Protein Biotinylation Visualized by a Complex Structure of Biotin Protein Ligase with a Substrate*

Biotin protein ligase (BPL) catalyzes the biotinylation of the biotin carboxyl carrier protein (BCCP) only at a special lysine residue. Here we report the first structure of BPL-BCCP complex crystals, which are prepared using two BPL mutants: R48A and R48A/K111A. From a detailed structural characterization, it is likely that the mutants retain functionality as enzymes but have a reduced activity to produce the reaction intermediate biotinyl-5′-AMP. The observed biotin and partly disordered ATP in the mutant structures may act as a non-reactive analog of the substrates or biotinyl-5′-AMP, thereby providing the complex crystals. The four crystallo-enzymes but have a reduced activity to produce the reaction intermediate biotinyl-5′-AMP. The observed biotin and partly disordered ATP in the mutant structures may act as a non-reactive analog of the substrates or biotinyl-5′-AMP, thereby providing the complex crystals. The four crystallographically independent BPL-BCCP complexes obtained can be classified structurally into three groups: the formation stages 1 and 2 with apo-BCCP and the product stage with biotinylated holo-BCCP. Residues responsible for the complex formation as well as for the biotinylation reaction have been identified. The C-terminal domain of BPL shows especially large conformational changes to accommodate BCCP, suggesting its functional importance. The formation stage 1 complex shows the closest distance between the carboxyl carbon of biotin and the special lysine of BCCP, suggesting its relevance to the unobserved reaction stage. Interestingly, bound ATP and biotin are also seen in the product stage, indicating that the substrates may be recruited into the product stage complex before the release of holo-BCCP, probably for the next reaction cycle. The existence of formation and product stages before and after the reaction stage would be favorable to ensure both the reaction efficiency and the extreme substrate specificity of the biotinylation reaction.

Biotin-dependent carboxylases constitute a ubiquitous family of enzymes that catalyze the transfer of carbon dioxide between metabolites using the biotin moiety as a carboxyl carrier (1, 2). The attachment of biotin to the biotin-dependent enzymes is catalyzed by the biotin protein ligase (BPL) in two steps (Reactions 1 and 2).

\[
\text{Biotin} + \text{ATP} \rightleftharpoons \text{biotinyl-5′-AMP} + \text{PP}_i \quad \text{REACTION 1}
\]

\[
\text{Biotinyl-5′-AMP} + \text{apo-BCCP} \rightarrow \text{holo-BCCP} + \text{AMP} \quad \text{REACTION 2}
\]

Firstly, BPL activates biotin at the expense of ATP to the reaction intermediate biotinyl-5′-AMP in which the carboxyl group of inert biotin is activated by the addition of an adenylate group. Subsequently, the biotin moiety of biotinyl-5′-AMP is transferred to the ε-amino group of a specific lysine residue of the target protein (e.g. the biotin carboxyl carrier protein (BCCP) subunit of acetyl-CoA carboxylase (3)). The biotinylated holo-BCCP subunit carries a covalently bound carboxyl unit between different active sites of the multienzyme biotin-dependent carboxylase complexes, which play essential roles in the fatty acid synthesis, the amino acid degradation, and the CO₂ fixation (4, 5). In bacteria and eukaryotes, biotinylation is essential to initiate the first step of fatty acid biosynthesis, which is catalyzed by acetyl-CoA carboxylase. Because human biotin-dependent carboxylases are the direct targets for the development of anti-obesity and anti-diabetes agents, structural information on these enzymes is important for drug discovery research (6). In archaea, the function of acetyl-CoA/propionyl-CoA carboxylases is different due to the absence of usual fatty acids in membranes: they act as CO₂ fixation enzymes in the modified 3-hydroxypropionate cycle to assimilate CO₂ into the cell (5). In addition, the structural basis of biotinylation is important to develop useful applications in protein engineering, such as the high affinity biotin tagging for protein purification (7) and the molecular imaging by quantum dots (8).

The understanding of the first step of BPL reaction at the atomic level was reached based on the crystal structures of BPL from Pyrococcus horikoshii OT3 (PhBPL) in complex with various biological ligands including biotinyl-5′-AMP (9) as well as on the crystal structure of BPL from Escherichia coli (EcBirA) in complex with the reaction intermediate analog biotinol-5′-AMP (10). However, scientists have not yet been able to resolve the structural details in the second step of biotinylation reaction in which the activated biotin is transferred to the special lysine of BCCP. This exceptionally selective post-translational modification makes understanding how the proteins BPL and
BCCP carry out the biotin transfer of particular interest. To elucidate the biotin transfer reaction, several biophysical or biochemical studies on the BPL-BCCP complex have been performed (e.g. an NMR study (11), mutagenesis studies (12, 13), and a chemical cross-linking study (14)). Despite these efforts, the structural mode of the biotin transfer reaction is not fully understood due to the absence of three-dimensional structures for the BPL-BCCP complex.

