Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 10396–10401, September 1998
Biochemistry

Evolution of an enzyme active site: The structure of a new crystal form of muconate lactonizing enzyme compared with mandelate racemase and enolase

MIRIAM SARAH HASSON*†, ILME SCHLICHTING*‡, JAVAD MOULAI*, KIRK TAYLOR§, WILLIAM BARRETT§, GEORGE L. KENYON¶, PATRICIA C. BABBITT¶, JOHN A. GERLT§‖, GREGORY A. PETSKO*,** AND DAGMAR RINGE*

*Departments of Biochemistry and Chemistry and the Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, MA 02254-9110; †Department of Biological Sciences, Purdue University, West Lafayette, IN 47907-1392; §Department of Chemistry and Biochemistry, University of Maryland, College Park, MD 20742; ¶Department of Pharmaceutical Chemistry, University of California, San Francisco, CA 94143-0446; and ‖Department of Biochemistry, University of Illinois, Urbana, IL 61801

Contributed by Gregory A. Petsko, June 25, 1998

ABSTRACT **Muconate lactonizing enzyme (MLE), a component of the β-ketoadipate pathway of *Pseudomonas putida*, is a member of a family of related enzymes (the "enolase superfamily") that catalyze the abstraction of the α-proton of a carboxylic acid in the context of different overall reactions. New untwinned crystal forms of MLE were obtained, one of which diffracts to better than 2.0-Å resolution. The packing of the octameric enzyme in this crystal form is unusual, because the asymmetric unit contains three subunits. The structure of MLE presented here contains no bound metal ion, but is very similar to a recently determined Mn^{2+}-bound structure. Thus, absence of the metal ion does not perturb the structure of the active site. The structures of enolase, mandelate racemase, and MLE were superimposed. A comparison of metal ligands suggests that enolase may retain some characteristics of the ancestor of this enzyme family. Comparison of other residues involved in catalysis indicates two unusual patterns of conservation: (*i*) that the position of catalytic atoms remains constant, although the residues that contain them are located at different points in the protein fold; and (*ii*) that the positions of catalytic residues in the protein scaffold are conserved, whereas their identities and roles in catalysis vary.**

The avalanche of newly available protein sequences can be sorted into groups of related proteins and analyzed in terms of family relationships. The analysis of a family of enzymes provides more information than a combination of separate studies of each member. A comprehensive study allows us to make predictions about some family members based on the properties of others. Comparisons of known enzyme structures can give us an "enzyme's eye view" of evolution: their similarities show us what is important to an evolving enzyme, whereas their differences give us an appreciation of the flexibility of an active site. Thus, the study of the evolution of enzyme active sites can illuminate the structural basis of catalysis in general. If we can identify those characteristics that are the most conserved as enzymes diverge, we will often have identified the factors that are most critical to catalytic activity. Crucial to any study that hopes to address these points are: (*i*) the identification of a group of enzymes that are unarguably homologous, but have diverged enough so that their similarities are likely to be indicative of important conserved characteristics; and (*ii*) high-quality three-dimensional structures of members of the family, to allow detailed comparisons of their active sites.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

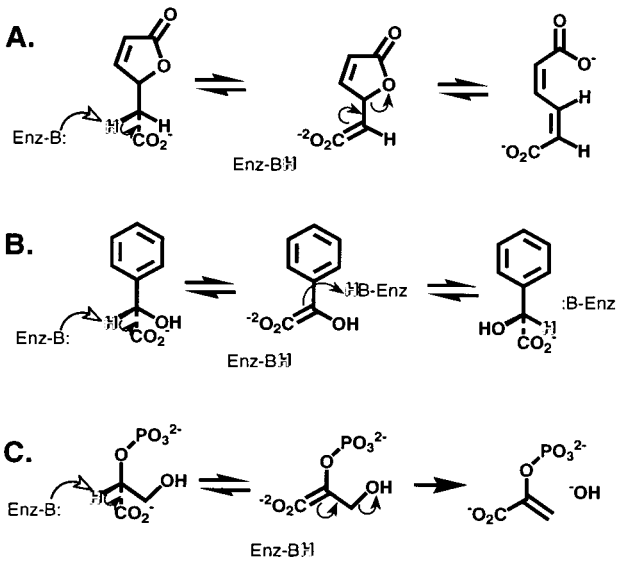

FIG. 1. Comparison of the mechanisms of the reactions catalyzed by MLE (*A*), MR (*B*), and enolase (*C*). The common step of proton abstraction is indicated by an open arrow. The abstracted proton is outlined.

The two enzymes mandelate racemase (MR) and muconate lactonizing enzyme (MLE) were a promising starting point for a group of highly diverged homologous enzymes (1). Their primary sequences, approximately 25% identical, are related but significantly different; whereas their three-dimensional structures are similar. The enzyme enolase has a more distant, but nevertheless clear, relationship to MLE and MR (2–4). Other enzymes in the "enolase superfamily" have been identified on the basis of sequence similarity (2, 5). The reactions catalyzed by these enzymes share the core chemical step of an abstraction of a proton from a carbon adjacent to a carboxylic acid (Fig. 1) and a requirement of a divalent metal ion. In contrast to many recognized families of enzymes whose members catalyze similar reactions on

Abbreviations: MLE, muconate lactonizing enzyme; MR, mandelate racemase.

