Crystal Structure of *Escherichia coli* Malate Synthase G Complexed with Magnesium and Glyoxylate at 2.0 Å Resolution: Mechanistic Implications[†,‡,§]

Bruce R. Howard,[||,⊥] James A. Endrizzi,[▽] and S. James Remington*[,||]

Institute of Molecular Biology and Departments of Chemistry and Physics, University of Oregon, Eugene, Oregon 97403, and Department of Molecular and Cell Biology, University of California, Berkeley, California 94720

Received November 1, 1999; Revised Manuscript Received January 3, 2000

ABSTRACT: The crystal structure of selenomethionine-substituted malate synthase G, an 81 kDa monomeric enzyme from *Escherichia coli* has been determined by MAD phasing, model building, and crystallographic refinement to a resolution of 2.0 Å. The crystallographic *R* factor is 0.177 for 49 242 reflections observed at the incident wavelength of 1.008 Å, and the model stereochemistry is satisfactory. The basic fold of the enzyme is that of a β8/α8 (TIM) barrel. The barrel is centrally located, with an N-terminal α-helical domain flanking one side. An inserted β-sheet domain folds against the opposite side of the barrel, and an α-helical C-terminal domain forms a plug which caps the active site. Malate synthase catalyzes the condensation of glyoxylate and acetyl-coenzyme A and hydrolysis of the intermediate to yield malate and coenzyme A, requiring Mg^{2+} . The structure reveals an enzyme−substrate complex with glyoxylate and Mg^{2+} which coordinates the aldehyde and carboxylate functions of the substrate. Two strictly conserved residues, Asp631 and Arg338, are proposed to provide concerted acid−base chemistry for the generation of the enol(ate) intermediate of acetyl-coenzyme A, while main-chain hydrogen bonds and bound Mg^{2+} polarize glyoxylate in preparation for nucleophilic attack. The catalytic strategy of malate synthase appears to be essentially the same as that of citrate synthase, with the electrophile activated for nucleophilic attack by nearby positive charges and hydrogen bonds, while concerted acid−base catalysis accomplishes the abstraction of a proton from the methyl group of acetyl-coenzyme A. An active site aspartate is, however, the only common feature of these two enzymes, and the active sites of these enzymes are produced by quite different protein folds. Interesting similarities in the overall folds and modes of substrate recognition are discussed in comparisons of malate synthase with pyruvate kinase and pyruvate phosphate dikinase.

The discovery of malate synthase (MS)[1] was reported in 1956 (*1*) and proved to be the link in closing the "modified" tricarboxylic acid or glyoxylate cycle (*2*). Krebs and Kornberg (*2*) described how malate synthase, together with isocitrate lyase, provides a pathway in intermediary metabolism which allows certain organisms to derive their carbon requirements from two-carbon compounds. Isocitrate lyase catalyzes the cleavage of isocitrate to succinate and glyoxylate, whereas the citric acid cycle would otherwise convert isocitrate to succinate and two molecules of carbon dioxide. Malate synthase can then condense glyoxylate with an acetyl group from acetyl-CoA, producing malate to replenish the pool of citric-acid-cycle intermediates.

The presence of the glyoxylate cycle has been established in a wide variety of organisms, but has been definitively shown to occur only in plants, fungi, some worms, and various unicellular organisms (*3, 4*). Although the reactions of this cycle are reported to have been observed in animals (*5*−*7*), these observations were based on immunocytochemistry or activities in crude extracts from various tissue preparations. To date, no gene or protein sequence for MS appears in the databases for any animal more complex than nematodes. Both monomeric and oligomeric forms of the enzyme have been characterized (*8*−*11*).

Malate synthase catalyzes the Claisen condensation of glyoxylate and acetyl-CoA to form a malyl-CoA intermediate which is subsequently hydrolyzed to release malate and CoA (Figure 1). The reaction, as with other Claisen enzymes such as citrate synthase (CS), proceeds with inversion of configuration at the methyl group of acetyl-CoA (*12, 13*). Although this is suggestive of a concerted reaction, the stepwise nature of the mechanism has been demonstrated for MS using a double-isotope fractionation procedure (*14*). It is therefore likely that an enol(ate) is first formed on acetyl-CoA which

[†] This work was supported by National Science Foundation Grant MCB-9728162.

[‡] Coordinates have been deposited with the Protein Data Bank, access code 1D8C. RCSB ID code RCSB009880.

[§] This paper is dedicated to the memory of the pioneer and forever inspirational Paul A. Srere (1925−1999).

* To whom correspondence should be addressed. Phone: (541) 346-5190. Fax (541) 346-5870. E-mail: jim@uoxray.uoregon.edu.

[||] University of Oregon.

[⊥] Current address: Department of Biochemistry, University of Utah, Salt Lake City, UT.

[▽] University of California.

[1] Abbreviations: MS, malate synthase; MSG, *Escherichia coli* G isozyme; CoA, coenzyme A; CS, citrate synthase; TIM, triosephosphate isomerase; CD, circular dichroism; PK, pyruvate kinase; PPDK, pyruvate phosphate dikinase; DTNB, 5,5′-dithiobis(2-nitrobenzoic acid); BME, β-mercaptoethanol; DTT, dithiothreitol; MAD, multiple-wavelength anomalous dispersion; PCR, polymerase chain reaction; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; CCD, charge-coupled device.

10.1021/bi992519h CCC: $19.00
Published on Web 02/23/2000

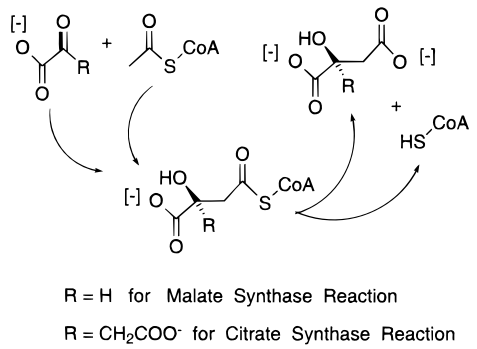

FIGURE 1: Malate synthase and citrate synthase reactions. The two substrates in each reaction are drawn on the left side of the figure and the products on the right. The intermediate in the reactions, malyl-CoA or citryl-CoA, which remains bound noncovalently to the enzyme until product release, is shown lower center.

subsequently attacks the electrophilic carbon of glyoxylate. This activation requires the enzymatic abstraction of a proton from the α-methyl group of the thioester, which is the rate-determining step. The pK_a in solution of the α-carbon of a thioester has been estimated to be 18−20 (*15, 16*), and proton abstraction from the methyl group has long been recognized to pose both a kinetic and an energetic challenge for the weak bases typically available in an enzyme (*17, 18*).

*Mechanistic Considerations.*The overall reaction carried out by malate synthase is very similar to that of citrate synthase; however, there is no sequence homology between these two families of enzymes. Citrate synthase, like malate synthase, catalyzes the transfer of an acetyl group from acetyl-CoA to the carbonyl carbon of oxaloacetate (Figure 1). Pig heart muscle CS has been studied extensively and is well characterized both biochemically and crystallographically. For reviews see refs *17* and *19*−*23*. The catalytic activity has also been probed in detail using site-directed mutagenesis (*24*−*28*) and a variety of other techniques including NMR and FTIR spectroscopy (*29*−*31*).

However, there are several lines of evidence that suggest that the reaction mechanisms of CS and MS are different. Unlike CS, MS requires Mg^{2+} for activity (*32*−*34*). Studies using fluoroacetyl-CoA as a substrate (*35, 36*) also point to mechanistic differences. CS produces only the (2*R*,3*R*) stereoisomer of fluorocitrate (*37*), requiring the enzyme to distinguish between the pro-*R* and pro-*S* hydrogens, whereas MS produces 3-fluoromalate with the two diastereomers (2*R*,3*R*) and (2*R*,3*S*) in approximately a 3:4 ratio (*38, 39*). It has been suggested that this difference in stereospecificity is due to the different nature of the intermediate formed. If CS produces a neutral enol intermediate, the *E*-enol is dictated, and this is calculated to be more stable than the *Z*-enol by about 4 kcal/mol (*35*), whereas the *Z*- and *E*-enolates are calculated to be roughly isoenergetic. These studies suggest that the citrate synthase reaction involves an intermediate that is more or less neutral in character, whereas the reaction of malate synthase proceeds via a charged enolate intermediate, possibly stabilized by Mg^{2+}.

In citrate synthase, the universally conserved Asp375 and His274 side chains (sequence numbering of the pig heart enzyme (*40*)) have been proposed to carry out the abstraction of a proton from acetyl-CoA in a concerted step to form a neutral enol intermediate (*41*), but the possibility of a partially charged intermediate stabilized by charge delocalization has also been discussed (*26*). Gerlt and Gassman (*42*) have proposed an "enolic" intermediate stabilized by a "short, strong" or low barrier hydrogen bond with His274. Interestingly, there are no conserved histidine residues in the known MS sequences, suggesting that histidine does not participate in the reaction.

In view of the overall similarity of the two enzymatic reactions, the lack of sequence similarity between the two enzymes, and differing requirements for divalent cations, we undertook the present crystallographic study to investigate the catalytic strategy of malate synthase and to compare this with that of citrate synthase. It was fascinating to discover that the basic catalytic strategies of the two enzymes appear to be essentially the same; however, all of the details with the exception of an aspartate residue as the putative catalytic base are different.

