Crystal Structure of Biotin Carboxylase in Complex with Substrates and Implications for Its Catalytic Mechanism*

Chi-Yuan Chou‡§1, Linda P. C. Yu‡, and Liang Tong‡2

From the ‡Department of Biological Sciences, Columbia University, New York, New York 10027 and the §Department of Life Sciences and Institute of Genome Sciences, National Yang-Ming University, Taipei 112, Taiwan

Biotin-dependent carboxylases are widely distributed in nature and have important functions in many cellular processes. These enzymes share a conserved biotin carboxylase (BC) component, which catalyzes the ATP-dependent carboxylation of biotin using bicarbonate as the donor. Despite the availability of a large amount of biochemical and structural information on BC, the molecular basis for its catalysis is currently still poorly understood. We report here the crystal structure at 2.0 Å resolution of wild-type Escherichia coli BC in complex with its substrates biotin, bicarbonate, and Mg-ADP. The structure suggests that Glu296 is the general base that extracts the proton from bicarbonate, and Arg338 is the residue that stabilizes the enolate biotin intermediate in the carboxylation reaction. The BC domain of BC is positioned closer to the active site, leading to a 2 Å shift in the bound position of the adenine nucleotide and bringing it near the bicarbonate for catalysis. Overall, this first structure of BC in complex with substrates offers unprecedented insights into the molecular mechanism for the catalysis by this family of enzymes.

Biotin-dependent carboxylases are widely distributed in nature and have important functions in fatty acid, cholesterol, and amino acid metabolism, gluconeogenesis, insulin secretion, and other cellular processes. These enzymes include acetyl-CoA carboxylase (ACC)3 (1–3), propionyl-CoA carboxylase (4, 5), methylcrotonyl-CoA carboxylase (6), and pyruvate carboxylase (7), and they have significant implications for human health. Mutations in propionyl-CoA carboxylase, methylcrotonyl-CoA carboxylase, and pyruvate carboxylase are linked to serious human diseases, such as propionic acidemia, lactic acidosis, and mental retardation, and the ACCs are attractive targets for the development of new therapeutic agents against diabetes, obesity, and other manifestations of the metabolic syndrome (8–10).

The biotin-dependent carboxylases generally contain three components: a biotin carboxylase (BC) activity that produces an activated carboxyl group covalently attached to biotin, a carboxyltransferase activity that transfers the carboxyl group to the acceptor molecule (acetyl-CoA, propionyl-CoA, pyruvate, and others), and a biotin carboxyl carrier protein (BCCP) that carries the biotin group, linked covalently to the side chain of a lysine residue. The carboxyltransferase component can differ greatly in sequence and structure, depending on the identity of the acceptor molecule for the carboxyl group. For example, there is no similarity between the carboxyltransferase components of ACC and pyruvate carboxylase (11–14). On the other hand, the BC activity is shared among all of these enzymes, and the BC components have significant amino acid sequence homology.

BC catalyzes the ATP-dependent carboxylation of biotin, and bicarbonate is the donor of the carboxyl group in this reaction (15, 16). This activity has been studied extensively over the years by kinetic and mutagenesis experiments, and structural information is available on the free enzyme and ATP complex of BC from several different carboxylases (13, 14, 17–24). The structure of BC contains three domains, the A, B, and C domains, and the active site is located at the interface between the B domain and the other two domains (Fig. 1A). Bacterial BC is a stable dimer (Fig. 1A), with an extensive, although rather hydrophilic, interface (21). In comparison, the BC component of eukaryotic ACC is monomorphic in solution and is catalytically inactive (19, 25).

Despite this large body of knowledge on BC, the binding modes of the biotin and bicarbonate substrates to this enzyme are still not known, and therefore the molecular basis for its catalysis is currently still poorly understood. We report here the crystal structure at 2.0 Å resolution of wild-type BC from Escherichia coli in complex with biotin, bicarbonate, and Mg-ADP. This first structure of BC in complex with its substrates offers unprecedented insights into the molecular mechanism for the catalysis by this family of enzymes.

