nature structural & molecular biology

Crystal structures of human and *Staphylococcus aureus* pyruvate carboxylase and molecular insights into the carboxyltransfer reaction

Song Xiang & Liang Tong

Pyruvate carboxylase (PC) catalyzes the biotin-dependent production of oxaloacetate and has important roles in gluconeogenesis, lipogenesis, insulin secretion and other cellular processes. PC contains the biotin carboxylase (BC), carboxyltransferase (CT) and biotin-carboxyl carrier protein (BCCP) domains. We report here the crystal structures at 2.8-Å resolution of full-length PC from *Staphylococcus aureus* and the C-terminal region (missing only the BC domain) of human PC. A conserved tetrameric association is observed for both enzymes, and our structural and mutagenesis studies reveal a previously uncharacterized domain, the PC tetramerization (PT) domain, which is important for oligomerization. A BCCP domain is located in the active site of the CT domain, providing the first molecular insights into how biotin participates in the carboxyltransfer reaction. There are dramatic differences in domain positions in the monomer and the organization of the tetramer between these enzymes and the PC from *Rhizobium etli*.

Pyruvate carboxylase (PC, EC 6.4.1.1) catalyzes the biotin-dependent carboxylation of pyruvate to produce oxaloacetate[1–4]. This enzyme was first identified more than 45 years ago in chicken and beef liver[5], and has since been found in most living organisms. In mammals, PC has crucial roles in gluconeogenesis in the liver and kidney, in lipogenesis and glyceroneogenesis in adipocytes, in glucose-induced insulin secretion by pancreatic beta cells, and in the biosynthesis of the excitatory neurotransmitter glutamate in astrocytes[3,4]. The functional importance of this enzyme is underscored by the fact that inherited PC deficiencies are linked to lactic acidemia, hypoglycemia, psychomotor retardation and even death[3,6]. Four single-site mutations, V145A, R451C, A610T and M743I, are currently known to be associated with these diseases in humans[6–9].

PC is a single-chain, multidomain enzyme of about 130 kD in eukaryotes and most bacteria (**Fig. 1a**), with highly conserved sequences (**Supplementary Fig. 1** online). The BC domain also shares sequence homology with BC domain or subunit in other biotin-dependent carboxylases such as acetyl-CoA carboxylase[10–12], whereas the CT domain is distinct[13,14]. PC enzymes are active only in the tetrameric form (α_4)[1,3,15]. In archaea and some bacteria, PC is made up of two subunits, one equivalent to the BC domain and the other to the CT+BCCP domain of single-chain PCs. These enzymes are also tetrameric, $(\alpha\beta)_4$.

The overall catalysis by PC proceeds in two steps. First, the BC domain catalyzes the carboxylation of biotin, which is covalently linked to the BCCP. Bicarbonate donates the carboxyl group, and ATP is hydrolyzed to ADP in this reaction. The CT domain then catalyzes the transfer of the activated carboxyl group to pyruvate to produce the oxaloacetate product. The BC domain requires divalent cations (Mg^{2+}) for binding of the ATP substrate and for catalysis. The CT domain contains a tightly bound cation—Mn^{2+} in mammalian PC and Zn^{2+} in yeast PC. Most PC enzymes are activated by acetyl-CoA and inhibited by aspartate, glutamate and α-ketoglutarate[3,16,17].

Structural information on the BC subunit of the *Aquifex aeolicus* PC[12], the CT (5S) subunit of the *Propionibacterium shermanii* transcarboxylase[13,18] and the CT domain of the CT+BCCP subunit of the *Vibrio cholerae* oxaloacetate decarboxylase Na^+ pump[14] are available. The structure of the BC subunit shares homology with that of the BC subunit of acetyl-CoA carboxylase[10,11]. The CT subunit contains a triose phosphate isomerase (TIM) barrel core and a long C-terminal extension that consists of several helices. The active site of the CT subunit is located in the middle of the TIM barrel core, where the divalent cation also binds[13].

We report here the crystal structures at 2.8-Å resolution of the full-length PC from *Staphylococcus aureus* and the C-terminal region (residues 494–1178, missing only the BC domain) of human PC. *Staphylococcus aureus* PC shares 49% amino acid sequence identity with human PC (**Supplementary Fig. 1**). The crystal structure of full-length PC from *Rhizobium etli* was recently reported[19]. There are notable differences in the organization of the tetramer and the relative positioning of the domains in the monomer between the *S. aureus* and *R. etli* enzymes, suggesting that the structures may represent distinct

Department of Biological Sciences, 1212 Amsterdam Avenue, Columbia University, New York, New York 10027, USA. Correspondence should be addressed to L.T. (ltong@columbia.edu).

Received 11 October 2007; accepted 23 January 2008; published online 24 February 2008; doi:10.1038/nsmb.1393

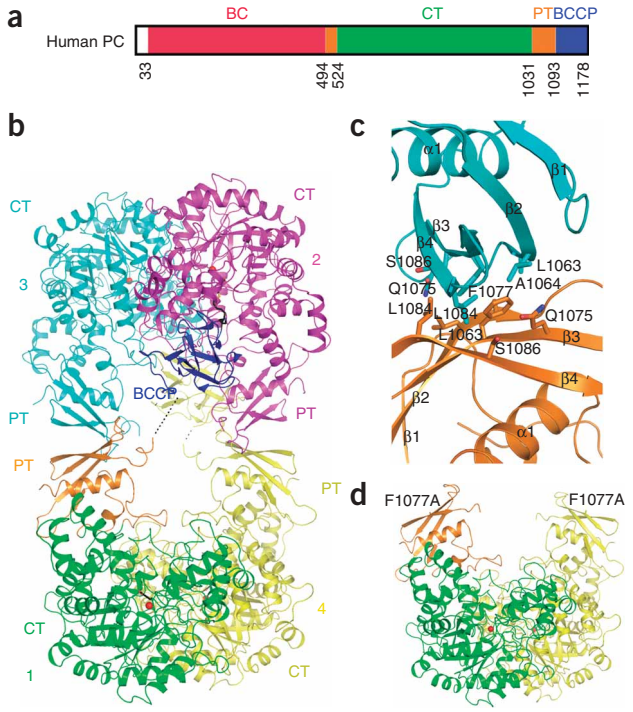

Figure 1 Structure of CT+PT+BCCP domain of human pyruvate carboxylase (PC). (**a**) Domain organization of human PC. BC, biotin carboxylase (red); CT, carboxyltransferase (green); PT, PC tetramerization (gold), which was identified from our studies; BCCP, biotin-carboxyl carrier protein (blue). (**b**) Schematic drawing of the structure of the CT+PT+BCCP domain tetramer of human PC. The domains in monomer 1 are given separate colors as in panel **a**, whereas the other three monomers are colored magenta, cyan and yellow. The biotin moiety on the BCCP molecule in blue is shown in black, pointed into the active site of CT. (**c**) Detailed interactions between the PT domains of monomers 1 and 3 in the tetramer interface of human PC. (**d**) Crystal structure of the dimer of the F1077A mutant of the CT+PT+BCCP domain of human PC. The mutation has disrupted the tetramer. All the structure figures were produced with Pymol[30].