MATERIALS AND METHODS

MSG Cloning, Overexpression, and Purification. Using the published sequence for glcB, the gene encoding MSG, an 81 kDa monomeric version of malate synthase from *E. coli* (*8*), was isolated by PCR. Primers were designed to each end of the open reading frame and used to amplify the gene (*43*) directly from *E. coli*, strain W3110 genomic DNA (Sigma). The resulting PCR fragment was ligated into a pET-28b expression vector (Novagen) using *Nco*I and *Xho*I restriction sites to link the enzyme to a C-terminal hexahistidine tag. This recombinant vector (pMSG-B) was used to transform a JM-109(DE3) cell line for overexpression of the protein. The expression of the histidine-tagged enzyme was induced with 1 mM IPTG and the culture allowed to grow for an additional 3.5 h. The cells were disrupted by a single passage through a prechilled French press, and the cellular debris was pelleted by centrifugation. The lysate was passed through a nickel-affinity column (Ni−NTA Agarose, Qiagen), and the bound malate synthase was then washed and subsequently eluted with a linear gradient of 5−100 mM imidazole. The fractions containing the most pure enzyme as judged by SDS PAGE and activity assays were combined, concentrated, and passed through a G-200 Sephadex sizing column. Activity assays were performed by following the loss of absorbance at 232 nm due to cleavage of the thioester bond of acetyl-coenzyme A during the course of the reaction as previously described (*10, 33*). The purified enzyme was then concentrated to 48 mg/mL in a solution containing 10 mM Tris buffer, pH 8.1, 10 mM $MgCl_2$, and 2 mM DTT. Aliquots were frozen in liquid nitrogen and stored at −80 °C.

Introduction of Selenomethionine. Selenomethionine was incorporated into the protein by growth of the same transformed cell line in minimal media, followed by addition of selenomethionine along with an amino acid cocktail designed to shut down methionine biosynthesis 25 min prior to induction (*44*). The selenomethionyl enzyme was purified as described for the native protein except for the gel filtration step. The enzyme, after elution from the nickel-affinity column, was already of sufficient purity and was concentrated to 32 mg/mL, in a solution of 20 mM imidazole, pH 8.0, 10 mM $MgCl_2$, and 4 mM DTT, then frozen, and stored as described above.

Both the native and selenomethionyl MSG were analyzed for molecular weight by matrix-assisted laser desorption

ionization mass spectroscopy (MALDI-MS), indicating a molecular weight of 81 407 for the native enzyme, which is close to the predicted M_r of 81 437. The M_r of the selenomethionyl-incorporated enzyme was found to be 82 427 for an increase of 1020 over the native (*45*). This corresponds to an incorporation level of 98.6% selenomethionine at the 22 methionine sites within the mature protein.

Crystallization. Crystals of MSG were grown in hanging drops by vapor diffusion against a well buffer containing ammonium sulfate as the precipitant. The well solution contained approximately 1.8 M $(NH_4)_2SO_4$, 100 mM imidazole, pH 8.0, and 50 mM sorbitol. The protein solution contained the enzyme at 24 mg/mL, 14 mM $MgCl_2$, and 7 mM DTT. Equal amounts of well solution and the protein solution were combined in the hanging drop and allowed to equilibrate with the well solution at room temperature for several days. This resulted in broccoli-shaped clusters of birefringent crystals unsuitable for diffraction studies. Single, diffraction-quality crystals were obtained by microseeding with this crystalline material 7−8 h after the drops were initially set up. The crystals grow over a period of approximately 5 days, reaching a maximal size of about 0.12 × 0.12 × 0.5 mm.

Data Collection and Reduction. The crystals were found to be very radiation sensitive, but the problem was circumvented by transfer of the crystals to a cryoprotectant, flash-freezing in liquid nitrogen, and collection of data at 100 K. The cryoprotectant consisted of 30% sorbitol, 2 M ammonium sulfate, 100 mM imidazole, pH 8.0, and 10 mM $MgCl_2$. MAD data from a selenomethionyl crystal measuring approximately 0.11 × 0.11 × 0.4 mm were collected on a Quantum4 CCD detector at beam line 5.0.2 at the Advanced Light Source in Berkeley, California. The images were processed and intensities integrated with the program DENZO and the data scaled and merged with SCALEPACK (*46*).

Phasing, Structural Solution, and Refinement. To find the positions of the selenium atoms, the dataset collected at the wavelength corresponding to the peak of the selenium anomalous signal was used as input for Shake-and-Bake v2.0 via the DREAR interface (*47*) with a resolution cutoff of 2.2 Å. One thousand trials were initiated, each beginning with 22 randomly placed selenium atoms, in accordance with the expected selenium substructure in the asymmetric unit. Each trial structure was subjected to 33 cycles of phase refinement and Fourier filtering. The distribution of the final values of the minimal function indicated a solution in 8 of the 1000 trials. A second set of trials was carried out, this time using a resolution cutoff of 2.4 Å, with 22 cycles of phase refinement and Fourier filtering in each trial. For this set, 3 out of 222 trials converged to a solution. A comparison of the solutions from the two sets of trials revealed that only the top 20 sites were in common. The heavy atom parameters of these 20 sites were then refined against the diffraction datasets collected at the three different wavelengths using data from 30 to 2.2 Å with the program SHARP (*48*). The phase estimates output from SHARP were used to calculate an electron density map using the CCP4-supported program FFT (*49, 50*), resulting in a map with clear solvent channels contrasting regions of clearly interpretable protein density. This map was then solvent flattened using DM (*50, 51*). The protein model was built into the experimental map on a Silicon Graphics workstation with the program O (*52*), and refined against the low-energy remote dataset with the TNT refinement package using data from 20 to 2.0 Å (*53*). Sequence alignments were carried out with Clustal W (*54*) and illustrated using ESPript (*55*). Secondary structure assignments were determined by DSSP (*56*). The structural comparisons with other known structures was carried out with Dali ver. 2.0 (*57*). ChemDraw (*58*), Freehand 7 (*59*), Molscript (*60*), and Bobscript (*61*) were used to produce figures.

Table 1: Data Collection and Processing Statistics[a]

	wavelength 1.008 Å	wavelength 0.9798 Å	wavelength 0.9794 Å
no. of observed refln	367 712	308 418	272 589
no. of unique reflns	93 570	70 844	70 556
resolution (Å)	20.0−2.00	30.0−2.20	30.0−2.20
highest resolution shell	2.07−2.00	2.28−2.20	2.28−2.20
completeness (%)	99.4 (98.0)	99.9 (99.7)	99.5 (98.2)
R_{sym} (%)	5.2 (30.7)	4.8 (20.5)	5.0 (22.6)
mosaicity (deg)	0.536	0.318	0.461
av $I/\sigma(I)$	24 (2.7)	29 (4.7)	27 (4.2)

[a] Values in parentheses indicate the statistics for the highest resolution shell. "Unique" reflections treat Friedel mates as independent observations for the purpose of MAD phasing.

Table 2. Phasing Statistics[a]

	wavelength 1.008 Å	wavelength 0.9798 Å	wavelength 0.9794 Å
isomorphous Cullis *R* factor		0.576	0.688
anomalous Cullis *R* factor	0.970	0.647	0.588
isomorphous phasing power		2.622	2.220
anomalous phasing power	0.9564	2.859	3.170
overall figure of merit from SHARP		0.632	
overall figure of merit after DM		0.769	

[a] These statistics are based on 20 selenium sites found with Shake-and-Bake and refined with SHARP. The figure of merit after solvent flattening with DM is also given.

RESULTS

Crystal Structure and Atomic Model. MSG crystallizes in the orthorhombic space group $P2_12_12_1$ with unit cell dimensions of $a = 73.8$ Å, $b = 88.7$ Å, and $c = 109.9$ Å. There is one molecule per asymmetric unit, resulting in a V_m of 2.20 Å^3/Da or a solvent content of approximately 44% (*62*). Data collection, heavy atom phasing, and refinement statistics are presented in Tables 1−3. The final model consists of a total of 709 amino acid residues, a magnesium ion, a glyoxylate ion, a molecule of sorbitol, 345 water molecules, and 5 sulfate ions. This includes residues 3−722 with the exception of a surface loop from 300 to 310 for which the electron density is weak and essentially uninterpretable. Although the C-terminal histidine tag is present in the crystal, it and the final residue in the native sequence are disordered and not included in the model. There is also insufficient electron density to define the conformations for 56 side chains which, as a result, have been truncated to some extent. The model does, however, include some side chains which have temperature factors ranging to 100 Å^2 for which the electron density suggests a dominant rotamer. A break also occurs in a $2F_o-F_c$ map contoured at one σ (0.29 e$^-$/Å^3) along the backbone between the Ca of Ser156 and the

