Structure-based rationalization of aldolase-catalyzed asymmetric synthesis

Junjie Liu, Grace DeSantis, and Chi-Huey Wong

Abstract: This paper describes a structure-based approach to elucidate the stereospecificity, including inversion of enantioselectivity, of the 2-deoxyribose-5-phosphate aldolase-catalyzed asymmetric aldol addition reaction using unnatural substrates designed for the total synthesis of epothilones. In addition, an aldolase variant with Ser-238 being altered for Asp was found to be 2.5 times more effective than the wild type in accepting the unphosphorylated substrate D-glyceraldehyde. A new H-bonding interaction between the Asp-238 carboxylate and the 3-hydroxyl of the substrate was identified and was used to rationalize the rate enhancements.

Key words: aldol reaction, 2-deoxyribose-5-phosphate aldolase, mutagenesis, inversion of enantioselectivity.

Résumé : Cette communication décrit une approche basée sur la structure d'élucider la stéréospécificité, y compris l'inversion d'énantiosélectivité, de la réaction d'addition aldolique asymétrique catalysée par l'aldolase de 2-désoxyribose-5-phosphate; cette étude fait appel à des substrats qui ne sont pas naturels et qui ont été développés pour la synthèse totale des épothilones. De plus, on a observé qu'une variante de l'aldolase, portant une Asp au lieu de la Ser-238, est 2,5 fois plus efficace que la forme naturelle pour accepter le substrat d-glycéraldéhyde non phosphorylé. On a identifié une nouvelle interaction par liaison hydrogène entre le carboxylate de l'Asp-238 et l'hydroxyle en position 3 du substrat; elle permet de rationaliser les augmentations de vitesses de réaction.

Mots clés : réaction aldolique, aldolase de 2-désoxyribose-5-phosphate, mutagénèse, inversion d'énantiosélectivité.

[Traduit par la Rédaction]

Introduction

Over the past 20 years, enzymes have been exploited as catalysts in large-scale organic synthesis. Recent advances in molecular genetics, protein engineering, and site-specific modification of enzymes have further expanded the scope of enzyme catalysis with regard to its synthetic application (1). Most of these new developments are based on the inspiring work of a few pioneers in the field, particularly that of Professor Bryan Jones (2). To date, many useful transformations based on enzymes have been developed on large scales for the chemical and pharmaceutical industries, and new developments based on biocatalysis continue to evolve (1, 3).

One subject of current interest to us is the use of enzyme-structure information to design new protein catalysts or new substrates for enzymatic reactions with altered stereoselectivity. The enzyme employed in this study is 2-deoxyribose-5-phosphate aldolase (DERA, EC 4.1.2.4), which catalyzes the aldol reaction between acetaldehyde and D-glyceraldehyde 3-phosphate to form D-2-deoxyribose-5-phosphate (4). The structure of this enzyme has been solved at the atomic level (~1 Å resolution) and the reaction mechanism has been studied in detail (Scheme 1) (5). The enzyme forms a Schiff base with acetaldehyde to generate an enamine intermediate, which then attacks the *Re*-face carbonyl of the incoming acceptor to form a new C–C bond with a new stereogenic center in the *R*-configuration. Both the Schiff base and the carbinol amine intermediates have been trapped and the structures have been determined (5). According to the structure, the C-3 phosphate of the acceptor substrate has H-bonding interactions with a hydrophilic site composed of Ser-238, Ser-239, and Gly-205, and the C-2 hydroxyl group has H-bonding interactions (5) through a water molecule with another hydrophilic site composed of Thr-170 and Lys-172. The C-2 hydrogen is positioned in a hydrophobic pocket (Fig. 1). The enzyme also accepts the corresponding unphosphorylated substrate and related structures with a much weaker affinity and activity (4).

Results and discussion

In an effort to find new substrates, we found that the enzyme catalyzes the aldol reaction with racemic 2-methyl-3-hydroxypropanal **1a** to give the pyranose **2a** (Scheme 2) (6). Characterization of the product by oxidative conversion to the lactone with Br_2 indicates that the enzyme exhibits a re-

Received 17 December 2001. Published on the NRC Research Press Web site at http://canjchem.nrc.ca on 29 May 2002.