conformational states of PC. Most importantly, whereas the biotin moiety is disordered in the *R. etli* PC structure, we observed two distinct conformations for BCCP-biotin in our structures of the human and *S. aureus* enzymes. In one conformation, BCCP-biotin is positioned in the active site of the CT domain, providing the first molecular insight into how this domain may participate in the carboxyltransfer reaction. In the other conformation, the biotin is located in a previously uncharacterized exo binding site, which may offer another mechanism for regulating the activity of some of the PC enzymes.

RESULTS

Structure determination

We initially attempted to study the full-length mammalian PC enzyme, but after bacterial expression the protein samples were aggregated in solution. We were able to crystallize the C-terminal region (residues 482–1178, missing only the BC domain) of human PC. The crystal structure was determined at 2.8-Å resolution by the molecular replacement method, using the structure of the 5S subunit of *P. shermanii* transcarboxylase[13] as the model. After refinement, electron density for BCCP was recognized and was modeled based on the structure of the BCCP subunit of *Escherichia coli* acetyl-CoA carboxylase[20].

To study the full-length PC enzyme, we screened through several different organisms and found that the *S. aureus* PC can be expressed in large quantities in *E. coli* and crystallized. The crystal structure was determined at 2.8-Å resolution by the molecular replacement method, using the structures of the C-terminal region of human PC and the BC subunit of *A. aeolicus* PC[12] as the models. BCCP and its biotin were placed in the electron density after refinement with the BC and CT domain models.

All the refined structures have excellent agreement with the crystallographic data and the expected bond lengths, bond angles and other geometric parameters. The majority of the residues (85%) in these structures lie in the most favored region of the Ramachandran plot, and 0.2–0.4% of the residues lie in the disallowed region (**Supplementary Discussion**).

A novel domain mediates tetramerization of human PC

The crystal structure of the C-terminal region of human PC shows a tetramer (**Fig. 1b**), consistent with our observations in solution by gel filtration (**Supplementary Fig. 2** online) and light scattering (data not shown). Unexpectedly, the structure reveals that the tetramerization is mediated by a previously uncharacterized domain, which lies outside of the CT and BCCP domains and is not recognized by sequence analysis. As our studies with full-length PC from *S. aureus* confirm that this domain is also central for its oligomerization (see below), we have named it the PC tetramerization (PT) domain (**Fig. 1a**).

The PT domain contains 92 residues and is made up of two segments of the PC polypeptide—a short segment of 30 residues just before the CT domain (residues 495–524) and a longer segment of 60 residues just after the CT domain (residues 1032–1093) (**Fig. 1a**). Residues from the C-terminal segment form a highly twisted four-stranded antiparallel β-sheet, and residues from the N-terminal segment form a long α-helix that is positioned in the concave face of this β-sheet (**Fig. 1c**).

The tetramer of the human CT+PT+BCCP domains is a dimer of dimers. The CT domain forms dimers between two monomers, mediated primarily by its TIM barrel fold, but there are differences to the dimers of the transcarboxylase 5S subunit[13] (**Supplementary Discussion**). The exposed, convex faces of the two PT domains in one dimer interact with their equivalents in the other dimer, forming the tetramer (**Fig. 1b**). Phe1077 and Leu1084 make the largest contributions to the buried surface area in this interface (**Fig. 1c**). Phe1077 is conserved as phenylalanine or tyrosine and Leu1084 shows conservative substitutions (methionine, valine, proline or alanine) among the PC enzymes (**Supplementary Fig. 1**), consistent with our observation that full-length *S. aureus* PC adopts a similar mode of tetramerization (see below).

To confirm the relevance of this tetramerization interface, we mutated residue Phe1077 in the PT domain to alanine or glutamic acid. Gel-filtration and light-scattering studies of the mutant proteins showed that they are dimeric in solution (**Supplementary Fig. 2**). To verify that the mutations have truly disrupted the tetramer, we crystallized the F1077A mutant and determined its structure at 3.0-Å resolution. The crystal contains only dimers of the mutant, but the mutation did not affect the overall conformation of the dimer (**Fig. 1d**).

To characterize the functional effects of mutations in this tetramer interface, we then introduced the F1077A and F1077E mutations into full-length human PC, and our kinetic studies showed that these

