Structural Definition of the Active Site and Catalytic Mechanism of 3,4-Dihydroxy-2-butanone-4-phosphate Synthase‡

Der-Ing Liao,*,§ Ya-Jun Zheng,|| Paul V. Viitanen,§ and Douglas B. Jordan*,⊥

Experimental Station, DuPont Central Research and Development, Wilmington, Delaware 19880, Stine-Haskell Research Center, DuPont Agricultural Products, Newark, Delaware 19714, and Experimental Station, Bristol-Myers Squibb Company, Route 141 and Henry Clay Road, Wilmington, Delaware 19880-0357

Received August 7, 2001; Revised Manuscript Received October 9, 2001

ABSTRACT: X-ray crystal structures of L-3,4-dihydroxy-2-butanone-4-phosphate synthase from *Magnaporthe grisea* are reported for the E−SO_4^{2-}, E−SO_4^{2-}−Mg^{2+}, E−SO_4^{2-}−Mn^{2+}, E−SO_4^{2-}−Mn^{2+}−glycerol, and E−SO_4^{2-}−Zn^{2+} complexes with resolutions that extend to 1.55, 0.98, 1.60, 1.16, and 1.00 Å, respectively. Active-site residues of the homodimer are fully defined. The structures were used to model the substrate ribulose 5-phosphate in the active site with the phosphate group anchored at the sulfate site and the placement of the ribulose group guided by the glycerol site. The model includes two Mg^{2+} cations that bind to the oxygen substituents of the C2, C3, C4, and phosphate groups of the substrate, the side chains of Glu37 and His153, and water molecules. The position of the metal cofactors and the substrate's phosphate group are further stabilized by an extensive hydrogen-bond and salt-bridge network. On the basis of their proximity to the substrate's reaction participants, the imidazole of an Asp99−His136 dyad from one subunit, the side chains of the Asp41, Cys66, and Glu174 residues from the other subunit, and Mg^{2+}-activated water molecules are proposed to serve specific roles in the catalytic cycle as general acid−base functionalities. The model suggests that during the 1,2-shift step of the reaction, the substrate's C3 and C4 hydroxyl groups are cis to each other. A cis transition state is calculated to have an activation barrier that is 2 kcal/mol greater than that of the trans transition state in the absence of the enzyme.

L-3,4-Dihydroxy-2-butanone-4-phosphate synthase (DS)[1] catalyzes the conversion of ribulose 5-phosphate to L-3,4-dihydroxy-2-butanone 4-phosphate (DHBP) and formate in a commitment step of riboflavin biosynthesis (*1*−*11*). A proposal for the chemical transformation on the enzyme involves a series of enolization, dehydration, ketonization, skeleton rearrangement, hydration, retro-aldol reaction, and protonation steps that account for the determined regiochemistry and stereochemistry (Figure 1) (*2*−*4*, *7*). The enolization steps are reminiscent of the triose phosphate isomerase (TIM)-catalyzed isomerization of D-glyceraldehyde 3-phosphate to dihydroxyacetone phosphate (*12*, *13*). The skeleton shift step is reminiscent of the xylose isomerase-catalyzed reactions (*14*−*18*), the acetohydroxy acid isomeroreductase-catalyzed reactions (*19*−*23*), and the 1-deoxy-D-xylulose 5-phosphate reductoisomerase-catalyzed reactions (*24*, *25*).

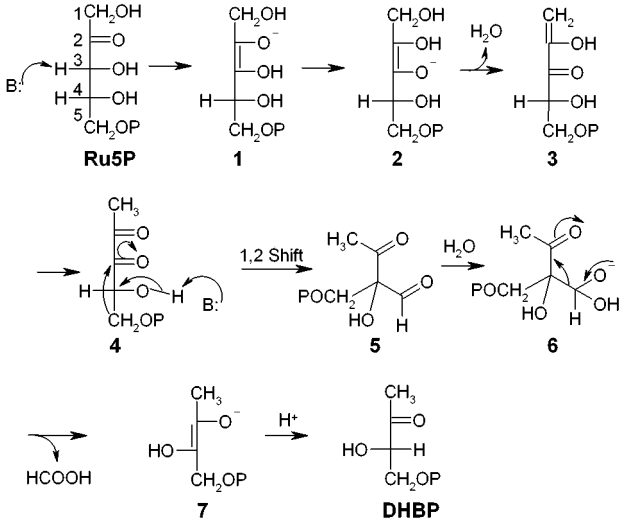

FIGURE 1: Reaction scheme for the DS-catalyzed reaction. The scheme that was formulated from the results of isotope labeling experiments is adapted from Volk and Bacher (*2*−*4*).

It has been suggested that sugar isomerases that act on nonphosphorylated substrates and have divalent metal ion requirements (usually Mg^{2+} or Mn^{2+}) follow a mechanism different from that of sugar isomerases that lack the metal requirement and act on phosphosugars (*14*). DS has a requirement for divalent metal (Mg^{2+}) and acts on a phosphosugar, so one might imagine that it could belong to either category or that it combines the catalytic features of both classes of sugar isomerases.

‡ Atomic coordinates of the enzyme complexes prepared for this work have been deposited in the Protein Data Bank as entries 1K49 (E−SO_4^{2-}, complex **I**), 1K4I (E−SO_4^{2-}−Mg^{2+}, complex **II**), 1K4L (E−SO_4^{2-}−Mn^{2+}, complex **III**), 1K4O (E−SO_4^{2-}−Mn^{2+}−glycerol, complex **IV**) and 1K4P (E−SO_4^{2-}−Zn^{2+}, complex **V**).

* To whom correspondence should be addressed. D.B.J.: e-mail, jordandbenz@yahoo.com. D.-I.L.: telephone, (302) 695-8332; fax, (302) 695-1351; e-mail, der-ing.liao@usa.dupont.com.

§ DuPont Central Research and Development.

|| DuPont Agricultural Products.

⊥ Bristol-Myers Squibb Company.

[1] Abbreviations: DS, L-3,4-dihydroxy-2-butanone-4-phosphate synthase; DHBP, L-3,4-dihydroxy-2-butanone 4-phosphate; MES, 2-morpholinoethanesulfonic acid; PEG, polyethylene glycol; Ru5P, ribulose 5-phosphate; TS, transition state.

10.1021/bi015652u CCC: $22.00
Published on Web 01/18/2002

Table 1: Data Collection Statistics[a]

	cation-free (**I**)	Mg^{2+} (**II**)	Mn^{2+} (**III**)	Mn^{2+} (**IV**)	Zn^{2+} (**V**)
resolution (Å)	30−1.50	30−0.98	30−1.60	30−1.10	30−1.00
highest-resolution shell (Å)	1.53−1.50/1.58−1.55[b]	1.00−0.98	1.63−1.60	1.12−1.10/1.18−1.16[b]	1.02−1.00
wavelength (Å)/ temperature (°C)	1.54/−168	1.00/−173	1.54/−168	1.00/−173	1.00/−173
space group	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$
cell dimensions *a*, *b*, *c* (Å)	53.70, 87.84, 43.96	53.89, 88.14, 43.99	53.37, 88.35, 43.75	53.61, 88.07, 43.85	53.72, 88.53, 43.99
no. of observations/ no. of unique reflections	110275/30209	864084/117559	106303/26458	336658/73399	686847/105091
completeness (%)	88.6 (36.0/65.7[b])	97.4 (70.6)	94.3 (70.6)	86.4 (38.1/66.2[b])	92.4 (60.7)
$I/\sigma(I)$	26.7 (3.5/5.7[b])	31.2 (2.9)	21.5 (2.0)	18.6 (2.1/3.2[b])	27.9 (2.0)
R_{merge}	0.042	0.086	0.060	0.078	0.054

[a] Each data set was collected on a single crystal. $R_{merge} = \sum_{hkl}[\sum_n |I(hkl,n) - \langle I(hkl)\rangle|]/\sum_{all\ observations} I$, where n is the number of observations for each unique reflection. The numbers in parentheses are for the highest-resolution shells. [b] The highest-resolution cell with a completeness of at least 60%.

DS may be combined with GTP cyclohydrolase II to form a single polypeptide (*11*) or the two functions may constitute two separate proteins (*1*−*10*), depending on the biological species of origin. Notably, GTP cyclohydrolase II catalyzes the other commitment step of riboflavin biosynthesis. The monofunctional DS is a homodimer of 23−24 kDa subunits. Recently, the X-ray structure of this enzyme from *Escherichia coli* was reported (*26*). The *E. coli* DS is an $\alpha+\beta$ protein, which appears to represent a novel fold among protein structures according to a VAST search (*27*) of currently available structures in the database. The two active sites of the *E. coli* enzyme are located at the dimer interface where the conserved residues among species (that lack clear structural roles) reside. Among those with potential roles in catalysis are four Glu, two Asp, two His, three Arg, one Cys, and two Thr residues. Three of the active-site residues are not defined by the electron density map (*26*). Also, even though the crystals of the *E. coli* DS contained 3−6 M formate (a DS-catalyzed reaction product) in their media, there is no evidence for its binding mode, thus removing a potential method for ruling in or out residues for having roles in catalysis. To obtain a better definition of the DS active site, we have turned to the monofunctional DS from *Magnaporthe grisea*, whose crystals are known to diffract to high resolution in the presence or absence of divalent metal cations (*28*). We report five high-resolution structures of the enzyme complexed with ligands that enable us to build views of the substrate and the divalent metal cofactor(s) in the DS active site and propose roles for active-site residues in the catalytic cycle.

