



Supporting Information

© Wiley-VCH 2006

69451 Weinheim, Germany

## Stereoselectivity and expanded substrate scope of an engineered PLP-dependent aldolase

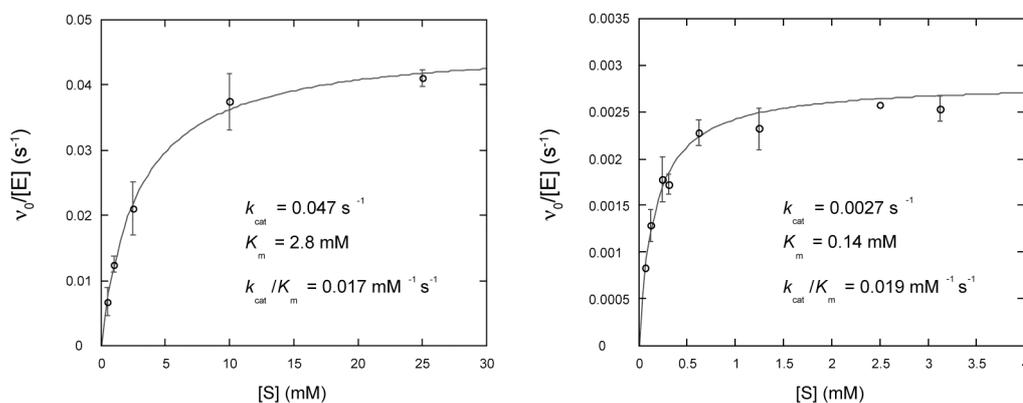
Florian P. Seebeck<sup>a</sup>, Angelo Guainazzi<sup>a</sup>, Celine Amoreira<sup>b</sup>, Kim K. Baldrige<sup>b</sup> and Donald Hilvert<sup>a\*</sup>

<sup>a</sup>Laboratorium für Organische Chemie, ETH Zürich, Hönggerberg HCI F337, CH-8093, Zürich, Switzerland <sup>b</sup> Organisch-Chemisches Institut, Universität Zürich, Winterthurerstrasse 190, CH-8057 Zürich, Switzerland

**Materials.** Chemicals were purchased from Sigma, Acros, Merck, Aldrich or Fluka unless noted otherwise and used without further purification. (2*R*,3*S*)- $\alpha$ -Methyl- $\beta$ -phenylserine and (2*R*,3*R*)- $\alpha$ -methyl- $\beta$ -phenylserine were synthesized following published procedures.<sup>[1-3]</sup>

**Production and purification of recombinant alrY265A.** Cloning of alrY265A and vector construction were described previously.<sup>[4]</sup> Calcium competent *Escherichia coli* cells (BL21) were transformed with pET22b-alrY265A.<sup>[4]</sup> Recombinant protein production was induced with 0.5 mM IPTG in cultures grown at 37 °C in LB medium. Cell lysis and Ni<sup>2+</sup>-ion affinity chromatography were performed following standard protocols (Quiagen). The resulting protein preparations were further purified by preparative gel filtration using a Superdex 75 (26/60) FPLC column from Pharmacia. Purity was assessed by SDS-PAGE. Protein concentration was determined by measuring the absorbance at 280 nm using the calculated molar extinction coefficient  $\epsilon_{280} = 46,370 \text{ M}^{-1} \text{ cm}^{-1}$ . The protein solutions (5 – 15 mg/ml alrY265A, 50 mM Tris-HCl, pH 8.0, 150 mM NaCl) were sterile filtered and stored at 4 °C.