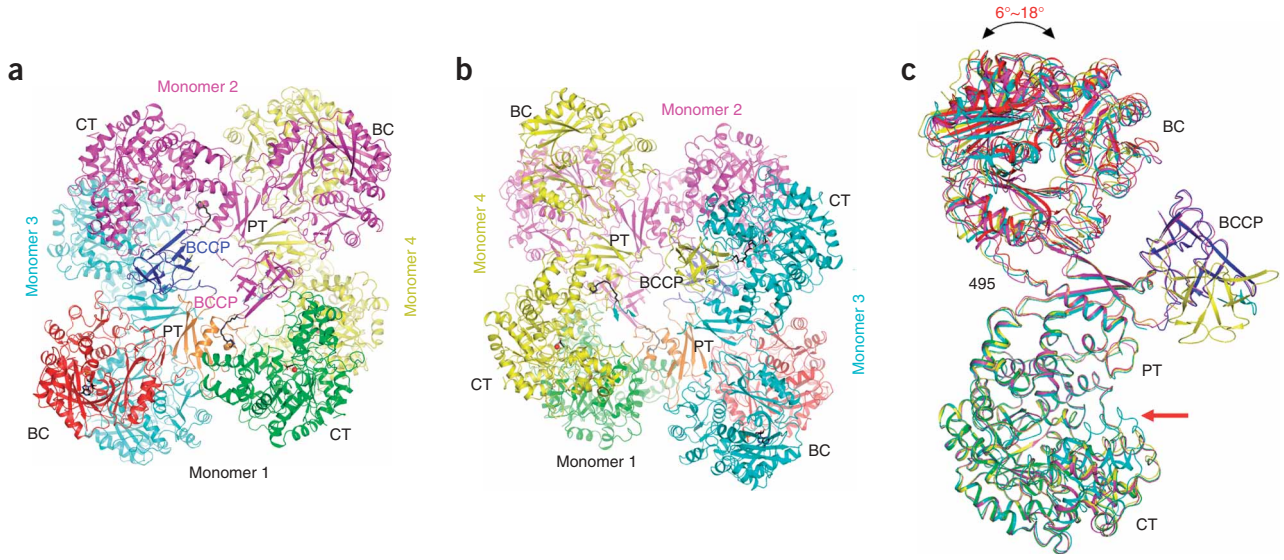

Figure 2 Structure of *S. aureus* pyruvate carboxylase (PC). (**a**) Schematic drawing of the structure of the *S. aureus* PC tetramer. The domains in monomer 1 are given separate colors as in **Figure 1a**, whereas the other three monomers are colored in magenta, cyan and yellow. The biotin moiety is shown as a stick drawing in black. (**b**) Schematic drawing of the structure of the *S. aureus* PC tetramer, after 180° rotation around the vertical axis to show the 'bottom layer'. (**c**) Overlay of the four monomers of the *S. aureus* PC tetramer. The red arrow points to the conformational difference in monomer 3 (cyan) in the C-terminal extension of its carboxyltransferase (CT) domain.

mutants are essentially inactive (**Supplementary Fig. 2**), consistent with biochemical data that PC is active only in its tetrameric form[1,3,15]. Overall, the mutagenesis and kinetic studies confirm the functional importance of the tetramerization interface observed in the structure of the CT+PT+BCCP domain of human PC.

Overall structure of the full-length PC from *S. aureus*

We next determined the crystal structure of full-length *S. aureus* PC at 2.8-Å resolution. The tetramer of *S. aureus* PC has a roughly square shape (**Fig. 2a**) and generally obeys 222 point-group symmetry. Two monomers (1 and 2) lie in the 'top layer' of the tetramer, along two parallel edges of the square, whereas the two monomers (3 and 4) in the 'bottom layer' lie along the other two edges of the square (**Fig. 2b**). The BC, CT and PT domains form dimers between monomers in different layers of the tetramer, with the BC and CT dimers located at alternate corners of the square, creating extensive contacts between the two layers. In contrast, there is little direct contact between the two monomers in the same layer of the tetramer, except for the exchange of their BCCP domains (**Fig. 2a**, see below). The BCCP domains are located near the middle of the square, consistent with predictions based on earlier avidin binding studies[1,3].

The crystallographic asymmetric unit contains a tetramer of the enzyme. The atomic model for each monomer contains essentially the entire molecule, except for residues 165–240 (domain B; residue numbers based on human PC) in the BC domain of monomers 2 and 4 and the BCCP domain of monomer 3 (where only the β-hairpin containing the biotin is ordered). The crystal was grown in the presence of ATP and oxaloacetate, and the binding of ATP and pyruvate (the decarboxylation product of oxaloacetate) is observed in the active sites of some of the monomers. The CT domain also contains a divalent cation in its active site, modeled as Mn^{2+} in the current structure, although its exact identity is not known.

Although the overall structures of the four monomers are similar, there are large differences in the relative positions of the BC and CT domains in each monomer (**Fig. 2c**). Superposition of the CT domains of the four monomers gives an r.m.s. deviation of about 0.4 Å for their equivalent Cα atoms, which also brings the PT domain of the four monomers into close superposition (with an r.m.s. deviation of 0.4 Å). Only residues 863–983 in the CT domain of monomer 3 in the bottom layer have large conformational differences, related to the presence of a BCCP domain in its active site. In contrast, the four BC domains show marked variations in their orientations, ranging from a 6° to 18° rotation hinged around residues 495 in the BC-PT domain boundary (**Fig. 2c**).

Conserved mode of tetramerization for human and *S. aureus* PC

The mode of tetramerization of full-length *S. aureus* PC (**Fig. 2a**) is essentially the same as that of the CT+PT+BCCP domain of human PC (**Fig. 1b**). This similarity can be recognized more easily if one views the CT+PT domains in the *S. aureus* tetramer in the same orientation as the equivalent domains in the human PC tetramer (**Fig. 3a**). In addition, the Y1077A mutant of full-length *S. aureus* PC is dimeric in solution and catalytically inactive (**Supplementary Fig. 2**). These data confirm that the human and *S. aureus* PC share a conserved mode of tetramerization.

Differences with the *R. etli* PC structure

There are notable differences in the organization of the tetramer between the structure of *S. aureus* PC and the recently reported *R. etli* PC structure, obtained in the presence of the activator ethyl-CoA[19], even though the two enzymes share 51% sequence identity (**Supplementary Fig. 1**). This difference can be readily recognized if the CT+PT domains of the *R. etli* PC are viewed in the same orientation as the equivalent domains in the human and *S. aureus* PC (**Fig. 3b**). The PT domains in the *R. etli* structure are not involved in tetramerization; the two PT domains in the bottom layer especially are located far from the tetramer interface (**Fig. 3b**). The exposed surface of the β-sheet in the four PT domains, including the conserved Phe1077 residue, is not in contact with other residues in the *R. etli* tetramer. This tetramer is stabilized solely by the BC and CT domain

