a** and **2a**. Bacteria expressing the enzymes RA95.5-8, K83M RA95.5-8, KN.1, and KN.4 produce the insoluble red Knoevenagel product **3a**, with the most active variants producing the most intense color (color gradation is more obvious by direct visual inspection). (B) The KN.4 mutations are shown mapped onto the structure of RA95.5-5. The location of the active site is indicated by a diketone inhibitor (shown as cyan sticks) covalently bound to Lys83.[8] (C) Spectroscopic analysis of buffered solutions (HEPES 25 mM, NaCl 100 mM, pH 7.5) of 5 $\mu$M **1a** (left) and **1b** (right), before (blue trace) and after (orange trace) addition of 10 $\mu$M KN.4. Similar spectra were obtained using RA95.5-8 (Figure S4). (D) Double-reciprocal plots for the KN.4-catalyzed reactions of **1a** (upper plot) and **1b** (lower plot) with **2a**.

After four rounds of directed evolution, the best variant, KN.4, contained 15 mutations (L9V, V12I, S21R, S58L, D61L, R97H, K98N, I107V, I113V, N135G, M182L, N183D, N204D, K210L, S214G). These substitutions include six residues that point into the active site and four more in the second shell of residues surrounding the binding pocket. Changes in the L1 (D61L, S58L), L6 (M182L, N183D) and L7 (K210L, S214G) loops, which bracket the active site, may favor reorganization of these highly flexible regions.[8] Four additional mutations, located in the first $\alpha$-helix (L9V, V12I) and the neighboring L3 (I113V) and L4 (N135G) loops, cluster around the putative acceptor binding site (see Figure 1B). The remaining mutations are in the hydrophobic core (I107V) or scattered over the protein surface (S21R, R97H, K98N, N204D).

Several lines of evidence suggest that the enzyme-catalyzed Knoevenagel reactions involve initial formation of a Schiff base adduct between the acceptor aldehyde and Lys83, followed by attack of **2a** on the iminium ion (Scheme 1C). First, mutation of Lys83, which is required for the aldol and Michael reactions catalyzed by RA95.5-8,[8,9] to methionine decreases the rate of the Knoevenagel reaction 2100 fold. Second, addition of either RA95.5-8 or KN.4 to a buffered solution of **1a** results in the appearance of a new absorption peak at 475 nm (Figure 1C), indicative of Schiff base formation. Based on spectroscopic studies of Schiff base adducts of **1a** in different solvents,[12] this species is probably protonated. In analogous experiments with **1b**, which possesses an additional $\alpha$-methyl group, intermediate formation was more pronounced judging from the higher intensity of the new signal and marked reduction of the original absorption band (Figure 1C). Lastly, the Schiff base between **1a** and KN.4 can be reduced by sodium borohydride and the resulting covalent adduct detected by ESI-MS (calcd 29 789.3 Da, found 29 789.8 Da; Figure S5).

RA95.5-8 and the KN.4 variant were kinetically characterized by varying the concentration of the acceptor at several fixed concentrations of the donor. The extinction coefficient for the Knoevenagel products **3a** and **3b** was estimated by spectroscopic analysis of the more soluble methoxyethyl ester analogues **3c** ($\Delta\varepsilon_{520nm}$ = 29 160 $M^{-1}$ $cm^{-1}$) and **3d** ($\Delta\varepsilon_{500nm}$ = 17 457 $M^{-1}$ $cm^{-1}$), respectively (see SI). The intersecting lines of the double-reciprocal plot obtained for the reaction of **1a** and **2a** (e.g., Figure 1D) are consistent with Schiff base formation being readily reversible. Although 45-fold less reactive than **1a**, **1b** is also an excellent substrate for the enzyme. In this case, the covalent Schiff base adduct is substantially populated, and the double-reciprocal plot gave a set of parallel lines (Figure 1D) consistent with a ping-pong mechanism. The same substrate-specific kinetic patterns were observed following directed evolution (Figures S2 and S3).

Steady-state parameters (Table 1) were obtained by fitting the kinetic data to either a simplified equation for a bi-substrate