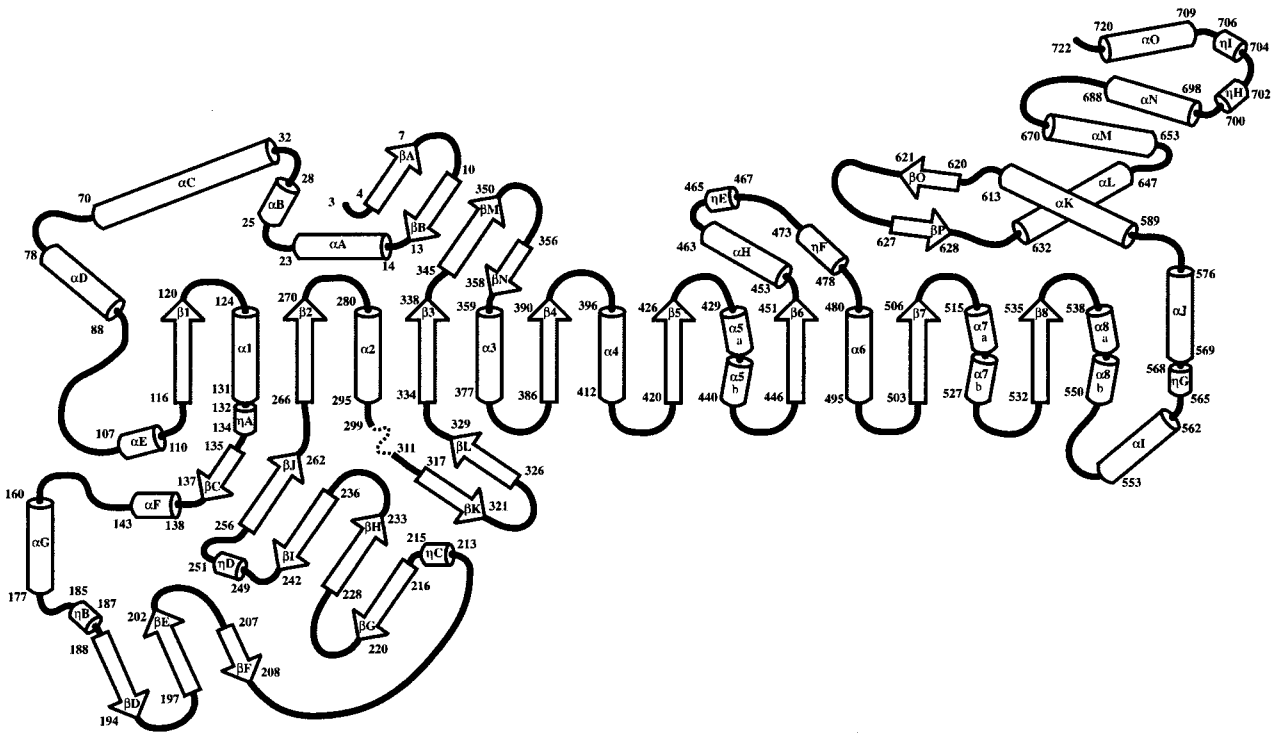

FIGURE 2: Topology diagram of malate synthase G. α-Helices are labeled with α,*β*-strands with *β*, and 3_{10}-helices with *η*. In addition, each secondary structural element is labeled with a number (for the *β*8/α8 barrel) or uppercase letter (the remaining elements of secondary structure). The chain trace is continuous from residue 3 to residue 722 with the exception of residues 300−310 (dashed) which form a disordered loop.

Table 3. Refinement and Final Model Statistics

wavelength (Å)	1.008	correlated *B*-factors (Å^2)	4.59
resolution (Å)	20.0−2.00	⟨*B*⟩ main chain atoms (Å^2)	29.0
initial model *R* factor (%)	38.4	⟨*B*⟩ side chain atoms (Å^2)	34.8
working *R* factor (%)	17.5	⟨*B*⟩ water molecules (Å^2)	38.8
R-free (5% of data reserved) (%)	27.0	*ϕ*/*ψ* angles	
final crystallographic *R* factor (%)	17.7	amount in most favored regions (%)	91.7
no. of non-hydrogen protein atoms	5331	amount in additional allowed regions (%)	8.2
no. of water molecules	345	amount in generously allowed regions (%)	0.2
deviations from ideality		amount in disallowed regions (%)	0
rms bond lengths (Å)	0.015		
rms bond angles (deg)	2.5		

carboxylate of Gly157. Although the temperature factors are high, the electron density in this region is sufficient to define the backbone trace. Having been verified by an omit map, this loop has been included in the final model. Finally, a glyoxylate molecule complexed with a magnesium ion, evidently scavenged from the growth medium as glyoxylate was not included in the crystallization mixture, is included in the model. Refinement statistics for the final model are given in Table 3. As defined by PROCHECK (*63*), there are no residues with disallowed backbone conformational angles and only one residue (Ser8) in the "generously allowed" region of *ϕ*/*ψ* space.

Overall Fold. The structure of malate synthase G is based on a *β*8/α8 barrel fold which was first seen in triosephosphate isomerase (TIM barrel) (*64*). Insertions within the fold of the barrel and extensions at both ends form additional domains and regions of secondary structure (Figures 2 and 3). The centrally located barrel domain is buttressed on one side by an N-terminal α-helical clasp (residues 3−88) which is linked to the first strand of the barrel by a long, extended surface loop (residues 89−116). On the opposite side of the barrel from this clasp is an α/*β* domain that is formed by two distinct inserts which emanate from within the sequence of the barrel (residues 135−262 and 296−333). The C-terminal domain of the enzyme (residues 589−722) consists of a five-helix plug which fits into the end of the barrel within the loops at the C-termini of the *β*-strands. This plug is connected to the barrel by a loop−helix−turn−helix−loop (residues 551−588). The main chain temperature factors in the turn and loop regions suggest that this connection is somewhat flexible. A sorbitol molecule was found wedged into the interface on the opposite side of the C-terminal plug from the active site cleft. A loop of residues 702−704 between the 3_{10}-helices H and I (Figure 2) form the sorbitol binding site against the long α-helix of the N-terminal clasp.

Active Site. The active site is located in a cleft at the interface between the C-terminal plug and the loops at the C-terminal ends of the *β*-strands of the TIM barrel. A loop consisting of residues 614−631 within the sequence of the C-terminal plug folds into a strand−*β*-turn−strand which has higher temperature factors than the rest of the domain and forms part of the interface of this domain with the TIM barrel. Asp631 at the base of this loop, the proposed active site base, is thus placed within 5.8 Å of the bound glyoxylate

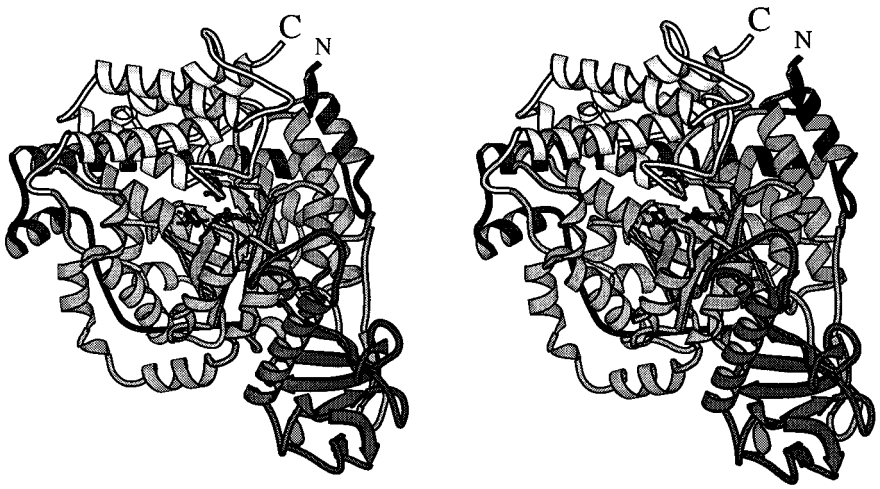

FIGURE 3: Stereo ribbon diagram of malate synthase G. The overall fold is shown looking into the active site cleft. Active site residues are represented in ball-and-stick fashion. The C-terminal plug is drawn in white.

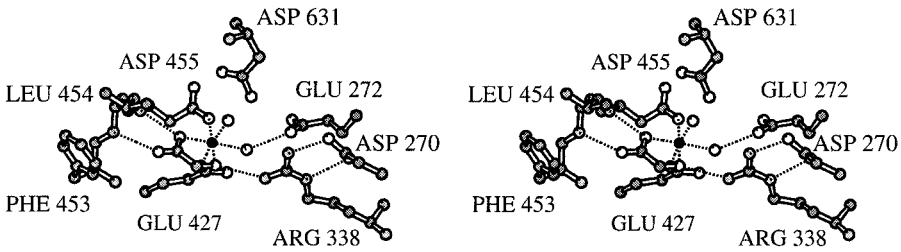

FIGURE 4: Stereoview of bound glyoxylate, octahedrally coordinated magnesium (dashed lines), and neighboring residues in the active site of MSG. Oxygen atoms are indicated by open spheres, while carbon and nitrogen are shaded.

molecule (Figure 4). This positioning of the active site is consistent with other TIM barrel containing enzymes which all have their respective active sites located within the loops at this end of the barrel (*65*). The active site was initially discovered during refinement when a strong positive peak in an $F_o - F_c$ difference map was noticed which was linked closely to Glu427 and Asp455. An inspection of this region revealed that the distances from the closest carboxylate oxygens to an atom modeled into the difference density were approximately 2.0 Å, consistent with a magnesium−oxygen bond. Given this interpretation, the remaining difference density became easily interpretable as two water molecules and a glyoxylate molecule coordinated to the magnesium ion in nearly perfect octahedral configuration. A check of the experimental MAD-phased map in this area shows clear density in the magnesium and glyoxylate binding site (Figure 5).

One of the substrate carboxylate oxygens coordinates the magnesium ion, but in addition the carboxylate of the substrate is stabilized by hydrogen bonds to the amides at the N-terminus of an α-helix (residues 453 and 454) which immediately follows the sixth strand of the TIM barrel. The aldehyde oxygen of the substrate, in addition to coordinating the magnesium ion, is hydrogen-bonded to Arg338 (see below).