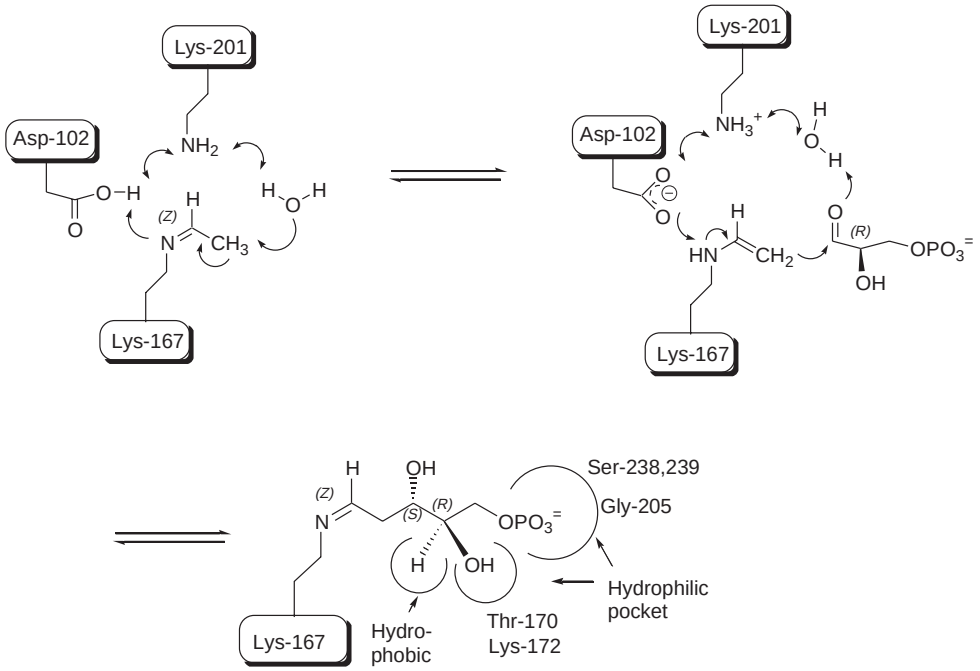
Dedicated to Professor J. Bryan Jones on the occasion of his 65th birthday.

J. Liu, G. DeSantis, and C.-H. Wong.[1] Department of Chemistry and the Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 N. Torrey Pines Road, La Jolla, CA 92037, USA.

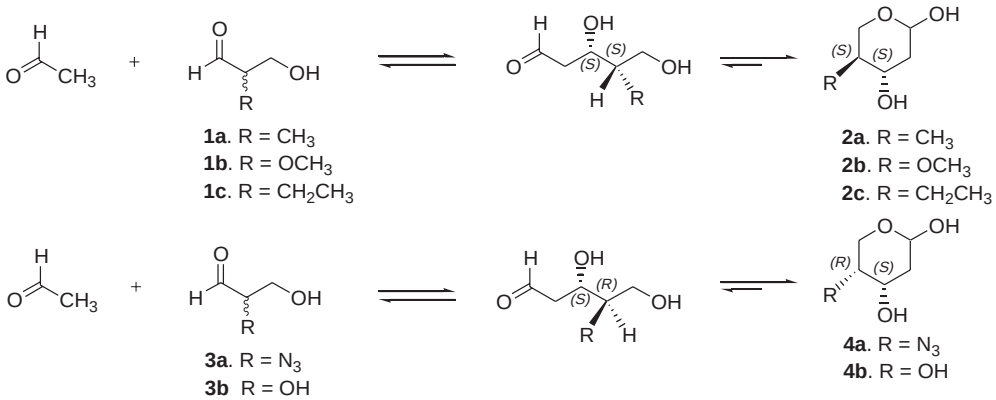
[1]Corresponding author (e-mail: wong@scripps.edu).

 DOI: 10.1139/V02-094

Scheme 1. Highlights of the mechanism of DERA catalyzed aldol addition reaction. Shown are those steps associated with C–H cleavage and C–C bond formation. Asp-102 and Lys-201 act as a proton relay system to assist the proton abstraction and carbonyl protonation mediated by the conserved water molecule. The enzyme side chains interacting with the phosphate group and the C-4 hydroxyl (corresponding to the C-2 OH of acceptor) have been identified and designated each as a hydrophilic pocket. The C-4 hydrogen is located in a small hydrophobic pocket without interaction with side chains.