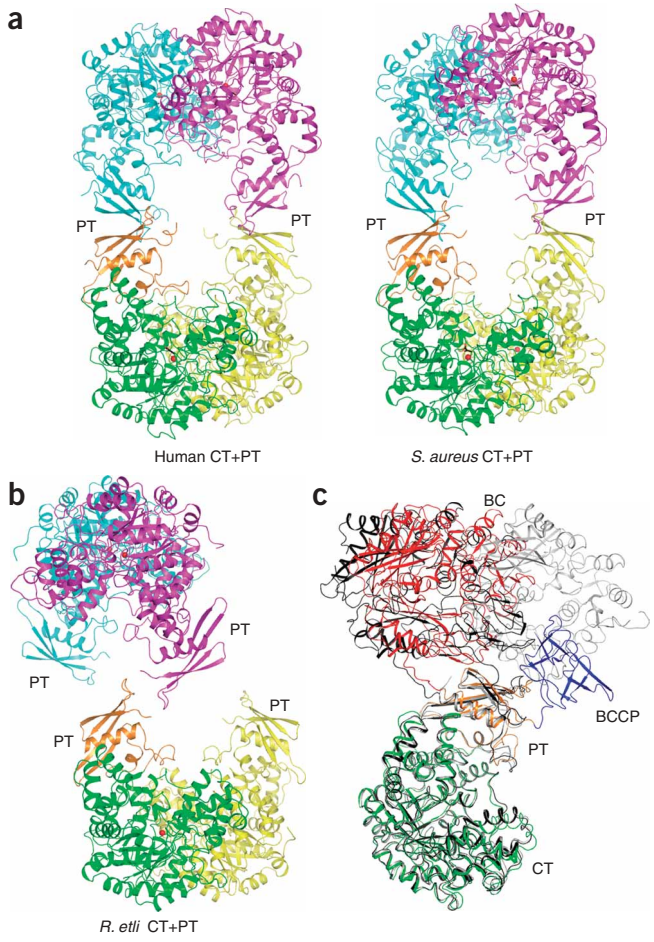

Figure 3 Dramatic differences in the tetramer interface of human and *S. aureus* pyruvate carboxylase (PC) compared to that of *R. etli* PC. (**a**) Schematic drawing of the tetramer of the carboxyltransferase and PC tetramerization (CT+PT) domain of human and *S. aureus* PC, viewed in the same orientation. (**b**) Schematic drawing of the four CT+PT domains in the *R. etli* tetramer[19], viewed in the same orientation as in panel **a** for the two monomers in green and yellow. (**c**) Overlay of the structures of monomer 1 in *S. aureus* PC (colored) and the monomer bound to ethyl-CoA (black) or the free monomer (gray) of *R. etli* PC[19].

dimers, which have roughly the same organization as those in *S. aureus* PC. The dramatic structural differences may be related to the conformational changes in the PC tetramer upon acetyl-CoA activation (see below).

There are also large differences in the domain positions in the monomer between the two structures (**Fig. 3c**). The conformation of the *R. etli* PC monomer bound to ethyl-CoA is actually more similar to that of the *S. aureus* PC monomer, in the absence of any activator. The position of the BC domain for the two monomers without ethyl-CoA in the *R. etli* PC tetramer is different from that in the *S. aureus* structure (**Fig. 3c**).

Molecular insights into the carboxyltransfer reaction

Our structures of human and *S. aureus* PC reveal for the first time how BCCP and its biotin can participate in the carboxyltransfer reaction in the CT domain. In *S. aureus* PC, the BCCP domain of monomer 4 is located in the active site of the CT domain of monomer 3 (**Figs. 2b** and **4a**). Notably, a similar conformation for the BCCP domain and its biotin is observed in the structure of human PC (**Fig. 4b**). In both structures, the biotin moiety is projected away from the BCCP and is located near the pyruvate substrate bound in the CT active site (**Fig. 4c**). Therefore, this conformation may represent that of BCCP during catalysis in the carboxyltransfer reaction, and this is the first time that BCCP has been observed in a catalytically competent conformation. The electron density for biotin is not as strong as that for the protein residues in the active site (**Supplementary Fig. 3** online), suggesting that it may be partly disordered in these structures.

In the active site of the CT domain, the BCCP domain interacts with residues 868–970, at the beginning of the C-terminal extension to the central TIM barrel core of the CT domain (**Figs. 4a**,**b**). Important hydrogen-bonding and hydrophobic interactions occur in this interface (**Supplementary Fig. 4** online). The β-hairpin of BCCP containing the biotin moiety is inserted deepest into the CT domain. The bound positions of biotin in the human and *S. aureus* structures are close to each other (**Fig. 4c**), supporting the biological relevance of the observed binding mode. The carbonyl oxygen in biotin makes a hydrogen bond with the side chain hydroxyl of conserved Ser911

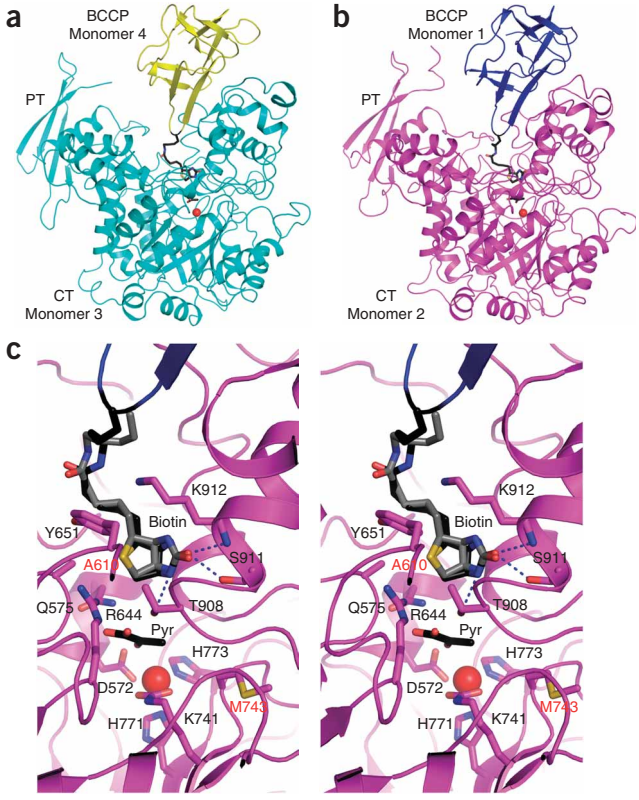

Figure 4 Binding of the biotin-carboxyl carrier protein (BCCP) domain in the active site of the carboxyltransferase (CT) domain. (**a**) Binding of the BCCP domain from monomer 4 (yellow) into the active site of the CT domain of monomer 3 (cyan) in *S. aureus* PC. The metal ion in the CT domain is shown as a red sphere. (**b**) Binding of the BCCP domain from monomer 1 (blue) into the active site of the CT domain of monomer 2 (magenta) in human PC. Note the similarity to the binding mode for the *S. aureus* enzyme in panel **a**. (**c**) Stereo drawing of the active site of the CT domain in human PC. The biotin moieties from the human (black) and *S. aureus* (gray) PC have essentially the same conformation. Hydrogen bonds from the biotin are indicated with dashed lines.

