Supporting Text for:

Origins of Catalysis by Computationally Designed Retroaldolase Enzymes

Jonathan K. Lassila[1], David Baker[2], and Daniel Herschlag[1,3]

[1]Department of Biochemistry, Stanford University, Stanford, California 94305, U.S.A.

[2]Department of Biochemistry, University of Washington, Seattle, Washington 98195, U.S.A.

[3]To whom correspondence should be addressed: Department of Biochemistry, Beckman Center, B400, Stanford University, Stanford, CA 94305-5307. Phone 650-723-9442. Fax: 650-723-6783. Email: herschla@stanford.edu

Binding of naphthaldehyde product to enzyme

The 6-methoxy-2-naphthaldehyde product can return to the active site and form a covalent iminium species with lysine (Figure S1). Formation of this iminium state is not part of the reaction pathway, and it requires free lysine, which is regenerated after loss of acetone in steps 4 and 5.

The fluorescent signal of 6-methoxy-2-naphthaldehyde exhibits an exponential decrease when incubated with enzyme, and the rate and extent of loss depends on the concentration of enzyme (Figure S2). Fluorescence was monitored as described in the main text, with excitation at 330 nm and emission at 452 nm.

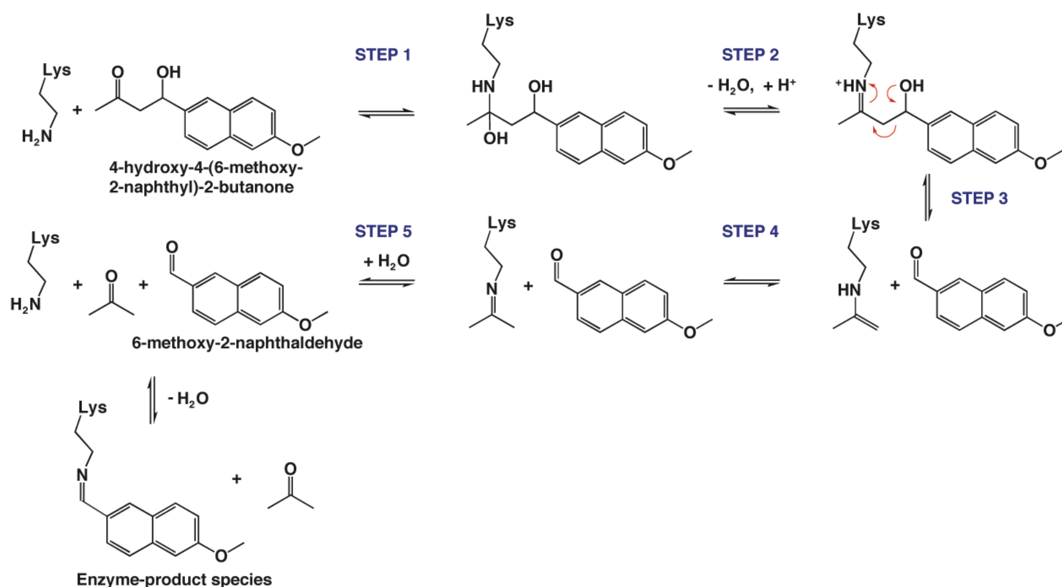


Figure S1. Formation of covalent iminium species with the product 6-methoxy-2-naphthaldehyde.