Revised Sequence Alignment of Malate Synthases. Although limited sequence identities among various malate synthase sequences have been reported (*8*), the structure and identification of active site residues caused us to revise the proposed sequence alignments and include more members of the family (see the discussion on conserved domains below). With regard to active site residues, Glu427 and Asp455 which bind the magnesium ion, Arg338 which makes a hydrogen bond with the aldehyde oxygen of glyoxylate, and the proposed catalytic base, Asp631, are conserved among 23 known malate synthase sequences. Figure 6 shows a synopsis of the alignment of these sequences and includes one example of each of the three major families ranked according to size (see below). Residues printed as white on black are conserved in all known sequences (the full alignment of all 23 known sequences of MS is available as Supporting Information). Additionally, Gly452 which lies immediately under the bound glyoxylate, Trp534 which sits adjacent to the glyoxylate ion, and makes up one side of the active site, and Ala633 which "hangs into" the active site above the glyoxylate are all strictly conserved. Indeed, of the 21 residues which are strictly conserved among these 23 sequences, 15 cluster in the local vicinity of the active site. Those not immediately near the active site are involved in secondary structure within the barrel or hydrophobic packing: Gly296 terminates helix α2 (see Figure 3), Pro538 initiates the α8 helix, and Phe406 packs between helix α4 and the side chains of Ile388 and Met422 of β-strands β4 and β5, respectively. Asp363 makes a hydrogen bond with the backbone amide of Val340, while Asp398 forms a hydrogen bond to the backbone amide of His394, both within the fold of the barrel. In fact, the only strictly conserved residue not found in the immediate vicinity of the active site or within the barrel itself is Trp645. This residue is packed between the 467−473 loop and the ends of helices αK and αL, beneath the loop which connects the C-terminal plug to the barrel.

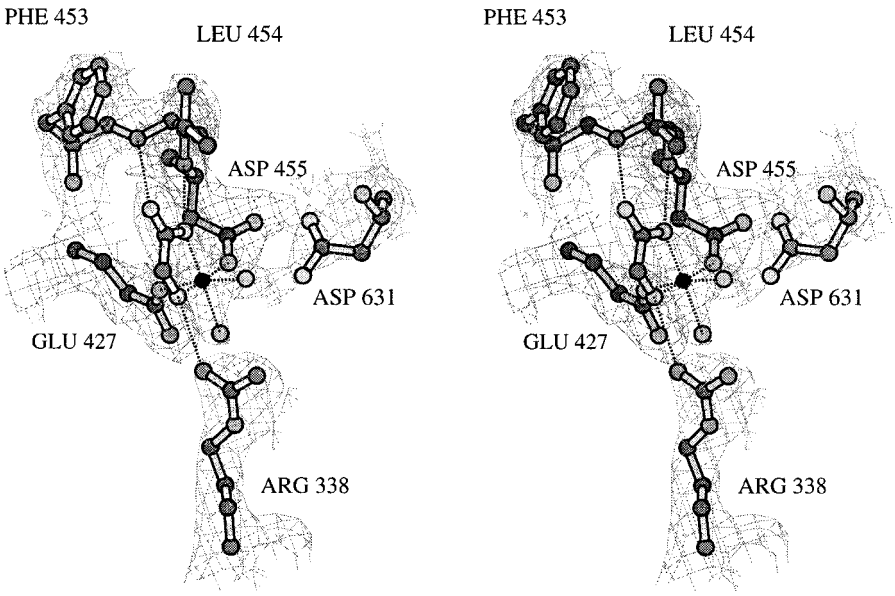

FIGURE 5: Experimental MAD-phased electron density map in the active site region. The contour cage of the map is at a level of 1 standard deviation or 0.344 $e^-/Å^3$ calculated to 2.2 Å resolution. Oxygen and nitrogen atoms are lightly shaded, while carbon is darkly shaded. Magnesium is black.

DISCUSSION

Two Critical Domains for Malate Synthase Activity. The sequence alignment of the known malate synthases, summarized in Figure 6, suggests three different families or types of malate synthase characterized by size: long, intermediate, and short. The enzyme from *Caenorhabditis elegans* fits into the intermediate family, but is the C-terminal 522 amino acids of a 1005-residue chimeric protein. The N-terminal 483 residues comprise an isocitrate lyase domain, and this is the only known example of the two enzymes unique to the glyoxylate cycle to be contained in the same polypeptide. Most malate synthases fall into the intermediate family, including those from higher plants and most known bacterial and fungal enzymes.

The first four sequences in the alignment belong to the long family and are all bacterial enzymes. This family is distinguished by the inclusion of an N-terminal extension which forms the N-terminal clasp that wraps around one side of the TIM barrel (see Figures 2 and 3). This extension is absent in all of the intermediate-sized malate synthases and in the single representative of the short family, the enzyme from the fungus *Laccaria bicolor*, which is the last sequence in the alignment. In addition to missing the N-terminal clasp, this enzyme seems also to be missing the inserted domain corresponding to residues 135−262 in MSG. This short malate synthase is only 419 residues long, and the sequence alignment for this enzyme is delayed until residue 255 in the MSG sequence. It therefore appears that both the N-terminal extension and the inserted domain found in the MSG structure are not crucial to the activity of the enzyme and can be deleted. The critical domains for malate synthase activity, it appears, are the barrel domain and the C-terminal plug which caps the active site, and which also contributes the proposed catalytic base (Asp631).

Possible Relative Mobility of the C-Terminal Domain. The C-terminal domain consists of residue 589 to the C-terminus of the protein. The loop formed by residues 577− 588 seems to be particularly flexible (Figure 2); the average temperature factors of the main chain atoms in residues 583−586 is 61.3 $Å^2$, while the average temperature factor for all main chain atoms in the model is 29.0 $Å^2$. The flexible connection may permit this domain to move relative to the rest of the enzyme. Such mobility would allow an opening of the active site cleft for substrate entrance and product release, and permit the enzyme to close and sequester the reactants from bulk solvent.

This proposed domain motion is consistent with results from low-angle X-ray scattering studies on trimeric malate synthase from baker's yeast. These studies showed that, upon substrate binding, there is a decrease in both the radius of gyration and the maximum particle diameter and a concomitant increase in the axial ratio of the enzyme (*66*). A conformational change upon substrate binding is also supported by substrate binding experiments as monitored by circular dichroism for both the yeast and maize enzymes (*67*, *68*). Proteolysis studies with the maize enzyme also support the notion of a flexible linkage which becomes rigid upon acetyl-CoA binding. Trypsin cleaves the maize enzyme into two domains of roughly 45 and 19 kDa as judged by SDS−PAGE, but cleavage is inhibited by addition of acetyl-CoA (*68*). An arginine−glycine bond is present in the maize enzyme in the position corresponding to residues 586−587 in the MSG sequence (see Figure 2). This is precisely the region with the highest temperature factors in the linkage to the C-terminal domain in MSG, and proteolysis of the maize enzyme by trypsin at this point would result in appropriately sized fragments.

It is interesting to note that after complete cleavage of the native protein into these two smaller fragments, the activity of the mixture is still 30% that of the intact enzyme. Finally, the fact that addition of sorbitol to the crystallization setups resulted in higher quality single crystals, combined with the observed position of sorbitol binding between the C-terminal domain and the long helix of the N-terminal α-helical clasp,

