# Crystal Structure of the Carboxyltransferase Domain of Acetyl–Coenzyme A Carboxylase

Hailong Zhang, Zhiru Yang,* Yang Shen,* Liang Tong†

Acetyl–coenzyme A carboxylases (ACCs) are required for the biosynthesis and oxidation of long-chain fatty acids. They are targets for therapeutics against obesity and diabetes, and several herbicides function by inhibiting their carboxyltransferase (CT) domain. We determined the crystal structure of the free enzyme and the coenzyme A complex of yeast CT at 2.7 angstrom resolution and found that it comprises two domains, both belonging to the crotonase/ClpP superfamily. The active site is at the interface of a dimer. Mutagenesis and kinetic studies reveal the functional roles of conserved residues here. The herbicides target the active site of CT, providing a lead for inhibitor development against human ACCs.

Acetyl–coenzyme A carboxylases (ACCs) catalyze the formation of malonyl–coenzyme A (CoA) and regulate fatty acid biosynthesis and oxidation (1–6). They are targets for the development of therapeutic treatments against obesity, type 2 diabetes, and microbial infections (7–10), in addition to being the site of action for commercial herbicides (11). In mammals, ACC1 is a cytosolic enzyme, and its production of malonyl-CoA is the committed step in the biosynthesis of long-chain fatty acids (1–4). In comparison, ACC2 is a mitochondrial enzyme, and its malonyl-CoA product regulates fatty acid oxidation by potently inhibiting the shuttle that transports long-chain acyl-CoAs from the cytosol to the mitochondria for oxidation (5, 6). Mice lacking ACC2 have a higher than normal rate of fatty acid oxidation and reduced body fat and body weight (8).

Mammalian, yeast, and most other eukaryotic ACCs are large multifunctional enzymes, containing the biotin carboxylase (BC) domain, the biotin carboxyl carrier protein (BCCP) domain, and the carboxyltransferase (CT) domain (Fig. 1A). BC catalyzes the adenosine triphosphate–dependent carboxylation of a biotin group covalently linked to a lysine residue in BCCP, and then CT catalyzes the transfer of the carboxyl group from biotin to acetyl-CoA to produce malonyl-CoA (Fig. 1B). In Escherichia coli and other bacteria, ACCs are multisubunit enzymes, with BC, BCCP, and two subunits for the CT (Fig. 1A). Crystal structures are available for the BC and BCCP subunits of E. coli ACC (4), but currently there is no structural information on the CT enzyme. The amino acid sequences of the CT domains are highly conserved among the eukaryotic multifunctional ACCs (fig. S1) but share very limited homology to the two subunits of the bacterial CTs (Fig. 1A). They do not share any recognizable homology with other proteins in the sequence database.

We expressed and purified the CT domain of yeast ACC, which constitutes the 90-kD fragment at the C terminus of the protein (Fig. 1A and fig. S1). The CT domain of yeast ACC shares 50% sequence identity with that of human ACCs. The structure of the free enzyme was determined at 2.7 Å resolution by the selenomethionyl single-wavelength anomalous diffraction method (Table 1 and fig. S2) (12). The

structure of CT in complex with CoA was determined at 2.7 Å resolution from a crystal grown in the presence of 2 mM acetyl-CoA (Table 1). Several attempts at determining the binding mode of the biotin substrate were unsuccessful, possibly because binding has low affinity. The Michaelis constant ($K_m$) of biocytin, an analog of biotin, for the E. coli CT has been reported to be about 8 mM (13). The atomic coordinates have been deposited at the Protein Data Bank, with accession codes 1OD2 and 1OD4.

The crystal structures show that each CT domain molecule is made up of two subdomains (N and C domains) that are intimately associated with each other (Figs. 1C and 2, A and B). This domain organization of CT is consistent with the fact that the N- and C-terminal halves show limited sequence homology to the β and α subunits of the bacterial CT, respectively (Fig. 1A). More than 50 residues at the N terminus and 30 residues at the C terminus (together with the His tag) are disordered in both structures.