Data deposition: The atomic coordinates for this muconate lactonizing enzyme structure have been deposited in the Protein Data Bank, Biology Department, Brookhaven National Laboratory, Upton, NY 11973 (PDB ID code 1BKH).

‡Present address: Max Planck Institut für Molekulare Physiologie, Abteilung Physikalische Biochemie, Rheinland Damm 201, D-44139 Dortmund, Germany.

**To whom reprint requests should be addressed at: Rosenstiel Basic Medical Sciences Research Center, Brandeis University, Waltham, MA, 02254-9110. e-mail: ringe@binah.cc.brandeis.edu.

different substrates, the enolase superfamily includes enzymes catalyzing a wide variety of reactions and performing diverse roles in metabolism.

The structure of MLE was originally solved from a twinned crystal form (6). Because of the twinning, the data from this crystal form were somewhat difficult to interpret, and for many years only a low resolution structure was available. Recently, the structure from this crystal form has been successfully refined to high resolution (7).

In this study the structure of MLE in a new, untwinned, crystal form is described at 2.2-Å resolution. The active sites of MLE, MR and enolase are compared and several new observations on the comparisons are presented. Although ligands to the metal ions are conserved among the three enzymes (4), no other catalytic residue is present in all three enzymes; that is, in no case are both the position and identity of a catalytic residue conserved. In some cases there is conservation of the position of catalytic atoms in the active sites, although the residues containing these atoms come from different points in the protein fold. In other cases residues important in catalysis are found at the same position in all three enzymes, although the identities of the residues are different. The conservation of position, but not the identity of catalytic residues, is an unusual observation, and hints that certain positions in a protein scaffold may have intrinsic qualities that predispose them to hold residues that play a role in catalysis.

EXPERIMENTAL PROCEDURES

Purification of MLE. MLE protein was expressed in the *Escherichia coli* strain JM105 from the plasmid pKMLE52 (kindly provided by L. N. Ornston, Yale University), which contains the *catB* gene (8), encoding MLE, in the pKK223–3 expression system (Pharmacia). The culture was induced in late logarithmic phase for 10–16 hr with 1 mM isopropyl β-D-thiogalactoside. Under these conditions, MLE accounts for approximately 5–15% of the soluble protein in the lysate, and a 1 liter culture yields 10–50 mg of purified protein. MLE was purified as described (9), with the following modifications. Cells were ruptured by passage through a French press. A protamine sulfate (1.25 mg/ml) precipitation step was added before ammonium sulfate precipitation. After DEAE-52 anion-exchange chromatography, the protein was purified on a Sephacryl S200 gel filtration column. MLE protein purified in this manner was >95% pure, as estimated from analysis by SDS/PAGE. Purified protein was stored in 33% ammonium sulfate at 4°C.

Crystallization. Initial conditions for crystallization were screened by using sparse matrix sampling (10). To obtain crystal forms I–IV, MLE was dialyzed overnight against 20 mM Tris·HCl, pH 7.4/0.1 mM $MnCl_2$/7 mM 2-mercaptoethanol/100 mM NaCl, and concentrated to 10 mg/ml by using a Centricon-30 (Amicon). Crystals were grown by hanging-drop vapor diffusion against a well solution of 2 M sodium formate/0.1 M sodium acetate (pH 4.6). Drops contained equal volumes (3 μl) of well solution and MLE, sometimes with the addition of 3 μl of buffer A. Variation in pH, temperature, or $MnCl_2$ concentration did not affect crystallization. Under these conditions, four crystal forms were obtained: (*i*) needle-like crystals, not analyzed further; (*ii*) plate-like crystals that diffracted to no better than 5 Å; (*iii*) star-like or octahedral crystals of space group $P4_22_12_1$, with unit cell dimensions $a = b = 100$ Å and $c = 164$ Å, which diffracted to no better than 5 Å; and (*iv*) rod-like crystals of space group $P2_12_12$, with unit cell dimensions $a = 165$ Å, $b = 117$ Å, and $c = 84$ Å, which can diffract to up to 2.4-Å resolution.

To obtain crystal form V, MLE (5–10 mg/ml) was dialyzed against 0.2 mM $MnCl_2$/7 mM 2-mercaptoethanol/50 mM Tris (pH 7), and *cis,cis*-muconate was added to a final concentration of 0.2 mM. Crystals were grown at room temperature by sitting-drop vapor diffusion against a well solution of 70 mM NaCl/70 mM sodium acetate (pH 5.2)/0.25% polyethylene glycol (average molecular weight 3,350). Drops contained equal volumes (10 μl) of well solution and MLE. Crystals of form V are rectangular prisms, occasionally with two opposite triangular faces, which diffract to better than 2.0-Å resolution. They are of space group *I*422, with unit cell dimensions $a = b = 136$ Å and $c = 265$ Å.

Data Collection. Form V crystals were kept at a temperature of 4°C during x-ray experiments with a stream of chilled air. Initial data, used to determine the space group and unit cell dimensions, were collected to 3-Å resolution on a Siemens X-100A multiwire proportional x-ray detector. X-rays were generated by an Eliot GX-6 rotating anode (30 kV × 30 mA) with a 0.5 mm focusing cup and a 0.3 mm collimator. The set of x-ray diffraction images was reduced to integrated indexed intensities and processed with XDS (11) to determine the unit cell dimensions, Laue symmetry, and space group. A second data set of up to 2.0-Å resolution, used for structure solution and refinement, was collected on an R-AXIS IIC image plate detector by using 0.3 mm collimated monochromatized CuK_α radiation from a Rigaku RU-200 rotating anode generator (50 kV × 150 mA). The diffraction images were reduced to integrated indexed intensities with the R-AXIS processing software PROCESS [T. Higashi, Rigaku Corporation (1990)] and MOSFLM [A. J. Wonocott, Imperial College (1980)]. The data set consists of 219,744 total observations (of 58,765 independent reflections) with an overall merging *R*-factor of 6.7% (Table 1).