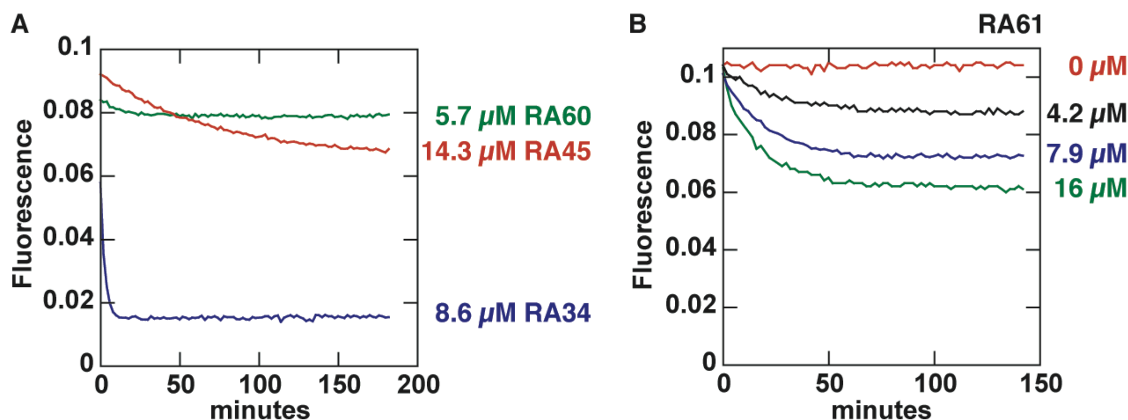


Figure S2. An exponential decay of fluorescence was observed when 6-methoxy-2-naphthaldehyde was incubated with enzyme. Standard assay conditions were used (see main text). **A.** Three variants at the concentrations indicated were incubated with 0.1 μM naphthaldehyde and monitored by fluorescence at 452 nm. Fluorescence background of enzyme and cuvette was not subtracted. **B.** Loss of fluorescence observed at the indicated concentrations of RA61, corrected for fluorescence background of enzymes and cuvettes. Initial naphthaldehyde concentration was 0.1 μM.

The fluorescence of bound 6-methoxy-2-naphthaldehyde is lower than that of the free species, but because of limitations in the solubility of some of the enzymes, saturating concentrations of enzyme were not attainable and we could not obtain a value for the fluorescence of the enzyme-bound naphthaldehyde species. For further analysis of RA61, we made the assumption that the fluorescence of the naphthaldehyde species is zero when it is completely bound to enzyme. The assumption is based on the observation that the Y78F/S87A double mutant could be concentrated slightly more and showed fluorescence loss of ~80% at achievable concentrations. While this assumption may not be correct and the enzyme-naphthaldehyde complex may show some fluorescence, the results and conclusions described here would not be significantly affected by a low level of fluorescence remaining for the enzyme-bound species, as discussed below.

The final observed fluorescence was converted to a concentration of free naphthaldehyde, using a standard curve. The final concentration of free naphthaldehyde was plotted against enzyme concentration, and the following binding equation was used to obtain a K_d value:

$$[N]_f = \frac{K_d[N]_t}{[E]_t + K_d}$$

In this equation, $[N]_f$ represents the final free naphthaldehyde concentration, $[E]_t$ and $[N]_t$ are the total enzyme and naphthaldehyde concentrations, respectively. Figure S3 shows results for RA61, giving a K_d value from curve fitting of 26 μM. Because the enzyme concentration could not be raised to obtain data points above the K_d, the value represents an approximate upper limit for K_d. If, for example, the naphthaldehyde product retained 20% of its fluorescent signal when fully bound to enzyme, the resulting K_d would be about 17 μM. If 40% of the fluorescence was retained, the resulting K_d would be about 10 μM.

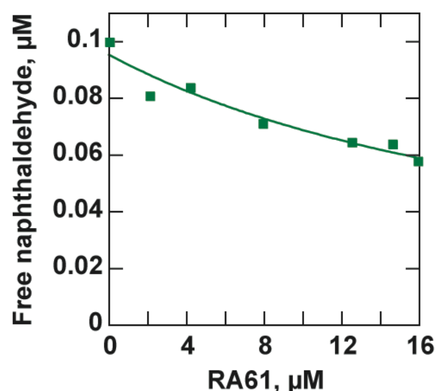


Figure S3. Binding of 6-methoxy-2-naphthaldehyde to RA61. Final apparent free naphthaldehyde concentration was plotted against enzyme concentration, and fit using a binding equation to give K_d = 26 μM. The initial concentration of naphthaldehyde was 0.1 μM.

The loss of fluorescence of 6-methoxy-2-naphthaldehyde in the presence of enzyme has implications for following reaction progress curves. As the product builds up in the reaction, it increasingly binds to free enzyme. Binding of naphthaldehyde to enzyme can reduce the apparent fluorescence signal as well as the amount of free enzyme available to react. Because the rate of fluorescence production is slowed as more product is formed, the loss of fluorescence can result in burst-like behavior during reaction time courses.

This phenomenon is pronounced in RA34, for which a rapid and extensive loss of fluorescence is observed when enzyme is incubated with 6-methoxy-2-naphthaldehyde (Figure S2A). As shown in Figure S4, RA34 also displays burst-like behavior in fluorescence time courses, in which the degree of curvature increases with increasing enzyme concentration.