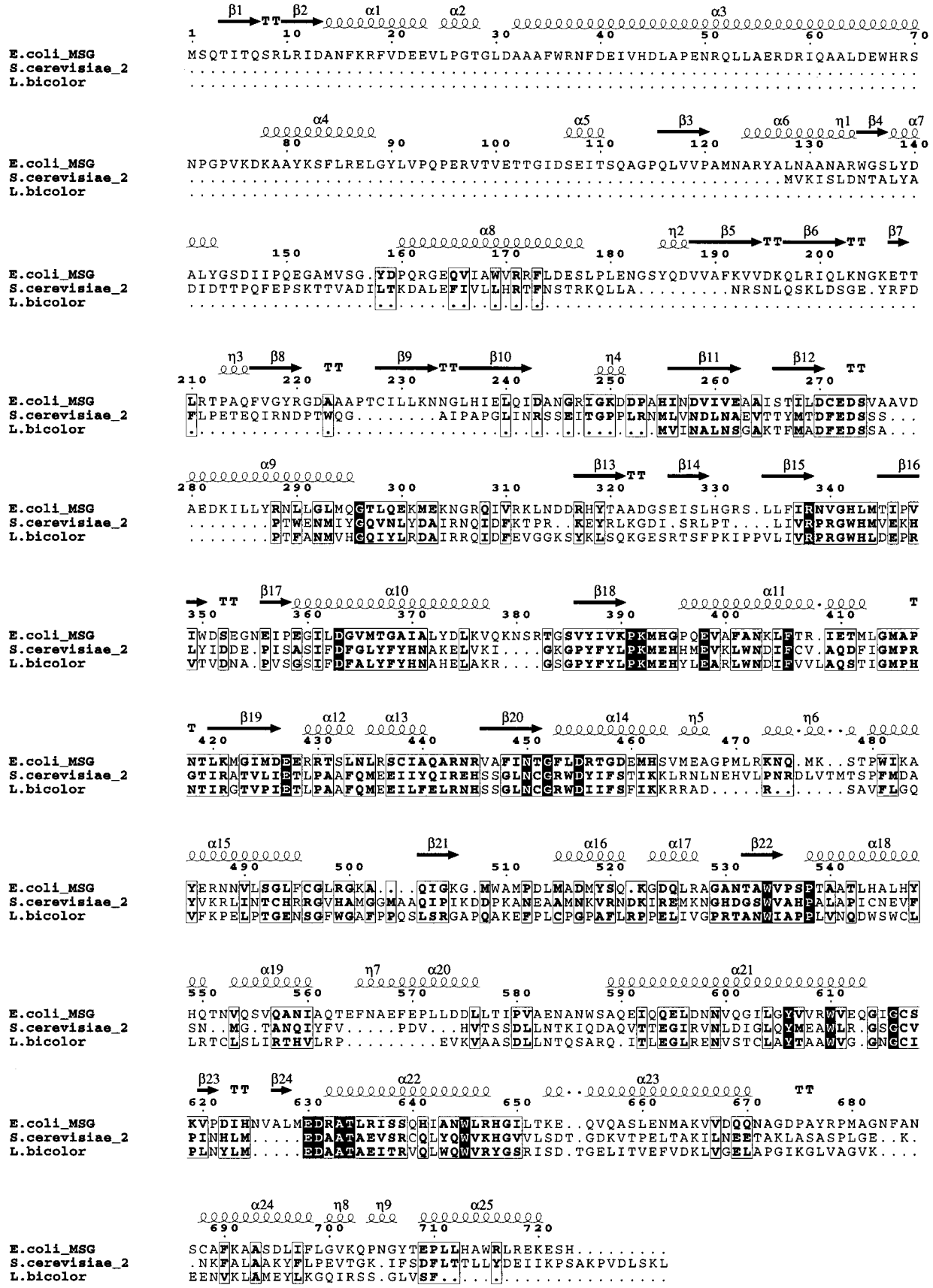

FIGURE 6: Synopsis of sequence alignment of 23 known malate synthases. The complete alignment is available as Supporting Information. The numbering corresponds to that found in the malate synthase G sequence (top). The secondary structure above the sequences also corresponds to that of MSG. Similarity is denoted by residues enclosed in a black box, while those residues which are strictly conserved in all 23 sequences are printed as white on a black background. The sequences are those of *E. coli*, MSG (*8*, long), *Saccharomyces cerevisiae*, MLS2 (*80*, intermediate), and *L. bicolor* (*81*, short).

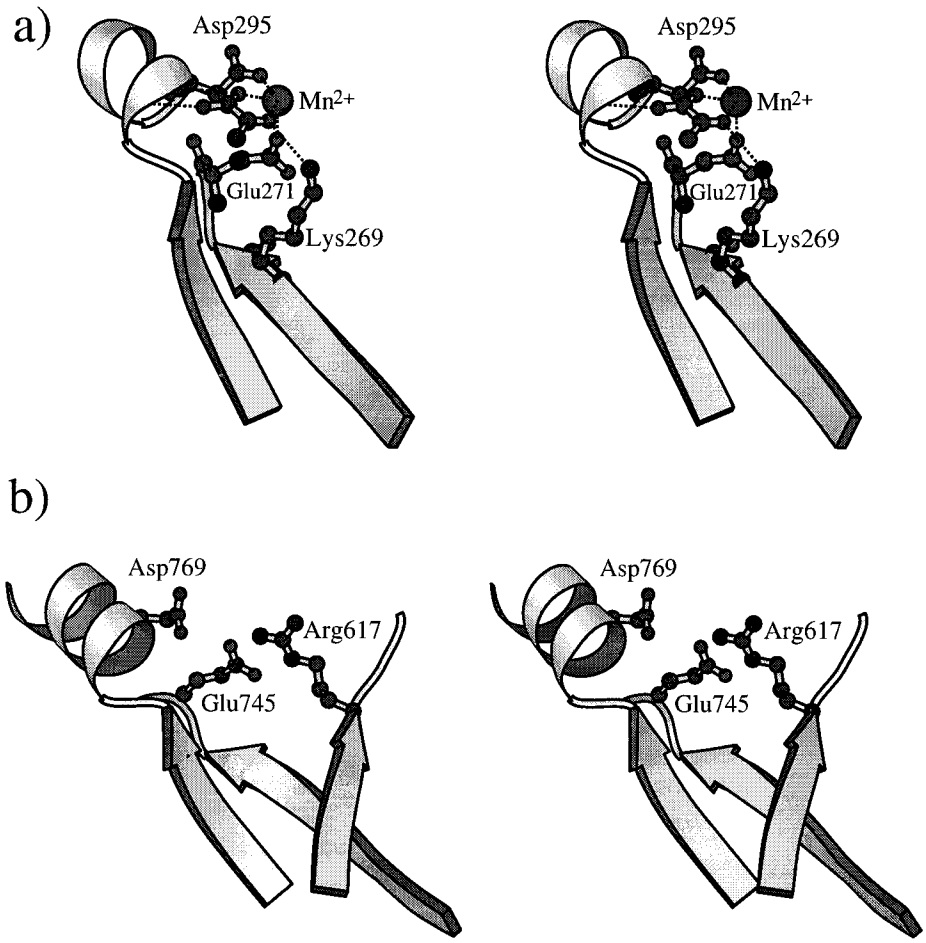

FIGURE 7: (a) Schematic representation of the pyruvate and manganese binding site in pyruvate kinase as seen in the crystal structure of the rabbit-muscle enzyme complex (*69*). (b) Schematic drawing of the proposed pyruvate binding site in pyruvate phosphate dikinase (*71*). The proposed divalent cation binding determinants, Asp769 and Glu745, are seen to be in positions in this structure identical to those in MSG, while Arg617 corresponds to Arg338.

is consistent with a model in which the C-terminal domain is mobile relative to the rest of the protein until restrained by sorbitol binding.

Structural Comparisons. Since MSG is centered around a TIM barrel, it is not surprising that a search using DALI indicates structural homology to other enzymes containing this motif. Two enzymes, however, stand out above the others as being structurally more closely related. These two enzymes are pyruvate kinase and pyruvate phosphate dikinase. It should be noted that pyruvate is a competitive inhibitor of malate synthase (*33*). As will be seen in the following, there are sufficiently great similarities among these enzymes that they may have diverged from a common ancestor.

Comparison to Pyruvate Kinase. The overall fold of the TIM barrel domain of pyruvate kinase (PK (*69*)) is very similar to that of MSG. This includes a helix following the sixth strand of the barrel corresponding to α-helix H in the MSG structure (see Figure 2). PK lacks an N-terminal extension preceding the barrel domain, but does have a C-terminal extension. Unlike MSG, however, the C-terminal domain formed by this extension folds into an α/β-structure which fits "under" the barrel, opposite from the active site. PK has a large insertion within the sequence of the barrel. Rather than the insert occurring after the first α-helix of the barrel fold, as in MSG, the extension follows the third strand of the barrel and forms a β-domain placed over the active site. It appears that this insert could provide a cap for the active site, analogous to the C-terminal plug in MSG. In the crystal structure of a complex of this enzyme with magnesium, potassium, and L-phospholactate, there are two tetramers per asymmetric unit (*70*). This inserted domain is observed in different positions relative to the TIM barrel domain, which results in different degrees of closure of the active site cleft which lies at the interface of these two domains. There is a >20° rotation of one domain relative to the other, from the most open to the most closed subunit in this structure.

In a comparison of substrate binding, the complex of rabbit-muscle PK with pyruvate and manganese (*69*) reveals a situation similar to that seen in the MSG/glyoxylate complex (Figure 7a). The carboxylate of the bound pyruvate molecule is hydrogen-bonded to the backbone amides of residues 294 and 295 at the N-terminus of the helix which immediately follows strand 6 of the TIM barrel. The manganese ligands of pyruvate kinase, Asp295, and Glu271, are structurally equivalent to Asp455 and Glu427 in MSG. Rather than an arginine donated from strand 3 of the barrel as in MSG, however, Lys269 extends from lower within the barrel on strand 5 to coordinate the ketone of the pyruvate molecule.