Solution of the Type V Crystal Structure by Molecular Replacement and Refinement. The type IV crystal structure was solved by molecular replacement using the program XPLOR (12), with the coordinates obtained from previously described MLE crystals as a starting model (6). The structure was refined by using the programs XPLOR and TNT (13). This model was used as the starting model for the molecular replacement solution of the type V crystal structure. Although MLE forms an octamer in solution, the assumption of a typical V_m (14) of 2.5 leads to the unexpected conclusion that there are three monomers in an asymmetric unit, given the unit cell and space group of the form V MLE crystals. The self-rotation function (version 3.0 of the program XPLOR) showed a noncrystallographic twofold axis at approximately $\phi = 120°$, $\psi = 0°$, $\kappa =$

Table 1. Analysis of the x-ray data set from a form V crystal used for structural solution

	Value at resolution										
Measurement	All	71–15 Å	−10 Å	−7.5 Å	−5 Å	−3.5 Å	−3 Å	−2.75 Å	−2.5 Å	−2.25 Å	−2.0 Å
Independent reflections	58,765	213	487	928	3,747	10,017	8,820	6,863	9,598	11,238	6,854
% of theoretical	70.0	84.2	91.5	93.5	95.1	95.8	95.4	92.6	90.2	70.9	27.8
Average intensity	26.3	88.1	100.9	79.0	44.5	61.6	29.1	16.2	11.4	8.5	7.0
σ	3.1	3.8	3.3	2.7	2.3	3.8	3.3	3.1	3.0	2.9	3.3

180°; this result would have been consistent with the presence of only two monomers in the asymmetric unit. A monomer of MLE was the starting model for the cross-rotation search, using the coordinates of the final wild-type model after refinement with noncrystallographic symmetry restraints. After Patterson-correlation refinement, three clear peaks were found. Each peak was related to the other two by approximately the same relationship found in the self-rotation function ($\phi = 120°$, $\psi = 0°$, $\kappa = 180°$). The translation function was run four times. The first time, one of the solutions from the Patterson-correlation refinement was used as input; for each subsequent run, input coordinates were the output coordinates from the previous run rotated by $\phi = 120°$, $\psi = 0°$, $\kappa = 180°$. Three distinct solutions were obtained. (*i*) A monomer near the origin of the unit cell resulted in a complete octamer when the crystallographic symmetry operators (fourfold and twofold axes) was applied. (*ii*) A monomer near the *z* axis resulted in one-half an octamer when the crystallographic fourfold symmetry axis was applied. If the other half of the octomer was set in place, it collided with solution *i*. This incorrect solution was found twice. (*iii*) A monomer near the *z* axis resulted in one-half an octamer when the crystallographic fourfold symmetry axis was applied. This time, if the other half of the octomer was set in place, it did not collide with the octamer from solution *i*; instead, it seemed to pack against it and establish reasonable crystal contacts. Therefore, solutions *i* and *iii* were used for refinement as monomers A and B, respectively. The other half of the dimer for solution *iii* (monomer C) was located by applying the known symmetry of MLE octamers. Thus, there are a total of three monomers per asymmetric unit, one a member of one octomer (monomer A), and two a member of a second octamer (monomers B and C). The packing of octamers in a unit cell is shown schematically in Fig. 2.

Rigid-body refinement and least-squares minimization were carried out in XPLOR. The programs FRODO (15) and O (16) were used to manually adjust the coordinates throughout the refinement. Noncrystallographic symmetry restraints were applied to backbone atoms during the initial stages of the refinement. Individual temperature factors were refined with restraints. The program WATERHUNTER was used to place water molecules.

The temperature factors of the manganese ions in the active sites reached 70–90 Å^2 after several rounds of refinement. Simulated annealing-refined omit maps, in which residues within 8 Å of each of the three manganese ions were deleted from the model, showed no electron density for any of the manganese ions (data not shown). A peak of density nearby was filled with a water molecule that forms hydrogen bonds with some of the metal ligands.

As the refinement progressed, residues at the N terminus, C terminus, and near a mobile loop (approximately residues 20–30) were deleted or modified as suggested by the $2F_o - F_c$ map or the $F_o - F_c$ map. The final model contained residues 4–19 and 31–372 for all three monomers. In addition, monomer A contained residues 20 and 21, although residue 20 was modeled as an alanine rather than an arginine, because the side chain was not clearly defined.

The most recently determined sequence for the *catB* gene that encodes this MLE (8) was used for the final model. The sequence was modified to replace E138 with valine, according to our own sequence data (K.T. and J.A.G., unpublished data), because valine was more consistent with the electron density in a $2F_o - F_c$ map.

The final *R*-factor for data to 2.1 Å is 15.6%. The final model contains 564 water molecules and 1,078 residues. The root-mean-square (rms) deviations from ideal geometry are 0.010 Å for bond lengths and 1.5° for bond angles. Of the residues other than proline or glycine, 93.2% are in the most favored regions of a Ramachandran plot and 6.6% are in additional allowed regions. Two residues, E225 from two of the monomers, were in generously allowed regions, and no residues were in disallowed regions.