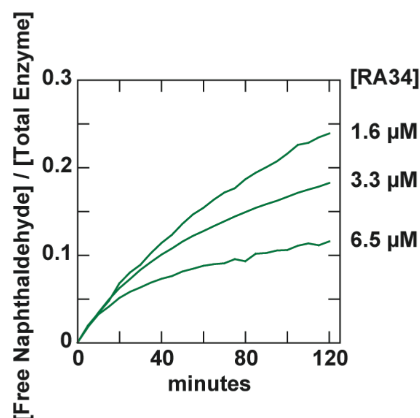


Figure S4. Reaction progress curves for RA34. Substrate (270 μM final concentration) was mixed with the indicated enzyme concentrations under standard conditions (see main text). The fluorescence signal was converted to free naphthaldehyde concentration with a standard curve, and the assumption was made (as discussed above) that the fluorescence of the enzyme-naphthaldehyde bound species is zero.

With RA61, only minor curvature is observed over several turnovers at low enzyme concentration (Figure S5), consistent with less pronounced rebinding of naphthaldehyde as seen in Figure S2. As with RA34, the degree of curvature is greater with higher enzyme concentration.

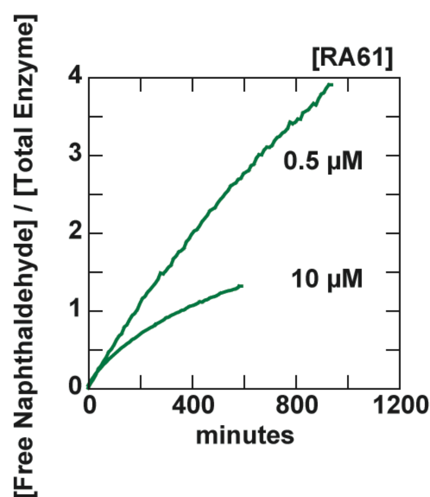


Figure S5. Reaction progress curves over multiple turnovers for RA61. Substrate (500 μM final concentration) was mixed with the indicated enzyme concentrations under standard conditions (see main text). The fluorescence signal was converted to free naphthaldehyde concentration with a standard curve, and the assumption was made (as discussed above) that the fluorescence of the enzyme-naphthaldehde bound species is zero.

Assignment of pK_a values to the catalytic lysine

As described in the main text, pH-rate profiles were determined for RA34, RA45, RA60, and RA61. Figure S6 shows the profiles for RA34, RA45, and RA60, while that of RA61 was included in the main text.

pH-rate profiles obtained under subsaturating conditions reflect ionizations of both the free enzyme and free substrate. Thus, although the pH-dependencies of multistep reactions can be complicated given the possibility of differential pH effects for different steps, under subsaturating conditions, these complications simplify to a consideration of the reactive forms of free enzyme and substrate and the rate-limiting transition state: ($v_o = k_{cat}/K_M$ [E_{active}][S_{active}]) (References S1-S3). Because the substrate 4-hydroxy-4-(6-methoxy-2-naphthyl)-2-butanone does not have titratable groups within the range of protein stability, observed pK_a values are expected to arise from the free enzyme. In particular, because the deprotonated, uncharged form of the lysine side chain is required for reaction, the reaction rate has a first-order dependence on the concentration of deprotonated amine under subsaturating conditions.

Because the deprotonated form of lysine is also required in the model for naphthaldehyde fluorescence loss shown in Figure S1, we tested the simple model that the observed pK_a in the retroaldol reaction with RA61 represents a simple titration of the enzyme by determining the pH-dependence of 6-methoxy-2-naphthaldehyde binding to RA61 (Figure S7). The resulting pK_a value of 6.9 for binding is within error of the value of 6.8 obtained from the retroaldol reaction pH-rate profile. These results suggest most simply that the pK_a observed in the pH-rate profile for RA61 arises from titration of the active site lysine.