The N and C domains share similar polypeptide backbone folds, with a central β-β-α superhelix (Fig. 2, A and B). A total of 127 equivalent Cα atoms can be superimposed to within 2.5 Å of each other between the two domains, and the root mean squares (rms) distance for these atoms is 1.3 Å. However, the amino acid sequence identity between these 127 residues is only 12%, underscoring the lack of sequence conservation
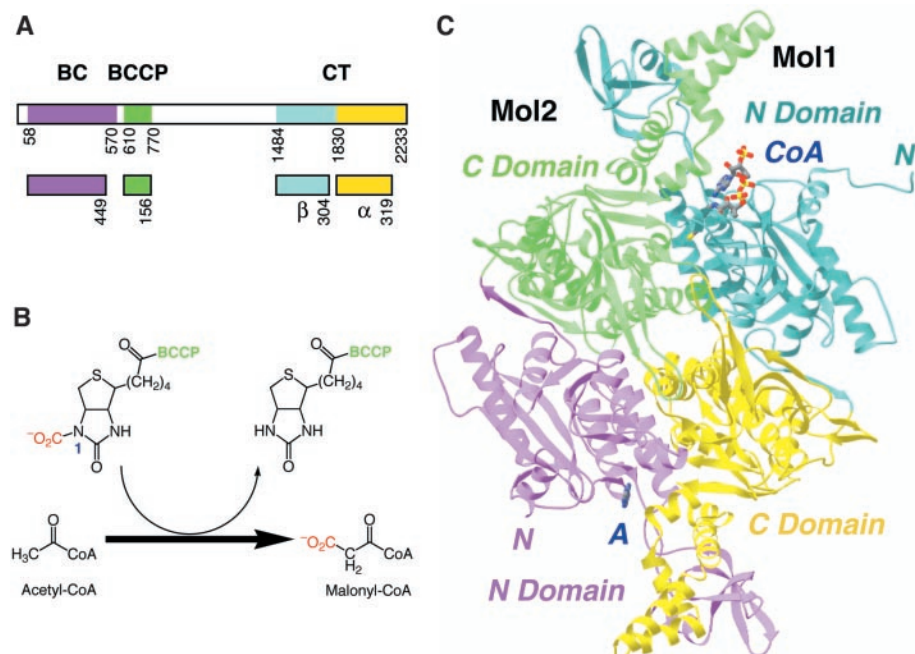


**Fig. 1.** Structures of ACCs. (**A**) Schematic drawing of the primary structures of eukaryotic multidomain ACC and bacterial multisubunit ACC. (**B**) The chemical reaction catalyzed by CT. The N1 atom of biotin is labeled. (**C**) Schematic drawing of the structure of the CT domain dimer of yeast ACC. The N and C domains of one monomer are colored cyan and yellow, whereas those of the other monomer are colored purple and green. The CoA molecule bound to one monomer is shown as a stick model. Only the adenine base was observed in the other monomer (labeled A). (C) was produced with Ribbons (22).

Department of Biological Sciences, Columbia University, New York, NY 10027, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: tong@como.bio.columbia.edu

between the two domains. The backbone fold of the two domains is similar to that of the crotonase/ClpP superfamily (Fig. 2C) (14–18), even though the amino acid sequence identity between CT domains and these other proteins is less than 14%. Several crotonase/ClpP proteins are acyl-CoA–dependent enzymes that catalyze various reactions for fatty acid β-oxidation (4, 15, 17, 18). However, there are substantial differences in the oligomerization state and the composition of the active site between CT and other crotonase family members. In addition, the domains in CT contain substantial insertions to the crotonase/Clp fold (Fig. 2, A to C), and some of these inserted segments are important for the oligomerization of the enzyme (Fig. 1C).

A dimer of the CT domain is observed in all the structures that we have determined so far (Fig. 1C), and the organization of this dimer is essentially the same in these different structures. About 5300 Å$^2$ of the surface area of each monomer is buried in the dimer interface, involving mostly residues that are highly conserved among the ACCs (fig. S1). A stable and conserved dimer is consistent not only with the x-ray data but also with gel-filtration and light-scattering studies in solution (19) as well as the $\alpha_2\beta_2$ stoichiometry of the bacterial CT subunits (4). The dimer is formed by the side-to-side arrangement of the two monomers, in such a way that the N domain of one molecule is placed next to the C domain of the other (Fig. 1C). The α6A-α6D insertion in the C domain of one monomer (Fig. 2B) is interdigitated between the β7A-β7D insertion (Fig. 2A) and the core of the N domain of the other monomer (Fig. 1C). The insertion between β4 and β5 of the C domain (Fig. 2B) also contributes to the formation of the dimer.