Sequence Alignment. Sequences were aligned by using the program PILEUP (17) as implemented in the Genetics Computer Group package (18). The sequences of several members of the family were used to improve the quality of the alignment (MR and MLE from *Pseudomonas putida*, enolase from *Saccharomyces cerevisiae*, glucarate dehydratase from *P. putida*, and the sequences of three other enzymes with MLE activity).

Structural Superposition. Superposition was performed and optimized by using the LSQ facility in the graphics program O (16). Subunit A of the MLE structure presented here was used for comparison with MR and enolase. The structures of enolase (subunit A of reference 1EBH) and MR (reference 1MNS) were retrieved from the Protein Data Bank. MLE coordinates determined from the I4 crystal form were kindly provided by Adrian Goldman (Center for Biotechnology, Turku, Finland) for comparison. Pairwise comparisons of the structures of each of the subunits in the two crystal forms of MLE were performed by using XPLOR.

Generation of Figures. Fig. 2 was rendered using the program RAYSHADE (L. Coffin and D. Debry; http://www-graphics.stanford.edu/~cek/rayshade/rayshade.html). Figs. 3–5 were rendered using RAYSCRIPT (E. Fontano, D. Peisach, and E. Peisach; http://www.sb.fsu.edu/software/rayscript.html), which takes input suitable for MOLSCRIPT (19) and converts the information to input suitable for RAYSHADE.

RESULTS AND DISCUSSION

A Crystal Form of MLE with Unusual Octamer Packing. A search for conditions that would promote the growth of a new crystal form of MLE was undertaken in an attempt to produce crystals that are not twinned, in contrast to previously characterized crystals (20), and diffract to high resolution. Several new crystal forms were identified and characterized. The structure of MLE in crystal form IV, which does not diffract well beyond 2.5-Å resolution, was determined by molecular replacement. The structure of crystal form V, which diffracts to beyond 2.0-Å resolution, was solved by molecular replacement by using the structure from form IV, and was used for all further studies.

Crystal form V of MLE has the unusual property of containing three monomers in an asymmetric unit, although the enzyme is an octamer. The crystal packing is shown schematically in Fig. 2*A*. One octamer is situated at the origin of the crystal lattice and is built up by the crystallographic symmetry elements (fourfold and twofold axes of symmetry) from one of the monomers in the asymmetric unit. Two additional octamers are translated down the crystallographic *z* axis (the fourfold axis of symmetry) and rotated by approximately 120°. These octamers are built up from the other two monomers in the asymmetric unit by the crystallographic fourfold axis. The fact that the neighboring octamers along the *z* axis are related to each other by approximately 120° is apparently fortuitous, and leads to the curious observation of an approximate threefold screw axis coincident with the crystallographic fourfold. It also explains why each of the three solutions of the rotation function was related to the other two by approximately the same noncrystallographic symmetry element ($\phi = 120°$, $\psi = 0°$, $\kappa = 180°$).

The packing of crystal form V of MLE is similar to the packing of a previously described MLE crystal form in space group *I*4 (6) and to the packing of crystalline MR in space group *I*422 (21) (Fig. 2). In all three crystal forms, the octamers are packed in two interdigitating rows as a result of body-centered packing. The main difference between them is the rotation of the octamers in each row relative to each other and to the crystallographic fourfold axis. The similarity in

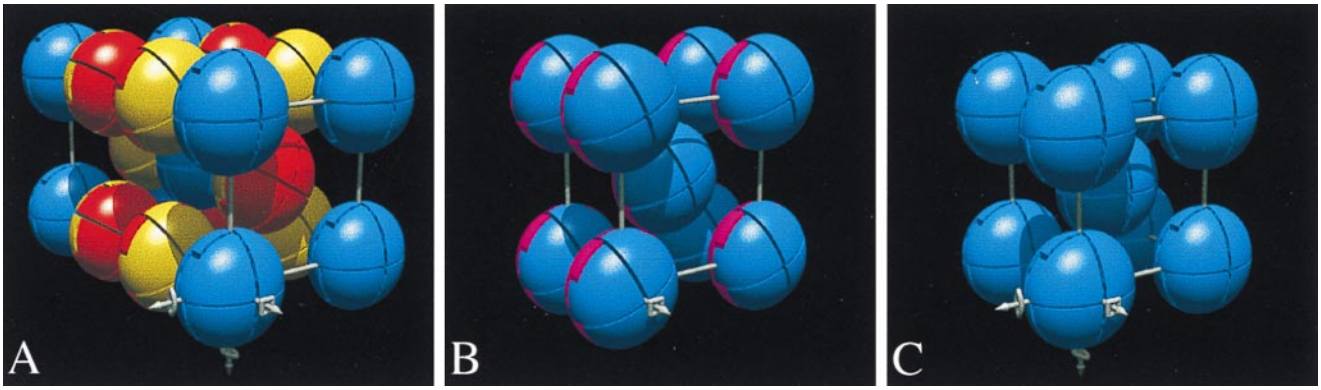

FIG. 2. Packing of octamers in a unit cell for three crystal forms. (*A*) Form V crystals of MLE, described in this study (space group *I*422). Monomer A is shown schematically in blue, monomer B in yellow, and monomer C in red. (*B*) An MLE crystal form, space group *I*4 (7). (*C*) MR crystals, space group *I*422.

packing is reflected in the relationship between unit cell lengths in the two MLE crystal forms. The *a* and *b* axes differ by only 2% (136 Å vs. 139 Å) and the *c* axes are related by a factor of 3 (265 Å vs. 84 Å).