EXPERIMENTAL PROCEDURES

Crystallization and Data Collection. Cloning, expression in *E. coli*, purification, and crystallization methods of the *M. grisea* DS have been reported (*28*). Divalent cation-free crystals were obtained in 24−30% PEG 5000 monomethyl ether, 0.2 M Li_2SO_4, and 0.1 M MES-NaOH (pH 6.0−6.5) by the hanging drop vapor diffusion method. The crystals are orthorhombic in space group $P2_12_12$ with one monomer per asymmetric unit. To prepare divalent cation-containing crystals for data collections of the complexes **II**, **III**, and **V**, divalent cations [200 mM $MgCl_2$, 200 mM $MnCl_2$, or 200 mM $Zn(acetate)_2$] were added to the crystals for 8−16 h in soaking solutions of the well solutions omitting Li_2SO_4 and 4% higher in the concentration of PEG 5000 monomethyl ether. These crystals were washed briefly (0.3−0.5 min) in the well solution enriched with 15% glycerol and without divalent metal salts before being flash-frozen in the cold stream (−173 or −168 °C) for data collection. A data set was collected on a $MnCl_2$-containing crystal (complex **IV**) that was grown under the same condition as the divalent cation-free crystals but with the addition of 20 mM $MnCl_2$ to the crystallization solution. This crystal stayed longer (∼5 min) than the other crystals in the cryostabilizing solution containing 15% glycerol before being frozen in the cold stream for data collection. The longer incubation period in the cryostabilizing solution, in addition to the lower divalent cation concentration in comparison to that for the crystals soaked with divalent cations, may explain why the complex **IV** crystal produced the only glycerol-containing data set (see Structure Determination and Refinement). The individual data sets were collected using a single crystal per complex. The 1.55 Å cation-free complex **I** and the 1.60 Å $MnCl_2$ complex **III** data were collected in house using an Raxis-IV imaging plate system with Cu Kα X-rays from a Rigaku rotating anode generator. Data for the 0.98 Å $MgCl_2$ complex **II**, the 1.16 Å $MnCl_2$−glycerol complex **IV**, and the 1.00 Å $Zn(acetate)_2$ complex **V** were collected at the Advanced Photon Source (5-ID beam line) using a MAR CCD detector. All of the data sets were processed using DENZO/SCALEPACK (*29*). Data collection statistics are listed in Table 1.

Structure Determination and Refinement. The structure of the cation-free DS (complex **I**) was determined by the molecular replacement method using the program AMoRe (*30*). The searching model was a partially refined model of the *E. coli* DS (*26*) with the side chains other than Cβ atoms removed. The sequence of *E. coli* DS is 42% identical with that of the *M. grisea* DS. The electron density map was displayed, and the model of DS was manually constructed on a Silicon Graphics system using the program O (*31*). The structure was refined using the program XPLOR (*32*) followed by the least-squares refinement program TNT (*33*). A portion of the data (10%) was omitted from the refinements for the purpose of calculating R_{free} (*34*). The divalent cation-containing structures (complexes **II**−**V**) were determined by the difference Fourier method using the refined cation-free DS model (complex **I**) but with all the water and sulfate molecules removed. For each divalent cation-containing structure (complexes **II**−**V**), after one cycle of rigid body

Table 2: Refinement Statistics

	cation-free (**I**)[a]	Mg^{2+} (**II**)	Mn^{2+} (**III**)	Mn^{2+} (**IV**)[a]	Zn^{2+} (**V**)
resolution (Å) of the structure (of the data used for the refinement)	1.55 (15.0−1.50)	0.98 (25.0−0.98)	1.60 (15.0−1.60)	1.16 (15.0−1.10)	1.00 (15.0−1.00)
no. of residues in the model	208	216	216	216	216
missing residues	1−10, 84−90, 226−233	1−11, 228−233	1−11, 228−233	1−11, 228−233	1−11, 228−233
no. of water molecules	224	278	209	237	211
no. of sulfates, divalent cations, glycerols	4, 0, 0	2, 2, 0	2, 2, 0	2, 1, 1	2, 5, 0
R_{work}/R_{free}	0.201/0.258	0.178/0.196	0.199/0.267	0.177/0.196	0.178/0.201
rmsd of bond lengths (Å)/bond angles (deg) from ideality	0.015/2.4	0.012/2.2	0.015/2.5	0.012/2.3	0.013/2.4
average *B* of main chain atoms (Å^2) (rmsd)	10.0 (2.7)	13.8 (1.1)	18.5 (2.1)	15.6 (1.2)	8.4 (1.3)
average *B* of side chain atoms (Å^2) (rmsd)	15.7 (5.3)	19.3 (3.5)	23.3 (4.3)	20.7 (3.6)	14.6 (4.5)

[a] The high-resolution data in complex **I** (1.55−1.50 Å) and complex **IV** (1.16−1.10 Å) are less than 60% complete but were included in the refinement and the calculation of the R_{work} and R_{free} values.

refinement and one cycle of least-squares positional and *B* factor refinement using TNT, the model was manually rebuilt. Water and divalent cations were added in the next three cycles of positional and *B* factor refinement; all of the refined parameters converged after four cycles. The electron density of a glycerol molecule became obvious for the 1.16 Å $MnCl_2$-containing structure (complex **IV**) at the end of cycle 4. A glycerol molecule was incorporated into the model, and two more cycles of refinement were performed. Since the resolutions of all of the data sets extend to ≥1.60 Å, temperature factors were refined without restraints.

Modeling Methods. The substrate D-ribulose 5-phosphate was modeled into the active site of the Mg^{2+}-containing enzyme (complex **II**) using Sybyl molecular modeling software (*35*). The active-site sulfate molecule was used to position the phosphate group of the substrate; the coordinates of the glycerol molecule of complex **IV** were transferred to the complex **II** structure and used to position the C2 carbonyl group and the C3 and C4 COH groups of the substrate. The active-site residues of complexes **II** and **IV** occupy nearly identical positions (see the Supporting Information for an overlay); the use of complex **II** for the modeling of substrate serves to position the two Mg^{2+} cofactors. The substrate model was energy minimized in the active site of DS to relieve bond and angle strains. To obtain a more accurate model of the substrate binding in the active site, the substrate was further refined by energy minimization using a quantum mechanical method, which bypasses the requirement of reliable molecular mechanics force field parameters in the conventional structure refinement. A hybrid density functional theory method [at the B3LYP/6-31G(d) level of theory], as implemented in Jaguar 4.1 (*36*), was used. During this calculation, only protein side chains of Asp41 (as modeled by formate), His153 (imidazole), Glu37 (acetate), Tyr94 (phenol), Glu174 (acetate), His136 and Asp99 (imidazole and formate, respectively), Cys66 (CH_3SH), two Mg^{2+} cations and their water ligands, and the substrate were included. The protein atoms, together with the metal ions and their water ligands, were fixed to their corresponding crystallographic positions; only the substrate was allowed to move. Working from the ribulose 5-phosphate model, we built the structures of the reaction product (DHBP) and its dienol tautomer (the last intermediate) and manually docked them into the enzyme active site using the phosphate binding site as an anchor. Potential interactions with the enzyme and the metal ions were optimized.

To investigate the rearrangement reaction (the 1,2-shift), we used a quantum chemistry approach. All quantum mechanical calculations were carried out using the Gaussian 98 program (*37*). Geometry optimizations were performed at both the HF/6-31G(d) and B3LYP/6-31G(d) levels of theory. Additional energy calculations were also computed at the MP2/6-31+G(d)//HF/6-31(d) level of theory. This procedure is similar to that previously used to investigate the mechanism of D-xylose isomerase (*15*).

Figures 2, 3, and 6 were generated using the program RIBBONS (*38*).