To gain insight into a complex array of interactions between BPL and BCCP at biotinylation, we co-crystallized and determined the crystal structures of BPL-BCCP from *P. horikoshii* OT3. For the experiments, we used the biotinyl domain of PhBCCP, the C-terminal 73-residue fragment of the 149-residue long hypothetical methylmalonyl-CoA decarboxylase gamma chain (PhBCCPΔN76). Although the co-crystallization of the wild-type PhBPL with PhBCCPΔN76 was unsuccessful, using PhBPL mutants carrying a single mutation R48A (PhBPL*), and a double mutation R48A/K111A (PhBPL**), successfully yielded the complex crystals. Here we report the crystal structures of PhBPL*, PhBPL*-biotinyl-5’-AMP, PhBPL**, PhBPL**-biotin-adenosine, PhBPL**-biotinol-5’-AMP, PhBCCPΔN76, and the complexes PhBPL*-biotin-adenosine PhBCCPΔN76 and PhBPL**-biotin-adenosine-holo-PhBCCPΔN76. The information obtained from these structures provides a good starting point to understand the structural basis of protein biotinylation.

**EXPERIMENTAL PROCEDURES**

**Protein Expression, Purification, Crystallization, and Data Collection**—The expression and purification of PhBPL, PhBPL*, PhBPL**, and PhBCCPΔN76 were performed as described elsewhere (15, 16). Crystals of the ATP liganded form of wild-type PhBPL (PhBPL-ATP) were prepared by adding 5 mM ATP to the previously reported crystallization condition for the unliganded
wild-type PhBPL (15). Crystals of the liganded forms of PhBPL mutants were prepared using the wild-type crystallization condition except for adding ligands: 5 mM ATP and 5 mM biotin for PhBPL*-biotinyl-5′-AMP and PhBPL**-biotin-adenosine; 5 mM biotinol-5′-AMP for PhBPL**-biotinol-5′-AMP. Other forms of crystals were prepared as described elsewhere (16). All data were collected at 100 K using synchrotron radiation on a Jupiter 210 charge-coupled device at the beamline BL26B1 of SPring-8, Japan (17). Diffraction data were processed and scaled with the HKL-2000 program suite (18). Data collection statistics are summarized in tables (Table 1 and supplemental Tables S1 and S2).

RESULTS AND DISCUSSION

Structures of PhBPL—To examine the effect of mutations R48A and R48A/K111A on the structure and reactivity of PhBPL, we determined the crystal structures of PhBPL* and PhBPL** and their liganded forms PhBPL*-biotinyl-5′-AMP, PhBPL**-biotin-adenosine, and PhBPL**-biotinol-5′-AMP (supplemental Table S1). Overall, all the structures were found isomorphous to the wild-type PhBPL (9). The crystals contain one dimer in the asymmetric unit. The mutants have essentially the same overall protomer conformation as the wild-type PhBPL: a Cα superposition between any pair of the protomers results in the root-mean-square deviation (r.m.s.d.) value of <1 Å. The liganded forms of crystals were prepared by a cocrystallization with ligands, to avoid crystal packing effects. The apo structures of PhBPL mutants show that the active site loop Gly45-Trp53 is disordered. In the liganded structures, the loop becomes ordered and closes over the bound substrates to form the main hole of the active site. This ligand-induced ordering of the active site loop is also observed in the structures of wild-type PhBPL (9). The cocrystallization of PhBPL* with ATP and biotin provided the PhBPL*-biotinyl-5′-AMP complex where the U-shaped biotinyl-5′-AMP was found in the bifurcated main hole of BPL*, indicating the retained functionality of the single mutant (Fig. 1A, and supplemental Table S3 Fig. S1). However, the same cocrystallization condition using PhBPL** resulted in a substrate complex where the active site hole was occupied by biotin and adenosine. Because ATP dominates the adenosine nucleotide species in the crystallization solution, it is likely that the modeled adenosine is an ordered part of ATP. The nucleotide binding mode in the double mutant was compared with that in the wild-type enzyme (supplemental Table S3 and Fig. S2): PhBPL-ATP (PDB ID 1X01) and PhBPL-ADP (PDB ID 1WNL). In the wild-type enzyme, the phosphate part of nucleotides is recognized by polar interactions with the basic residues Arg48, Arg51, Lys111, and Arg231. Thus the mutations R48A and K111A would induce the disordering of the triphosphate moiety of ATP and tend to prevent it from reacting with biotin in the mutant crystals. Most likely, this reduced reactivity of bound ATP allowed the successful cocrystallization of single and double mutants with BCCP. As described in the later section of BPL-BCCP complex, PhBPL** can produce holo-BCCP, indicating the retained functionality of the double mutant. However, the lowest activity in the reaction intermediate production may hinder the formation of biotinyl-5′-AMP-ligated crystals in the double mutant. The binding mode of the reaction intermediate in the double mutant could be estimated from the PhBPL**-biotinol-5′-AMP structure, which was determined from a cocystal with the reaction intermediate analog biotinol-5′-AMP (supplemental Table S3 and Fig. S1). The conformation of bound biotinol-5′-AMP in PhBPL** is quite similar to that of bound biotinyl-5′-AMP in PhBPL* as well as in wild-type PhBPL, suggesting that the mutations used are not essentially defective in the formation of biotinyl-5′-AMP.