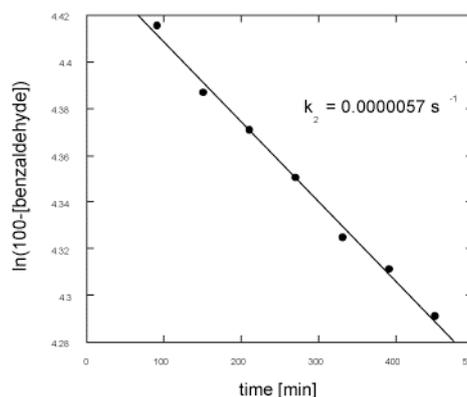
**Kinetic Assays.** All kinetic measurements were performed in 100 mM Hepes buffer, pH 8.0, at 30 °C. The assay mixture contained 50  $\mu\text{M}$  pyridoxal 5'-phosphate. Initial rates were determined by monitoring the appearance of benzaldehyde spectrophotometrically at 279 nm ( $\epsilon_{279 \text{ nm}} = 1,400 \text{ M}^{-1} \text{ cm}^{-1}$ ) using a Lambda series UV-Visible spectrophotometer (Perkin-Elmer) equipped with a thermoelectric cuvette holder. Measurements were made in quartz cuvettes with a 1 cm path length. Initial rates were corrected for the background reaction. Kinetic parameters  $k_{\text{cat}}$  and  $K_{\text{m}}$  were calculated from the initial rates using concentrations of  $\alpha$ -methyl- $\beta$ -phenylserine ranging from 0.1 mM to 25 mM (Figure 1). To prevent potential product inhibition by alanine, the reaction mixture was further supplemented with D-amino acid oxidase (0.19 U) and catalase (1175 U) to convert D-alanine into pyruvate, which does not interfere with the assay.



**Figure S1.** alrY265A assays with (2*R*,3*S*)- $\alpha$ -methyl- $\beta$ -phenylserine (left) or (2*R*,3*R*)- $\alpha$ -methyl- $\beta$ -phenylserine (right).

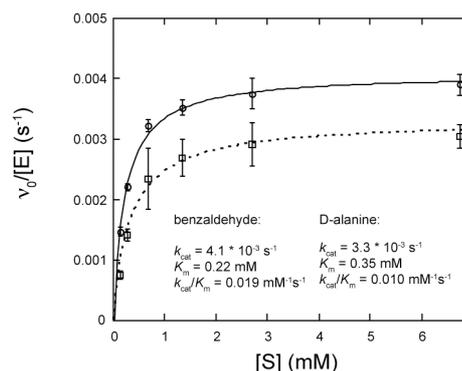
**Determination of background rate.** The rate constant of the pyridoxal-dependent retroaldol reaction of (2*R*,3*S*)- $\alpha$ -methyl- $\beta$ -phenylserine in the absence of enzyme was determined as previously described.<sup>[4, 5]</sup> The concentration of (2*R*,3*S*)- $\alpha$ -methyl- $\beta$ -phenylserine and pyridoxal hydrochloride was 0.1 M; the pH was adjusted to  $8.0 \pm 0.1$ ; the temperature of the reaction was maintained at  $30 \pm 0.1$  °C. Instead of D<sub>2</sub>O, H<sub>2</sub>O/D<sub>2</sub>O (9/1) was used as solvent. <sup>1</sup>H NMR spectra were obtained with a Bruker 500 MHz NMR spectrometer. The reaction was followed for 8 h collecting 16 scans every 60 min. The appearance of benzaldehyde was monitored at  $\delta$  9.93 (1H, s, OCHPh) and  $\delta$  = 7.92 PPM (2H, d, J = 0.017, Ph). The two data sets were pooled to determine the product concentration, and  $k_2$  was determined by plotting  $\ln(100 - [\text{benzaldehyde}])$  vs. time (Figure. 2). The resulting rate constant ( $k_2 = 1.0 \times 10^{-6} \text{ s}^{-1}$ ) is roughly 20-fold slower than in the case of  $\beta$ -phenylserine ( $k_2 = 2.3 \times 10^{-5} \text{ s}^{-1}$ ).<sup>[4]</sup>

**Figure S2.** Pyridoxal catalyzed retroaldol cleavage of (2*R*,3*S*)- $\alpha$ -methyl- $\beta$ -phenylserine.



**AlrY265A catalyzed conversion of  $\alpha$ -methyl- $\beta$ -phenylserine to D-alanine.** To determine if reprotonation of the alrY265A-bound carbanionic intermediate proceeds stereospecifically, cleavage of (2*R*,3*R*)- $\alpha$ -methyl- $\beta$ -phenylserine was monitored by D-alanine dependent oxidation of NADH using D-amino acid oxidase and lactate dehydrogenase as coupling enzymes. Kinetic measurements were performed at 30 °C in 100 mM Hepes buffer, pH 8.0, with 100  $\mu$ M NADH ( $\epsilon_{340\text{nm}}$ :  $6.2 \times 10^3 \text{ M}^{-1}\text{cm}^{-1}$ ) and 40 units/ml lactate dehydrogenase. The resulting kinetic data were subsequently correlated with the production of benzaldehyde under identical conditions (Figure 3).