The structure of the CoA complex shows that the active site of the enzyme is at the interface of the dimer, with approximately equal contributions from the N and C domains of the two monomers (Figs. 1C and 3A). The active site is located in a cavity between the small β sheets (with strands β5, β7, β9, and β11) of the β-β-α superhelix of the two domains (Fig. 3A). Above these two β sheets, the α6 helices of the two domains form two walls, providing additional binding surfaces for the CoA and biotin substrates as well as restricting their directions of approach (Fig. 3, A and B). Residues in this active site are generally well conserved among the various CT domains (fig. S1).

The CoA molecule is mostly associated with the N domain of one molecule in the dimer (Fig. 3A). The N1 and N6 atoms of the adenine base are recognized by hydrogen bonds with the
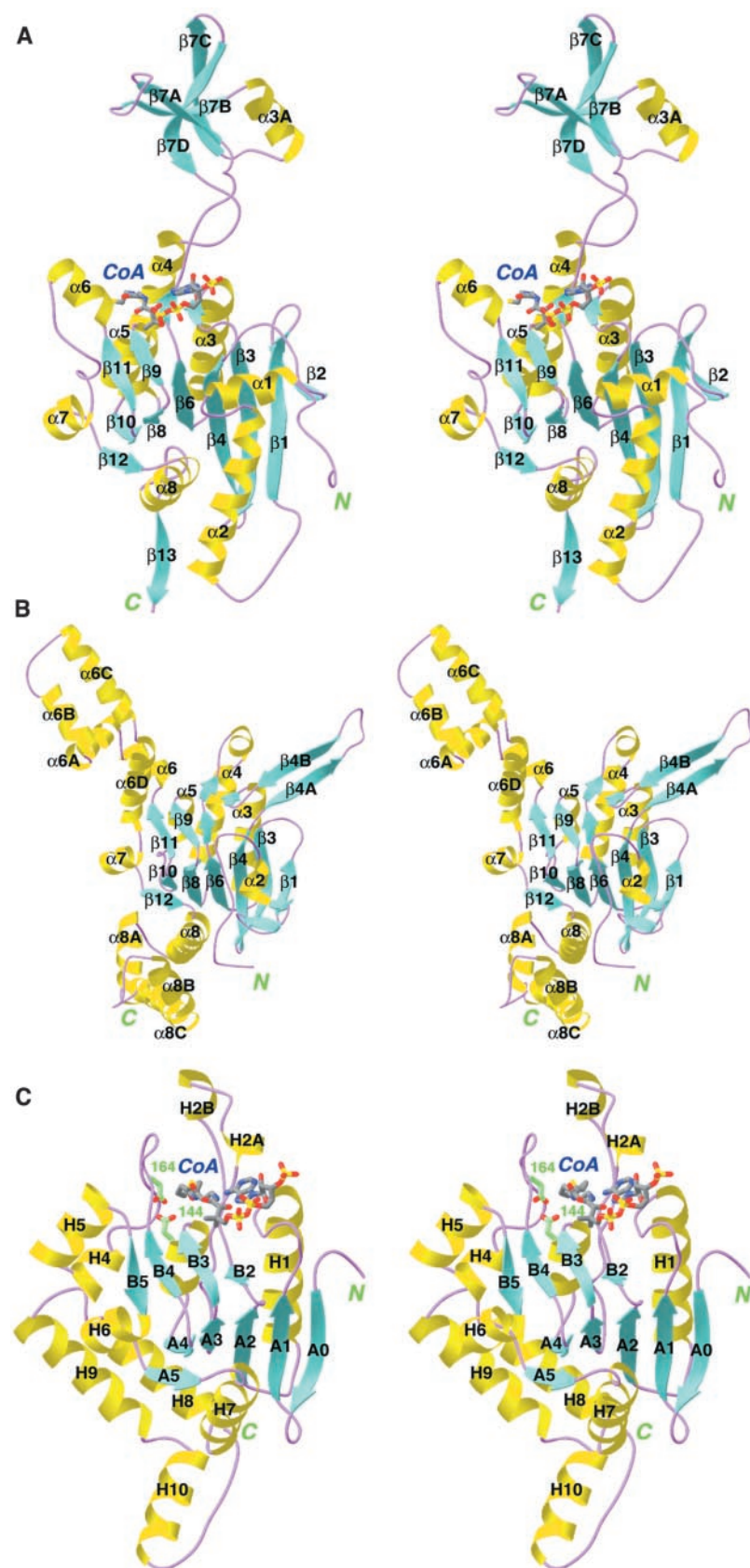