Structures of PhBCCPΔN76—We designed and expressed a truncated version of PhBCCP lacking N-terminal 76 residues and containing 73 C-terminal amino acids from Val77 (PhBCCPΔN76, residues 76–149), to improve its handling and crystallizability. It is well known that the C-terminal half of biotinyl domain is expressed as a stable protein, which can be biotinylated normally both in vivo and in vitro (28–33).
Located upstream of the biotinyl domain sequence are proline/alanine-rich sequences of varying lengths, which have been proposed to act as flexible linkers (34). The crystal structure of PhBCCPΔN76 (form I; modeled residues 76–149) has been solved by the molecular replacement using BCCP from EcBCCP (PDB ID 1BDO) as a search model and

**FIGURE 2.** Amino acid sequence alignment with secondary structural assignments. The secondary structural elements are shown above the sequence: coils for helices, arrows for strands, and T for turns. The conserved residues are highlighted. A, comparison between BCCP homologs. The amino acid sequence of PhBCCP is aligned with those of BCCP from E. coli (1BDO 20) and P. shermanii (1DD2 37), and lipoyl domains from Bacillus stearothermophilus (1LAB 39), E. coli (1QJO 49), Neisseria meningitidis (1GJX 50), Azotobacter vinelandii (1GHJ, 2-oxoglutarate dehydrogenase complex 51), and A. vinelandii (1IYU, pyruvate dehydrogenase complex 52). Amino acid residues interacting with PhBPL are indicated as orange (C subunit) and green (D subunit) triangles; direct or water-mediated hydrogen bonds (2.2–3.5 Å), electrostatic interactions (3.5–5.0 Å), and van der Waals interactions (2.65–4.0 Å) are included.

B, comparison between BPL homologs. The amino acid sequence of PhBPL is aligned with those of BPL from E. coli (1HXD 53) and M. tuberculosis (2CGH, Q. Ma and M. Wilmanns, unpublished result), and LplA from Thermoplasma acidophilum (2ART 54) and LipB from M. tuberculosis (1W66 25). Amino acid residues responsible for the PhBCCP recognition are indicated by orange (A subunit) and green (B subunit) triangles; direct or water-mediated hydrogen bonds (2.2–3.5 Å), electrostatic interactions (3.5–5.0 Å), and van der Waals interactions (2.65–4.0 Å) are included.
refined at a resolution of 1.55 Å (supplemental Table S2). The second form of crystal structure was determined by molecular replacement using the form I structure and refined at a resolution of 1.55 Å (form II; modeled residues 80–112 and 117–149) in the form I crystal, there is one protomer in the asymmetric unit. The r.m.s.d. values from a Cα superposition between the observed and calculated structure factors, respectively. The second form of crystal structure was determined by molecular replacement using the coordinates of BCCP and many other BCCP structures lack the βββ “thumb” loop, which is observed in EcBCCP (20); the two symmetric halves of all biotinyl domain structures are connected by a type I ’β-turn, whereas the lipoyl domains adopt a type I conformation for this turn. PhBCCP may be an evolutionarily more primitive form as compared with EcBCCP, because PhBCCP is more symmetric in terms of the duplication on hammerhead motifs and lacks the thumb loop providing additional recognition for the biotin moiety.

**Structures of PhBPL-PhBCCP Complex**—The complex structures of PhBPL*-biotin-adenosine-PhBCCPΔN76 and PhBPL**-biotin-adenosine-holo-PhBCCPΔN76 were determined by molecular replacement using the coordinates of PhBPL*, PhBPL**, and PhBCCPΔN76 structures and refined at resolutions of 2.7 and 2.0 Å, respectively (Fig. 1C and Table 1). These two forms of PhBPL-PhBCCP complex crystals are nearly isomorphous and both have a 2:2 heterotetramer in the asymmetric unit. The PhBPL-PhBCCP association involves the formation of a large intermolecular β-sheet, which is solvent-exposed on one side to house the biotinyl-5'-AMP. Notably, the bifurcated main holes of both the complexes are occupied by biotin and adenosine. As mentioned in the first section, we suppose that the modeled adenosine is an ordered part of ATP. The B-factors for the

---

**TABLE 1**

Crystallographic data collection and refinement statistics of PhBPL-PhBCCP complexes

Values in parentheses correspond to the highest resolution shell.

| Data collection | PhBPL*-biotin-adenosine-PhBCCPΔN76 | PhBPL**-biotin-adenosine-holo-PhBCCPΔN76 |
|-----------------|------------------------------------|------------------------------------------|
|                | P2_1                               | P2_1                                      |
| Space group     |                                    |                                          |
| Unit-cell parameters |                                  |                                          |
| a, b, c (Å)     | 69.57, 63.49, 74.72                | 69.85, 63.12, 75.64                      |
| β (°)           | 93.7                               | 95.9                                      |
| Resolution range (Å) | 50.0–2.70 (2.80–2.70) | 50.0–2.00 (2.07–2.00)                     |
| No. of unique reflections | 15,564 (1378)                   | 42,152 (3935)                            |
| Redundancy      | 3.1 (2.8)                          | 3.2 (3.0)                                |
| Completeness (%)| 95.1 (86.6)                        | 95.8 (90.0)                              |
| (I/σ(I))       | 6.9 (2.2)                          | 9.0 (2.5)                                |
| R_merge (%)     | 10.7 (32.0)                        | 8.7 (31.3)                               |