Although the packing of the two MLE crystal forms is similar, the specific contacts between octamers in the crystals differ. The crystal contacts in the *I*4 form are tenuous (7), whereas those in the *I*422 form described here are more intimate. Most of the contacts occur between subunit B in the octamers situated on the fourfold cell axis and subunit B in the octamers situated on the body-centered fourfold axis (subunit B is depicted in yellow in Fig. 2). Substantial contacts also occur between subunit A and B (blue and yellow, respectively, in Fig. 2) in adjacent octamers along the fourfold axes. A few contacts are present between subunit A along the cell axis and subunit C (red in Fig. 2) along the body-centered axis. Interestingly, no contacts at all occur between adjacent C subunits along the fourfold axes.

Structure of MLE. A monomer of MLE is composed of three domains: (*i*) an N-terminal $\alpha + \beta$ domain, (*ii*) a central α/β barrel missing the final helix, and (*iii*) a C-terminal domain that replaces the final helix of the barrel. The model of MLE presented here is similar to the structure of MLE determined from the *I*4 crystal form (7). Pairwise comparisons of the structures of each of the subunits in the two crystal forms were performed; the two subunits in the I4 crystal form and the three subunits of the *I*422 crystal form. No α-carbon position differs by more than 1 Å between any two of the subunits. The rms deviation in corresponding α-carbon positions between subunits of the two structures is 0.2 Å in each case, and the rms deviation for all atom positions is 0.7–0.8 Å. This is comparable to the rms deviations of α-carbon and all atom positions between subunits of the same structure, which are 0.1 Å and 0.6 Å, respectively, for the *I*4 structure and 0.2 Å and 0.6–0.7 Å, respectively, for the *I*422 structure.

A simulated annealing-refined omit map of the active site of MLE shows that the structure presented here has no detectable metal ion bound (data not shown). The absence of metal ion is probably because 0.1 mM $MnCl_2$ was present in the crystallization drop, whereas in the crystallization mix that produced MLE crystals in the *I*4 crystal form, 2 mM $MnCl_2$ was present (20). The absence of the metal ion causes little perturbation of the structure (Fig. 3). The only metal ligand whose position is significantly different in the structures with and without metal is an indirect ligand, E250, whose terminus is slightly displaced. The only other residue whose position differs between the structures is K169, which may be the catalytic base. This residue differs not only between the two crystal forms but also between subunits in the same crystal form. The B factor of the terminal nitrogen is high, ranging between 47 and 115. Perhaps the position of this residue is fixed only upon binding of substrate.

Comparison of MLE to Enolase and MR. Several enzymes related to MLE have been identified by searching the database of known sequences (2, 5). The sequences of MLE and yeast enolase are about 15% identical, a level of similarity that is only marginally significant, and not sufficient for recognition through a database search. Similarity between MLE and enolase has been noted (2–4). When MR, MLE, and several related proteins were aligned with enolase by using the program PILEUP, the direct metal ligands were aligned with no gaps introduced into any of the sequences in the intervening region, consistent with a recently reported superposition of enolase, MR, and MLE (2, 4, 5). Beyond one metal ligand and a proline, there is no sequence identity between MR, MLE, and enolase in this region. However, the pattern of polar and nonpolar residues are similar, suggesting that the alignment in this region could be used as the basis for a structural superposition of MLE and enolase. The superposition was accomplished by using the LSQ facility of the graphics program O, by least-squares minimization of the distances between matched α-carbons. The list of matches (initially residues 224–250 in MLE with residues 295–321 in enolase) was expanded by

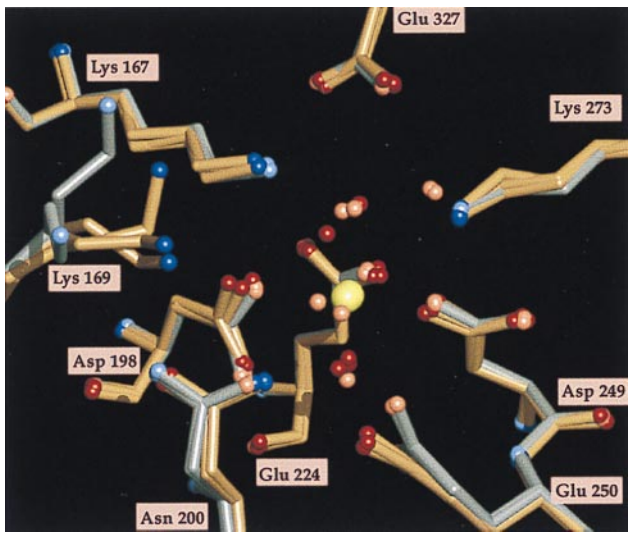

FIG. 3. Comparison of the active sites of MLE structures. The structure of the three subunits of MLE determined in this study from the I422 crystal form are shown in brown, with oxygen atoms depicted in blue and nitrogen atoms in red. The structure of the two subunits of MLE determined from the I4 form (7) are shown in gray, with oxygen atoms depicted in light blue, nitrogen atoms in pink, and manganese ions in yellow.