**Fig. 2.** Structures of the domains of CT. (**A**) Schematic drawing in stereo of the structure of the N domain (residues 1484 to 1824) of CT. The secondary structure elements are labeled. The inserted elements to the crotonase fold are named by appending a letter, for example β7A. (**B**) Schematic drawing of the structure of the C domain (residues 1825 to 2202) of CT. Secondary structure elements that are equivalent to those in the N domain are given the same name. (**C**) Schematic drawing of the structure of crotonase monomer in complex with octanoyl-CoA (15). The catalytic residues Glu[144] and Glu[164] are shown in green. Produced with Ribbons (22).

main chain of residues immediately after β7 in the N domain (Fig. 3A). The phosphate groups of CoA are located near the side chains of Arg[1731], Lys[2034'] and Arg[2036'] (the residue numbers with primes indicate the second monomer). The pantotheine arm lies on the surface of the small β sheet in the N domain, and the thiol group is placed in the cavity between the two domains (Fig. 3A). Although acetyl-CoA was used in the crystallization, there was no electron density for the acetyl group, and only the CoA molecule is included in the current atomic model. In the other active site of the dimer, only the electron density for the adenine base of the coenzyme was observed (Fig. 1C), suggesting that the rest of the CoA molecule is disordered in this binding site. The binding mode of CoA in CT shows general similarity to that observed in crotonase (fig. S3).

Based on the location of CoA, it is likely that the biotin substrate is mostly associated with the small β sheet in the C domain of the other monomer in the active site (Fig. 3A). This is partly supported by observations with the *E. coli* biotin ligase/repressor BirA, where the biotin molecule is bound on the surface of a β sheet (20). If this is true, two domains with similar backbone folds in CT apparently recognize rather different chemical entities.

There are only minor changes in the conformation of the core of the monomer or the organization of the dimer when CoA is bound to the enzyme. The rms distance between 1200 equivalent Cα atoms of the dimers of the free enzyme and CoA complex is 0.4 Å.

To characterize the functional roles of the conserved residues in the active site, we mutated many of them and determined the kinetic parameters of the mutants for the reverse reaction, which transfers the carboxyl group of malonyl-CoA to biotin (Fig. 1B and table S1). The largest effect on the catalytic activity was observed with the mutation of the Arg[1954] residue, in the C domain (Fig. 3A), which led to a 75-fold increase in the $K_m$ for malonyl-CoA and a 300-fold decrease in the overall $k_{cat}/K_m$ ($k_{cat}$, turnover number) of the enzyme (table S1). Based on our structures, it is likely that this residue is important for recognizing the carboxyl group of malonyl-CoA. This residue may also recognize the carboxyl group of carboxybiotin (Fig. 1B), which is consistent with the hypothesis that the biotin substrate may be mostly associated with the C domain. Of the charged side chains that interact with the phosphate groups of CoA, mutation of Arg[1731] gave rise to a 14-fold increase in the $K_m$ for malonyl-CoA (table S1). Kinetic results with the Arg[2036'] mutant suggest that it makes only minor contributions to the binding of CoA (table S1).