MATERIALS AND METHODS

Crystals of wild-type E. coli BC in complex with biotin, bicarbonate, and Mg-ADP were obtained at room temperature by
the sitting drop vapor diffusion method. The protein solution was at 10 mg/ml concentration, which also included 5 mM Mg-ADP, 100 mM biotin, 20 mM bicarbonate, and 3.9% (w/v) sorbitol as an additive. The reservoir solution contained 23% (w/v) polyethylene glycol 3350 and 0.12 mM Li$_2$SO$_4$. Crystals were soaked overnight in a solution of 2.5 mM Mg-ADP, 200 mM biotin, 10 mM bicarbonate, 23% (w/v) polyethylene glycol 3350, 0.05 mM Li$_2$SO$_4$, and 2.1% (w/v) sorbitol, cryo-protected in this solution supplemented with 12% (v/v) glycerol, and flash-frozen in liquid nitrogen.

X-ray diffraction data were collected at 100 K at the National Synchrotron Light Source beamline X4C on a Mar CCD (x-ray wavelength 0.979 Å). The diffraction images were processed and scaled with the HKL package (26). The crystal belongs to space group $P_2_1_2_1_2_1$, with cell parameters of $a = 83.3$, $b = 106.2$, and $c = 121.5$ Å. There is a dimer in the asymmetric unit.

The structure was solved by the molecular replacement method with the program Phaser (27), using the structure of the ATP complex of the E288K mutant of BC (Protein Data Bank code 1DV2) as the search model (18). Manual rebuilding of the structure model was performed with O (28), and structure refinement was carried out with the programs CNS (29) and Refmac (30). The data processing and refinement statistics are summarized in Table 1.

The E296A and R338A mutants of BC were produced by using the QuikChange kit (Stratagene) and verified by DNA sequencing. The enzyme activity of the wild-type and mutant BC was determined spectrophotometrically at 340 nm, following a published protocol (31). The reaction mixture contained 100 mM HEPES (pH 8.0), 5 mM ATP, 8 mM MgCl$_2$, 40 mM biotin, 0.2 mM NADH, 0.5 mM phosphoenolpyruvate, 7 units of lactate dehydrogenase, 4.2 units of pyruvate kinase, and varying concentrations of bicarbonate. For assays that varied biotin concentration, the reaction mixture contained 100 mM HEPES (pH 7.4), 0.5 mM ATP, 8 mM MgCl$_2$, 40 mM bicarbonate, 0.2 mM NADH, 0.5 mM phosphoenolpyruvate, 7 units of lactate dehydrogenase, and 4.2 units of pyruvate kinase. To maintain the ionic strength of the solution, KCl was added to make the total concentration of biotin and KCl 200 mM.

The amount of carboxybiotin produced by BC was determined by a $^{14}$C bicarbonate fixation assay (31). The reaction was carried out in a total volume of 0.5 ml at 25°C. The mixture contained 5 mM ATP, 8 mM MgCl$_2$, 100 mM biotin, 70 mM KHCO$_3$, 100 mM HEPES (pH 8.0), and 3.5 nCi (0.875 nCi) of an NaH$^{14}$CO$_3$ solution. The reaction was initiated by the addition of enzyme. After 1 h of incubation, the reaction was stopped by the addition of 3.6 $\mu$L of a saturated KOH solution. Then 200 $\mu$L of a 2.0 $\times$ M BaCl$_2$, 0.1 $\times$ M Ba(OH)$_2$ solution were added to precipitate the carbonate. The tubes were mixed and sat for 2 min and then were centrifuged at 13,000 rpm for 5 min. A 100-$\mu$L aliquot from the supernatant was transferred to a scintillation vial containing 3 ml of scintillation mixture and counted by a liquid scintillation analyzer (TRI-Carb 2900TR; PerkinElmer).