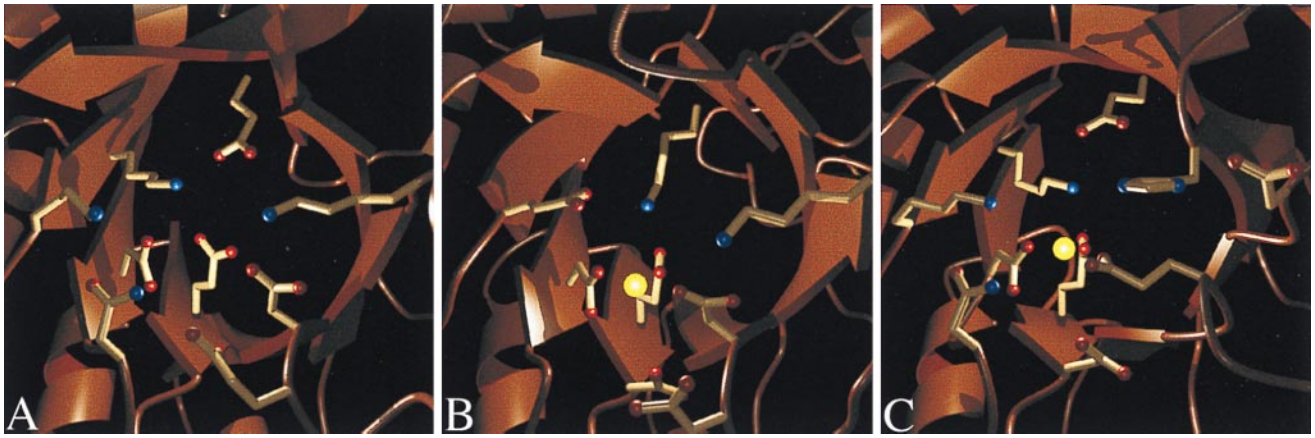

FIG. 4. Active sites of MLE (*A*), MR (*B*), and enolase (*C*).

adding corresponding α-carbons of proximate secondary structural elements in the two proteins. There is a correspondence between many elements of secondary structure in the three enzymes. For MLE and enolase, the rms deviation in the position of 171 matched α-carbons is 2.0 Å. For MLE and MR, the rms deviation in the position of 325 matched α-carbons is 1.7 Å. The active sites of MR, MLE, and enolase, oriented according to the structural superposition, are shown in Fig. 4; their similarities are striking. All three active sites are found at the C-terminal end of the barrel [as is true for all known α/β barrel enzymes (22)]. Each active site can effectively be divided in half, with the metal ions and their ligands near the bottom in this view and other residues involved in catalysis clustered near the top.

Comparison of Metal Ligands. Whereas MR and MLE both require a divalent metal ion for activity, enolase requires two metal ions, bound to a high-affinity and a low-affinity site (23–25). The metal ligands in MR, MLE, and the high-affinity enolase site are quite similar (Fig. 4) (4, 7, 21). The three direct protein ligands to the metal ion are structurally equivalent in each of the three proteins (D195, E221, and E247 in MR; D198, E224, and D249 in MLE; D246, E295, and D320 in enolase). In MR, E222, a residue that immediately follows one of the direct ligands in the primary sequence, acts as an indirect ligand through a water molecule (21). In MLE, E250, the residue following a different ligand, is an appropriate distance from the metal to act as an indirect ligand (7). The residue in MLE analogous to E222 in MR, as well as the residue in MR analogous to E250 in MLE, are glutamine residues with their side chains bending away from the active site. In enolase, the analogous residues are D296 and D321, both of which interact with the metal ion. The interaction is mediated through one water molecule in the case of D296, two in the case of D321. It seems likely that a common ancestor of these enzymes might have had two indirect ligands, similar to modern enolase. As MR and MLE evolved from this ancestor, only one of the two ligands retained its role.

A second metal ion binds to enolase in the presence of substrate or inhibitor and is coordinated by two oxygens of the substrate, two oxygens of S39, and two water molecules (24, 25). In the superimposed structures of enolase and MLE, one of the terminal oxygens of E250 in MLE is in the same position as one of the water ligands in enolase. One of the MLE subunits in the structure presented here, and both subunits in the structure of MLE with a metal ion bound (7), contain a water molecule corresponding to the other water ligand in enolase. Therefore, if MLE binds a complex of substrate and a second metal ion, as does enolase, one might expect that E250 would be a ligand of that metal ion.

Residues Involved in Catalysis. Because the reactions catalyzed by MR, MLE, and enolase share the chemical step of abstraction of the α-proton of a carboxylate, two common features are necessary in their active sites. First, a catalytic base is required to abstract the proton. Second, something (a functional group and/or a bound metal ion) must stabilize the negatively charged transition state. An overlay of the portion of the three active sites containing these features is shown in Fig. 5.

As a racemase, MR catalyzes a similar reaction in the forward and backward directions. Two catalytic bases, K166 and H297, are employed to accomplish this task (21, 26, 27). From their positions in the active site, it seems that D270 assists H297 in proton abstraction, and that together they form a histidine/aspartate dyad that can more properly be considered the second catalytic base. The catalytic base in enolase is K345 (28), which is homologous to D270 in MR. The catalytic base in MLE has not yet been identified, but it seems likely to be either K169 (homologous to K166 in MR) or K273 (homologous to D270 in MR). In any case, the catalytic bases of MLE and enolase are found at either of two conserved positions, those of K166 or D270 in MR, which also act as catalytic bases. However, although the positions of the catalytic bases are conserved, the identities of the residues at those positions are not. That is, the catalytic base in enolase, K345, is homologous to a component of the catalytic base in MR, D270. This sort of conservation, of position but not of identity of a catalytic base, is unexpected. A second type of conservation, that of the location of the catalytic atoms relative to the metal ion is also apparent; the functional groups of the catalytic bases H297 in MR and K345 in enolase are in the same position in the two active sites.