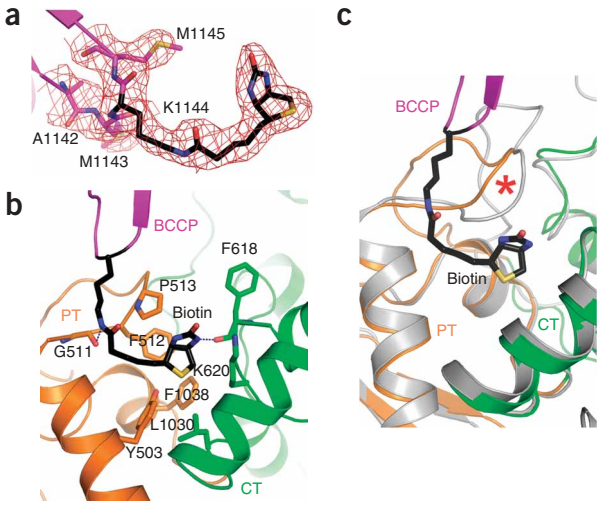

Figure 5 Binding of biotin-carboxyl carrier protein (BCCP)-biotin in a previously uncharacterized exo pocket at the interface between the pyruvate carboxylase (PC) tetramerization (PT) and carboxyltransferase (CT) domains of *S. aureus* PC. (**a**) Final $2F_o - F_c$ electron density at 2.8-Å resolution for the biotin group bound in the exo pocket. The contour level is at 1σ. (**b**) Detailed interactions between biotin and the novel exo pocket at the interface between the PT and CT domains. (**c**) Conformational differences between the human PC (gray) and *S. aureus* PC (colored) in the exo binding site for biotin, for residues just after the long helix in the PT domain (highlighted with a red star).

(residue 879 in *S. aureus* PC; residue numbers in *S. aureus* PC are given in parentheses throughout) and the main chain amide of Lys912 (880), and the sulfur atom is located in a pocket formed by conserved residues Gln575 (545), Ala610 (580), Arg644 (614) and Tyr651 (621) (**Fig. 4c**). The N1 atom of biotin is located within 4.7 Å of the methyl group of the pyruvate substrate, suggesting that it is in the appropriate position for catalysis, possibly a direct transfer of the carboxyl group to pyruvate. In the current structure, in its noncarboxylated form, the N1 atom makes a hydrogen bond with the side chain hydroxyl group of the strictly conserved Thr908 (876).

Our structures suggest that the CT domain may need to undergo a conformational change for its C-terminal extension before BCCP binding and catalysis. Only the CT domain of monomer 3 of the *S. aureus* PC tetramer has a BCCP domain in its active site. In the other three monomers, the C-terminal extension of the CT domain (residues 863–983) assumes a more open conformation (**Fig. 2c**) and is not positioned correctly for binding BCCP in the active site. In human PC, conformational changes for a smaller region in the C-terminal extension, residues 875–885, are also observed in the monomer bound to BCCP (**Supplementary Fig. 5** online).

The carboxylate group of pyruvate shows bidentate ion-pair interactions with the side chain of Arg644 (614), and its carbonyl group is located within 3.5 Å of the divalent cation. The binding mode of pyruvate was modeled on the basis of chemical interaction considerations as well as comparisons to the transcarboxylase 5S subunit (**Supplementary Fig. 6** online)[13]. Residues Asp572 (542), His771 (741), His773 (743) and carbamylated Lys741 (712) are the ligands to the cation in human PC, whereas the electron density for the carbamyl group on this lysine residue in the *S. aureus* structure is much weaker (**Supplementary Fig. 3**). The binding mode of the cation is similar to that observed in the active site of the 5S subunit[13] (**Supplementary Fig. 6**). However, the ion-pair interactions between pyruvate and Arg644 seem to be unique to PC. This residue is conserved in all PC enzymes, as well as the 5S subunit, but its side chain is pointed away from the substrate in that structure[13] (**Supplementary Fig. 6**).

Binding of BCCP in an exo pocket

The *S. aureus* PC structure revealed a second conformation for the BCCP domain. The BCCP domain of monomer 1 is placed near the CT and PT domains of monomer 2 in the top layer of the tetramer, and vice versa (**Fig. 2a**). The biotin moiety is well ordered and has a clearly defined electron density (**Fig. 5a**). It is located in a small pocket at the interface between the PT and CT domains of the other monomer, near the C-terminal end of the long helix in the PT domain (**Fig. 5b**). There are many intimate contacts between biotin and the enzyme in this exo binding pocket. The convex face of biotin is placed against the side chains of Phe512 (481), Pro513 (482) and Phe1038 (1006), and the side chain of Lys620 (590) is located over the concave face of biotin (**Fig. 5b**). The sulfur atom of biotin is pointed toward the side chains of Tyr503 (472), Thr1023 (991), Tyr1027 (995) and Leu1030 (998). The N1 atom of biotin is recognized via a hydrogen bond to the main chain carbonyl oxygen of residue Phe618 (588), suggesting that this exo pocket can accommodate only noncarboxylated biotin. The NZ atom of Lys1144, in the amide bond with biotin, makes a hydrogen bond with the main chain carbonyl of Gly511 (480).

Whereas this exo binding site is located near the C terminus of the long helix in the PT domain, the binding site for ethyl-CoA is located near the N terminus of this helix[19]. Residues Phe512 (481) and Pro513 (482) in this exo pocket are conserved in a subset of bacterial PC enzymes but are not conserved in mammalian PCs (**Supplementary Fig. 1**), and the human PC has a conformational difference in this region (**Fig. 5c**). This previously unknown exo binding site may present another mechanism for regulating the activity of some of the PC enzymes.

Binding of ATP in the active site of the BC domain

Electron density for ATP was observed in the active site of the BC domain in monomers 1 and 3 of *S. aureus* PC (**Supplementary Fig. 7** online). The adenine base is recognized by specific hydrogen bonds to its N1, N6 and N7 atoms, and several hydrophobic residues contact both of its faces (**Supplementary Fig. 7**). The triphosphate group is located in a highly charged and hydrophilic environment. Domain B of the BC domain assumes a closed conformation, but there are large differences with the domain B observed for the BC subunit of *E. coli* acetyl-CoA carboxylase[21] (**Supplementary Fig. 7**). In the other two monomers, ATP is absent and the entire domain B is disordered. There is no clear correlation between ATP binding in the BC domain and the conformation of BCCP in the current structure.

DISCUSSION

Our structural information on human and *S. aureus* PC has substantial implications for the catalysis by this enzyme. The structures reveal that the BCCP domains of the two monomers in the same layer of the tetramer are swapped, such that the BCCP domain of one monomer actually participates in catalysis in the CT domain of the other monomer (**Fig. 6a**). This provides the molecular basis for the tetramerization requirement of the enzyme, as the BCCP domain cannot reach the active site of the CT domain in the monomeric or dimeric form of the enzyme. A similar conclusion was reached from studies on the *R. etli* enzyme[19].