Comparison to Pyruvate Phosphate Dikinase. Like pyruvate kinase, pyruvate phosphate dikinase (PPDK) also shares

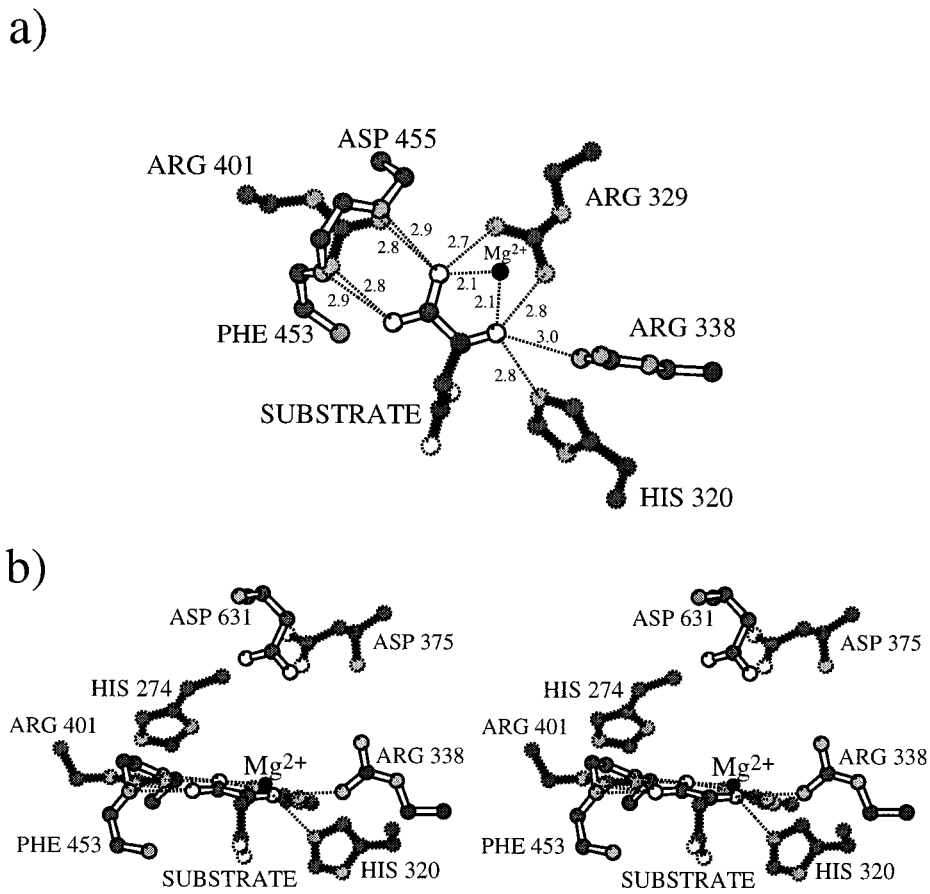

FIGURE 8: (a) "Top" view of an active site overlay of the citrate synthase (*74*) and malate synthase complexes based on the carbonyl and carboxylate functions of the electrophilic substrate. Atom spheres and bonds in citrate synthase are stippled. Bonds unique to the citrate synthase complex are shown in black, while those in the malate synthase complex are shown in white for clarity; Arg421 in citrate synthase, which is contributed by the neighboring molecule to bind the second carboxylate group, is not shown. Hydrogen bonds or metal−oxygen bonds are depicted as thin dotted lines, with distances shown in Å. Oxygen atoms are unshaded, nitrogen atoms lightly shaded, and carbon atoms darkly shaded. (b) "Side" view of the above.

a TIM barrel fold very similar to that of MSG (*71*). PPDK also has a helix corresponding to αH in MSG following strand 6 of the barrel, and has an insert (albeit shorter) after strand 3 of the barrel. PPDK has a large N-terminal extension which comprises two distinct domains: an N-terminal nucleotide-binding domain followed by a phosphohistidine domain. The phosphohistidine domain is thought to couple the N-terminal nucleotide-binding domain and the C-terminal PEP/pyruvate-binding domain by swiveling from one to the other around two flexible peptide linkers (*71, 72*). Although no substrate complex is available for the PPDK structure, and no divalent metal was included in the crystallization, PPDK appears to be more closely related to MSG than is pyruvate kinase. Figure 7b is a schematic representation of the N-terminus of the helix following strand 6 of the barrel where pyruvate is presumed to bind (*71*). PPDK has not only identical aspartate and glutamate ligands, but in addition an arginine extending from strand 3 corresponding to Arg338 in MSG (compare to Figure 4).

Structural Comparison with Citrate Synthase. Although the reaction of citrate synthase is very similar to that of malate synthase as noted earlier, the overall folds of these enzymes are completely different. Dimeric citrate synthase (monomeric versions have not been reported) is almost entirely α-helical with each active site formed by a cleft between the large and small domains of one subunit near the dimer interface. This enzyme undergoes a large conformational change, in which the smaller domain rotates away from the main body of the dimer by approximately 18° to open the active site for substrate entrance and product release (*73*).

Comparisons of Malate Synthase and Citrate Synthase Active Sites. The crystal structure of the CS/oxaloacetate/amidodethia/CoA complex (PDB ID code 1CSH) was used for structural comparison (*74*). An overlay of the 2-keto acid portions of the electrophilic substrates in each reaction was performed by hand. This permitted straightforward comparison of the interactions which bind and activate the substrates within the active sites of these two enzymes.

Activation of the Electrophilic Substrate. When the 2-keto acid portions of oxaloacetate in the CS complex are superimposed with the same portion of glyoxylate in the MSG complex, interesting mechanistic implications emerge (Figure 8). The terminal nitrogen atoms of the guanidinium group of Arg401 in citrate synthase (CS-Arg401) overlay closely with the backbone amides of residues 454 and 455 in the MSG complex. The bound magnesium ion in malate synthase is close to the axis of the NE−CZ bond of CS-Arg329 but closer to the substrate, nevertheless providing a positively charged coordination site analogous to the terminal nitrogens of CS-Arg329. The nitrogen of CS-His320 which hydrogen bonds to the ketone oxygen of oxaloacetate is placed 2.5 Å from the position of the nitrogen in MSG-Arg338 which forms a hydrogen bond to the aldehyde

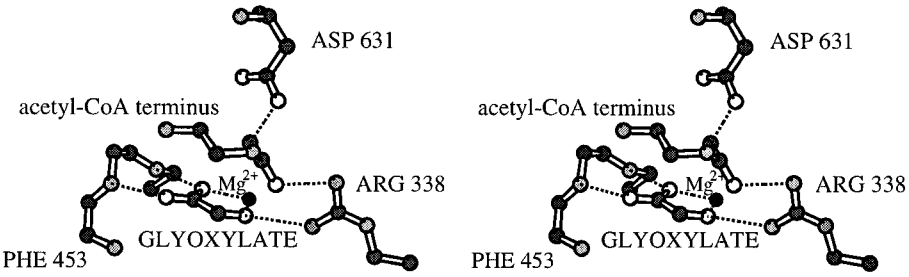

FIGURE 9: Hypothetical model of acetyl-CoA terminal atoms bound in the active site of MSG. Oxygen is shown as open spheres, nitrogen lightly shaded, and carbon darkly shaded.

oxygen of glyoxylate. Although these hydrogen bond donors to the carbonyl oxygen are 50° apart in the overlay, the angles between the two ligands in each individual complex are similar: the magnesium−oxygen−arginine angle in MSG is 103°, while the analogous arginine−oxygen−histidine angle in CS is 96°. Both modes of substrate coordination are expected to withdraw electrons from the substrate carbonyl carbon, activating it for nucleophilic attack.

*Enolization Reaction.*The rate-limiting step in both CS and MSG is the enolization of the nucleophilic substrate. How MSG is able to catalyze the deprotonation of the terminal acetyl group of acetyl-CoA is a major focus of this structural study. Soaking experiments with acetyl-CoA or analogues have so far failed to provide direct crystallographic evidence for the mode of acetyl-CoA binding; however, active site comparisons suggest plausible roles for various side chains. The residues in CS which carry out the proton abstraction are Asp375 and His274 acting in concert (*24*, *26*, *41*). The overlay of the *substrates* of MSG and CS bring the MSG-Asp631 carboxylate to within 1.3 Å of the CS-Asp375 carboxylate (Figure 8b). This close proximity suggests that MSG-Asp631 acts as the catalytic base in the reaction.

On the other hand, several observations have led to proposals in the literature that a sulfhydryl is the catalytic base in the MS reaction. MS is inactivated and aggregates upon exposure to X-rays, a process that can be reversed by addition of DTT (*75*). Derivatization with DTNB also inactivates the enzyme (*76*). These observations can be rationalized on the basis of the crystal structure. There is a highly conserved cysteine residue (Cys617 in the MSG sequence; see Figures 2 and 6) which is in the flexible active site loop of the C-terminal domain. If this cysteine were to form an intermolecular disulfide bond, this would prevent the loop from folding into the active site in a productive manner, misplacing the proposed catalytic Asp631. Likewise, derivatization with DTNB would also inactivate the enzyme by sterically preventing the proper closure of the active site for catalysis. Thus, the evidence for a sulfhydryl base appears to be circumstantial, and in any event, there are no strictly conserved cysteine residues in the known sequences.

An inspection of the active site overlay reveals no candidate for a residue in a position analogous to that of CS-His274, the catalytic acid in the CS reaction. There are two possibilities, either that one is not required and the enolization reaction is not accomplished by concerted acid/base catalysis on this enzyme or that some other group fulfills the function of CS-His274. On the basis of the following considerations, we prefer the latter. Given the present MSG model, the only candidate for a hydrogen bond donor or charged group to stabilize the enol(ate) intermediate is the nitrogen of the MSG-Arg338 guanidinium group that is not involved in hydrogen bond formation. In Figure 8b, this nitrogen can be seen to be positioned at approximately the same level above the reactive carbonyl of the electrophilic substrate, but opposite CS-His274. This positioning nevertheless suggests that MSG-Arg338 could act in a fashion analogous to that of CS-His274, either by protonating the enol oxygen or by stabilizing an enolate intermediate by virtue of its positive charge. This placement also suggests that Arg338 could act as a bridge between the two substrates in the MSG reaction. It could both stabilize the enol(ate) in the first step of the reaction and also help stabilize (in conjunction with the magnesium ion) the oxyanion resulting from the condensation step, thereby taking the place of the opposing histidines (His320 and His274) in citrate synthase. This interaction would also seem to be important to steer the reactive enol(ate) for proper reaction with the electrophilic substrate and avoidence of unwanted side reactions.

Direct participation of Mg^{2+} in the enolization step seems to be excluded by the tight coordination of the cation with glyoxylate. The enolate oxygen would probably have to replace a solvent molecule (Wat 1002) in the Mg coordination sphere, which is sterically disallowed by the active site configuration of the present model.