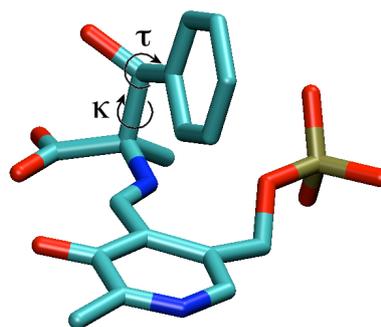
**Figure S3.** Stereochemistry of quinoid reprotonation. Kinetic assay with alrY265A and (2*R*,3*R*)- $\alpha$ -methyl- $\beta$ -phenylserine monitoring either the production of benzaldehyde (solid line) or the production of D-alanine (dashed line).



**Computational Methods.** The initial starting structure was obtained from the protein data bank (1L6F) and modified by mutating Tyr265 to alanine. Hydrogens were added to both the protein and the ligand using Babel,<sup>[6]</sup> and the PDB2PQR<sup>[7]</sup> tool used to prepare the protein (heavy atom position checks, assignment of PARSE<sup>[8]</sup> atomic charges and radii). Hybrid docking and *ab initio* methods were used to determine optimal fit of the aldimine in the binding site. *Ab initio* calculations were performed using GAMESS<sup>[9]</sup> at the B3LYP<sup>[10]</sup>/DZV+(2d,p)<sup>[11]</sup> level of theory ( $\epsilon = 4$ ), with CHELPG charges<sup>[12]</sup> and standard radii computed for subsequent binding energy studies. Initial docking was performed with the aid of QMView tools,<sup>[12]</sup> and grids were generated by varying the torsional angles corresponding the C $\alpha$ -

C $\beta$  and C $\beta$ -C $\gamma$  bonds of the intermediates ( $\kappa$  and  $\tau$ , respectively; Fig 4) in 30° increments. The aldimines of all four  $\alpha$ -methyl- $\beta$ -phenylserine diastereomers bound at the enzyme active site revealed that, while both 2R diastereomers are accommodated, neither the 2S,3R nor 2S,3S isomer fits without clashes with active site residues. Electrostatic binding energies were computed for each pose in the grid of 2R,3R and 2R,3S structures using APBS (Adaptive Poisson Boltzmann Solver),<sup>[14]</sup> providing an energy map for each. Parameter specifications for APBS were set as follows: counter ion concentration = 0.03 M, dielectric constant  $\epsilon = 2$  for the protein and  $\epsilon = 78.4$  for the surrounding water, grid spacing = 0.3 Å, solvent radius = 1.4 Å, T = 298.14 K. Molecular dynamics (MD) was performed for 1.2 ns on the lowest energy structures predicted from the hybrid *ab initio* APBS computations, using the following specifications: generalized AMBER force field<sup>[15]</sup> for the ligand and the parm99 force field<sup>[16]</sup> for protein atoms, generalized Born solvation model,<sup>[17]</sup> Langevin dynamics<sup>[18]</sup> with collision frequency 5.0 ps<sup>-1</sup>, temperature of 300 K, and a non-bonded energy cutoff of 16.0 Å. The key torsion angles and the ligand backbone were constrained using a force constant of 10.0 kcal/mol/rad<sup>2</sup>, using a soft-square function.<sup>[15]</sup> The resulting trajectories were analyzed and the representative 2R,3S and 2R,3R complex structures selected using the cluster analysis technique.<sup>[19]</sup> Energy grid maps were constructed for these structures using the hybrid *ab initio* APBS method to predict accessible regions within the active site for the 2R,3S and 2R,3R aldimines. For each isomer, the general areas of accessible conformational flexibility were determined. For the 2R,3S isomer, this corresponds to  $\kappa$  values of  $-138.8^\circ \pm 30^\circ$ , with the energetically most favorable structure in the middle of this range. Structures with the phenyl ring plane parallel to that of the aldimine provide the lowest energy form within each region, since  $\tau$  torsional angle movements towards the methyl group or towards the carboxylate group increase the energy. For the 2R,3R isomer, the same active site regions are accessible, but the rotamer with  $\kappa = -108.8^\circ$ , which places the aryl ring closer to His166, is predicted to have the lowest energy in this case. The most favorable conformation for each isomer places the  $\beta$ -hydroxyl group of the substrate within hydrogen bonding distance of the phosphate oxygen. In each case, molecular orbitals calculated by QM reveal that the orbital associated with this hydrogen bond is fairly deep (HOMO-12, 2R3S and HOMO-12, 2R3R). All complex surface visualizations necessary for analysis were performed with VMD.<sup>[20]</sup>