RESULTS AND DISCUSSION

Structures. The structure of *M. grisea* DS was determined by using molecular replacement with the *E. coli* DS as the searching model. The asymmetric unit of the *E. coli* DS crystals contains a dimer (*26*); that of each of the *M. grisea* DS crystals contains a monomer, which upon applying the symmetry function clearly defines a dimeric protein. The five complexes (**I**−**V**) of *M. grisea* DS have good statistics for data collection (Table 1) and refinement (Table 2). The electron density provides a clear assignment for most of the residues and ligands, particularly those associated with the active site (Figure 2a). More residues are defined in the *M. grisea* DS model than in that of the *E. coli* DS (Table 2). Importantly, all of the residues in the active site are well-defined in the new structures, whereas they were not in the *E. coli* structure (*26*). The *M. grisea* DS shares the same fold as the *E. coli* DS. A comparison of Cα atoms between the two proteins (*E. coli* DS vs complex **I**) indicates a 1.0 Å rms difference for the 193 residues used in the alignment (see the Supporting Information for an overlay of the Cα atoms). Comparisons of Cα atoms among the five new structures indicate a rms difference of less than 0.5 Å. The location of the active site as defined by the ligands (SO_4^{2-}, glycerol, and Mn^{2+}) is at the dimer interface in the vicinity of the residues conserved among species (Figure 2b). Access of the ligands to the active site is likely through movement of two nearby loops of the protein's periphery (Figure 2b) where the *B* factors of the Cα atoms are 2−8-fold greater than those of the others (see the Supporting Information). One of these loops (residues 33−37) is not seen in the electron density map of the *E. coli* DS that lacks active-site ligands (*26*). The other, occupied by residues 84−90, has an orientation in DS−SO_4^{2-}−Mn^{2+}−glycerol complex **IV** that differs significantly from that of complexes **II**, **III**, and

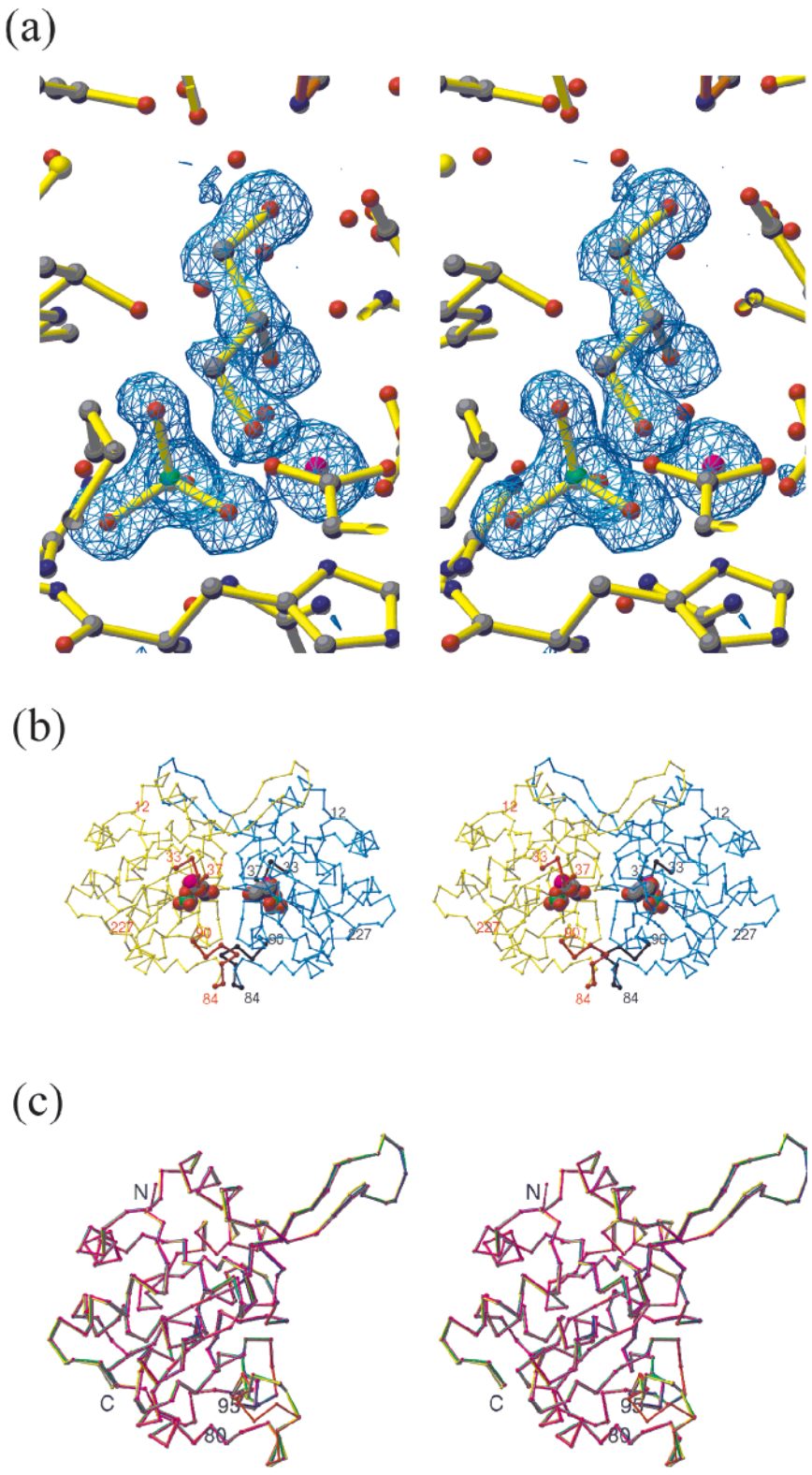

FIGURE 2: Structure of DS. (a) Stereoview of the $F_o - F_c$ omit map of glycerol, sulfate, and Mn^{2+} in the active site of DS (complex **IV**). The map is contoured at 3.5σ. (b) Stereoview of the homodimer of DS with the locations of glycerol, sulfate, and Mn^{2+} indicated with space-filled atoms. The two loops (residues 33−37 and residues 84−90) are highlighted. Their Cα atoms have *B* factors 2−8-fold greater than that of the average *B* factor (see the Supporting Information). The loops are proposed to move, allowing for substrate entry and egress. (c) Stereoview of an overlay of Cα traces of complexes **I** (magenta), **II** (green), **III** (blue), **IV** (red), and **V** (yellow). The rmsd between the Cα atoms of complex **III** and any of the other complexes, except complex **IV**, is 0.1 Å. The rmsd between the Cα atoms of complex **III** and complex **IV** is higher (0.5 Å) because of the differences in the loop of residues 84−90.

V (Figure 2C and Supporting Information) and has insufficient electron density for making residue assignments in complex **I** (Table 2). Otherwise, the Cα traces are similar among the five complexes as noted above and seen in an overlay (Figure 2C). Another view of the way ligands access the DS active site through movement of the two loops is

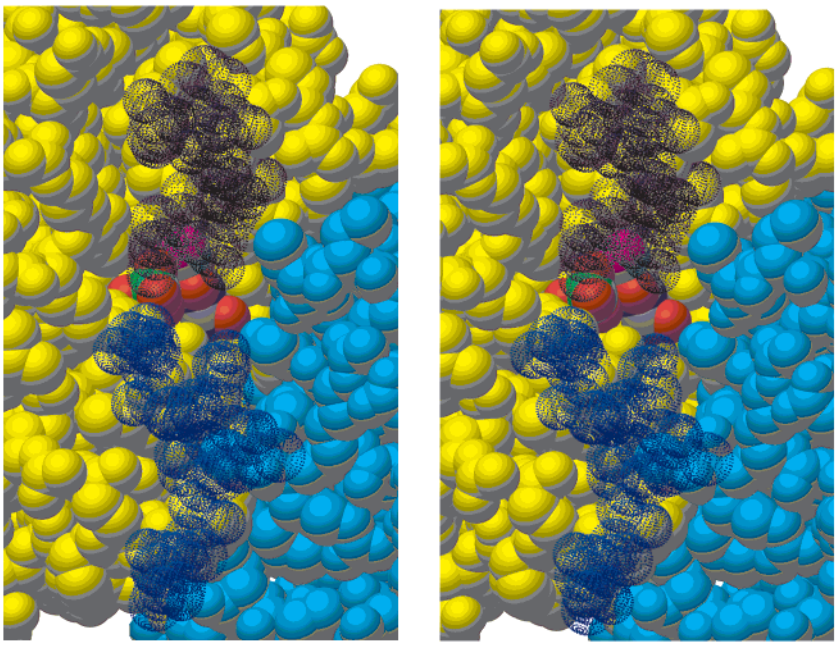

FIGURE 3: Ligand access to the active site of DS. In the stereoview of the complex **IV** active site, the atoms of one subunit are shown in yellow opaque space fill; the other subunit is shown in cyan opaque space fill. The loop containing residues 33−37 is shown in black dotted space fill. The loop containing residues 84−90 is shown in blue dotted space fill. The glycerol, sulfate, and Mn^{2+} are shown in opaque space fill and colored according to atom type as shown in Figure 2b. Hydrogen atoms are not shown. Movement of the two loops, the Cα atoms of which have high *B* factor values, is proposed to facilitate substrate entry and egress. As shown, in the closed state, the two loops bury the ligands.

shown in Figure 3. In the closed state, the two loops bury the ligands.