Proposed Mechanism of Malate Synthase. The overlay of the MSG/magnesium/glyoxylate complex with the CS/oxaloacetate/AMX complex provides an interesting comparison of similar interactions in both active sites, but with different sets of functional groups. We propose that, in both enzymes, the electrophilic substrate is polarized for nucleophilic attack in the ground state by hydrogen bonds and nearby positive charges. Mg^{2+} is essential for this step in MS, and is positioned by strictly conserved residues Glu427 and Asp455. We further propose that, in MSG, Arg338 is the partner to Asp631 in the enolization step. Modeling the terminal seven non-hydrogen atoms of acetyl-CoA into the structure of the malate synthase G complex shows that displacement of solvent molecules in the active site would provide room for this group without serious steric interference with neighboring atoms, although the geometry is not ideal (Figure 9). A comparison with the enolization partners in the CS complex (Asp375 and His274) reveals that the distance between the proposed analogous partners in the malate synthase complex is 2 Å shorter: 4.3 Å vs 6.3 Å. It is therefore likely that when acetyl-CoA binds to the enzyme, the C-terminal plug is displaced slightly relative to that seen in the current complex, positioning Asp631 more appropriately for proton abstraction. Malate synthase binds acetyl-CoA in the absence of the second substrate, but does not catalyze exchange of protons between solvent and acetyl-

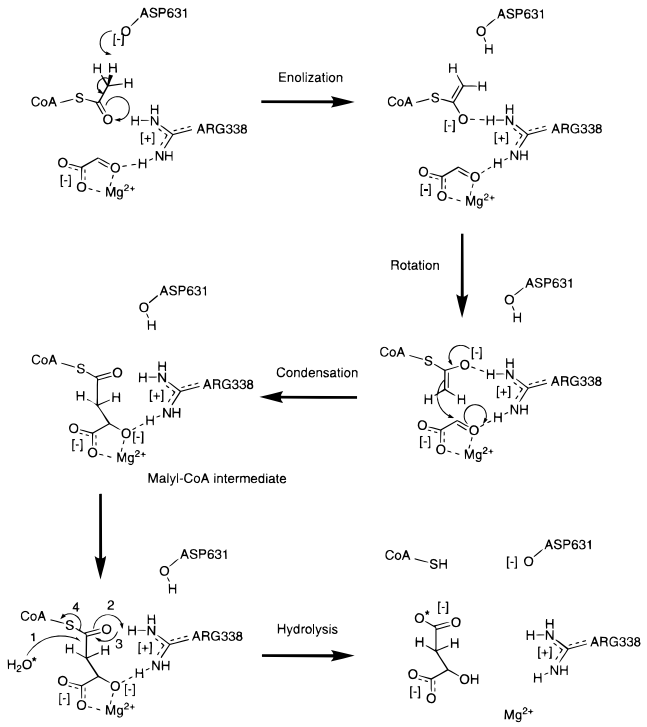

FIGURE 10: Proposed mechanism of malate synthase G.

CoA at an appreciable rate unless Mg^{2+} and a substrate analogue such as pyruvate or oxaloacetate are present (*33*). Therefore, both substrates must be present for the active site geometry to attain the configuration required for the activation of acetyl-CoA. This is also consistent with our observation that acetyl-CoA analogues do not appear to bind to the molecule in the present crystal form.

The proposed mechanism for malate synthase, in which an enolate intermediate is formed, is depicted in Figure 10. In the first step, Asp631 acts as the catalytic base to deprotonate the terminal acetyl group of acetyl-CoA, forming the enolate intermediate, during which the developing negative charge on the enolate oxygen is stabilized by the adjacent positive charge on Arg338. This is consistent with the observation that MS shows little discrimination with regard to the pro-*R* versus the pro-*S* hydrogens in fluoroacetyl-CoA (*35*, *36*), resulting in either a *Z*- or *E*-enolate intermediate. A model of charge−charge stabilization in this enolization step rather than proton donation to form the neutral intermediate is also supported when one considers the groups neighboring Arg338. Asp270 forms a bidentate salt bridge to the guanidinium group, and the carboxylate of Glu272 is within 3.4 Å in a direction orthogonal to the Asp−Arg salt bridge (Figure 4). These close negative charges would stabilize a positive charge on Arg338, making proton donation to the enolate oxygen highly unlikely.

Once this high-energy intermediate is formed, it could then rotate around the oxyanion−guanidinium hydrogen bond to bring the methylene carbon into range for nucleophilic attack on the 2-carbon of glyoxylate. In the CS reaction, His320 has been proposed to simultaneously donate a proton to the ketone oxygen of oxaloacetate as it condenses with the enol of acetyl-CoA to form citryl-CoA (*24*, *41*). For reasons discussed above, such a role for MSG-Arg338 seems unlikely. The nearby presence of the charged guanidinium moiety of Arg338 and Mg^{2+} would suggest that the resulting oxyanion is stabilized until product release when protons become available from the solution or the (presumably activated) water molecule which hydrolyzes the thioester.

Solution studies are consistent with an oxyanion intermediate. Malyl-CoA is a very poor substrate of the enzyme (*33*), with the maximum rate of hydrolysis of 1/1000 that of the physiological reaction. One possible explanation for this, along the lines of a suggestion by Lill et al. (*77*), is that the 2-hydroxyl of malyl-CoA must be deprotonated for this compound to bind to the Mg/enzyme complex. Such an anionic form would not be easily formed in solution.

Thus, we propose that while CS appears to stabilize a primarily neutral enol intermediate via hydrogen bonding to neutral His274, MSG stabilizes first an enolate and then an oxyanion malyl-CoA intermediate via the positively charged Arg338 and Mg^{2+}. However, the CS and MS reactions both appear to involve concerted chemistry on the part of two side chains on opposite sides of the active site cleft. In the absence of a structure of a bound malyl-CoA analogue, it is not possible to provide evidence for the groups that might participate in the hydrolysis reaction.

Short, Strong Hydrogen Bonds versus Weak Acids in the Enzymatic Reaction. Citrate synthase has become one focus of a debate on the importance of "short, strong" (donor−acceptor distance $<\sim$2.5 Å) hydrogen bonds between pK_a matched groups in enzymatic reactions (*18*, *42*). Specifically, it has been proposed that the enol intermediate generated in the CS reaction forms an exceptionally strong hydrogen bond with neutral His274, and that the strength of this interaction is sufficient to lower the energy of the intermediate by >8 kcal/mol to account for the observed kinetics of proton abstraction. The pK_a of a neutral histidine is about 14 (*42*), and it would be anionic when deprotonated. In another proposal, the developing negative charge on the enolate could be delocalized by this interaction, and theoretical electrostatic calculations have demonstrated that this delocalization would also be sufficient to account for the observed rate enhancement without invoking an unusually strong hydrogen bond (*26*). These proposals have been controversial, and several crystallographically observed examples of very short hydrogen bonds on enzymes have been shown not to have the very high energies of formation suggested by the strong hydrogen bond hypothesis (*74*, *78*). As reviewed by Perrin and Nielson, evidence for very strong hydrogen bonds in enzymatic catalysis is at best inconclusive (*79*).

If an enolate intermediate in MSG is indeed stabilized by a formal positive charge on Arg338, this interaction would appear to be fundamentally different from either of those proposed for citrate synthase (or other enzymes) as the overall charge on the enolate−Arg pair is zero. On the other hand, both neutral His274 and positively charged Arg338 are very weak acids and from this point of view could play similar roles in the reaction. Resolution of this dichotomy will undoubtedly yield new insight into the chemistry of carbon acids in enzymatic reactions.

CONCLUSIONS

Malate synthase represents a clear example of both divergent and convergent evolution in enzyme mechanisms. A case for divergent evolution is argued by the close similarity in structure and mode of substrate recognition to

other TIM barrel enzymes, particularly pyruvate kinase and pyruvate phosphate dikinase. These enzymes are based on a common fold and bind similar substrates, and yet have evolved different extensions and insertions to complete the architecture of their active sites to serve their different metabolic roles. At the same time the enzyme is a clear example of convergent evolution in a functional sense. The two very different protein folds found in malate synthase and citrate synthase have evolved to converge to catalyze a very similar reaction, in which carbon−carbon bond formation involves an electrophilic substrate activated by nearby positive charges and hydrogen bonds, while the nucleophilic substrate is deprotonated by an aspartate residue acting in concert with a weakly acidic partner.

ACKNOWLEDGMENT

We thank Karen Kallio for expert technical assistance and Marc-André Elsliger and Dale Tronrud for assistance with computer software.