**Refinement**

|                | PhBPL*-biotin-adenosine-PhBCCPΔN76 | PhBPL**-biotin-adenosine-holo-PhBCCPΔN76 |
|----------------|------------------------------------|------------------------------------------|
|                |                                    |                                          |
| Resolution (Å) |                                    |                                          |
| Proteins atoms |                                    |                                          |
| Ligand atoms   |                                    |                                          |
| Water oxygen atoms |                            |                                          |
| R_merge (%)     |                                    |                                          |
| R_free (%)      |                                    |                                          |
| Mean B-value (Å²)|                                    |                                          |
| Estimated coordinate error (Å) |                                    |                                          |
| r.m.s.d. Bond lengths (Å) |                                    |                                          |
| r.m.s.d. Bond angles (°) |                                    |                                          |
| r.m.s.d. Ramachandran plot |                                    |                                          |
| r.m.s.d. Favorable (%) |                                    |                                          |
| r.m.s.d. Additional (%) |                                    |                                          |
| r.m.s.d. Generous (%) |                                    |                                          |
| PDB code | 2EJF                                      | 2EJG                                      |

* R_merge = \( \sum \left| b(h) - \langle b(h) \rangle \right| / \sum b(h) \), where \( b(h) \) is the i-th observation of the intensity of reflection h and \( \langle b(h) \rangle \) is the mean value of all \( b(h) \).

* R_free = \( \sum |F_{\text{calc}}| - |F_{\text{calc}}|\) / \( \sum |F_{\text{calc}}| \), where \( F_{\text{calc}} \) and \( F_{\text{calc}} \) are the observed and calculated structure factors, respectively.

* Rmerge is the Rmerge for a subset of 5% of the reflections that were omitted from refinement.
**TABLE 2**

Intermolecular hydrogen bonds in the PhBPL-PhBCCP complexes

| A          | B          | PhBPL* biotin-adenosine-PhBCCPΔN76 | PhBPL** biotin-adenosine-holo-PhBCCPΔN76 |
|------------|------------|-----------------------------------|------------------------------------------|
| Met1 N     | Asp40 O   | 3.15                              | 2.72                                     |
| Gly3 N     | Asp36 O   | 3.01                              | 2.82                                     |
| Leu5 O     | Tyr15 O   | 3.36                              | 3.03                                     |
| Lys2 N    | Glu27 O   | 3.36                              |                                          |
| Lys4 N    | Glu27 O   |                                    |                                          |
| Thr3 O    | Glu17 O   |                                    |                                          |
| Thr3 Oγ1  | Glu17 N   | 2.79                              | 2.81                                     |
| Gly30 O   | Glu17 N   | 3.06                              | 3.19                                     |
| Arg11 O   | Glu17 N   | 3.29                              | 3.15                                     |
| Arg11 Nγ1 | Glu17 O   |                                    |                                          |
| Val6 N    | Tyr15 O   | 3.16                              | 2.98                                     |
| Val6 O    | Tyr15 N   | 3.06                              | 2.81                                     |
| Tyr13 N   | Val13 O   | 2.88                              | 2.82                                     |
| Tyr13 O   | Val13 N   | 3.15                              | 2.93                                     |
| Tyr15 Oγ1 | Leu1 N    | 2.99                              | 3.11                                     |
| Gln17 N   | Arg11 O   |                                    | 3.16                                     |
| Gln17 Nγ2 | Thr1 O    | 2.93                              | 2.97                                     |
| Gln37 Nγ2 | Thr1 Oγ2  | 2.93                              | 2.97                                     |
| Gly20 O   | Met1 N    | 2.90                              | 2.82                                     |
| Asp40 O   | Glyγ N    | 2.74                              | 2.84                                     |
| Lys17 N   | Leu16 N   | 2.64                              | 2.76                                     |
| Leu16 N   | Asp36 Nγ1 | 2.39                              | 2.77                                     |
| Asn30 O   | Asp36 Nγ2 | 2.77                              | 3.20                                     |
| Val17 O   | Gly119 N  | 2.89                              | 3.02                                     |
| Gly119 N  | Gly119 O  | 2.82                              | 2.97                                     |
| Gly119 O  | Gly119 N  | 3.24                              | 2.97                                     |
| Lys120 N  | Gly119 O  | 3.49                              | 3.49                                     |
| Lys120 Nγ2| Gly119 Oγ2|                                    | 3.22                                     |
| Phe210 O  | Biocytin115 Nγ | 3.17                      | 2.94                                     |
| Gly211 O  | Lys115 Nγ |                                    | 3.27                                     |
| Tyr227 N  | Gly218 O  |                                    | 3.21                                     |
| Asp229 O  | Gly218 O  |                                    | 3.21                                     |
| Gly228 O  | Gly119 N  |                                    | 3.21                                     |
| Gly228 O  | Gly119 N  |                                    | 3.21                                     |
| Gly228 O  | Gly119 N  |                                    | 3.21                                     |
| Met1 N    | Glyγ N    |                                    | 3.21                                     |
| Glu117 Oγ2| Asn118 Nγ2| 3.20                              | 2.76                                     |
| Lys115 Nγ | Glu117 N  | 2.93                              | 2.97                                     |
| Gly117 Oγ2| Glu117 Oγ2| 2.97                              | 2.97                                     |
| Gly117 O  | Glu117 O  | 2.87                              | 2.80                                     |
| Gly117 O  | Glu117 N  | 2.85                              | 2.73                                     |
| Gly146 O  | Met1 N    | 3.20                              | 2.98                                     |
| Gly146 O  | Lys115 Nγ | 3.20                              | 2.98                                     |
| Gly146 Oγ2| Lys115 Nδ |                                    | 2.82                                     |