The second requirement for catalysis, stabilization of the negatively charged transition state, is accomplished by interactions with the carboxylate oxygens of the substrate. One

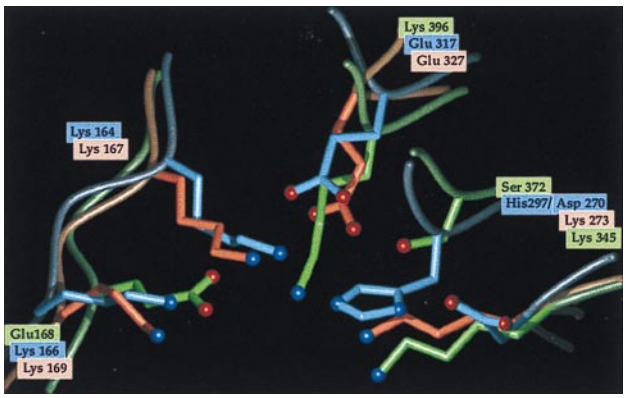

FIG. 5. Comparison of residues involved in catalysis in MR (blue), MLE (pink), and enolase (green).

carboxylate oxygen is probably stabilized by functioning as a metal ligand in all three cases. This oxygen is also stabilized by binding to a lysine. In MR, the Lys is K164; in MLE, it is probably the homologous residue K167. In contrast, in enolase the role is played by K396. The terminus of K396 is in the same position in the active site as the termini of K164 in MR and K167 in MLE. However, K396 of enolase is homologous not to K164 of MR but to a residue with slightly different role, E317. The second carboxylate oxygen in the transition state is evidently stabilized though a short, strong hydrogen bond [also termed a low-barrier hydrogen bond (29)] with E317 in MR (30). The homologous residue in MLE, E327, probably plays the same role. In enolase, there is no residue that plays a role analogous to E327. Instead, the other carboxylate oxygen is stabilized by acting as a ligand to the second metal ion in enolase (25). Thus, the transition-state-stabilizing residues are conserved in the same unusual way as the catalytic bases; their position in the active site is conserved, but their identities are not (K396 in enolase, E317 in MR, and E327 in MLE). In this case, their roles as stabilizers are also slightly different.

In enolase, E168 plays an important role in catalysis, probably aiding in both steps of the reaction (28, 31, 32). One of the catalytic bases in MR, K166, and the structurally analogous K169 in MLE and E168 in enolase are positioned on one of the β-strands that make up the barrel. The positions of the α-carbons of this strand are quite similar in the three enzymes (Fig. 5). However, in enolase, this strand runs in the opposite direction. It is the sole known exception to the rule that all known α/β barrels are composed of eight parallel strands (33). Yet, although the orientation of this strand is reversed, the position of E168 in enolase is completely analogous to the positions of K166 in MR and K169 in MLE. It is as if, as the enzymes evolved, the position of these catalytic residues has outlasted the orientation of the element of secondary structure that holds them.

Why might the position of catalytic residues be conserved, whereas their identity and roles change during evolution? It may be a historical accident, as it were; perhaps it is easier for enzymes to evolve through the changing of roles and identities of residues at certain positions in the active site than for residues at new positions to be recruited. Or perhaps, some positions in an enzyme may be especially amenable to holding catalytic residues. This has been noted previously in a general way. For instance, if a catalytic residue is at the end of a helix, the helix dipole moment may aid in catalysis (34). The active site of all known α/β barrel enzymes is at the C-terminal end of the barrel (22), perhaps to take advantage of the distinct electrostatic field formed by the protein backbone of the barrel (35).

In summary, two types of conservation are apparent upon comparison of the catalytic residues of MLE, MR, and enolase. (*i*) The position of catalytic atoms relative to the metal ion can be identical, even when those atoms belong to residues that arise from different points in the protein scaffold. (*ii*) The position of catalytic residues in the protein scaffold can be identical, even when the identity of those residues and their particular roles in catalysis vary. Although these modes of conservation have not often been noted in enzyme families, they may become more apparent when the large number of sequences and structures currently becoming available are examined with these possibilities in mind. Indeed, the first type of conservation has been described in the thioredoxin family (36); both may be evident in the family of thiamin diphosphate-dependent enzymes (37). Further exploration should include dissection of the reasons for the conservation of position but not identity of catalytic residues, and determination of whether similar observations can be made in other enzyme families.

We thank Adrian Goldman who kindly provided MLE coordinates for comparison and Daniel Peisach who helped with the generation of figures. This research was supported by National Institutes of Health Grants GM40570 (to J.A.G., G.L.K., and G.A.P.) and GM26788 (to G.A.P. and D.R.), Fellowship DRG-1194 of the Cancer Research Fund of the Damon Runyon–Walter Winchell Foundation (to M.S.H.), and in part by a grant from the Lucille P. Markey Charitable Trust. I.S. thanks the Alexander von Humboldt Gesellschaft for support.