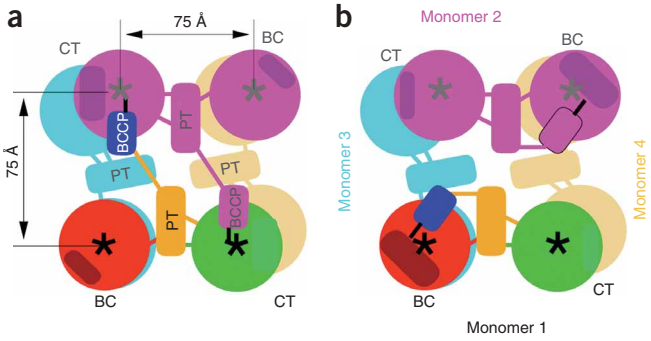

Figure 6 Possible conformations of the biotin-carboxyl carrier protein (BCCP) domain during catalysis. (**a**) A cartoon for the pyruvate carboxylase (PC) tetramer with BCCP in the active site of the carboxyltransferase (CT) domain. The biotin moiety is indicated with a thick black line extending from the BCCP domain. The dark shaded areas in the biotin carboxylase (BC) and CT domains represent subdomains (domain B in BC and the C-terminal extension in CT) that can undergo conformational changes during catalysis. Stars indicate the location of the active sites. (**b**) A cartoon for the PC tetramer with BCCP in the active site of the BC domain. Additional states of the tetramer, with the BCCP domains (randomly) distributed between the BC and CT domains of the four monomers, are also possible, as there may not be any coordination among them. A similar mechanism has been proposed on the basis of the structure of *R. etli* PC[19].

Our observation of the binding of BCCP in the active site of the CT domain has provided the first molecular insights into how biotin participates in the carboxyltransfer reaction (**Fig. 4c**), and it suggests that a conformational change in the C-terminal extension of the CT domain may be important for BCCP binding and catalysis (**Fig. 6a**). The molecular basis for the role of BCCP in the biotin carboxylase reaction remains to be established. On the basis of the structure of the *S. aureus* PC, it is most likely that BCCP-biotin becomes carboxylated in the active site of the BC domain from the same monomer (**Fig. 6b**). The BC and CT active sites are separated by about 75 Å in the current structure (**Fig. 6a**), suggesting that simply swinging the biotin arm on BCCP is not sufficient for biotin to reach both active sites. The BCCP domain itself must also undergo translocation during catalysis (**Fig. 6b**). Residues 1094–1108 in the N-terminal region of BCCP, outside the core of the domain, may serve as the flexible linker for this transition between the two active sites (**Fig. 6b**).

In contrast, the biotin moiety together with the β-hairpin containing the biotinylated lysine residue are disordered in the two BCCP domains that are observed in the *R. etli* PC tetramer[19]. Moreover, the two BCCP domains are located far from the active site of the CT (or BC) domain. Therefore, a change in the position of the BCCP domain is needed before the structure of the *R. etli* enzyme can become catalytically competent.

The structural information also provides an explanation for the molecular basis of some disease-causing mutations. The A610T mutation is located in the biotin binding site in the CT domain (**Fig. 4c**), and the bulkier threonine side chain would disallow the binding of the biotin substrate. The M743I mutation is located in the active site of the CT domain (**Fig. 4c**) and may disrupt substrate binding by the enzyme[13]. The R451C mutation is located in a small pocket that contains a cluster of four strictly conserved arginine residues (Arg398, Arg451, Arg453 and Arg1085), where, in the structure of the *R. etli* PC, the 3′-phospho-ADP portion of ethyl-CoA is bound[19] (**Supplementary Fig. 8** online). Our kinetic studies showed that the R451C mutant is much less sensitive to acetyl-CoA activation than the wild-type enzyme (**Supplementary Fig. 8**), which may explain its disease-causing effects.

In the *S. aureus* PC tetramer, the BC, CT and PT domains all contribute to the oligomerization interface. Although the buried surface area for the PT domain (300 Å^2) is much smaller than that for the BC and CT domains (1,200 Å^2 and 1,400 Å^2, respectively), our mutagenesis studies showed that the PT domain is crucial for the tetramerization of *S. aureus* and human PC. In contrast, the PT domain makes only minor contributions to the tetramer interface in *R. etli* PC (**Fig. 3b**). As the *R. etli* PC structure was determined in the presence of ethyl-CoA, it may be possible that binding of the activator has alleviated the requirement of the PT domain for tetramerization. Our kinetic studies showed that, although the Y1077A mutant of the *S. aureus* PC is dimeric and essentially inactive in the absence of acetyl-CoA, it becomes fully active in the presence of 1 mM acetyl-CoA (**Supplementary Fig. 8**).

An asymmetric tetramer was observed in the structure of *R. etli* PC, in complex with ethyl-CoA, in apparent agreement with the half-sites reactivity of PC enzymes[19]. Whereas our structure of the *S. aureus* PC is symmetric overall, the BCCP domains do not obey this symmetry, which may also contribute to the half-sites reactivity. Further evidence of the asymmetry in the *R. etli* PC structure is provided in that ethyl-CoA is bound only to two monomers in the tetramer, and the other two monomers show a large conformational difference[19] (**Fig. 3c**). Interestingly, the conformation of the *R. etli* PC monomer in complex with ethyl-CoA is actually more similar to that of the *S. aureus* PC monomer, in the absence of CoA (**Fig. 3c**). It remains to be determined what conformational changes will be induced by acetyl-CoA binding in the *S. aureus* enzyme.

Hepatic gluconeogenesis is elevated in diabetic patients, and enzymes in this pathway, including PC, are attractive targets for drug discovery[4,22]. Our studies define the atomic structures of this enzyme and provide a good starting point for the design and development of compounds that can modulate the activity of this important enzyme. Our observation of BCCP binding in the active site of CT provides molecular insights into how BCCP participates in the carboxyltransfer reaction, which should also help us to understand the catalysis by biotin-dependent carboxylases in general.