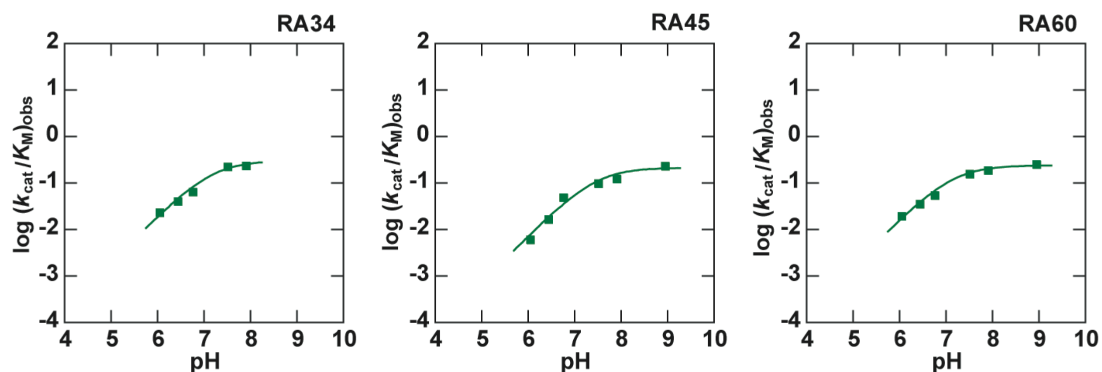


Figure S6. pH-rate profiles for RA34, RA45, and RA60. The profile for RA61 is included in the main text. Fitting the curve (see Materials and Methods) gave pK_a values of 7.5 for RA45 and 7.2 for RA60. Errors are estimated to be about ± 0.2 units.

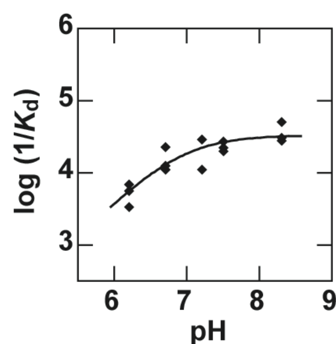


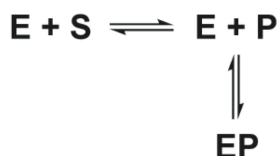
Figure S7. pH-binding profile for RA61 with 6-methoxy-2-naphthaldehyde. Note that the K_d values determined are subject to the uncertainties described in the prior section, *Binding of naphthaldehyde product to the enzyme*. The fit gives a pK_a value of 6.9, well within error of the pK_a value obtained from the pH-rate profile for RA61 of 6.8. The error is estimated to be about ±0.5 units.

Estimation of enantioselectivity of RA61

The enantioselectivity of RA61 was estimated by following a reaction time course to about 40% of completion. If the enzyme reacts with significant selectivity for one enantiomer over the other, a measurable slowing in the time course should be evident as the favored enantiomer becomes fully depleted at 50% of overall completion.

The retroaldol reaction of 4-hydroxy-4-(6-methoxy-2-naphthyl)-2-butanone was followed at low substrate concentration to avoid large effects on fluorescence from the naphthaldehyde product rebinding. Aliquots were taken from the reaction mixture and assayed separately to confirm full activity of the enzyme after the reaction time course.

The reaction time course was compared with simulated curves calculated using different degrees of enantioselectivity. Simulated curves were produced using KinTek Explorer (Reference S4) with the following simple model for subsatrating kinetics and product rebinding:



Enantiomeric preference was modeled by assuming equal concentrations of two reactants with 3-fold, 6-fold, and 10-fold differences in rate constant.

As shown in Figure S8A, when the measured value of K_d = 26 μM for product rebinding is used, the reaction progress curve fits best to a simulated progress curve with no enantioselectivity. The low degree of curvature suggests a preference of less than 3-fold for one enantiomer over another. If lower K_d values are used, greater curvature results and the reaction progress curve becomes even more inconsistent with the extent of curvature that would be seen if the enzyme were enantioselective.

To evaluate whether the degree of estimated enantioselectivity is dependent on the model for product rebinding and our measured K_d value, we modeled curves in which product rebinding is completely neglected (Figure S8B). When the loss of naphthaldehyde fluorescence in the presence of enzyme is not accounted for at all, the reaction progress curve still remains consistent with low enantioselectivity. The observed curvature in the reaction progress curve fits well to simulated curves with between 3-fold and 6-fold enantioselectivity, but does not fit well to a simulated curve with 10-fold enantioselectivity. This observation suggests that the enzyme does not exhibit enantioselectivity of more than 10-fold, regardless of the kinetic model used for product rebinding to the enzyme.