The amount of ADP produced was measured to determine the stoichiometry of carboxybiotin and ADP formation. The reaction was the same as that for the carboxybiotin assay except that no $^{14}$C bicarbonate was added. Upon mixing the reaction components, an aliquot of 100 $\mu$L was removed immediately and after a 1-h incubation. The amount of ATP in each aliquot was measured by end point analysis, in a total volume of 1.0 ml with 10 units of hexokinase, 5 units of glucose-6-phosphate dehydrogenase, 0.5 mM glucose, 0.4 mM NADP$^+$, 8 mM MgCl$_2$, and 100 mM HEPES (pH 8.0). The amount of ADP produced was determined from the difference between the amount of ATP at the beginning and after the 1-h incubation.

Crystals of the E296A mutant were obtained by the sitting drop vapor diffusion method at room temperature. The crystallization condition was essentially the same as that for the wild-type BC, and the crystal was soaked overnight in a solution containing 2.5 mM Mg-ADP, 200 mM biotin, 10 mM bicarbonate, 23% (w/v) polyethylene glycol 3350, 0.05 mM Li$_2$SO$_4$, and 2.1% (w/v) sorbitol. X-ray diffraction data were collected at the X29A beamline at the National Synchrotron Light Source. The structure was solved with the program COMO (32), and the structure refinement followed the protocol described above for the wild-type enzyme. The crystallographic statistics are summarized in Table 1.

### RESULTS AND DISCUSSION

**Overall Structure of BC in Complex with Substrates**—To ensure the binding of the biotin and bicarbonate substrates, we included high concentrations of both compounds (100 and 20 mM, respectively) as well as 5 mM Mg-ADP in the co-crystallization experiments. After many attempts, we were able to produce crystals of the wild-type BC from *E. coli* in complex with biotin, bicarbonate, and Mg-ADP and determined its structure at 2.0 Å resolution (Table 1). The refined atomic model has excellent agreement with the crystallographic data and the expected bond lengths and bond angles. About 92% of the residues are in the most favored region of the Ramachandran plot, and none of the residues are in the disallowed region.

The two monomers of the BC dimer have essentially the same conformation (Fig. 1A), with r.m.s. distance of 0.25 Å between equivalent C$_\alpha$ atoms (excluding residues in the B domain, which have weak electron density). The overall structure of the BC monomer in this complex is similar to other structures of *E. coli* BC reported earlier. For example, the r.m.s. distance between equivalent C$_\alpha$ atoms of this structure and those in the structure of the active site mutant (E288K) in complex with ATP (18) is 0.4 Å. The BC dimer in the two structures is also similar, with an r.m.s. distance of 0.45 Å. The B domain

<table>
<thead>
<tr>
<th>Protein</th>
<th>Wild-type BC</th>
<th>E296A mutant</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P2_1_2_1_2_1</td>
<td>P2_1_2_1_2_1</td>
</tr>
<tr>
<td>Cell parameters, a, b, c (Å)</td>
<td>83.3, 106.2, 121.5</td>
<td>81.3, 114.7, 122.2</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>30-2.0 (2.07-2.00)</td>
<td>30-1.9 (1.97-1.90)</td>
</tr>
<tr>
<td>Rmerge (%)</td>
<td>6.3 (34.3)</td>
<td>6.9 (37.0)</td>
</tr>
<tr>
<td>Rf (Å)</td>
<td>21.6 (4.3)</td>
<td>26.2 (5.2)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99 (100)</td>
<td>100 (100)</td>
</tr>
<tr>
<td>No. of reflections</td>
<td>69,308</td>
<td>86,281</td>
</tr>
<tr>
<td>R factor (%)</td>
<td>18.2 (19.2)</td>
<td>19.8 (21.3)</td>
</tr>
<tr>
<td>Free R factor (%)</td>
<td>21.6 (25.0)</td>
<td>22.6 (25.9)</td>
</tr>
<tr>
<td>r.m.s. deviation in bond angles (degrees)</td>
<td>0.009</td>
<td>0.009</td>
</tr>
</tbody>
</table>

* The numbers in parentheses are for the highest resolution shell.
has noticeable differences in its position in the two structures and is excluded from these comparisons (see below).