Kinetic and chemical modification studies of *E. coli* CT have suggested that a cysteine residue may act as a general base to initiate the carboxylation of acetyl-CoA (4). Our structure does not show a cysteine residue in the active
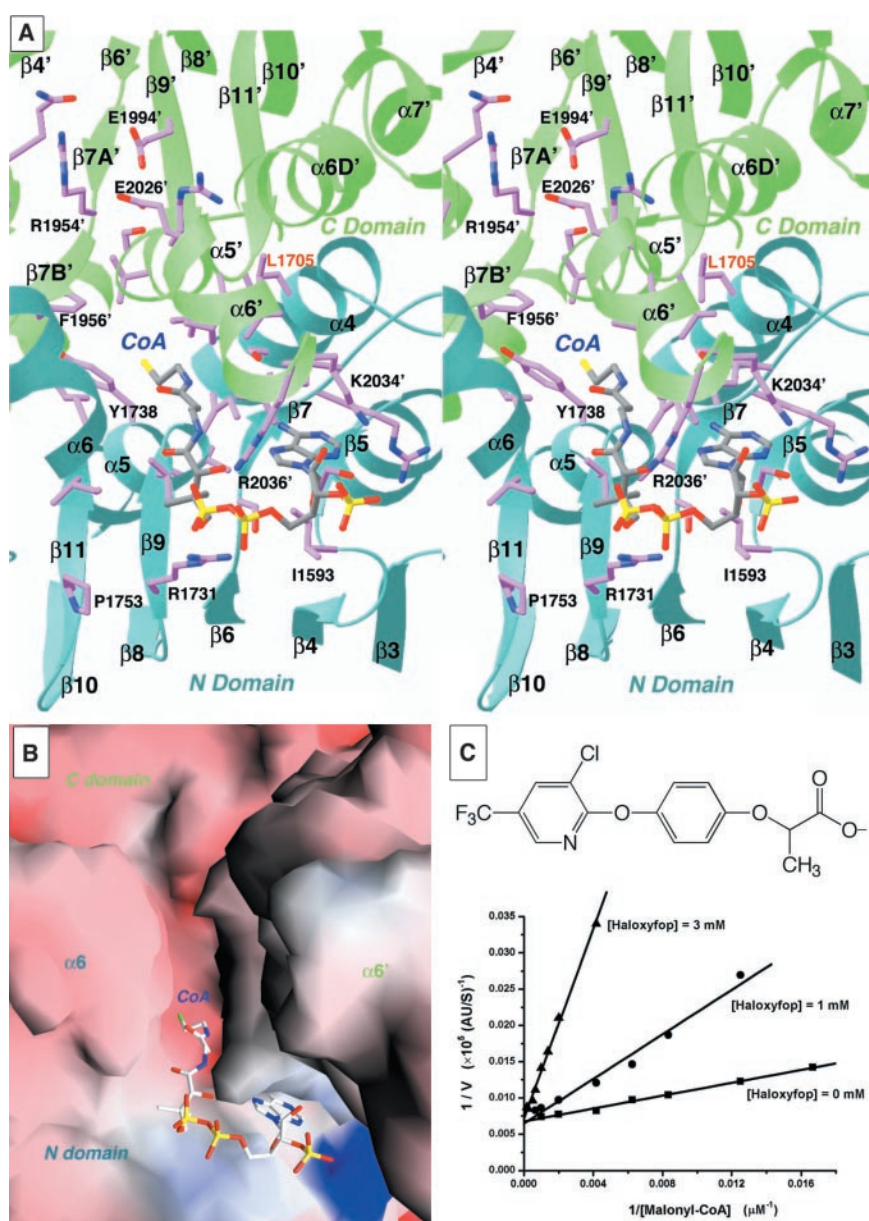


**Fig. 3.** The active site of CT and the binding mode of CoA. (**A**) Schematic drawing in stereo of the active site of CT. The N domain is shown in cyan, and the C domain of the other monomer is shown in green. The side chains of residues in the active site are shown in purple. The prime (') in the labels indicates the C domain of the other monomer of the dimer. (**B**) Molecular surface of the active site region of CT. The side chain of Lys[1764] (in helix α6, 15 Å from the active site) has been removed to facilitate the viewing of the active site. (**C**) Chemical structure of haloxyfop and the double reciprocal plot showing the competitive inhibition of wild-type yeast CT by haloxyfop. (A) was produced with Ribbons (22), and (B) with Grasp (23).

site (Fig. 3A). In addition, we have mutated those residues in the active site that could function as a general base (Tyr[1738], Glu[1994], and Glu[2026]), and kinetic studies showed that these residues are not required for catalysis (table S1). Therefore, the structural and mutagenesis studies support a mechanism in which the N1 atom of biotin itself functions as the general base (Fig. 1B) (3). This contrasts with the crotonase, where an acidic side chain of the enzyme is required for catalysis (Fig. 2C).