**Table 1. Steady-State Kinetic Parameters for the Enzymatic Knoevenagel Reactions**[a]

| catalyst | reaction | $k_{cat}$ ($s^{-1}$) | $K_1$($\mu$M) | $K_{2a}$ (mM) | $k_{cat}/(K_1K_{2a})$ ($M^{-2}$ $s^{-1}$) | $k_{cat}/k_{uncat}$ (M) | $1/K_{TS}$[b] ($M^{-1}$) |
|---|---|---|---|---|---|---|---|
| RA95.5-8 | **1a** + **2a** | 0.14 ± 0.01 | 7.9 ± 0.7 | 8.3 ± 1.3 | $2.2 \times 10^6$ | 345 | $5.3 \times 10^9$ |
| RA95.5-8 | **1b** + **2a** | 0.48 ± 0.07 | 106 ± 20 | 29 ± 5 | $1.5 \times 10^5$ | $5.3 \times 10^4$ | $1.7 \times 10^{10}$ |
| KN.4 | **1a** + **2a** | 1.9 ± 0.1 | 19 ± 2 | 1.7 ± 0.2 | $5.9 \times 10^7$ | $4.6 \times 10^3$ | $1.4 \times 10^{11}$ |
| KN.4 | **1b** + **2a** | 7.1 ± 1.2 | 74 ± 16 | 21 ± 4 | $4.6 \times 10^6$ | $7.8 \times 10^5$ | $5.1 \times 10^{11}$ |

[a]Measured at 29 °C in 25 mM HEPES, 100 mM NaCl, pH 7.5. The estimated errors reflect the standard deviations of the data global fit. Under the same conditions, the second-order constant for the uncatalyzed Knoevenagel condensation of **1a** and **2a**, $k_{uncat}$, was $4.1 \times 10^{-4}$ $M^{-1}$ $s^{-1}$; for the reaction of **1b** with **2a**, $k_{uncat}$ = $9.1 \times 10^{-6}$ $M^{-1}$ $s^{-1}$. [b]Chemical proficiency: $1/K_{TS} = [k_{cat}/(K_1K_{2a})]/k_{uncat}$.[15]



DOI: 10.1021/jacs.6b00816
J. Am. Chem. Soc. 2016, 138, 6972–6974

reaction (**1a**) or a ping-pong mechanism (**1b**) (see SI).[13] The mutations introduced into RA95.5-8 by directed evolution improved the chemical proficiency ($[k_{cat}/(K_1K_{2a})]/k_{uncat}$) for both reactions ca. 30-fold, largely owing to increases in turnover number. Although the $\alpha$-methyl substituent in **1b** increases the Michaelis constant for both substrates, possibly reflecting an altered binding mode for the Schiff base intermediate, it does not preclude high turnover. Thus, the evolved catalyst KN.4 promotes the normally slow reaction of **1b** and **2a** with a $k_{cat}$ of 7.1 $s^{-1}$, which corresponds to an effective molarity ($k_{cat}/k_{uncat}$) of $7.8 \times 10^5$ M. KN.4 is also $>10^5$-fold more proficient than bovine serum albumin, a generic protein that contains lysine residues adjacent to a hydrophobic pocket.[14] Moreover, its nucleophilic lysine is $>10^8$ times more effective than *n*-butylamine or simple organocatalysts such as piperidine or L-proline at promoting the Knoevenagel reaction of **1a**/**1b** and **2a** (see the SI).

Enzymatic Knoevenagel condensations are rare in natural biosynthetic pathways; to date, only one enzyme is known to catalyze an intramolecular Knoevenagel reaction as its physiological function.[16] In mechanistic terms, though, the KN.4-catalyzed transformation is reminiscent of a lysine-assisted carbonyl condensation step in the multistep sequence catalyzed by porphobilinogen synthase.[17] This and other Schiff-base-dependent activities of the "aldolase" superfamily highlight the versatility of amine catalysis in the $(\beta\alpha)_8$ barrel scaffold utilized by the artificial retro-aldolase.[18]

The Knoevenagel condensation is a fundamental, broadly applied carboligation in organic chemistry. Knoevenagel adducts are often electrophilic intermediates in one-pot multicomponent reactions that afford complex, potentially chiral molecules.[19] Such transformations could become biocatalytically accessible if tailored enzymes, such as KN.4, were combined with other enzymatic activities to create innovative cascades. Though promiscuous Knoevenagel reactions with activated aldehydes have been reported for several proteases and esterases,[20] these transformations typically require high catalyst loading (e.g., >1:500)[21] and largely organic reaction media. The high reactivity and evolvability of RA95.5−8 suggest that it and related computationally designed retro-aldolases might be excellent starting points for catalyzing asymmetric versions of the Knoevenagel reaction based on the dynamic kinetic resolution of $\alpha$-substituted aldehydes.[22] Because the chemistry of iminium ions and enamines is broad in scope but mechanistically simple, this approach could also conceivably be applied to other condensations such as oxime or hydrazone formation[23,24] or to tune the photochemical properties of conjugated carbonyl compounds.[25]

The simple active sites and robust scaffolds of computationally designed enzymes represent a promising source of promiscuous activity. Functional diversification of these proteins by laboratory evolution has the potential to provide access to tailored active sites able to promote the extensive repertoire of known amine-catalyzed reactions with higher rates and selectivities than simple organocatalysts. By extending this approach to other catalytic motifs, it may be possible to enhance and diversify the chemistry of many de novo enzymes.