Binding Modes of Bicarbonate, Biotin, and Mg-ADP—The crystallographic analysis showed that biotin (Fig. 1B), bicarbonate (Fig. 1C), and Mg-ADP (Fig. 1D) have clearly defined electron density, suggesting that they are well ordered in the active site of BC. Two of the oxygen atoms (O1 and O2) of bicarbonate are recognized by bidentate ion pair interactions with the side-chain guanidinium group of Arg292 (in the H9252 N loop in the C domain), and the O2 atom also has a hydrogen-bonding interaction with the main-chain amide of Val295 (Fig. 1E). The third oxygen atom (O3) of bicarbonate is located within 2.5 Å of the side-chain carboxylate group of Glu296. This oxygen therefore is likely to carry the proton of bicarbonate, and Glu296 probably functions as the general base during the catalysis to extract this proton from bicarbonate (see below).

Biotin is placed over the side chains of Tyr82 (in the A domain) and Val295 (in the H9252 N loop in the C domain) in the active site of BC (Fig. 1E). The carbonyl oxygen atom in the ureido ring has hydrogen-bonding interactions with the side chain of Arg238 (in strand β17 in the C domain), which also shows interactions with the bicarbonate. Unexpectedly, the structure reveals that the N1’ atom of biotin is located within 2.7 Å of the O1 atom of bicarbonate. This close approach probably helps to position the biotin group in the active site, and also has important implications for the catalysis by BC (see below).

The biotin concentration used in the crystallization experiment (100 mM) is about 3-fold higher than the \( K_m \) value based on our kinetic assays (Table 2), although an earlier study suggested that the \( K_m \) for biotin is about 130 mM (31). The crystallographic analyses showed that the average temperature factor for biotin (43 Å²) is comparable with that of ADP, suggesting that biotin is present in high occupancy in the active site.

The magnesium ion is coordinated by six ligands arranged in an octahedral fashion (Fig. 1E), including a terminal oxygen atom from the α- and β-phosphate groups of ADP, an oxygen...
from the side-chain carboxylate groups of Glu276 (strand β15 in the C domain) and Glu288 (strand β16) and two water molecules. The β-phosphate cannot rotate freely for MgADP binding, which may explain the fact that positional isotope exchange was not observed in kinetic studies (15). The bound position of ADP observed here is significantly different from that of ATP bound to the E288K mutant (18) as well as that of AMPPNP bound to Staphylococcus aureus BC (24) (see below).

The Active Site of BC—Earlier studies showed that the B domain of BC undergoes a large movement to close over the active site upon ATP binding (18). In our structure, the B domain is placed even closer to the active site, corresponding to roughly a 7° rotation compared with that in the ATP complex (Fig. 2A). Even with this additional movement, the active site of BC in the current complex, especially the biotin molecule, is still exposed to the solvent (Fig. 2B). In the reaction with the natural BCCP-biotin substrate, BCCP may help cover up this active site, which may explain why BCCP-biotin is a much better substrate for BC than biotin alone (16). The electron density for the B domain is weaker than that for the other domains in our structure, and this conformational flexibility of the B domain could also be stabilized upon the binding of BCCP.

The further closure of the B domain in our structure has a dramatic impact on the bound position of the adenine nucleotide (Fig. 2C). The adenine base interacts primarily with residues in the B domain, especially a hydrogen bond between its N6 atom and the main-chain carbonyl oxygen of residue Lys202 (Fig. 2C). Due to the further closure of the B domain in our structure, this carbonyl oxygen moves by 2.2 Å as compared with that in the structure of the ATP complex (18), and the bound position of ADP is also shifted by about 2 Å relative to that of ATP (Fig. 2C). This further closure of the B domain brings the phosphate group of the adenine nucleotide closer to the bound position of the bicarbonate group (Fig. 2C) and therefore should make the BC enzyme more competent for catalysis. Residues in the glycine-rich loop (GGGGGRG) in the B domain have very weak electron density, and they may become better ordered in the presence of ATP and interact with its phosphate group (18).

The ATP complex of E. coli BC was obtained with a kinetically inactive E288K mutant (18), which may have affected the conformation of the B domain and the binding mode of ATP. The structure of wild-type S. aureus BC in complex with AMP-PNP was reported recently (24). A comparison of our structure with that complex showed that there are still differences in the conformation of the B domain, such that the positions of the adenine base and the ribose are different by about 1.3 Å between the two structures.