Scheme 2. Enantioselective aldol reaction catalyzed by DERA. A reversal of enantioselectivity was observed with acceptor substrates with small hydrophobic substituents at C-2 as indicated in the conversions of **1a–c** to **2a–c**, respectively.



versal of enantioselectivity in the aldol reaction, accepting the L-aldehyde as an acceptor. Similarly, reverse enantioselectivity was observed with the 2-methoxy and 2-ethyl derivatives **1b** and **1c**, respectively. Alternatively, when the racemic 2-azido or 2-hydroxy derivative was used (**3a** and **3b**, respectively), enantioselectivity that followed that of the natural reaction was observed, i.e., the D-enantiomer was converted to the pyranose product (**4a** and **4b**). In all cases, no change in facial selectivity was observed (6). The equilibrium of the reaction is driven by the formation of the pyranose ring in favor of condensation (Fig. 2). With the enzyme structure now available, the results can be rationalized such that the 2-methyl, 2-methoxy, or 2-ethyl groups is bound to the hydrophobic pocket and the C-2 hydrogen is in the hydrophilic pocket. If the methyl, methoxy, or ethyl group was in the hydrophilic pocket, it would have a close contact with the water molecule and the carbonyl oxygen of Thr-170, resulting in a less favorable binding. Since the azide or the hydroxy group is considered to be hydrophilic, it fits into the hydrophilic pocket (Scheme 3).

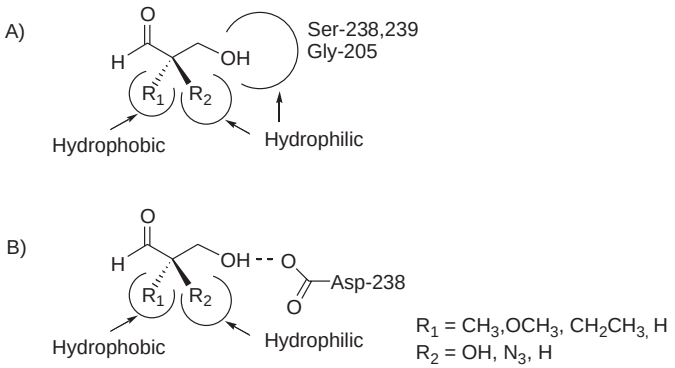
In another effort to engineer the enzyme to accept unphosphorylated neutral substrates more effectively, Ser-238 was exchanged for Asp. It was anticipated that this change would still retain the hydrophilic nature of the C-3

Table 1. Kinetic parameters of wild-type DERA and the Ser238Asp variant in the retroaldol reaction.

Enzyme[a]	Substrate[b]	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m (s^{-1} mM^{-1})
WT	DERP	68 ± 1	0.64 ± 0.01	106 ± 2
	DER	0.107 ± 0.005	57 ± 7	0.0020 ± 0.0003
Ser238Asp	DERP	0.58 ± 0.05	61 ± 11	0.01 ± 0.001
	DER	0.21 ± 0.01	39 ± 6	0.005 ± 0.0009

Note: Enzymes and kinetic parameters were determined according to the procedures previously described (5)
[a]WT, wild-type DERA.
[b]DERP, D-2-deoxyribose-5-phosphate; DER, D-2-deoxyribose.

Scheme 3. (*a*) Acceptor model of wild-type DERA; (*b*) acceptor model of the Ser238Asp variant.



binding pocket, but substrates with a negative charge in the C-3 position would be disfavored. Indeed, the Ser238Asp-mutant enzyme is 2.5 times more active (based on the k_{cat}/K_m value of the retroaldol reaction) than the wild-type to D-glyceraldehyde as a substrate, as indicated in Table 1. Molecular modeling indicates that the C-3 hydroxy hydrogen forms a 2.9–3.2 Å hydrogen bond with the Asp-238 carboxylate, accounting for the increase in reactivity (Fig. 3).

In summary, this study has demonstrated that with a well-defined enzyme structure, one can effectively alter the substrate specificity of an enzyme by protein engineering or substrate design. Methyl pyranose **2a** has been used as a key synthon in the total synthesis of the anti-cancer agents epothilones (6, 7).[2] Work is in progress to further utilize the structural information to design new substrates for DERA-catalyzed aldol addition reactions.

Acknowledgment

We thank the National Institutes of Research for support of this research.

Reference

1. For a recent special issue on this subject, see: Nature (London), **409**, 226 (2001).
2. J.B. Jones. Tetrahedron, **42**, 3351 (1986); G. DeSantis and J.B. Jones. Curr. Opin. Biotechnol. **10**, 324 (1999).
3. R.N. Patel. *In* Stereoselective biocatalysis. Dekker, New York. 2000, and refs. cited therein.
4. T.D. Machajewski and C.-H. Wong. Angew. Chem. Int. Ed. **39**, 1352 (2000), and refs. cited therein.
5. A. Heine, G. DeSantis, J.G. Luz, M. Mitchell, C.-H. Wong, and I.A. Wilson. Science (London), **294**, 369 (2001).
6. For details, see: J. Liu and C.-H. Wong. Angew. Chem. Int. Ed. **41**, 1404 (2002).
7. K.M. Koeller and C.-H. Wong. Nature (London), **409**, 232 (2001).