Ligands. Each of the five DS structures has more than one SO_4^{2-} molecule per DS monomer (Table 2), but only one is associated with the active site; the other SO_4^{2-} molecules are located far from the active site at the surface of the protein. All of the divalent metal cations in complexes **II**−**IV** (two, two, and one divalent cation, respectively) are associated with the active site, and four of the five Zn^{2+} atoms of complex **V** are in the active site. Only complex **IV** has a glycerol molecule, and it is located in the active site. Active-site contacts with the ligands are described in detail in the schematics of Figure 4. As might have been expected for the polar ligands, their protein contacts are entirely through hydrogen bonds and salt bridges in complexes **I**−**III** and **V**; in complex **IV**, there is one contact between the CD1 methyl carbon of Ile172 and a methylene group of glycerol (C−C distance = 3.4 Å). Among species of DS, Ile172 can be replaced with a Leu or Phe residue.[2]

The SO_4^{2-} molecules occupy similar positions in the active sites of complexes **I**−**V** with a rms difference of less than 0.3 Å among their respective atom positions. In each of the five complexes, the active-site SO_4^{2-} molecule is secured in place through interactions with the guanidinium groups of Arg36 and Arg150, the backbone NH groups of His153 and Thr154, and the hydroxyl group of Thr154. Arg36, Arg150, His153, and Thr154 are strictly conserved among species of DS.[2] In complexes **II**−**V** that contain divalent metal cation(s) in the active site, M1 (the circled metal cation of Figure 4) interacts with an oxygen atom of the SO_4^{2-} molecule. The strictly conserved Arg114 and Thr107 residues do not interact with the sulfate dianion. Instead, they serve structural roles at the dimer interface in a hydrogen-bond network that extends to active-site residues (Figure 5). The guanidinium group of Arg114 shares hydrogen bonds with carbonyl oxygen atoms of the strictly conserved Gly63 of the same subunit and Ser62 of the adjacent subunit. The hydroxyl group of Ser62 (replaced with Arg, Lys, and Thr among species) is hydrogen bonded to the carboxyl group of the strictly conserved, active-site Glu174 of the same subunit. The hydroxyl group of Thr107 hydrogen bonds to the carboxyl group of the strictly conserved, active-site Glu39 of the adjacent subunit. The single and double bonds of the SO_4^{2-} molecules were refined as being in similar orientations among the complexes as indicated in Figure 4.

Two divalent metal cations occupy similar positions (less than 0.1 Å rms difference among their positions) in the active sites in complexes **II**, **III**, and **V** (Figure 4). M1 (the circled divalent cation of Figure 4) is replaced by a water molecule in the cation-free complex **I**. M2 (the boxed divalent cation of Figure 4) is replaced by a water molecule in DS−SO_4^{2-}−Mn^{2+}−glycerol complex **IV**. The Mg^{2+} and Mn^{2+} cations are hexacoordinate; the Zn^{2+} cations are pentacoordinate. The occurrence of a single divalent cation in complex **IV** may be a result of the high K_m value for the divalent metal (1 mM for Mg^{2+}) (*10*), the lower concentration of divalent metal in the cocrystallization medium (20 mM) in comparison to that in the crystal soaking medium (200 mM) of the other complexes, and the longer incubation period for the complex (5 vs 0.5 min) in divalent metal-free cryoprotection medium (see Experimental Procedures). The longer incubation period in the glycerol-containing cryoprotection medium also seems to explain the presence of the glycerol molecule in complex **IV**. In complex **IV**, two of the coordination sites of the Mn^{2+}

[2] The comparisons of DS sequences refer to the alignment of 19 species of DS reported previously (*26*) and shown in the Supporting Information. The term "strictly conserved" means that the residue is identical in the sequence alignment.

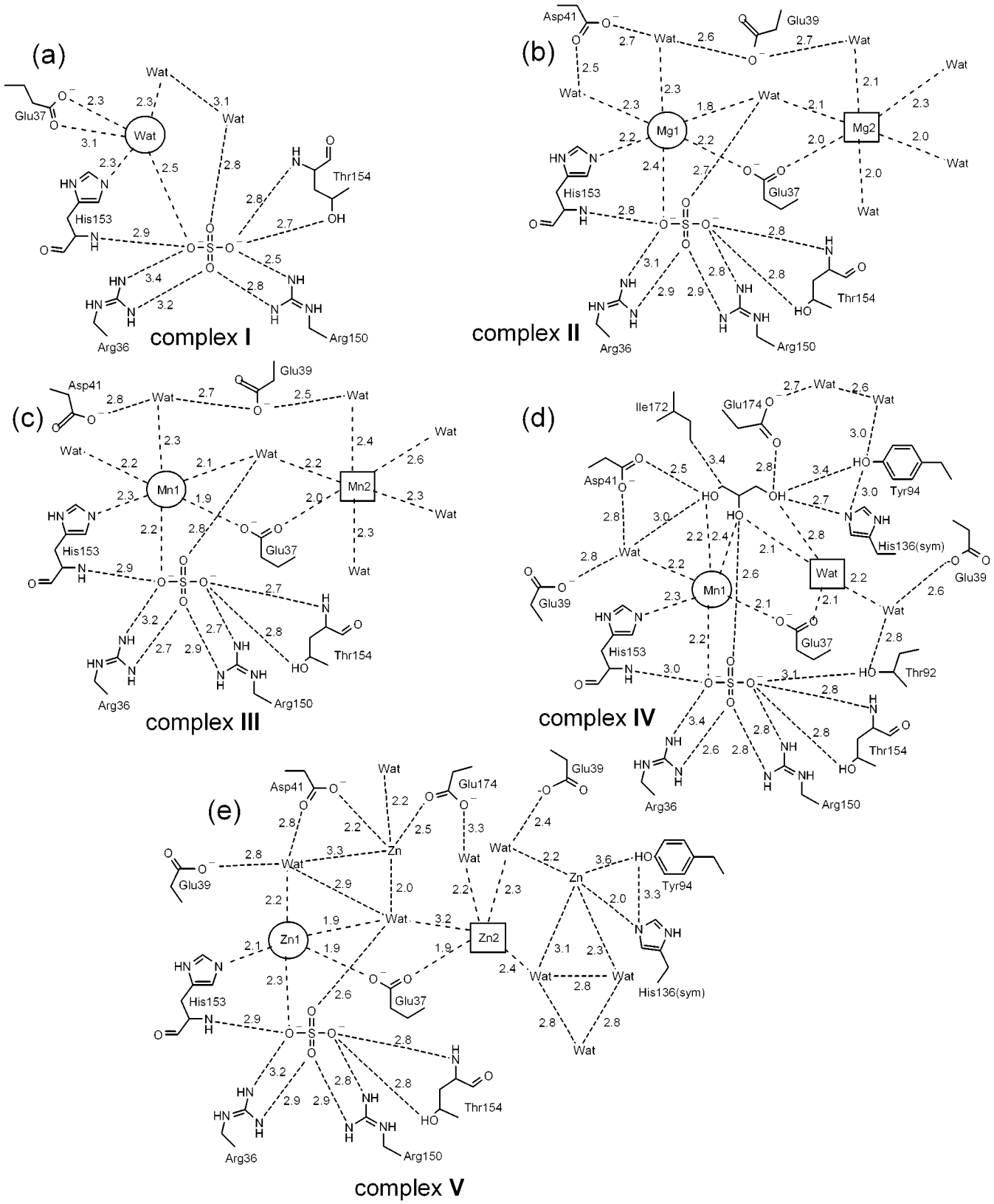

FIGURE 4: Schematics of ligand interactions in the active site of DS: (a) DS−SO_4^{2-} (complex **I**), (b) DS−SO_4^{2-}−Mg^{2+} (complex **II**), (c) DS−SO_4^{2-}−Mn^{2+} (complex **III**), (d) DS−SO_4^{2-}−Mn^{2+}−glycerol (complex **IV**), and (e) DS−SO_4^{2-}−Zn^{2+} (complex **V**). Distances are in angstroms. Residues with the notation (sym) are from the second monomer of the dimer. The circled and boxed divalent metal cations or water molecules occupy similar respective positions among the five complexes. Wat represents a water molecule.

cation are occupied by hydroxyl groups of glycerol and one is occupied by the sulfate dianion, and this has implications for the binding mode of substrate ribulose 5-phosphate.