Herbicides that target the CT domain are powerful inhibitors of plastid ACC and can kill

sensitive plants by shutting down fatty acid biosynthesis (11). Mutagenesis and biochemical studies showed that an Ile residue in the CT domain plays an important role in determining the sensitivity of the wheat enzyme to the commercial herbicides (11). An Ile→Leu mutation renders the wheat enzyme resistant to the herbicides, and plant ACCs that are insensitive to the herbicides have a Leu residue at this position. This residue is equivalent to Leu[1705] in yeast ACC (fig. S1), which is located on helix α4 of the N domain, deep at the bottom of the active site cavity and also in the dimer interface

**Table 1.** Summary of crystallographic information.

|  | Free enzyme | CoA complex |
| --- | --- | --- |
| Space group | $C2$ | $P2_1$ |
| Maximum resolution (Å) | 2.7 | 2.7 |
| Number of observations | 533,916 | 168,936 |
| $R_{merge}$ (%)* | 7.6 (18.1) | 5.2 (11.6) |
| Resolution range for refinement | 30–2.7 Å | 30–2.7 Å |
| Number of reflections | 113,103 | 59,546 |
| Completeness (%) | 94.5 | 93.2 |
| $R$ factor (%)† | 22.6 (29.9) | 22.6 (29.6) |
| Free $R$ factor (%) | 26.2 (34.8) | 27.9 (33.4) |
| rms deviation in bond lengths (Å) | 0.009 | 0.009 |
| rms deviation in bond angles (°) | 1.4 | 1.4 |

*$R_{merge} = \sum_h \sum_i |I_{hi} - \langle I_h \rangle| / \sum_h \sum_i I_{hi}$. The numbers in parentheses are for the highest resolution shell.
†$R = \sum_h |F_h^o - F_{ch}^c| / \sum_h F_h^o$.

(Fig. 3A). Leu[1705] is unlikely to contact the substrate directly as it is about 8 Å from the thiol group of CoA (Fig. 3A). Thus, mutagenesis and structural information suggest that the herbicides target the active site of CT.

Our kinetic experiments showed that the herbicide haloxyfop is a competitive inhibitor of yeast CT with respect to the substrate malonyl-CoA (Fig. 3C), which is consistent with herbicide binding at the active site. The inhibition of yeast CT by haloxyfop is very weak, with an inhibition constant ($K_i$) of about 0.5 mM (Fig. 3C). The herbicide is a poor inhibitor of the L1705I mutant of yeast CT as well (table S1)(19). These observations, together with those on the apicoplast ACC of *Toxoplasma gondii* (21), indicate that there are additional structural determinants for the binding of these compounds to the active site of CT.

Based on the structural, biochemical, and mutagenesis observations, it may be possible that part of the herbicide is bound in the cavity between the two domains in the active site (Fig. 3A). This cavity extends from the thiol group of CoA to the side chains of Leu[1705] and Ser[1708] and is mostly hydrophobic in nature. The proximity of the inhibitor to Leu[1705] is consistent with its importance in determining the herbicide sensitivity of plant ACCs. The carboxyl group of the inhibitor may mimic the carboxyl group in the malonyl-CoA substrate of the enzyme.

The successful development of inhibitors against the active site of the CT domain of plant ACCs holds promise for the development of inhibitors against other CT domains, especially that of human ACC2. The control of herbicide sensitivity by a single amino acid in the wheat enzyme demonstrates that it could also be possible to identify highly selective inhibitors of mammalian enzymes. For example, inhibitors that target human ACC2 while having only minor effects on ACC1 may prove beneficial for controlling body weight (8). Our structural information about the CT domain provides a starting point for understanding the catalysis by this enzyme as well as for designing and optimizing its inhibitors.