## ■ ASSOCIATED CONTENT

### Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b00816.

Complete experimental procedures, including Schemes S1 and S2, Table S1, and Figures S1−S5, with sequence information and additional kinetic data (PDF)

## ■ AUTHOR INFORMATION


### Corresponding Author

*hilvert@org.chem.ethz.ch


### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

The authors are grateful to the Swiss National Science Foundation (SNSF) and the ETH Zurich for generous support of this work.

## ■ REFERENCES


(1) Bornscheuer, U. T.; Huisman, G. W.; Kazlauskas, R. J.; Lutz, S.; Moore, J. C.; Robins, K. *Nature* **2012**, *485*, 185−194.

(2) Kries, H.; Blomberg, R.; Hilvert, D. *Curr. Opin. Chem. Biol.* **2013**, *17*, 221−228.

(3) Kiss, G.; Çelebi-Ölçüm, N.; Moretti, R.; Baker, D.; Houk, K. N. *Angew. Chem., Int. Ed.* **2013**, *52*, 5700−5725.

(4) 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.; Hilvert, D.; Houk, K. N.; Stoddard, B. L.; Baker, D. *Science* **2008**, *319*, 1387−1391.

(5) 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.

(6) Obexer, R.; Studer, S.; Giger, L.; Pinkas, D. M.; Grütter, M. G.; Baker, D.; Hilvert, D. *ChemCatChem* **2014**, *6*, 1043−1050.

(7) Bjelic, S.; Kipnis, Y.; Wang, L.; Pianowski, Z.; Vorobiev, S.; Su, M.; Seetharaman, J.; Xiao, R.; Kornhaber, G.; Hunt, J. F.; Tong, L.; Hilvert, D.; Baker, D. *J. Mol. Biol.* **2014**, *426*, 256−271.

(8) Giger, L.; Caner, S.; Obexer, R.; Kast, P.; Baker, D.; Ban, N.; Hilvert, D. *Nat. Chem. Biol.* **2013**, *9*, 494−498.

(9) Garrabou, X.; Beck, T.; Hilvert, D. *Angew. Chem., Int. Ed.* **2015**, *54*, 5609−5612.

(10) List, B. *Angew. Chem., Int. Ed.* **2010**, *49*, 1730−1734.

(11) Li, W.; Li, R.; Yu, X.; Xu, X.; Guo, Z.; Tan, T.; Fedosov, S. N. *Biochem. Eng. J.* **2015**, *101*, 99−107.

(12) Jimenez-Sanchez, A.; Rodriguez, M.; Metivier, R.; Ramos-Ortiz, G.; Maldonado, J. L.; Reboles, N.; Farfan, N.; Nakatani, K.; Santillan, R. *New J. Chem.* **2014**, *38*, 730−738.

(13) Segel, I. H. *Enzyme Kinetics: behavior and analysis of rapid equilibrium and steady-state enzyme systems*; Wiley: New York, 1993.

(14) Carter, D. C.; Ho, J. X. *Adv. Protein Chem.* **1994**, *45*, 153−203.

(15) Radzicka, A.; Wolfenden, R. *Science* **1995**, *267*, 90−93.

(16) Balakrishnan, B.; Chen, C.-C.; Pan, T.-M.; Kwon, H.-J. *Tetrahedron Lett.* **2014**, *55*, 1640−1643.

(17) Jaffe, E. K. *Bioorg. Chem.* **2004**, *32*, 316−325.

(18) Gerlt, J. A.; Raushel, F. M. *Curr. Opin. Chem. Biol.* **2003**, *7*, 252−264.

(19) *Multicomponent Reactions*; Zhu, J., Bienaymé, H., Eds.; Wiley-VCH: Weinheim, 2005.

(20) Miao, Y.; Rahimi, M.; Geertsema, E. M.; Poelarends, G. J. *Curr. Opin. Chem. Biol.* **2015**, *25*, 115−123.

(21) Wang, C.-H.; Guan, Z.; He, Y.-H. *Green Chem.* **2011**, *13*, 2048−2054.

(22) Lee, A.; Michrowska, A.; Sulzer-Mosse, S.; List, B. *Angew. Chem., Int. Ed.* **2011**, *50*, 1707−1710.

(23) Dirksen, A.; Hackeng, T. M.; Dawson, P. E. *Angew. Chem., Int. Ed.* **2006**, *45*, 7581−7584.

(24) Dirksen, A.; Dirksen, S.; Hackeng, T. M.; Dawson, P. E. *J. Am. Chem. Soc.* **2006**, *128*, 15602−15603.

(25) Smith, S. O. *Annu. Rev. Biophys.* **2010**, *39*, 309−328.




DOI: 10.1021/jacs.6b00816
J. Am. Chem. Soc. 2016, 138, 6972−6974