Compared with the structure of the ATP complex, there are only a few conformational changes in the binding sites for biotin and bicarbonate (Fig. 2C) (18). The largest difference is in the side chain of Gln294, whose conformation in the ATP complex would clash with the bicarbonate. Smaller conformational differences are seen for the
side chains of Asn$^{290}$, Arg$^{292}$, and Glu$^{296}$. On the other hand, Arg$^{338}$ has essentially the same conformation in the two structures (Fig. 2C).

**Molecular Basis for Catalysis by BC**—To derive further insights into the catalysis by BC, we modeled the binding mode of ATP into the active site of the current structure, by manually superimposing the AMP portion of the ATP in the E288K complex (18) onto the AMP portion of ADP in our current structure. We did not make any further adjustments in the conformation of the $\beta$- and $\gamma$-phosphate groups. Modeling the binding mode of ATP based on that of AMPPNP in the *S. aureus* BC structure (24) produced a very similar result. In this model of the ATP complex, the magnesium ion is coordinated by a terminal oxygen atom from the $\alpha$- and $\gamma$-phosphates (Fig. 3A). A second magnesium ion coordinates the $\beta$- and $\gamma$-phosphates (13, 24). After the $\gamma$-phosphate is transferred to bicar-
bonate, the two metal ions are likely to remain associated with ADP (Fig. 3B), one between the α- and β-phosphate (Fig. 1E) and one with the β-phosphate.

Most importantly, the O1 atom of bicarbonate is located about 4 Å from and directly below the γ-phosphorus atom (Fig. 3A), suggesting that this oxygen is in the appropriate location for initiating a nucleophilic attack on the phosphorus atom. As discussed above, it is possible that an additional conformational rearrangement of the B domain can further reduce the distance between ATP and bicarbonate and thereby facilitate the catalysis. This rearrangement may also reduce the surface exposure of the active site region.

Our structural observations provide unprecedented molecular insights into the catalysis by BC. Earlier studies have suggested the presence of a carboxyphosphate intermediate, although it has not been observed experimentally for the BC reaction (15, 16). Our structure shows that the O1 atom of bicarbonate can initiate the nucleophilic attack on the γ-phosphate of ATP (Fig. 3A), which would lead to the formation of carboxyphosphate (Fig. 3B). Glu296 is the best candidate for the general base that deprotonates the bicarbonate. The pKₐ of bicarbonate is normally 10.3, which would make it difficult for a Glu side chain to extract its proton. However, the bicarbonate interacts with several positively charged side chains of BC, including Arg338 and Arg892 (Fig. 1E), which should reduce its apparent pKₐ value. In addition, the Glu296 side chain interacts with the side chain of Glu211 (Fig. 1E), which should disfavor ionization and increase the apparent pKₐ value of Glu296. It is known from studies of other proteins that the pKₐ of an acidic side chain can be >9 in special cases (33). Overall, the combined effects of the local environment of the BC active site probably make the pKₐ values of Glu296 and bicarbonate closer to each other, facilitating the deprotonation of bicarbonate to initiate the reaction.

Decomposition of the unstable carboxyphosphate produces CO₂ and PO₄³⁻. One of the oxygen atoms on the phosphate group is expected to be located close to the N1 atom of biotin, since the attacking O1 atom of bicarbonate is within 2.7 Å of this N1 atom (Fig. 1E). Therefore, the phosphate group is probably the general base that extracts the proton from the N1 atom (15). This leads to the enolization of the biotin ring, and the oxygen is stabilized by interaction with the Arg338 side chain (Fig. 3B). Finally, attack on the N1 atom by CO₂ leads to the formation of carboxybiotin (Fig. 3B). Overall, the structural observations define the molecular basis for the catalysis and provide the first direct experimental evidence in support of the kinetic mechanism proposed earlier (15).