Ligand atoms are comparable to those of other atoms in the crystal structure; averaged B-factors for the ligand atoms and all atoms are 43.7 Å² and 42.1 Å² in the single mutant complex and 24.8 Å² and 28.4 Å² in the double mutant complex, respectively (Table 1 and supplemental Table S3). Judging from this fact, it is likely that the bound ligands have some biological role in the reaction mechanism rather than a mutation/crystalization artifact from the binding of low affinity ligands. In both the complexes, although the binding mode of the biotin and the adenine ring is similar to that of the corresponding part of biotinyl-5’-AMP seen in the other liganded forms, the ribose ring shows distinct conformation when compared with that of biotinyl-5’-AMP (supplemental Fig. S3). This suggests that the mutations tend to prevent ATP from reacting with biotin in the complex crystals. However, because the single mutant can provide the biotinyl-5’-AMP liganded form and the biotinylated holo-PhBCCP is found in the double mutant complex, the functionality of the mutants would be retained. Therefore, the snapshots fortunately captured at points before and after the biotinylation of PhBCCP may shed light on the interaction of proteins and the target protein/lysine residue specificity of BPL.

The association of two 1:1 complexes of PhBPL-PhBCCP to make the 2:2 complex is mediated by the PhBPL dimer interface analogous to one observed in the wild-type PhBPL (9). At the PhBPL-PhBCCP interface, buried solvent-accessible surface area is ~900 Å² per protomer, which is comparable to the PhBPL dimer interface area of 1030 Å² in the complex. The intermolecular interface is mainly hydrophobic (over 60% of interface atoms are non-polar) and defined by a number of hydrogen bonds (Table 2). The electrostatic potential on the surfaces of proteins shows a heterogeneous charge distribution that reflects a good charge complementarity between the interacting surfaces (Fig. 3). This indicates that electrostatic interactions are important for the complex formation in addition to a pronounced molecular surface complementarity. The structures of the apo and holo forms of PhBCCPΔN76 in free and complex states are generally similar, suggesting that binding to PhBPL and biotinylation causes few significant changes in the overall fold of the biotinyl domain, except for the β4-β5 hairpin turn. In contrast, structures of PhBPL* and PhBPL** show more extensive local conformational changes. Although the catalytic domain is relatively similar to the one in free state, the C-ter-
shift toward the active site of PhBPL due to both the local deformation of the turn and the rigid-body intermolecular relocation. In the formation stage 1, the strictly conserved Glu112 of PhBCCP makes a direct or water-mediated hydrogen bond with the main-chain nitrogen of the target lysine Lys115, which may be a key interaction to promote the local conformational change of the hairpin turn. The interatomic distance between the carboxyl carbon of biotin and the Cα of Lys115 is ~11.9 Å in the formation stage 2 and the product stage, whereas it is ~8.4 Å in the formation stage 1. This indicates that the formation stage 1 is more favorable for the biotinylation reaction to take place (supplemental Table S3).

Recognition of Biotinyl Domain by PhBPL—In the PhBPL:PhBCCP complex, the biotinyl domain interacts with PhBPL mainly in two regions (Table 2 and Figs. 2A and 5). The first cluster comprises its strand β5 and the loop residues that precede it. The second cluster contains the strands β4 and β6, including the biotinylation target Lys115. On the other hand, PhBPL residues interacts with the biotinyl domain mainly in six regions: the β3, β4 active site loop, the N-terminal part of the α2 helix, the β4, β5 turn involving the invariant residue Asn103, the C-terminal half of the strand β6, the β5–β5 loop, and the β11–β12 loop (Table 2 and Figs. 2B and 5). Although these residue clusters are remote in the primary sequence, the tertiary fold brings them into close proximity. Interestingly, the β3 strand in PhBCCP and the β5–β6 turn in PhBPL show distinct pattern of interactions depending upon the stage of reaction (Table 2 and Fig. 2). Particularly, in the β5–β6 turn of PhBPL, only Trp101 participates in the PhBCCP recognition at the formation stage 2 or the product stage. However, additional β5–β6 turn residues Pro102–Asp104 of PhBPL are also buried by the biotinyl domain at the formation stage 1, indicating a significant structural difference.

Because BPL can biotinylate the biotinyl domains from diverse sources (28), key molecular recognition in the BPL-BCCP complex formation should be common. Thus, conserved residues among BCCP orthologs (Fig. 2A), Met87, Gly109, Gly112, Glu114, Lys115, and Met116 may have unique roles in the interaction with BPL. The role of the target lysine Lys115 is obviously crucial. Several mutagenesis studies provided an indication of the importance of the other conserved residues. For example, in EcBCCP, the Glu to Lys substitution at Glu119, which corresponds to Glu112 in PhBCCP, makes it virtually inactive as a substrate for BPL (40), and the mutations at either of the methionine residues flanking the specific lysine severely reduce the ability of biotinyl domains to accept biotin moieties (41). In addition, a mutation at the site equivalent to Met87 in the biotinyl domain of
Crystal Structure of BPL-BCCP Complex