METHODS

Protein expression and purification. We subcloned the CT+PT+BCCP domain of human PC (residues 482–1178) into vector pET28a (Novagen). The plasmid was co-transformed into *E. coli* BL21Star (DE3) cells with a compatible plasmid that carries the bacterial biotin ligase (*birA*) gene. Cells were cultured in LB medium containing 50 mg l^{-1} kanamycin and 34 mg l^{-1} chloramphenicol, and induced with 0.3 mM IPTG for 12–14 h at 20 °C. 30 min before induction, 0.1 mM $MnCl_2$ and 15 mg l^{-1} of D-biotin were added to the medium. The protein was purified by nickel-agarose affinity chromatography (Qiagen) and gel-filtration chromatography (S-300, GE Healthcare). Purified protein was concentrated to 8 mg ml^{-1} in a buffer containing 20 mM Tris (pH 7.5), 50 mM NaCl, 5% (v/v) glycerol and 2 mM DTT, flash-frozen in liquid nitrogen and stored at –80 °C. The N-terminal histidine tag was not removed for crystallization.

The cloning, expression and purification procedures for the full-length human PC (residues 20–1178, excluding the mitochondrial targeting sequence) and the full-length PC from *Staphylococcus aureus* followed similar protocols. Purified proteins were concentrated to 20 mg ml^{-1} in 20 mM Tris (pH7.5), 5% (v/v) glycerol, 2mM DTT and 250 mM NaCl (for human PC) or 200 mM NaCl (for *S. aureus* PC).

Table 1 Data collection and refinement statistics

	CT+PT+BCCP domains of human PC	F1077A mutant of human PC	Full-length *S. aureus* PC
Data collection			
Space group	$P2_1$	$P3_121$	$P2_1$
Cell dimensions			
a, *b*, *c* (Å)	81.2, 173.3, 118.2	107.5, 107.5, 524.5	96.4, 256.1, 126.0
α, β, γ (°)	90, 95.9, 90	90, 90, 120	90, 109.5, 90
Resolution (Å)	30–2.8 (2.9–2.8)	30–3.0 (3.11–3.0)	30–2.8 (2.9–2.8)
R_{merge} (%)	7.2 (41.3)	12.5 (41.9)	8.7 (41.6)
$I / \sigma I$	17.2 (2.0)	16.2 (3.7)	14.3 (3.1)
Completeness (%)	94 (67)	89 (91)	99 (97)
Redundancy	3.6 (3.0)	5.4 (5.5)	3.4 (3.0)
Refinement			
Resolution (Å)	30–2.8	30–3.0	30–2.8
Number of reflections	71,717	58,177	131,856
R_{work} / R_{free} (%)	21.6 / 27.1	19.0 / 23.6	21.5 / 26.8
No. atoms			
Protein	19,774	18,512	34,293
Ligand/ion	43	4	150
Water	—	—	—
B-factors			
Protein	94.3	46.8	68.1
Ligand/ion	83.3	86.6	84.3
Water	—	—	—
r.m.s. deviations			
Bond lengths (Å)	0.010	0.010	0.008
Bond angles (°)	1.2	1.4	1.1

The numbers in parentheses are for the highest-resolution shell. One crystal was used for each data collection.

All the proteins purified by the above procedures are fully biotinylated, as demonstrated by an avidin binding gel-shift assay (data not shown).

We carried out analytical gel-filtration experiments using a Superose-12 10/30 column (GE Healthcare), with a buffer containing 20 mM Tris (pH 7.5) and 200 mM NaCl, injecting 0.5 mg of protein onto the column. Static light-scattering experiments were performed on a Dawn Eos instrument (Wyatt Technology), using 0.1 mg of protein on a Shodex PROTEIN KW-804 column and the same buffer as used for the analytical gel-filtration experiment.

Protein crystallization. Plate-like crystals of the CT+PT+BCCP domain of human PC were obtained at room temperature after microseeding in a sitting-drop setup. The reservoir solution contained 0.8% (w/v) PEG3350 and 90 mM $MnCl_2$. Before crystallization setup, we added 5 mM of oxaloacetic acid (Sigma) and 10 mM of N,N-dimethyl-decylamine-N-oxide (DDAO; Hampton Research) to the protein. The crystals were cryoprotected by transferring them to the reservoir solution supplemented with 20% (v/v) glycerol (in 5% steps). There is a tetramer in the asymmetric unit.

Diamond-shaped crystals of the F1077A mutant of the CT+PT+BCCP domain of human PC were obtained at room temperature in a sitting-drop setup. The reservoir solution contained 15% (w/v) PEG3350 and 200 mM ammonium tartrate. The crystals were cryoprotected by transferring them to the reservoir solution supplemented with 25% (v/v) ethylene glycol. There are two dimers in the asymmetric unit.

Rod-like crystals of *S. aureus* PC were obtained in a sitting-drop setup at room temperature. The reservoir solution contains 20% (w/v) PEG3350 and 200 mM ammonium tartrate. Before crystallization setup, 5 mM each of ATP and oxaloacetic acid were added to the protein solution. The crystals were cryoprotected by transferring them to the reservoir solution supplemented with 25% (v/v) ethylene glycol. There is a tetramer in the asymmetric unit.

Data collection and structure determination. X-ray diffraction data for the CT+PT+BCCP domain of human PC were collected at the National Synchrotron Light Source (NSLS) beamline X29A on an ADSC q315 CCD detector (X-ray wavelength 1.0809 Å). Diffraction data for the F1077A mutant and the *S. aureus* PC were collected at NSLS beamline X4A on an ADSC quantum 4 CCD detector (X-ray wavelength 0.9795 Å). The diffraction images were processed with the program HKL[23], and the data-processing statistics are summarized in **Table 1**.

The structure of the CT+PT+BCCP domain of human PC was solved by molecular replacement with Molrep[24], using a search model constructed from the structure of the 5S subunit of *P. shermanii* transcarboxylase (PDB code 1RQH)[13]. The model contains the dimer of 5S, but with all of the amino acid side chains truncated to alanine. Manual rebuilding to modify the model was done in O[25], and structure refinement was carried out with CNS[26] and Refmac[27].

The structure of the F1077A mutant was solved by molecular replacement with Molrep, with the wild-type structure as the search model. The crystals are hemihedrally twinned, and twinned refinement was carried out with CNS.

The structure of the full-length *S. aureus* PC was determined by molecular replacement with the program COMO[28]. The structures of the CT+PT+BCCP domain of human PC and the BC subunit of *A. aeolicus* PC[12] were used as the search models. Model rebuilding was done in O, and structure refinement was carried out with CNS and Refmac.

Mutagenesis and activity assays. The mutants were made with the QuikChange kit (Stratagene) and verified by sequencing. We expressed and purified the mutant proteins following the same protocol as that used for the wild-type protein.