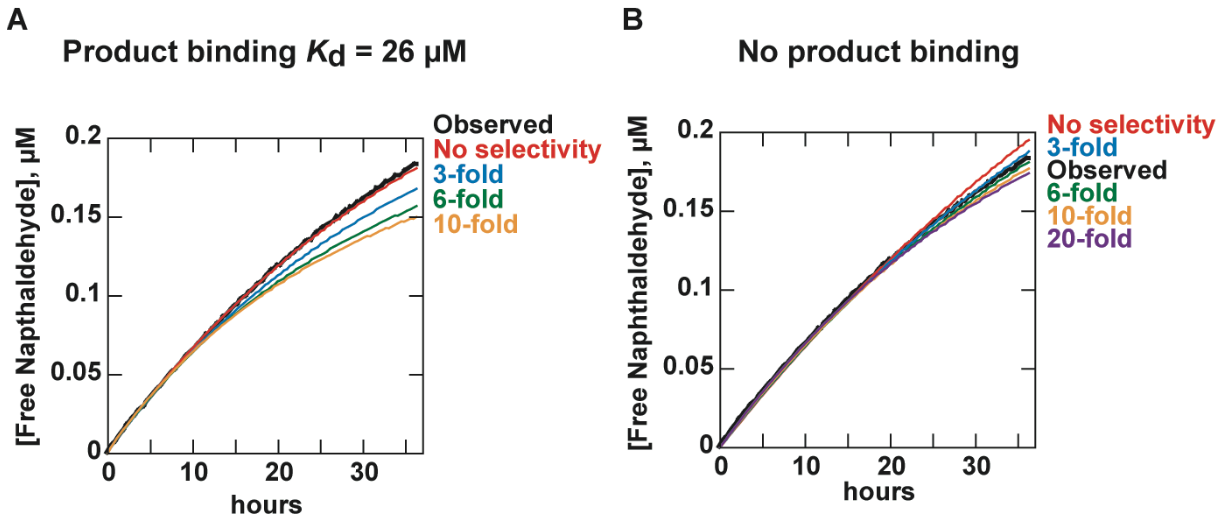


Figure S8. Observed reaction progress curve for RA61 (black points) compared to simulated curves for varying levels of preference for one enantiomer over another. Enzyme concentration was 10 μM and total substrate concentration was 0.5 μM. Free naphthaldehyde concentration was calculated using a standard curve for fluorescent signal and assuming, as described earlier, a zero fluorescence for the enzyme-product species. Observed data is indicated in black curves, while simulated curves are colored as labeled. **A.** Curves simulated using the simple model for subsaturating kinetics and product rebinding with a K_d value of 26 μM. **B.** Curves simulated with a simple model for subsaturating kinetics and completely neglecting the effect of fluorescence loss in the presence of enzyme.

Catalytic contributions of shifted lysine pK_a calculated using different Brønsted values

Catalytic contributions of the shifted lysine pK_a were calculated with different Brønsted slope values (Table S1). For each enzyme, the difference between the measured pK_a and the pK_a of the lysine side chain, 10.6, was used to calculate the shift in maximal rate constant as follows:

$$\Delta(\log(k_{cat}/K_M)_{\max}) = \beta(\Delta pK_a)$$

Given the values for maximal rate constant for the lowered pK_a values, observed second order rate constants at pH 7.5 were then calculated using this expression to account for the quantity of deprotonated amine at pH 7.5:

$$(k_{\text{cat}}/K_{\text{M}})_{\text{obs}} = \frac{(k_{\text{cat}}/K_{\text{M}})_{\text{max}}}{1+10^{\text{p}K_{\text{a}}-\text{pH}}}$$

The ratio of observed rate constants at pH 7.5 for an amine of pK_a 10.6 and an amine of the measured enzymatic pK_a value was taken as the catalytic contribution of the pK_a shift.

Table S1. Values for the catalytic contribution of lysine pK_a shifts at pH 7.5, calculated using different Bronsted values.

Retroaldolase	Fold Effect of pK_a shift using different Brønsted values				
	0.2	**0.4**	**0.54**	**0.6**	**0.8**
RA45	151	36	**14**	9	2
RA60	175	37	**13**	8	2
RA61	182	32	**10**	6	1

References

S1. Cook PF, Cleland WW (2007) Enzyme kinetics and mechanism. (Garland Science, New York)

S2. Fersht A (1999) Structure and mechanism in protein science (WH Freeman and Company, New York)

S3. Knowles JR (1976) Intrinsic pK_a values of functional groups in enzymes: Improper deductions from pH-dependence of steady-state parameters. *CRC Critical Reviews in Biochemistry* 4:165-173.

S4. Johnson KA, Simpson ZB, Blom T (2009) Global Kinetic Explorer: A new computer program for dynamic simulation and fitting of kinetic data. *Analytical Biochemistry* 387:30-41.