[2] General enzymatic synthesis: To a 100 mL buffer solution (0.1 M KH_2PO_4, pH 7.5) containing 0.1 M acceptor aldehyde and 0.3 M donor acetaldehyde was added 3000 units of DERA. The resulting solution was stirred in the dark for 3–6 days under argon. The reaction was quenched by addition of 2 volumes of acetone. The mixture was then stirred at 0°C for 1 h and centrifuged to remove the precipitated enzyme. The aqueous phase was concentrated in vacuo, and the residue was purified by flash chromatography (silica, 1:2 to 4:1 EtOAc–hexane) to yield 31% of **2c**. Characterized by its lactone form: $[\alpha]_D$ = 30.2 (*c* 0.7, $CHCl_3$). IR (film) (cm^{-1}): 3424.6, 2970.0, 1719.2, 1190.4, 1114.0, 1055.2. ^{1}H NMR (600 MHz, $CDCl_3$) δ: 4.43 (dd, *J* = 4.4, 11.4 Hz, 1H), 3.97 (dd, *J* = 7.9, 11.4 Hz, 1H), 3.96 (m, 1H), 2.85 (dd, *J* = 5.7, 17.1 Hz, 1H), 2.74 (d, *J* = 3.9 Hz, 1H), 2.55 (dd, *J* = 6.2, 17.6 Hz, 1H), 1.79 (m, 1H), 1.64 (m, 1H), 1.35 (m, 1H), 1.01 (t, *J* = 7.1 Hz, 3H). ^{13}C NMR (150 MHz, $CDCl_3$) δ: 171.07, 69.21, 67.81, 42.32, 38.03, 21.63, 11.23. HRMS *m*/*z* calcd. for $C_7H_{12}O_3$ (M^+): 144.0786; found: 167.0678 ([M + Na]$^+$). D-aldehyde product; 3% yield. Characterized by its lactone form: $[\alpha]_D$ = –1.0 (*c* 0.6, $CHCl_3$). ^{1}H NMR (600 MHz, $CDCl_3$) δ: 4.37 (t, *J* = 10.9 Hz, 1H), 4.25 (dd, *J* = 4.9, 11.0 Hz, 1H), 4.22 (br s, 1H), 2.73 (dd of AB, *J* = 3.1, 18.0 Hz, 1H), 2.68 (dd of AB, *J* = 4.4, 18.4 Hz, 1H), 2.37 (br s, 1H), 1.86 (m, 1H), 1.48 (m, 1H), 1.35 (m, 1H), 1.00 (t, *J* = 7.4 Hz, 3H). ^{13}C NMR (150 MHz, $CDCl_3$) δ: 170.30, 68.91, 64.50, 39.28, 39.18, 19.60, 11.25. **4a**; 47% yield. ^{1}H NMR (500 MHz, $CDCl_3$) δ major isomer): 5.14 (dt, *J* = 2.6, 7.7 Hz, 1H), 4.22 (m, 1H), 4.13 (dd, *J* = 10.1, 11.9 Hz, 1H), 3.71 (dd, *J* = 4.8, 11.8 Hz, 1H), 3.58 (ddd, *J* = 2.9, 4.8, 9.9 Hz, 1H), 2.93 (d, *J* = 5.2 Hz, 1H), 2.10 (ddd, *J* = 2.6, 10.2, 19.3Hz, 1H), 1.92 (dt, *J* = 3.3, 16.8 Hz); minor isomer: 5.32 (q, *J* = 3.3 Hz, 1H), 4.24 (m, 1H), 4.10 (dd, *J* = 2.6, 12.5 Hz, 1H), 3.84 (dd, *J* = 4.8, 12.1 Hz, 1H), 3.77 (q, *J* = 3.3 Hz, 1H), 1.98 (ddd, *J* = 3.0, 9.9, 12.8 Hz), 1.88 (dt, *J* = 4.0, 13.2 Hz, 1H). ^{13}C NMR (125 MHz, $CDCl_3$) δ major isomer: 92.28, 67.12, 58.97, 56.83, 35.14; minor isomer: 91.74, 64.92, 61.07, 60.63, 35.33.