Although there are six Asp or Glu residues strictly conserved among species of DS that potentially serve as chelators for the metal cofactor(s), Glu37 is the sole residue

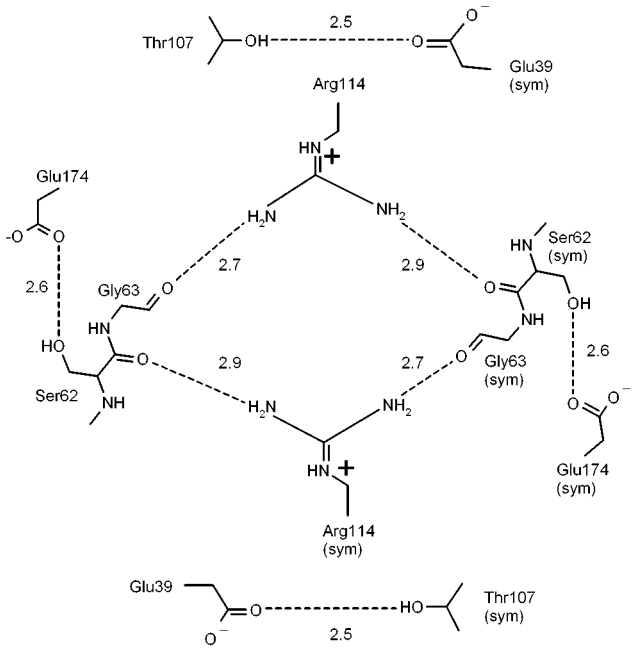

FIGURE 5: Hydrogen-bond network involving conserved residues at the dimer interface of DS. The network that connects the subunits at the dimer interface extends to the strictly conserved active-site residues Glu39 and Glu174. Residues with the notation (sym) are from the second monomer of the dimer. Distances are in angstroms. The shortest distance among complexes **I**−**V** for each interaction is shown. The distances are as much as 0.3 Å longer in the individual structures with the exception of the distance between Glu174 and Ser62 in complex **I**, which is 3.8 Å.

with this role (Figure 4). The strictly conserved Asp32 carboxyl group clearly has a role in stabilizing the position of the strictly conserved Arg36 guanidinium group. Likewise, the carboxyl group of the strictly conserved Glu155 forms a salt bridge with the guanidinium group of the strictly conserved Arg150. The carboxyl group of the strictly conserved Glu39 has a structural role in hydrogen bonding to the strictly conserved Thr107 of the adjacent subunit (Figure 5), as well as a role in securing the metal cofactors through indirect hydrogen bonds to the metals through water molecules (Figure 4). Among species of DS, Asp99 is replaced with Glu but not other residues; its role is in forming an Asp−His dyad (with the strictly conserved His136) that serves a general acid-base function during catalysis (see Catalysis). The remaining, strictly conserved residue with a carboxyl group, Glu174, is proposed to serve as a general acid-base functionality in promoting catalysis (see Catalysis). M1 (the circled metal cation of Figure 4) also interacts with the imidazole group of the strictly conserved His153.

Even though multiple residues from both monomers at the dimer interface are conserved among species, only one residue from the second monomer (His136) is in direct contact with ligands in the active site (in Figure 4d, His136 with the hydroxyl group of glycerol; in Figure 4e, His136 with the Zn^{2+} cation). His136 is paired with Asp99 from the same monomer in forming an Asp−His dyad. Because of its importance in modeling the substrates in the active site, a stereoview of the DS−SO_4^{2-}−Mn^{2+}−glycerol complex is shown in Figure 6a.

Catalysis. Using the active-site sulfate dianion and glycerol molecule as guides, the substrate D-ribulose 5-phosphate was modeled into the active site of DS−SO_4^{2-}−Mg^{2+} complex **II** (after transferring the glycerol to it from complex **IV**) to provide a view of the Michaelis complex (Figure 6b). Though only one divalent metal cation (Mn1, the circled one in Figure 4d) is observed in the active site of the DS−SO_4^{2-}−Mn^{2+}−glycerol complex, we kept the two Mg^{2+} cations in the model of the substrate complex because the second Mg^{2+} cation (the boxed one in Figure 4b) is without steric conflicts in the model and has a potential role in stabilizing the substrate and reaction intermediates. Certainly, complexes **II**, **III**, and **V** have the two metal cofactors occupying similar positions in the active site, and the X-ray structures of the mechanistically related enzymes, xylose isomerase and acetohydroxy acid isomeroreductase, have two metal ions in their active sites. However, because of the presence of a single Mn^{2+} (Mn1) with a water molecule occupying the position of the second metal cofactor in complex **IV**, the proposal for two metal cofactors in the DS active site is qualified by the possibility that a water molecule satisfies the role of Mg2 in our model. According to our model, shown in schematic form in Figure 7, one Mg^{2+} cation (Mg1, the circled one) is hexacoordinate and binds to the substrate via one of its phosphate group oxygen atoms and its C3 and C4 hydroxyl groups. The three remaining coordination sites of the Mg^{2+} cation are occupied by the imidazole group of His153, the carboxyl group of Glu37, and a water. The other Mg^{2+} cation (Mg2, the boxed one in Figure 7) is also hexacoordinate and binds to the carboxyl group of Glu37 and the C2 carbonyl oxygen atom and the C3 hydroxyl group of the substrate with the remaining ligand sites occupied by water molecules. In turn, the substrate's C2 carbonyl oxygen is hydrogen bonded to the hydroxyl group of Tyr94 and the imidazole group of the His136−Asp99 dyad from the second subunit. The two metal ions are bridged by the Glu37 carboxyl group and the C3 hydroxyl group of the substrate. The substrate's C1 hydroxyl group is near the thiol group of Cys66. Its C4 hydroxyl group is hydrogen bonded to the carboxyl group of Asp41. Other interactions with the substrate's phosphate group are as found for the sulfate dianion of complexes **I**−**V** with salt-bridge interactions with Arg150 and Arg36 and hydrogen-bonding interactions with the hydroxyl group of Thr154 and two backbone amide NH groups. Similarly, the extensive hydrogen-bond network in the active site is retained. With the exceptions of Tyr94, which is replaced only with Phe among DS species, and Asp99, which is replaced only with Glu among DS species, the residues depicted in Figure 7 are strictly conserved among species, and their positions are rather static among the five X-ray structures (see the Supporting Information).

Our proposed catalytic mechanism has the initial enolization step mediated by the carboxyl group of Glu174 acting as a general base in abstracting the substrate's C3 hydrogen atom in concert with the donation of a proton to the C2 carbonyl oxygen atom by the imidazolium group of His136 (Figure 8). Since the C3 hydroxyl group of substrate binds to both Mg^{2+} ions, it seems likely that the hydroxyl group is rather acidic, so it is drawn in its deprotonated form. In principle, the C3 hydroxyl group can be deprotonated either before or after the enolization. Tyr94 is replaced with Phe in several DS proteins, so its interactions with the substrate and intermediates are considered nonessential. The resulting enolate **2** is stabilized by hydrogen-bond interactions between the substrate's C2 hydroxyl group, the imidazole group of