**Figure S4.** A model of the PLP-(2R,3S)- $\alpha$ -methyl- $\beta$ -phenylserinyl aldimine. The two torsional angles  $\tau$  and  $\kappa$  indicated were varied in 30° increments. Position as shown corresponds to  $\kappa=0$ ,  $\tau=0$ .



His166 can potentially interact with the bound aldimine intermediate. Its preferred tautomeric form reflects a favorable hydrogen-bonding interaction between N $\delta$  of the imidazole ring and Arg219 (3.20 Å) that is present in native alr.<sup>[21]</sup> The alternative tautomer, in which N $\delta$  rather than N $\epsilon$  bears the hydrogen, was examined briefly to assess whether potentially more favorable interactions with the substrate might arise. In the calculations, this was not the case. With both aldimine diastereomers, the favorable hydrogen bond between the substrate alcohol and the phosphate group of the cofactor was maintained in preference to an interaction with His166.

## References

- [1] U. Schöllkopf, *Tetrahedron* **1983**, *39*, 2085.
- [2] U. Schöllkopf, U. Groth, W. Hartwig, *Justus Liebigs Ann. Chem.* **1981**, 2407.
- [3] U. Schöllkopf, U. Groth, K.-O. Westphalen, C.Z. Deng, *Synthesis* **1981**, 969.
- [4] F. P. Seebeck, D. Hilvert, *J. Am. Chem. Soc.* **2003**, *125*, 10158.
- [5] K. Tatsumoto, A. E. Martell, *J. Am. Chem. Soc.* **1978**, *100*, 5549.
- [6] <http://openbabel.sourceforge.net>
- [7] T. J. Dolinsky, J. E. Nielsen, J. A. McCammon, N. A. Baker, *Nucleic Acids Res.* **2004**, *32*, W665.
- [8] D. Sitkoff, K. A. Sharp, B. Honig, *Biophys. Chem.* **1994**, *51*, 397.
- [9] M. W. Schmidt, K. K. Baldrige, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis, J. A. Montgomery, *J. Comp. Chem.* **1993**, *14*, 1347.
- [10] C.T. Lee, W.T. Yang, and R. G. Parr, *Phys. Rev.* **1988**, *B37*, 785.
- [11] T.H.Dunning, Jr., P.J.Hay, In *Methods of Electronic Structure Theory*, H.F.Shafer III, Ed. Plenum Press, N.Y. 1977, pp 1-27.
- [12] C. M. Breneman, K. B. Wiberg, *J. Comp. Chem.* **1990**, *11*, 361.
- [13] K. K. Baldrige, J. P. Greenberg, *J. Molec. Graph.* **1995**, *13*, 63.
- [14] N. A. Baker, D. Sept, M. J. Holst, J. A. McCammon, *IBM J. Res. Dev.* **2001**, *45*, 427.
- [15] D.A. Case, T.E. Cheatham, III, T. Darden, H. Gohlke, R. Luo, K.M. Merz, Jr., A. Onufriev, C. Simmerling, B. Wang and R. Woods, *J. Comp. Chem.* **2005**, *26*, 1668.
- [16] J.M. Wang, P. Cieplak, P.A. Kollman, *J. Comp. Chem.* **2000**, *21*, 1049.
- [17] M. Born, *Z. Physik* **1920**, *1*, 45.
- [18] R.M. Levy, M. Karplus, J.A. McCammon, *Chem. Phys. Lett.* **1979**, *65*, 4.
- [19] A.W.F. Edwards, L.L. Cavalli-Sforza, *Biometrics* **1965**, *21*, 362.
- [20] W. Humphrey, A. Dalke, K. Schulten, *J. Molec. Graphics* **1996**, *14*, 33.
- [21] J.P. Shaw, G.A. Petsko, D. Ringe, *Biochemistry* **1997**, *36*, 1329.