SUPPORTING INFORMATION AVAILABLE

Expanded version of Figure 6. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Wong, D. T. O., and Ajl, S. J. (1956) *J. Am. Chem. Soc. 78*, 3230−3231.
2. Kornberg, H. L., and Krebs, H. A. (1957) *Nature 179*, 988−991.
3. Cioni, M., Pinzauti, G., and Vanni, P. (1981) *Comp. Biochem. Physiol. 70B*, 1−26.
4. Holmes, R. P. (1993) *Biochim. Biophys. Acta 1158*, 47−51.
5. Benevides, J. M., Tremblay, G. C., and Hammen C. S. (1989) *Comp. Biochem. Physiol. 94B*, 779−782.
6. Goodman, D. B. P., Davis, Walter L., and Jones, Ruth G. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 1521−1525.
7. Davis, W. L., and Goodman, David B. P. (1992) *Anatom. Rec. 234*, 461−468.
8. Molina, I., Pellicer, M., Badia, J., Aguilar, J., and Baldoma, L. (1994) *Eur. J. Biochem. 224*, 541−548.
9. Woodcock, E., and Merrett, M. J. (1978) *Biochem. J. 173*, 95−101.
10. Durchschlag, H., Biedermann, G., and Eggerer, H. (1981) *Eur. J. Biochem. 114*, 255−262.
11. Khan, A. S., Van Driessche, E., Kanarek, L., and Beeckmans, S. (1993) *Protein Exp. Purif. 4*, 519−528.
12. Cornforth, J. W., Redmond, J. W., Eggerer, H., Buckel, W., and Gutschow, C. (1969) *Nature 221*, 1212−1213.
13. Luthy, J., Retey, J., and Arigoni, D. (1969) *Nature 221*, 1213−1215.
14. Clark, J. D., O'Keefe, S. J., Knowles, J. R. (1988) *Biochemistry 27*, 5961−5971.
15. Gerlt, J. A., Kozarich, J. W., Kenyon, G. L., and Gassman, P. G. (1991) *J. Am. Chem. Soc. 113*, 9667−9669.
16. Kresge, J. A. (1991) *Pure Appl. Chem. 63*, 213−221.
17. Remington, S. J. (1992) *Curr. Opin. Struct. Biol. 2*, 730−735.
18. Cleland, W. W., Frey, P. A., and Gerlt, J. A. (1998) *J. Biol. Chem. 273*, 25529−25532.
19. Srere, P. A. (1972) *Curr. Top. Cell. Regul. 5*, 245−283.
20. Spector, L. B. (1972) *Citrate Cleavage and Related Enzymes*, 3rd ed., Vol. 7, Academic Press, New York.
21. Weitzman, P. D. J., and Danson, M. J. (1976) *Curr. Top. Cell. Regul. 10*, 161−204.
22. Beeckmans, S. (1984) *Int. J. Biochem. 16*, 341−351.
23. Wiegand, G., and Remington, S. (1986) *Annu. Rev. Biophys. Biophys. Chem. 15*, 97−117.
24. Alter, G. M., Casazza, J. P., Wang, Z., Nemeth, P., Srere, P., and Evans, C. T. (1990) *Biochemistry 29*, 7557−7563.
25. Zhi, W., Srere, P. A., and Evans, C. T. (1991) *Biochemistry 30*, 9281−9286.
26. Evans, C. T., Kurz, L. C., Remington, S. J., and Srere, P. A. (1996) *Biochemistry 35*, 10661−10672.
27. Kurz, L. C., Drysdale, G. R., Riley, M. C., Evans, C. T., and Srere, P. A. (1992) *Biochemistry 31*, 7908−7914.
28. Kurz, L. C., Nakra, T., Stein, R., Plungkhen, W., Riley, M., Hsu, F., and Drysdale, G. R. (1998) *Biochemistry 37*, 9724−9737.
29. Kurz, L. C., Shah, S., Crane, B. R., Donald, L. J., Duckworth, H. W., and Drysdale, G. R. (1992) *Biochemistry 31*, 7899−7907.
30. Kurz, L. C., Ackerman, J. J., and Drysdale, G. R. (1985) *Biochemistry 24*, 452−457.
31. Kurz, L. C., and Drysdale, G. R. (1987) *Biochemistry 26*, 2623−2627.
32. Dixon, G. H., Kornberg, H. L., and Lund, P. (1960) *Biochim. Biophys. Acta 41*, 217−233.
33. Eggerer, H., and Klette, A. (1967) *Eur. J. Biochem. 1*, 447−475.
34. Stern, J. R. (1961) *The Enzymes*, 2nd ed., Vol. 5, Academic Press, New York.
35. O'Hagan, D., and Rzepa, H. S. (1994) *J. Chem. Soc., Chem. Commun.* 2029.
36. O'Hagan, D., and Rzepa, H. S. (1997) *Chem. Commun. 7*, 645−652.
37. Fanshier, D. W., Gottwald, L. K., and Kun, E. (1964) *J. Biol. Chem. 239*, 425.
38. Keck, R., Haas, H., and Retey, J. (1980) *FEBS Lett. 114*, 287.
39. Marletta, M. A., Srere, P. A., and Walsh, C. (1981) *Biochemistry 20*, 3719.
40. Bloxham, D. P., Parmelee, D. C., Kumar, S., Wade, R. D., Ericsson, L. H., Neurath, H., Walsh, K. A., Titani, K. (1981) *Proc. Natl. Acad. Sci. U.S.A. 78*, 5381.
41. Karpusas, M., Branchaud, B., and Remington, S. J. (1990) *Biochemistry 29*, 2213−2219.
42. Gerlt, J. A., and Gassman, P. G. (1993) *Biochemistry 32*, 11943−11952.
43. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. (1988) *Science 239*, 487−491.
44. Van-Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L., and Clardy, J. (1993) *J. Mol. Biol. 229*, 105−124.
45. Barofsky, L. Personal communication, Mass Spectrometry Facility, Oregon State University, Corvallis, OR.
46. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymology 276*, 307−326.
47. Weeks, C. M., and Miller, R. *J. Appl. Crystallogr.* (in press).
48. La Fortelle, E. d., and Bricogne, G. (1997) *Methods Enzymol. 276*, 472−494.
49. Read, R. J. a. S., A. J. (1988) *J. Appl. Crystallogr. 21*, 490−495.
50. Collaborative Computational Project, IV. (1994) *Acta Crystallogr. D50*, 760−763.
51. Cowtan, K. (1994) *Joint CCP4 ESF-EACBM Newslett. Protein Crystallogr. 31*, 34−38.
52. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
53. Tronrud, D. E., Ten Eyck, L. F., and Matthews, B. W. (1987) *Acta Crystallogr. A43*, 489−503.
54. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res. 22*, 4673−4680.
55. Gouet, P., Courcelle, E., Stuart, D., and Metoz, F. (1999) *Bioinformatics 15* (4), 305−308.
56. Kabsch, W., and Sander, C. (1983) *Biopolymers 22*, 2577−2637.
57. Holm, L., and Sander, C. (1993) *J. Mol. Biol. 233*, 123−138.
58. CambridgeSoft Corp., 875 Massachusetts Ave., Cambridge, MA 02139.
59. Macromedia, Inc., 600 Townsend St., San Fransisco, CA 94103.
60. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.

61. Esnouf, R. M. (1997) *J. Mol. Graphics 15*, 132−134.
62. Matthews, B. W. (1968) *J. Mol. Biol. 33*, 491−497.
63. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
64. Banner, D. W., Bloomer, A. C., Petsko, G. A., Phillips, D. C., Pogson, C. I., and Wilson, I. A. (1975) *Nature 255*, 609−614.
65. Branden, C., and Tooze, J. (1991) *Introduction to Protein Structure*, Garland Publishing Inc., New York.
66. Durchschlag, H. a. Z., P. (1977) *Biochem. Biophys. Res. Commun. 75*, 394−400.
67. Schmid, G., Durchschlag, H., Biedermann, G., Eggerer, H., and Jaenicke, R. (1974) *Biochem. Biophys. Res. Commun. 58*, 419−426.
68. Beeckmans, S., Khan, A. S., Kanarek, L., and Van Driessche, E. (1994) *Biochem. J. 303*, 413−421.
69. Larsen, T. M., Laughlin, L. T., Holden, H. M., Rayment, I., and Reed, G. H. (1994) *Biochemistry 33*, 6301−6309.
70. Larsen, T. M., Benning, M. M., Wesenberg, G. E., Rayment, I., and Reed, G. H. (1997) *Arch. Biochem. Biophys. 345*, 199−206.
71. Herzberg, O., Chen, C. C. H., Kapadia, G., McGuire, M., Carroll, L. J., Noh, S. J., and Dunaway-Mariano, D. (1996) *Proc. Natl. Acad. Sci. U.S.A. 93*, 2652−2657.
72. McGuire, M., Huang, K., Kapadia, G., Herzberg, O., and Dunaway-Mariano, D. (1998) *Biochemistry 37*, 13463−13474.
73. Remington, S. J., Wiegand, G., and Huber, R. (1982) *J. Mol. Biol. 158*, 111−152.
74. Usher, K. C., Remington, S. J., Martin, D. P., and Drueckhammer, D. G. (1994) *Biochemistry 33*, 7753−7759.
75. Durchschlag, H., and Zipper, P. (1988) *FEBS Lett. 237*, 208−212.
76. Durchschlag, H., and Zipper, P. (1981) *Z. Naturforsch. 36c*, 516−533.
77. Lill, U., Kollmann-Koch, A., Bibinger, A., and Eggerer, H. (1991) *Eur. J. Biochem. 198*, 767−773.
78. Fraser, M. E., Strynadka, N. C. J., Bartlett, P. A., Hanson, J. E., and James, M. N. G. (1992) *Biochemistry 31*, 5201−5214.
79. Perrin, C. L., and Nielson, J. B. (1997) *Annu. Rev. Phys. Chem. 48*, 511−544.
80. Fernandez, E., Fernandez, M., and Rodicio, R. (1993) *FEBS Lett. 320*, 271−275.
81. Kim, S. J., and Podila, G. K. (1997) Biological Sciences, Michigan Technological University, 1400 Townsend Dr., Houghton, MI 49931.

BI992519H