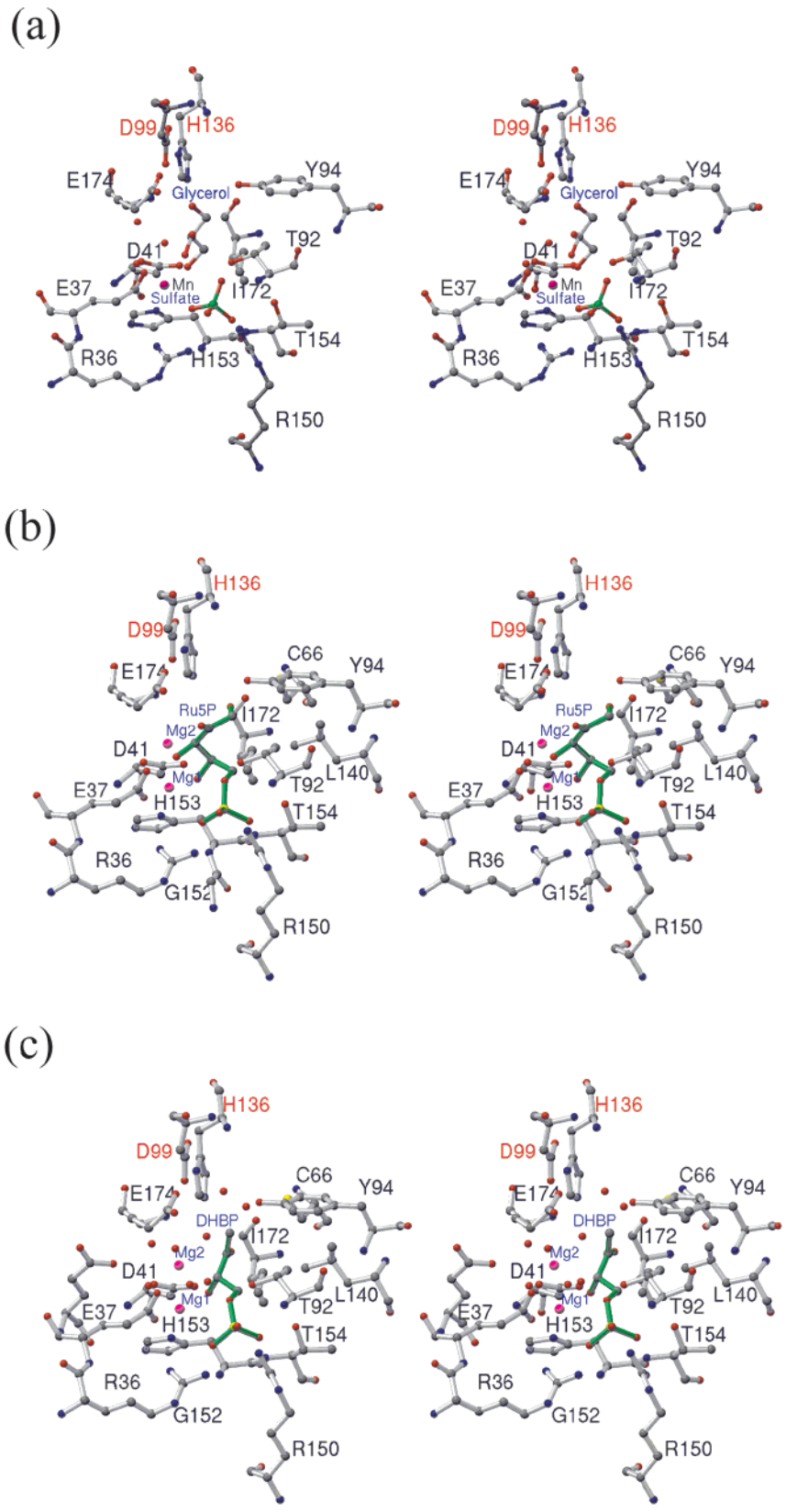

FIGURE 6: Stereoviews of the active site of DS. (a) The DS−SO_4^{2-}−Mn^{2+}−glycerol structure (complex **IV**). (b) A model of the substrate ribulose 5-phosphate (Ru5P) in the active site. (c) A model of the product (3*S*)-3,4-dihydroxy-2-butanone 4-phosphate (DHBP) in the active site. Residues labeled in red are from the second monomer of the dimer. Mg1 corresponds to Mg1 (the circled Mg atom) in Figure 4b; Mg2 corresponds to Mg2 (the boxed Mg atom) in Figure 4b. Water molecules (unlabeled and in red) are shown only in panels a and c. Mg^{2+} and Mn^{2+} cations are shown in magenta.

the His136−Asp99 dyad from the second subunit, and possibly the hydroxyl group of Tyr94. Subsequent dehydration of intermediate **2** is assisted by the thiol group of Cys66 acting as a general acid. The resulting enol **3** goes to the

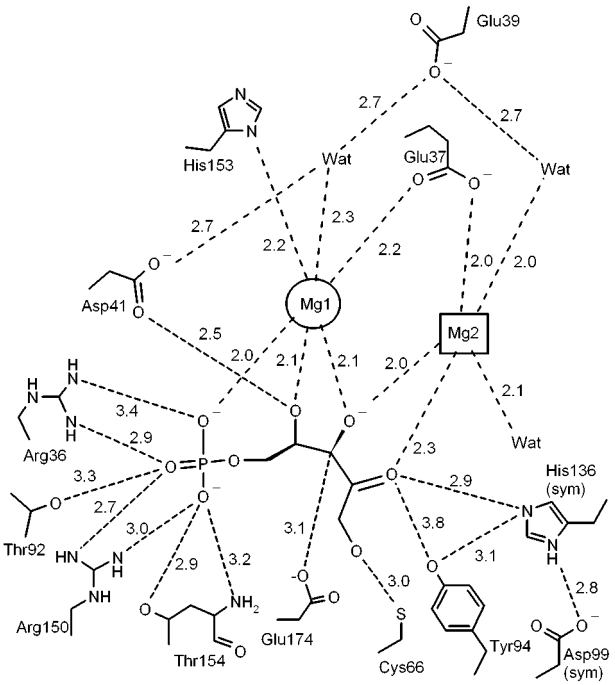

FIGURE 7: Schematic of the model of the substrate ribulose 5-phosphate in the active site. Distances are in angstroms. Residues with the notation (sym) are from the second monomer of the dimer. The circled and boxed Mg^{2+} cations occupy positions similar to those indicated in Figure 4. Mg1 (the circled one) corresponds to Mg1 of panels b and c of Figure 6; Mg2 (the boxed one) corresponds to Mg2 of panels b and c of Figure 6. With the exceptions of Tyr94, which is replaced only with Phe among DS species, and Asp99, which is replaced only with Glu among DS species, the residues that are depicted are strictly conserved among species.

diketone **4** via an acid-base catalyzed ketonization process, with the C2 hydroxyl group deprotonated by the imidazole of the His136−Asp99 dyad and proton donation to C1 through the carboxyl group of Glu174 either directly or indirectly (as assisted by nearby water molecules). The 1,2-skeleton rearrangement of **4** to generate **5** is initiated by the deprotonation of the C4 hydroxyl group by the carboxyl group of Asp41. Subsequent hydration of **5** by a Mg^{2+}-associated water molecule and the retro-aldol reaction of the aldehyde intermediate **6** with proton donation from the imidazole of the His136−Asp99 dyad yield the enolate form of 3,4-dihydroxy-2-butanone 4-phosphate, intermediate **7**. Once the formate leaves the active site, the intermediate **7** shifts toward the Mg1. In this rearranged active site, the two Mg^{2+} ions are bridged again by a water molecule, and intermediate **7** binds only to the circled Mg^{2+}. Now, the side chain carboxylate of Glu174 is close enough to interact with the C2 hydroxyl group. Apparently, there is no appropriate acid group nearby to protonate the enolate other than several crystallographic water molecules. Therefore, we propose that the final product is generated by protonation of the enolate by a Mg^{2+}-activated water molecule, giving the correct stereochemistry for (3*S*)-3,4-dihydroxy-2-butanone phosphate. The protonation can only occur at the correct face of **7** because that is where water molecules and hydrophilic residues reside; the other face of **7** is shielded from water molecules by the methyl groups of Leu140 and Ile172 (Figure 6c).

The arrangement of the substrate's C3 and C4 hydroxyl groups in our model suggests that the 1,2-skeleton rearrangement occurs with a cis orientation of the hydroxyl groups. We investigated the 1,2-skeleton rearrangement theoretically using quantum chemistry methods and 3-hydroxy-2-butanone as the model compound. The transition state structures were calculated using quantum mechanical methods. Either trans or cis transition states are possible (Figure 9a). In the absence of enzyme and solvent, the trans TS is expected to be favored due to electrostatic repulsion between the oxygen atoms. A schematic drawing of the relative energies is shown in Figure 9b. As expected, the trans transition state has a lower activation barrier, and the difference is ~2 kcal/mol. A cis transition state has also been suggested for the xylose isomerase-catalyzed reaction based on a high-resolution crystal structure (*18*). According to the present and previous theoretical calculations on model reactions (*15*), the hydride transfer occurs with a smaller barrier [16.6 kcal/mol at the MP2/6-31+G(d) level of theory; see ref *15*] than the migration of a carbon center (24.2 kcal/mol at the same level of theory; see Figure 9b). On the basis of the rates of catalysis, the free energy barrier for the 1,2-hydride transfer in the interconversion of xylose and xylulose catalyzed by the *Streptomyces* spp. enzyme is 16.8 kcal/mol (*39*), while the free energy barrier for the DS reaction from *E. coli* is 19.3 kcal/mol at 25 °C [k_{cat} of 0.040 s^{-1} (*10*)] and that for the DS reaction from *M. grisea* 19.2 kcal/mol at 25 °C [k_{cat} of 0.046 s^{-1} (*28*)], suggesting that the 1,2-shift step is rate-limiting to the reactions catalyzed by the enzymes. At the present time, there is no experimental evidence as to which transition state DS prefers. Factors that control the catalytic efficacy of DS and the exact role of each active-site residue need to be investigated experimentally before the catalytic mechanism can be fully elucidated. Our proposed catalytic mechanism provides a framework for further mechanistic investigations.

CONCLUSIONS

The five structures determined for this work complement one another. Among the structures, the active-site residues, metal cofactors, and sulfate group have only small positional changes. This adds a measure of confidence for the model of ribulose 5-phosphate in the active site (Figures 6b and 7) and the surmised roles of the active-site residues in catalysis (Figure 8). The major difference occurring among the structures, a reorientation of the loop of residues 84−90, implicates a route for substrate delivery to the active site, which is buried by the protein when occupied by ligands. Having definition of more residues in the structures presented here than in the structure of the *E. coli* DS allows for additional residues to be ruled in or out for having roles in catalysis and substrate recognition. A second loop containing residues 33−37 that is not defined by the electron density in the *E. coli* DS structure implicates another opening for substrate access to the active site, complementary to the loop of residues 84−90. The strictly conserved residues of the *E. coli* DS that have clear structural roles (one Ala, four Gly's, one Pro, and one Ser) maintain such roles in the *M. grisea* DS.