FIGURE 4. Superimposed stereo representations showing multiple conformations of BPL-BCCP complexes. A, structural superposition of four crystallographically independent BPL-BCCP complexes: the A and C subunits (formation stage 2, orange) and the B and D subunits (formation stage 1, blue) of PhBPL*-biotin-adenosine*-PhBCCP*-ΔN76, the A and C subunits (product stage, magenta) and the B and D subunits (formation stage 1, green) of PhBPL**-biotin-adenosine*-holo-PhBCCP*-ΔN76. The C-terminal domain (residues 188–235) and the biotinyl domain are not included in the calculation of superposition. In the double mutant complex, the biocytin residue in C subunit, the target lysine Lys115 in D subunit, the bound biotin and adenosine in B subunit are shown as stick models. B, conformational changes associated with the product formation in the BPL-BCCP complex. Structures before (magenta) and after (blue) the biotinylation reaction in the double mutant complex are compared. The residues biocytin, Lys115, and Tyr227 are depicted as stick models.

human propionyl-CoA carboxylase (Met641 → Lys) dramatically reduces the rate of biotinylation by EcBirA (32). Similarly, important roles should be assigned to the conserved residues of PhBPL (Fig. 2B): Trp101–Asp104 for directing of target Lys115 in the biotinyl domain. When the substrate Lys115 enters into the active site of PhBPL, it forms a hydrogen bond with Asn103 in the formation stage 1 (Fig. 5A). Notably, these two residues are invariant even in the level of protein family, among biotinyl and lipoyl domains as well as among BPL and LplA and LipB enzymes (Fig. 2). The electrostatic interaction between the substrate Lys115 and the conserved Asp104 of PhBPL would also be important. Other conserved PhBPL residues, Leu116 and Glu118, on the β6 strand associate with the β5 strand of PhBCCP, including a conserved residue Met116. Correlated interaction of these two strands may determine the settlement of PhBPL and PhBCCP during the biotinylation process. Only a moderate conservation of interaction residues in the level of protein family, except for the invariant Asn103 in PhBPL and Lys115 in PhBCCP, may indicate that main-chain interactions dominate the PhBPL-PhBCCP association.

Implications for the Reaction Mechanism—From the present results, it appears that the observed biotin and partly disordered ATP in the mutant structures act as a non-reactive analog of substrates or biotinyl-5′-AMP, thereby allowing us to visualize the snapshots of protein-protein complexes. Unfortunately, the non-hydrolyzable intermediate analog biotinol-5′-AMP did not provide analyzable complex crystals. To further characterize the observed three stages of the PhBPL-PhBCCP complex, structures of bound ligands are compared among various crystal forms (supplemental Table S3). Although the wild type and two mutants of PhBPL bind biotinyl-5′-AMP or biotinol-5′-AMP in essentially the same conformation, the ligand in the wild-type seems to be more flexible at the ribose-5′-phosphate part, which would reflect the higher reactivity of biotinyl-5′-AMP in the wild type. The averaged B-factor of ligand atoms is 33.8 Å² at the ribose-5′-phosphate part and 16.3 Å² at the remaining part of the PhBPL-biotinyl-5′-AMP structure (PDB ID 1WQW), whereas they are 16.2 Å² and 11.2 Å² at the respective parts in PhBPL*-biotinyl-5′-AMP, and 21.6 Å² and 15.6 Å² at the respective parts in PhBPL**-biotinol-5′-AMP. From a structural comparison, the interatomic distance between the N9 atom of adenosine moiety and the carboxyl carbon of biotin moiety (biotin-adenine ring distance) is longer in the wild-type structure: 6.9 Å in the wild-type, 6.6 Å in the single mutant, and 6.5 Å in the double mutant. This fact may indicate that the biotin-adenine ring distance is a good indicator of the reactivity of the intermediate biotinyl-5′-AMP. We also calculated the biotin-adenine ring distance as well as the interatomic distance...
between the carboxyl carbon of biotin and the C\(^{9280}\) of Lys115 in BCCP (biotin-BCCP distance), in the BPL-BCCP complex structures. Interestingly, the biotin-adenine ring distance and the biotin-BCCP distance show a complementary relationship: 7.2 and 8.4 Å in the formation stage 1, 6.7 and 12.0 Å in the formation stage 2, and 6.9 and 11.7 Å in the product stage, respectively. Assuming that the observed biotin and partly disordered ATP in the mutant complex structures mimic the biotinyl-5\(^{11032}\)-AMP in the formation stages, this analysis of interatomic distance may indicate the higher reactivity of biotinyl-5\(^{11032}\)-AMP in the formation stage 1. Due to the limited resolution of 2.7 Å in the single mutant complex, the structural difference between the formation stage 2 and the product stage is unclear. Because the product stage shows a medium value in the biotin-adenine ring distance, it is not conclusive whether the putatively recruited ATP and biotin probably for the next reaction cycle in the wild-type PhBPL-holo-PhBCCP complex can react to produce biotinyl-5\(^{11032}\)-AMP or not. However, the fact that the co-crystallization of wild-type PhBPL with ATP, biotin, and PhBCCP\(\Delta N76\) did not provide the expected PhBPL-biotinyl-5\(^{11032}\)-AMP-holo-PhBCCP\(\Delta N76\) complex crystals may indicate that the recruited ATP and biotin to the product stage complex is not reactive until the dissociation of holo-BCCP from BPL or the product releasing and the biotinyl-5\(^{11032}\)-AMP formation is coupled. This hypothesis is reasonable in terms of the efficiency of the biotinylation reaction; the assumption of biotinyl-5\(^{11032}\)-AMP formation in the product stage means the occurrence of competition with the formation stage 2, which would be unfavorable for the complete biotinylation of BCCP molecules in solution due to the product inhibition.