The catalytic activity of full-length PC was measured at room temperature following a published protocol[29]. The reaction mixture contains 100 mM Tris (pH 7.5), 200 mM NaCl, 5 mM $MgCl_2$, 2 mM ATP, 50 mM bicarbonate, varying concentrations of pyruvate or acetyl-CoA, 0.2 mM NADH and 8 μM of human PC or 0.1 μM of *S. aureus* PC and 5 units of malate dehydrogenase (Sigma).

Accession codes. Protein Data Bank: Coordinates and structure factors have been deposited with accession codes 3BG3 (human PC), 3BG9 (F1077A mutant of human PC) and 3BG5 (*S. aureus* PC).

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS

We thank A. Heroux and H. Robinson for setting up the X29A beamline; R. Abramowitz and J. Schwanof for setting up the X4A beamline at the NSLS; L. Yu for help with crystallization and kinetic experiments; D. Parisotto for help with kinetic assays; W.W. Cleland for helpful discussions. This research is supported in part by a grant from the US National Institutes of Health (DK67238) to L.T.

Published online at http://www.nature.com/nsmb
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions

1. Attwood, P.V. The structure and the mechanism of action of pyruvate carboxylase. *Int. J. Biochem. Cell Biol.* **27**, 231–249 (1995).
2. Wallace, J.C., Jitrapakdee, S. & Chapman-Smith, A. Pyruvate carboxylase. *Int. J. Biochem. Cell Biol.* **30**, 1–5 (1998).
3. Jitrapakdee, S. & Wallace, J.C. Structure, function and regulation of pyruvate carboxylase. *Biochem. J.* **340**, 1–16 (1999).
4. Jitrapakdee, S., Vidal-Puig, A. & Wallace, J.C. Anaplerotic roles of pyruvate carboxylase in mammalian tissues. *Cell. Mol. Life Sci.* **63**, 843–854 (2006).
5. Utter, M.F. & Keech, D.B. Formation of oxaloacetate from pyruvate and CO_2. *J. Biol. Chem.* **235**, 17–18 (1960).
6. Robinson, B.H. Lactic acidemia and mitochondrial disease. *Mol. Genet. Metab.* **89**, 3–13 (2006).
7. Carbone, M.A. *et al.* Amerindian pyruvate carboxylase deficiency is associated with two distinct missense mutations. *Am. J. Hum. Genet.* **62**, 1312–1319 (1998).
8. Wexler, I.D. *et al.* Molecular characterization of pyruvate carboxylase deficiency in two consanguineous families. *Pediatr. Res.* **43**, 579–584 (1998).

9. Carbone, M.A. & Robinson, B.H. Expression and characterization of a human pyruvate carboxylase variant by retroviral gene transfer. *Biochem. J.* **370**, 275–282 (2003).
10. Cronan, J.E. & Jr & Waldrop, G.L. Multi-subunit acetyl-CoA carboxylases. *Prog. Lipid Res.* **41**, 407–435 (2002).
11. Tong, L. Acetyl-coenzyme A carboxylase: crucial metabolic enzyme and attractive target for drug discovery. *Cell. Mol. Life Sci.* **62**, 1784–1803 (2005).
12. Kondo, S. *et al.* Structure of the biotin carboxylase subunit of pyruvate carboxylase from *Aquifex aeolicus* at 2.2 Å resolution. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 486–492 (2004).
13. Hall, P.R. *et al.* Transcarboxylase 5S structures: assembly and catalytic mechanism of a multienzyme complex subunit. *EMBO J.* **23**, 3621–3631 (2004).
14. Studer, R. *et al.* Crystal structure of the carboxyltransferase domain of the oxaloacetate decarboxylase Na. pump from *Vibrio cholerae*. *J. Mol. Biol.* **367**, 547–557 (2007).
15. Attwood, P.V., Johannssen, W., Chapman-Smith, A. & Wallace, J.C. The existence of multiple tetrameric conformers of chicken liver pyruvate carboxylase and their roles in dilution inactivation. *Biochem. J.* **290**, 583–590 (1993).
16. Cazzulo, J.J. & Stoppani, A.O.M. The regulation of yeast pyruvate carboxylase by acetyl-coenzyme A and L-aspartate. *Arch. Biochem. Biophys.* **127**, 563–567 (1968).
17. Scrutton, M.C. & White, M.D. Pyruvate carboxylase. Inhibition of the mammalian and avian liver enzymes by a-ketoglutarate and L-glutamate. *J. Biol. Chem.* **249**, 5405–5415 (1974).
18. Carey, P.R., Sonnichsen, F.D. & Yee, V.C. Transcarboxylase: one of nature's early nanomachines. *IUBMB Life* **56**, 575–583 (2004).
19. St. Maurice, M. *et al.* Domain architecture of pyruvate carboxylase, a biotin-dependent multifunctional enzyme. *Science* **317**, 1076–1079 (2007).
20. Athappilly, F.K. & Hendrickson, W.A. Structure of the biotinyl domain of acetyl-coenzyme A carboxylase determined by MAD phasing. *Structure* **3**, 1407–1419 (1995).
21. Thoden, J.B., Blanchard, C.Z., Holden, H.M. & Waldrop, G.L. Movement of the biotin carboxylase B-domain as a result of ATP binding. *J. Biol. Chem.* **275**, 16183–16190 (2000).
22. Moller, D.E. New drug targets for type 2 diabetes and the metabolic syndrome. *Nature* **414**, 821–827 (2001).
23. Otwinowski, Z. & Minor, W. Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
24. Vagin, A. & Teplyakov, A. An approach to multi-copy search in molecular replacement. *Acta Crystallogr. D Biol. Crystallogr.* **56**, 1622–1624 (2000).
25. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
26. Brunger, A.T. *et al.* Crystallography & NMR System: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905–921 (1998).
27. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255 (1997).
28. Jogl, G., Tao, X., Xu, Y. & Tong, L. COMO: A program for combined molecular replacement. *Acta Crystallogr. D Biol. Crystallogr.* **57**, 1127–1134 (2001).
29. Modak, H.V. & Kelly, D.J. Acetyl-CoA-dependent pyruvate carboxylase from the photosynthetic bacterium *Rhodobacter capsulatus*: rapid and efficient purification using dye-ligand affinity chromatography. *Microbiology* **141**, 2619–2628 (1995).
30. DeLano, W.L. The Pymol manual. (DeLano Scientific, San Carlos, CA, 2002).