**References and Notes**
1. A. W. Alberts, P. R. Vagelos, in *The Enzymes*, P. D. Boyer, Ed. (Academic Press, New York, 1972), vol. 6, pp. 37–82.
2. S. J. Wakil, J. K. Stoops, V. C. Joshi, *Annu. Rev. Biochem.* **52**, 537 (1983).
3. J. R. Knowles, *Annu. Rev. Biochem.* **58**, 195 (1989).
4. J. E. Cronan Jr., G. L. Waldrop, *Prog. Lipid Res.* **41**, 407 (2002).
5. J. D. McGarry, N. F. Brown, *Eur. J. Biochem.* **244**, 1 (1997).
6. R. R. Ramsay, R. D. Gandour, F. R. van der Leij, *Biochim. Biophys. Acta* **1546**, 21 (2001).
7. J. W. Campbell, J. E. Cronan Jr., *Annu. Rev. Microbiol.* **55**, 305 (2001).
8. L. Abu-Elheiga, M. M. Matzuk, K. A. H. Abo-Hashema, S. J. Wakil, *Science* **291**, 2613 (2001).
9. J. M. Lenhard, W. K. Gottschalk, *Adv. Drug Delivery Rev.* **54**, 1199 (2002).
10. K. L. Levert, G. L. Waldrop, J. M. Stephens, *J. Biol. Chem.* **277**, 16347 (2002).
11. O. Zagnitko, J. Jelenska, G. Tevzadze, R. Haselkorn, P. Gornicki, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 6617 (2001).
12. W. A. Hendrickson, *Science* **254**, 51 (1991).
13. S. E. Polakis, R. B. Guchhait, E. E. Zwergel, M. D. Lane, *J. Biol. Chem.* **249**, 6657 (1974).
14. M. M. Benning *et al.*, *Biochemistry* **35**, 8103 (1996).
15. C. K. Engel, M. Mathieu, J. P. Zeelen, J. K. Hiltunen, R. K. Wierenga, *EMBO J.* **15**, 5135 (1996).
16. J. Wang, J. A. Hartling, J. M. Flanagan, *Cell* **91**, 447 (1997).
17. M. M. Benning, T. Haller, J. A. Gerlt, H. M. Holden, *Biochemistry* **39**, 4630 (2000).
18. A. M. Mursula, D. M. F. van Aalten, J. K. Hiltunen, R. K. Wierenga, *J. Mol. Biol.* **309**, 845 (2001).
19. H. Zhang, Z. Yang, Y. Shen, L. Tong, data not shown.
20. L. H. Weaver, L. Kwon, D. Beckett, B. W. Matthews, *Prot. Sci.* **10**, 2618 (2001).
21. J. Jelenska, A. Sirikhachornkit, R. Haselkorn, P. Gornicki, *J. Biol. Chem.* **277**, 23208 (2002).
22. M. Carson, *J. Mol. Graphics* **5**, 103 (1987).
23. A. Nicholls, K. A. Sharp, B. Honig, *Proteins* **11**, 281 (1991).
24. We thank R. Abramowitz and C. Ogata for setting up the X4A beamline, R. Khayat and X. Tao for help with data collection at the synchrotron, C. Frank for technical support, P. Gornicki for helpful discussions, and Columbia University (L.T.) for financial support.

# The Pentacovalent Phosphorus Intermediate of a Phosphoryl Transfer Reaction

Sushmita D. Lahiri,[1] Guofeng Zhang,[2] Debra Dunaway-Mariano,[2]* Karen N. Allen[1]*

Enzymes provide enormous rate enhancements, unmatched by any other type of catalyst. The stabilization of high-energy states along the reaction coordinate is the crux of the catalytic power of enzymes. We report the atomic-resolution structure of a high-energy reaction intermediate stabilized in the active site of an enzyme. Crystallization of phosphorylated β-phosphoglucomutase in the presence of the Mg(II) cofactor and either of the substrates glucose 1-phosphate or glucose 6-phosphate produced crystals of the enzyme–Mg(II)–glucose 1,6-(bis)phosphate complex, which diffracted x-rays to 1.2 and 1.4 angstroms, respectively. The structure reveals a stabilized pentacovalent phosphorane formed in the phosphoryl transfer from the C(1)O of glucose 1,6-(bis)phosphate to the nucleophilic Asp8 carboxylate.

The unique ability of enzymes to synchronize multiple interactions at specific sites with reactants, intermediates, and products defines the reaction pathway and is the source of the enormous rate acceleration accomplished by enzymes ($10^7$- to $10^{19}$-fold) (1–9). The study of the reaction pathway as it proceeds within the protein walls continues to be the focal point of bioorganic chemistry. Kinetic investigations of

enzyme catalysis on natural substrates and the study of unreactive mimics of the ground state and transition state have contributed greatly to our understanding of the steric and electronic structures of reaction coordinate species (10–15). An actual image of each chemical species formed is, however, the ultimate goal.

X-ray crystallography is the conduit to such images (16, 17). If a chemical species is