The structures obtained provide a putative scenario for the biotin transfer reaction. The apo BPL exists with the disordered active site loop, which enlarges the entrance to the active site. In the presence of ATP and biotin, BPL forms the reaction intermediate biotinyl-5\(^{11032}\)-AMP and the side hole to accommodate the biotinyl domain of BCCP. In the formation stages 1 and 2, the negative charges of the BPL molecule from the conserved acidic residues (Asp\(^{104}\) and Asp\(^{229}\) in PhBPL) and phosphate group of biotinyl-5\(^{11032}\)-AMP may attract the target lysine of apo BCCP (Lys115 in PhBCCP) to make the BPL-BCCP complex, which is observed in the present crystal structures of PhBPL\(^{**}\)-biotin-adenosine-holo-PhBCCP\(\Delta N76\) (chains B and D in the formation stage 1) and PhBPL\(^{*-}\)-biotin-adenosine-PhBCCP\(\Delta N76\) (chains B and D in the formation stage 1; chains A and C in the formation stage 2). Once PhBPL-biotinyl-5\(^{11032}\)-AMP and the apo biotinyl domain are associated, the complex should proceed to the reaction stage producing the holo-biotinyl domain. The nucleophilic attack on the carbonyl carbon of activated biotin by the ε-amino group of the substrate lysine results in the transfer of biotin onto the apo-biotinyl domain and the concomitant formation of AMP. The side chain of the substrate lysine must be precisely positioned into the active site of BPL to bring the electron donating...
nitrogen close enough to the reactive carbonyl carbon of biotinyl-5′-AMP to permit the chemical reaction, which is in agreement with a mutational study reporting that a translocation of the lysine by one place to either side in the exposed β4-β5 turn abolishes biotinylation (41, 42). Thus, the biotinyl domain functions as a protein scaffold that displays the biotin-accepting lysine to BPL. On the other hand, it is likely that the conserved residues Asn103 and Asp104 of PhBPL form a catalytic dyad, in which Asn103 is important for the precise positioning of the target lysine by forming a hydrogen bond, and Asp104 acts as a general base to deprotonate the ε-amino group of the lysine to render it sufficiently nucleophilic to attack the carbonyl carbon of the valeric acid side chain of biotin. After the reaction stage, the biotinylated BCCP may not be released immediately but maintained in the product stage, which is observed in the chains A and C of the PhBPL*+biotin-adenosine-holo-PhBCCPΔN76 structure. At the product stage, the peptide group of the biocytin residue is recognized by BPL with a hydrogen-bonding network involving the main-chain atoms of Phe210 and Tyr227 in PhBPL, and the biocytin side chain is packed by van der Waals interactions with Met114 in PhBCCP, and Pro76, Trp101, Phe210, Gly211, Arg212, Ile226, and Tyr227 in PhBPL, which allowed us to observe the biocytin residue clearly in the electron density map.

A superposition of the before and after biotinylated subunits (Fig. 4A) provides insights into the unobserved reaction stage. Unfortunately, it is clear that both the formation stages 1 and 2 do not provide suitable geometry for the nucleophilic attack by the special lysine of BCCP, although a slight remodeling of the formation stage 1 seems to allow the reaction. In the chains B and D of the PhBPL*+biotin-adenosine-holo-PhBCCPΔN76 structure in the formation stage 1, we observed the closest distance of 6.1 Å between the Nε of Lys115 and the
bion carboxyl carbon, which is the putative site for the target carbonyl carbon of biotinyl-5’-AMP. In the putative reaction stage, the carbonyl carbon of biotinyl-5’-AMP should be located just adjacent to the ε-amino group of target lysine, which may be provided by an additional structural change of either the active site loop of BPL or the target β-turn of BCCP toward the active site. It is possible that the restoring mutations at Arg48 and Lys111 or the binding of physiological intermediate biotinyl-5’-AMP instead of the observed biotin and partly disordered ATP provides this putative conformational change. In this case, the current formation stage 1 in the mutants would correspond to the reaction stage in the wild type. Further investigation is required to clarify this point. In addition, it should be noted that multiple conformations are observed in the present crystal structures (Fig. 4A), suggesting that the free energy level of these conformations are comparable. The existence of formation and product stages before and after the reaction stage would be favorable to ensure both the reaction efficiency and the extreme substrate specificity in the biotinylation reaction; the high entropic energy barrier due to the strict substrate specificity is resolved by adding the intermediate steps, and the multistep reaction sequence structurally avoid the error of reaction.