1. Neidhart, D. J., Kenyon, G. L., Gerlt, J. A. & Petsko, G. A. (1990) *Nature (London)* **347,** 692–694.
2. Babbitt, P. C., Hasson, M. S., Wedekind, J. E., Palmer, D. R. J., Reed, G. H., Rayment, I., Ringe, D., Kenyon, G. L. & Gerlt, J. A. (1996) *Biochemistry* **35,** 16489–16501.
3. Lebioda, L. & Stec, B. (1988) *Nature (London)* **333,** 683–686.
4. Wedekind, J. E., Reed, G. H. & Rayment, I. (1995) *Biochemistry* **34,** 4325–4330.
5. Babbitt, P. C., Mrachko, G. T., Hasson, M. S., Huisman, G. W., Kolter, R., Ringe, D., Petsko, G. A., Kenyon, G. L. & Gerlt, J. A. (1995) *Science* **267,** 1159–1161.
6. Goldman, A., Ollis, D. L. & Steitz, T. A. (1987) *J. Mol. Biol.* **194,** 143–153.
7. Helin, S., Kahn, P. C., Guha, B. L., Mallows, D. G. & Goldman, A. (1995) *J. Mol. Biol.* **254,** 918–941.
8. Houghton, J. E., Brown, T. M., Appel, A. J., Hughes, E. J. & Ornston, L. N. (1995) *J. Bacteriol.* **177,** 401–412.
9. Meagher, R. B. & Ornston, L. N. (1973) *Biochemistry* **12,** 3523–3530.
10. Jancarik, J. & Kim, S. H. (1991) *J. Appl. Crystallogr.* **24,** 409–411.
11. Kabsch, W. (1993) *J. Appl. Crystallogr.* **26,** 795–800.
12. Brünger, A. T. (1992) XPLOR: A System for X-Ray Crystallography and NMR (Yale Univ. Press, New Haven, CT).
13. Tronrud, D. E., Ten Eyck, L. F. & Matthews, B. W. (1987) *Acta Crystallogr. A* **43,** 489–501.
14. Matthews, B. W. (1968) *J. Mol. Biol.* **33,** 491–497.
15. Jones, T. A. (1978) *J. Appl. Crystallogr.* **11,** 268–272.
16. Jones, T. A. (1991) *Acta Crystallogr. A* **47,** 110–119.
17. Feng, D. F. & Doolittle, R. F. (1987) *J. Mol. Evol.* **25,** 351–360.
18. Devereux, J., Haberli, P. & Smithies, M. (1984) *Nucleic Acids Res.* **12,** 387–395.
19. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24,** 946–950.
20. Goldman, A., Ollis, D., Ngai, K. L. & Steitz, T. A. (1985) *J. Mol. Biol.* **182,** 353–355.
21. Neidhart, D. J., Howell, P. L., Petsko, G. A., Powers, V. M., Li, R. S., Kenyon, G. L. & Gerlt, J. A. (1991) *Biochemistry* **30,** 9264–9273.
22. Farber, G. K. & Petsko, G. A. (1990) *Trends Biochem. Sci.* **15,** 228–234.
23. Poyner, R. R. & Reed, G. H. (1992) *Biochemistry* **31,** 7166–7173.
24. Wedekind, J. E., Poyner, R. R., Reed, G. H. & Rayment, I. (1994) *Biochemistry* **33,** 9333–9342.
25. Larsen, T. M., Wedekind, J. E., Rayment, I. & Reed, G. H. (1996) *Biochemistry* **35,** 4349–4358.
26. Powers, V. M., Koo, C. W., Kenyon, G. L., Gerlt, J. A. & Kozarich, J. W. (1991) *Biochemistry* **30,** 9255–9263.
27. Landro, J. A., Kallarakal, A. T., Ransom, S. C., Gerlt, J. A., Kozarich, J. W., Neidhart, D. J. & Kenyon, G. L. (1991) *Biochemistry* **30,** 9274–9281.
28. Poyner, R. R., Laughlin, L. T., Sowa, G. A. & Reed, G. H. (1996) *Biochemistry* **35,** 1692–1699.
29. Cleland, W. W. & Kreevoy, M. M. (1994) *Science* **264,** 1887–1890.
30. Gerlt, J. A. & Gassman, P. G. (1993) *Biochemistry* **32,** 11943–11952.
31. Brewer, J. M., Robson, R. L., Glover, C. V., Holland, M. J. & Lebioda, L. (1993) *Proteins* **17,** 426–434.
32. Reed, G. H., Poyner, R. R., Larsen, T. M., Wedekind, J. E. & Rayment, I. (1996) *Curr. Opin. Struct. Biol.* **6,** 736–743.
33. Lebioda, L., Stec, B. & Brewer, J. M. (1989) *J. Biol. Chem.* **264,** 3685–3693.
34. Hol, W. G., van Duijnen, P. T. & Berendsen, H. J. (1978) *Nature (London)* **273,** 443–446.
35. Raychaudhuri, S., Younas, F., Karplus, P. A., Faerman, C. H. & Ripoll, D. R. (1997) *Protein Sci.* **6,** 1849–1857.
36. Martin, J. L. (1995) *Structure* **3,** 245–250.
37. Hasson, M. S., Muscate, A., McLeish, M. J., Polovnikova, L. S., Gerlt, J. A., Kenyon, G. L., Petsko, G. A. & Ringe, D. (1998) *Biochemistry* **37,** 9918–9930.