The proposed mechanism for the DS-catalyzed reaction shares features with those of both types of sugar isomerases (those that act on phosphorylated or nonphosphorylated

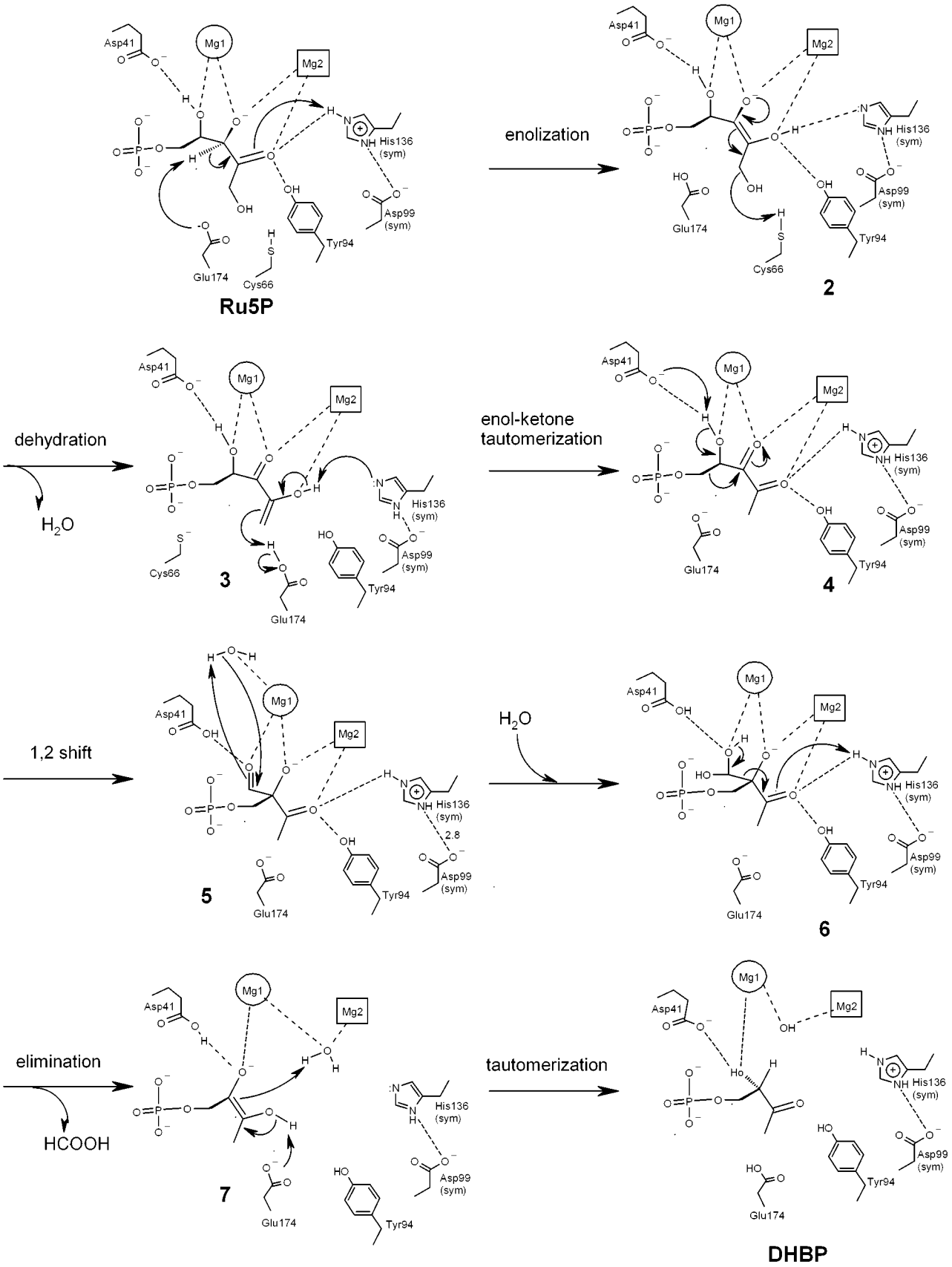

FIGURE 8: Proposed mechanism of the DS-catalyzed reaction. Intermediates **2**−**7** are similar to those in Figure 1.

substrates) to the extent that the DS isomerization step is thought to be catalyzed by a carboxylate group acting as a general base in abstracting a proton from a CH group in concert with an imidazol(e)(ium) group protonating the adjacent carbonyl oxygen of the substrate (as in the reaction catalyzed by triose phosphate isomerase) and that a metal cofactor is thought to be employed in the DS active site to secure the substrate (as in the reaction catalyzed by xylose

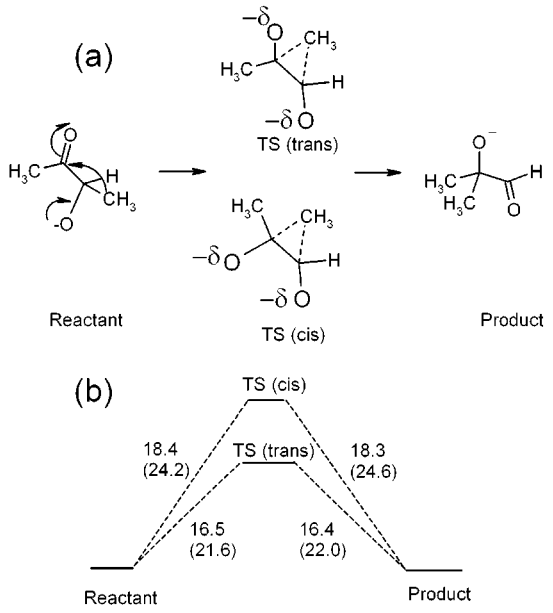

FIGURE 9: Comparison of cis and trans transition states for the 1,2-shift. (a) As modeled using 3-hydroxy-2-butanone. (b) Transition states for the 1,2-shift at the B3LYP/6-31G(d) and the MP2/6-31+G(d) levels of theory. Activation energy values are in kilocalories per mole. The MP2 values are in brackets. TS represents the transition state.

isomerase). DS also shares mechanistic features with acetohydroxy acid isomeroreductase-, 1-deoxy-D-xylulose 5-phosphate reductoisomerase-, and xylose isomerase-catalyzed reactions for the group transfer step by having a requirement for metal cofactor(s) for stabilizing intermediates.

The serendipitous discovery of the glycerol complex (DS−SO_4^{2-}−Mn^{2+}−glycerol, complex **IV**), through a modification of the procedure used for soaking crystals in glycerol-enriched cryoprotectant, and the ability to soak divalent metal cations into the crystals might suggest that the crystals of *M. grisea* DS are suitable for soaking in substrate or inhibitor solutions in determining structures of additional DS−ligand complexes in an efficient manner. The ability to obtain many protein−inhibitor structures is crucial to an iterative, structure-based design program for optimizing inhibitors. DS and the other enzymes of the riboflavin biosynthetic pathway (*1*, *40*−*43*) are attractive targets for antibiotic discovery efforts because humans are auxotrophic for riboflavin (vitamin B_2) and obtain it through nutrition, whereas pathogens, particularly the Gram-negative bacteria and certain fungi, are incapable of acquiring sufficient riboflavin from their hosts and must biosynthesize it. *M. grisea* is the disease vector of rice blast (*44*), which is responsible for enormous perennial losses in the largest human food staple. The disease is effectively controlled by active-site inhibitors of enzymes that belong to the fungal melanin biosynthesis pathway, scytalone dehydratase (*45*−*55*) and trihydroxynaphthalene reductase (*56*−*60*), but there remains a need for blasticides with lower application rates. That the residues of the DS active site are highly conserved among species implies that the structures of the *M. grisea* enzyme could be useful for the design of antibiotics that act on disease agents other than *M. grisea*.

ACKNOWLEDGMENT

This is DuPont CR&D Contribution 8219.