**Molecular Discrimination between Biotinyl and Lipoyl Domains**—The biotinyl domains of biotin-dependent enzymes and the lipoyl domains of 2-oxo acid dehydrogenase multienzyme complexes have homologous structures, but the post-translational modification of the target lysine residue in each domain is correctly selected by BPL and lipoyl protein ligases LplA/LipB, respectively. Although the reactions catalyzed by these ligases are analogous, the BPL and LplA/LipB bear poor sequence conservation (Fig. 2B), suggesting an evolutionarily distant relationship among the family members. BPLs from various sources have been found to recognize/biotinylate various BCCPs from different organisms (2, 12, 14, 42–45). Of all the PhBCCPΔN76 residues that make interactions with PhBPL, only the MKM motif at the tip of the hairpin loop is well exposed to the solvent. Therefore, the discrimination of this Lys115 from the other four lysine residues in PhBCCPΔN76 may be understood from the PhBPL-PhBCCPΔN76 structures showing that only the exposed β-turn can fit into the side hole connecting to the biotinyl-5’-AMP binding main hole of PhBPL (Fig. 1C). The local sequence surrounding the receptive lysine within the context of a folded BCCP should be important for the BPL recognition, given that the lipoyl domains containing DKA, DKV, and AKA motifs at the tip of the corresponding hairpin loop cannot accept biotinylation. Also it is reported that BPL does not biotinylate a biotinyl domain with DKA replacing the MKM motif (41) and S. tokodaii BCCP, which has the MKS motif does not serve as a substrate for the heterologous BPL or EcBirA (46). However, the MKM motif of the biotinyl domain by itself does not necessarily specify biotinylation, because it has been shown that the two methionines flanking the specific lysine can be replaced by leucines (32, 47) and that replacing the DKA sequence at the hairpin loop of lipoyl domain with MKM does not make it a substrate for BPL (48). Therefore, there must be key differences between the biotinyl and lipoyl domains that allow the relevant protein ligases to distinguish them for the different purposes of post-translational modification. To understand how discrimination between lipoylation and biotinylation is achieved, we superposed the selected biotinyl and lipoyl domains to the biotinyl domain in the PhBPLΦ-biotin-adenosine-holo-PhBCCPΔN76 complex. The biotinyl domains are well overlaid and show good shape/charge complementarities at the BPL-BCCP interfaces. The β2-β3 thumb loop of EcBCCP does not collide with the BPL structural elements (Fig. 6A). Although the PhBPL-lipoyl domains show overall shape complementarities at their interfaces, the loop β1-β2 of lipoyl domains tend to interfere with the N terminus of α2 of PhBPL (Fig. 6B). Additionally, the charge complementarities at the interfaces are destroyed in the putative complexes with lipoyl domains. The position of conserved positively charged Arg93 of PhBCCP is occupied by negatively charged Glu in the lipoyl domain, and the position of conserved nonpolar Gly is occupied by negatively charged Asp or Glu in the lipoyl domain. The negatively charged aspartic side chain in the motif DVK or DKA of the lipoyl domain instead of nonionic methionines in BCCP also prevents its recognition by the side hole of BPL. Because the attraction of BCCP would be affected by long range electrostatic interactions, we suggest that the difference in surface charge complementarities is one of the factors contributing to the molecular discrimination between the biotinyl and lipoyl domains. The other difference comes from diversity in the type of the target β-turn to be modified. In BCCP, the turn loop contains two residues and has the type I’, whereas in the lipoyl domain, it consists of four residues and has the type I. Therefore, the conformation of the β-turn in the lipoyl domain is different from that in the biotinyl domain, which would affect the positioning of the specific lysine to the biotin moiety (Fig. 6C).

In conclusion, it is apparent that the recognition of the biotinyl domain is a rather complex process. This is in contrast with most other post-translational modifications in which the primary structure surrounding the target residue can be of crucial importance. BPL and BCCP have several mobile parts that are essential for the biotinylation reaction in multiple stages: BCCP binding (formation stages 1 and 2), catalysis (reaction stage), and product release (product stage). Learning how these factors are choreographed to accomplish the protein biotinylation with

---

**FIGURE 6. Molecular discrimination between biotinyl and lipoyl domains.** A, overlay of the B and D subunits (formation stage 1, marine) of double mutant complex with the other biotinyl domains: EcBCCP in magenta and the 1.3 S subunit of P. shermani in green. The target lysine residue and the bound biotin/adenosine are depicted as stick models. B, overlay of the B and D subunits (formation stage 1, marine) of double mutant complex with the lipoyl domains from A. vinelandii (green) and E. coli (magenta). The target lysine residue and the bound biotin/adenosine are depicted as stick models. C, comparison of the hairpin turn that displays the target lysine residue. The B and D subunits (formation stage 1; BPL in cyan and BCCP in magenta) of double mutant complex are overlaid with the lipoyl domain from A. vinelandii (yellow). The bound biotin/adenosine and the target lysine residues (Lys115 of PhBCCP and Lys19 of lipoyl domain) are depicted as stick models.
Crystal Structure of BPL-BCCP Complex

extreme substrate specificity remains an unmet challenge for investigators.

Acknowledgments—We thank the staff of RIKEN Genomic Science Center for providing plasmids, Dr. M. Miyano for providing PH1284, the staff of beamline BL26B1 of SPring-8 for assistance during x-ray experiments, and Dr. D. Beckett of the University of Maryland for providing biotinol-S'-AMP.

REFERENCES