The pKₐ for the proton on the N1 atom of free biotin is 17.4 (34). The interaction between the ureido oxygen and Arg338 in the active site of BC is expected to reduce this pKₐ value, which should facilitate the extraction of this proton by the phosphate group (the pKₐ of HPO₄²⁻ is 12.7).

Mutagenesis Studies Support the Structural Observations—Our structural analyses identify Glu296 and Arg892 as two of the residues with important roles for catalysis by BC. The R338S

---

4 C.-Y. Chou and L. Tong, unpublished results.
is equivalent to an Arg residue in propionyl-CoA carboxylase and methylcrotonyl-CoA carboxylase, whose mutation is linked to diseases in humans (35).

Earlier studies of BC have shown that mutation of active site residues could cause a misalignment of the substrates, such that ATP hydrolysis is no longer coupled to biotin carboxylation (31, 35). We determined the ratio of ATP hydrolysis to carboxylbiotin formation for wild-type BC and the E296A and R338A mutants. The ratio is 1.08 for the wild-type enzyme, suggesting good coupling between the two steps of the catalysis. For the E296A and R338A mutants, the ratio is 2.5 and 6.2, respectively, suggesting a decoupling between the two steps. The ratio for the R338A mutant is similar to that observed earlier for the R338Q and R338S mutants (35). Therefore, the loss in the overall biotin carboxylase activity of the E296A and R338A mutants is larger than the values presented in Table 2, which are determined from monitoring ATP hydrolysis. The larger ratio for the Arg338 mutants than for the Glu296 mutant may be consistent with the importance of Arg338 in stabilizing the transition-state of the biotin carboxylation step.

For both the E296A and R338A mutants, the possibility exists that some of the observed activity could be due to the endogenous wild-type E. coli BC, co-purifying with the His-tagged mutant protein by forming a heterodimer with it. However, the mutant proteins were highly overexpressed, producing about 20–30 mg of purified protein/liter of E. coli culture, which should minimize the amount of heterodimers with endogenous BC in the purified sample. Moreover, earlier studies have shown that a heterodimer containing one wild-type BC monomer and one inactive, mutant BC monomer is essentially inactive (40). Therefore, the contribution of the endogenous BC to the activity of both mutants is probably very small (if any).

Mutations of other residues in the active site region have also been shown to have deleterious effects. These studies demonstrate the functional importance of Glu211, Lys238, Asn290, and Arg292 in the biotin and bicarbonate binding region (Fig. 1E) (31, 36) as well as residues in the ATP binding region (18, 37). Overall, our structural observations are consistent with and can explain these results from the mutagenesis and kinetic studies.

Implications for Catalysis by Related Enzymes—The BC component is shared among most of the biotin-dependent carboxylases. The residues that play important roles in substrate binding and/or catalysis as identified from this structure are strictly conserved among these enzymes (Fig. 5). The two residues that have van der Waals interactions with biotin (Tyr82 and Val295) are also highly conserved. Tyr82 has a conservative substitution to Trp in the eukaryotic ACCs, and human methylcrotonyl-CoA carboxylase is the only enzyme that contains a Cys residue at this position (Fig. 5). Overall,
Structural Studies of Catalysis by Biotin Carboxylase

our structure of the substrate complex of *E. coli* BC is likely to have strong implications for the catalysis by these other enzymes, and they probably have conserved substrate binding modes and catalytic mechanisms.

BC belongs to the ATP-grasp superfamily of enzymes, which catalyzes the ATP-dependent formation of an amide bond between a carboxylate compound and an amine compound (41, 42). These enzymes share a similar overall fold and similar binding modes for ATP. However, the chemical identities of their carboxylate and amine substrates are highly divergent, and the sequence and structural conservation for the binding sites of these substrates is much lower. For example, the residues that are important for binding biotin in BC are not conserved in *E. coli* carbamoyl phosphate synthetase (CPS; Fig. 5) (43). On the other hand, Arg<sup>292</sup> of BC is conserved in CPS and serves that the bicarbonate binding site in BC does not exist in CPS.

## Acknowledgments

We thank R. Abramowitz and J. Schwanof for helpful discussions, and C. Huang for careful reading of the manuscript.

## References