SUPPORTING INFORMATION AVAILABLE

Figures show an overlay of Cα traces of the *E. coli* DS and the *M.* grisea DS, the *B* factors of the Cα atoms for complexes **I**−**V**, the change in a loop orientation in complex **IV** in comparison to the other complexes, an overlay of the active-site residues for complexes **I**−**V**, and a sequence alignment of 19 species of DS. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Bacher, A. (1991) Biosynthesis of flavins, in *Chemistry and Biochemistry of Flavoenzymes* (Müller, F., Ed.) Vol. 1, pp 215−259, CRC Press, Boca Raton, FL.
2. Volk, R., and Bacher, A. (1988) *J. Am. Chem. Soc. 110*, 3651−3653.
3. Volk, R., and Bacher, A. (1990) *J. Biol. Chem. 265*, 19479−19485.
4. Volk, R., and Bacher, A. (1991) *J. Biol. Chem. 266*, 20610−20618.
5. Richter, G., Volk, R., Krieger, C., Lahm, H. W., Roethlisberger, U., and Bacher, A. (1992) *J. Bacteriol. 174*, 4050−4056.
6. Richter, G., Krieger, C., Volk, R., Kis, K., Ritz, H., Götze, E., and Bacher, A. (1997) *Methods Enzymol. 280*, 374−382.
7. Götze, E., Kis, K., Eisenreich, W., Yamauchi, N., Kakinuma, K., and Bacher, A. (1998) *J. Org. Chem. 63*, 6456−6457.
8. Richter, G., Kelly, M., Krieger, C., Yu, Y., Bermel, W., Karlsson, G., Bacher, A., and Oschkinat, H. (1999) *Eur. J. Biochem. 261*, 57−65.
9. Hümbelin, M., Griesser, V., Keller, T., Schurter, W., Haiker, M., Hohmann, H.-P., Ritz, H., Richter, G., Bacher, A., and Van Loon, A. P. G. M. (1999) *J. Ind. Microbiol. Biotechnol. 22*, 1−7.
10. Picollelli, M. A., Viitanen, P. V., and Jordan, D. B. (2000) *Anal. Biochem. 287*, 347−349.
11. Herz, S., Eberhardt, S., and Bacher, A. (2000) *Phytochemistry 53*, 723−731.
12. Knowles, J. R. (1991) *Nature 350*, 121−124.
13. Price, N. C., and Stevens, L. (1989) *Fundamentals of Enzymology*, 2nd ed., pp 216−224, Oxford University Press, Oxford, U.K.
14. Farber, G. K., Glasfeld, A., Tiraby, G., Ringe, D., and Petsko, G. A. (1989) *Biochemistry 28*, 7289−7297.
15. Zheng, Y.-J., Merz, K. M., Jr., and Farber, G. K. (1993) *Protein Eng. 6*, 479−484.
16. Lavie, A., Allen, K. N., Petsko, G. A., and Ringe, D. (1994) *Biochemistry 33*, 5469−5480.
17. Allen, K. N., Lavie, A., Farber, G. K., Glasfeld, A., Petsko, G. A., and Ringe, D. (1994) *Biochemistry 33*, 1481−1487.
18. Whitlow, M., Howard, A. J., Finzel, B. C., Poulos, T. L., Winbourne, E., and Gilliland, G. L. (1991) *Proteins: Struct., Funct., Genet. 9*, 153−173.
19. Chunduru, S. K., Mrachko, G. T., and Calvo, K. C. (1989) *Biochemistry 28*, 486−493.
20. Aulabaugh, A., and Schloss, J. V. (1990) *Biochemistry 29*, 2824−2830.
21. Biou, V., Dumas, R., Cohen-Addad, C., Douce, R., and Pebay-Peyroula, E. (1997) *EMBO J. 16*, 3405−3415.
22. Thomazeau, K., Dumas, R., Halgand, F., Forest, E., Douce, R., and Biou, V. (2000) *Acta Crystallogr. D56*, 389−397.
23. Dumas, R., Biou, V., Halgand, F., Douce, R., and Duggleby, R. D. (2001) *Acc. Chem. Res. 34*, 399−408.
24. Arigoni, D., Sagner, S., Latzel, C., Eisenreich, W., Bacher, A., and Zenk, M. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 10600−10605.
25. Takahashi, S., Kuzuyama, T., Watanabe, H., and Seto, H. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 9879−9884.
26. Liao, D.-I., Calabrese, J. C., Wawrzak, Z., Viitanen, P. V., and Jordan, D. B. (2001) *Structure 9*, 11−18.
27. Gibrat, J. F., and Bryant, S. H. (1996) *Curr. Opin. Struct. Biol. 6*, 377−385.

28. Liao, D.-I., Viitanen, P. V., and Jordan, D. B. (2000) *Acta Crystallogr. D56*, 1495−1497.
29. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
30. Navaza, J. (1994) *Acta Crystallogr. A50*, 157−163.
31. Jones, T. A., Zou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
32. Brünger, A. T., Krukowski, J., and Erickson, J. (1990) *Acta Crystallogr. A46*, 585−593.
33. Tronrud, D. E., Ten Eyck, L. F., and Matthews, B. W. (1987) *Acta Crystallogr. A43*, 489−501.
34. Brünger, A. T. (1992) *Nature 355*, 472−475.
35. *Sybyl Molecular Modeling Software*, version 6.6 (1999) Tripos, Inc., St. Louis.
36. *Jaguar 4.1* (2000) Schrodinger, Inc., Portland, OR.
37. *Gaussian 98* (1998) Gaussian, Inc., Pittsburgh, PA.
38. Carson, M. (1997) *Methods Enzymol. 277*, 493−505.
39. Van Bastelaere, P. B. M., Vangrysperre, W., and Kersters-Hilderson, H. (1991) *Biochem. J. 278*, 285−292.
40. Jordan, D. B., Bacot, K. O., Carlson, T. J., Kessel, M., and Viitanen, P. V. (1999) *J. Biol. Chem. 274*, 22114−22121.
41. Persson, K., Schneider, G., Jordan, D. B., Viitanen, P. V., and Sandalova, T. (1999) *Protein Sci. 8*, 2355−2365.
42. Zheng, Y.-J., Viitanen, P. V., and Jordan, D. B. (2000) *Bioorg. Chem. 28*, 89−97.
43. Liao, D.-I., Wawrzak, Z., Calabrese, J. C., Viitanen, P. V., and Jordan, D. B. (2001) *Structure 9*, 399−408.
44. Howard, R. J., and Valent, B. (1996) *Annu. Rev. Microbiol. 50*, 491−512.
45. Lundqvist, T., Rice, J., Hodge, C. N., Basarab, G. S., Pierce, J., and Lindqvist, Y. (1994) *Structure 2*, 937−944.
46. Basarab, G. S., Steffens, J. J., Wawrzak, Z., Schwartz, R. S., Lundqvist, T., and Jordan, D. B. (1999) *Biochemistry 38*, 6012−6024.
47. Jordan, D. B., Zheng, Y.-J., Locket, B. A., and Basarab, G. S. (2000) *Biochemistry 39*, 2276−2282.
48. Jordan, D. B., Basarab, G. S., Steffens, J. J., Schwartz, R. S., and Doughty, J. G. (2000) *Biochemistry 39*, 8593−8602.
49. Basarab, G. S., Jordan, D. B., and Zheng, Y.-J. (2000) *Org. Lett. 2*, 1541−1544.
50. Chen, J. M., Xu, S. L., Wawrzak, Z., Basarab, G. S., and Jordan, D. B. (1998) *Biochemistry 37*, 17735−17744.
51. Jordan, D. B., Lessen, T., Wawrzak, Z., Bisaha, J. J., Gehret, T. C., Hansen, S. L., Schwartz, R. S., and Basarab, G. S. (1999) *Bioorg. Med. Chem. Lett. 9*, 1607−1612.
52. Basarab, G. S., Jordan, D. B., Gehret, T. C., Schwartz, R. S., and Wawrzak, Z. (1999) *Bioorg. Med. Chem. Lett. 9*, 1613−1618.
53. Wawrzak, Z., Sandalova, T., Steffens, J. J., Basarab, G. S., Lundqvist, T., Lindqvist, Y., and Jordan, D. B. (1999) *Proteins: Struct., Funct., Genet. 35*, 425−439.
54. Jennings, L. D., Wawrzak, Z., Amorose, D., Schwartz, R. S., and Jordan, D. B. (1999) *Bioorg. Med. Chem. Lett. 9*, 2509−2514.
55. Jennings, L. D., Rayner, D. R., Jordan, D. B., Okonya, J., Basarab, G. S., Amorose, D. K., Anacelerio, B. M., Gehret, T. C., Lee, J. K., Schwartz, R. S., and Whitmore, K. A. (2000) *Bioorg. Med. Chem. 8*, 897−907.
56. Andersson, A., Jordan, D. B., Schneider, G., and Lindqvist, Y. (1996) *Structure 4*, 1161−1170.
57. Thompson, J. E., Basarab, G. S., Andersson, A., Lindqvist, Y., and Jordan, D. B. (1997) *Biochemistry 36*, 1852−1860.
58. Liao, D.-I., Basarab, G. S., Gatenby, A. A., Valent, B., and Jordan, D. B. (2001) *Structure 9*, 19−27.
59. Liao, D.-I., Basarab, G. S., Gatenby, A. A., and Jordan, D. B. (2000) *Bioorg. Med. Chem. Lett. 10*, 491−494.
60. Jordan, D. B., Basarab, G. S., Liao, D.-I., Johnson, W., Winzenberg, K., and Winkler, D. A. (2001) *J. Mol. Graphics Modell. 19*, 434−447.

BI015652U