Re-refinement of *Plasmodium falciparum* orotidine 5′-monophosphate decarboxylase provides a clearer picture of an important malarial drug target

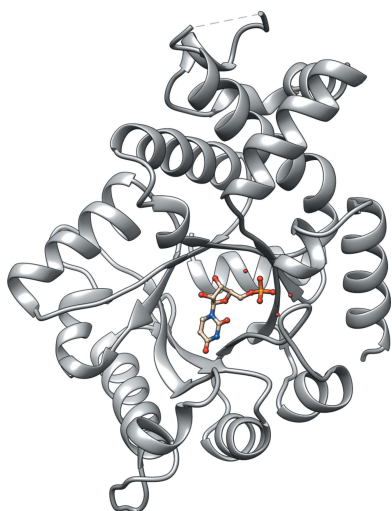Walter R. P. Novak,[a] Korbin H. J. West,[a] Lucy M. D. Kirkman[b] and Gabriel S. Brandt[b]*

[a]Chemistry Department, Wabash College, 301 West Wabash Avenue, Crawfordsville, IN 47933, USA, and [b]Chemistry Department, Franklin and Marshall College, PO Box 3003, Lancaster, PA 17604, USA. *Correspondence e-mail: gbrandt@fandm.edu

The development of antimalarial drugs remains a public health priority, and the orotidine 5′-monophosphate decarboxylase from *Plasmodium falciparum* (PfOMPDC) has great potential as a drug target. The crystallization of PfOMPDC with substrate bound represents an important advance for structure-based drug-design efforts [Tokuoka *et al.* (2008), *J. Biochem.* **143**, 69–78]. The complex of the enzyme bound to the substrate OMP (PDB entry 2za1) would be of particular utility in this regard. However, re-refinement of this structure of the Michaelis complex shows that the bound ligand is the product rather than the substrate. Here, the re-refinement of a set of three structures, the apo enzyme and two versions of the product-bound form (PDB entries 2za1, 2za2 and 2za3), is reported. The improved geometry and fit of these structures to the observed electron density will enhance their utility in antimalarial drug design.

1. Introduction

Protein structures, particularly those with substrates bound, have been invaluable in helping to illuminate the details of enzyme mechanisms and the design of novel inhibitors (Petsko & Ringe, 2004). In order to provide insight into the workings of enzymes, protein structures must be interpreted in the light of, and are limited by, their ability to capture atomic-level detail (Wlodawer *et al.*, 2018). The enzyme orotidine 5′-monophosphate decarboxylase (OMPDC) is involved in the synthesis of pyrimidine nucleotides and is essential in malaria parasites (Rathod & Reyes, 1983). Unlike their human hosts, who have a salvage pathway, malaria parasites rely entirely on *de novo* nucleotide synthesis, so OMPDC has been proposed as a potential therapeutic target (Cassera *et al.*, 2011). In addition, the ability of OMPDC to efficiently catalyze decarboxylation without a cofactor has attracted the interest of mechanistic enzymologists (Richard *et al.*, 2018; Lewis *et al.*, 2017; Reyes *et al.*, 2015; Vardi-Kilshtain *et al.*, 2013; Lee & Houk, 1997). For both of these reasons, the structure of OMPDC from *Plasmodium falciparum* (PfOMPDC) with the substrate orotidine 5′-monophosphate (OMP) bound would be of considerable scientific interest. The series of structures with PDB codes 2za2, 2za1 and 2za3 represents an attempt to capture the enzyme alone, with the OMP substrate and with its product uridine 5′-monophosphate (UMP), respectively (Tokuoka *et al.*, 2008).

Inspection of these models reveals significant imperfections in terms of backbone geometry, density fit and ligand selection (Table 1). The *PDB-REDO* server is often able to make

Table 1
Structure solution and refinement.

Condition	OMP added		Apo form		UMP added	
Re-refinement (Y/N)	N	Y	N	Y	N	Y
PDB code	2za1	6dsq	2za2	6dsr	2za3	6dss
Resolution range (Å)	38.14–2.65		39.30–2.70		38.27–2.65	
Completeness (%)	98.9		99.6		99.0	
σ Cutoff	0.00		0.00		0.00	
No. of reflections, working set	19268		18297		19658	
No. of reflections, test set	1877		1802		1939	
Final R_{cryst} (%)	21.0	18.3	21.0	19.5	20.4	21.0
Final R_{free} (%)	29.2	22.4	30.9	20.1	28.7	24.3
No. of non-H atoms						
Protein	5109	4998	4964	4979	5098	5260
Ligand	48	42	0	0	42	42
Water	118	54	38	33	47	56
Total	5275	5094	5002	5012	5187	5358
R.m.s. deviations						
Bonds (Å)	0.010	0.004	0.007	0.008	0.007	0.006
Angles (°)	1.8	0.7	1.4	1.0	1.4	0.8
Average B factors (Å^2)						
Protein	54	53	56	53	52	51
Ligand	73	52	—	—	48	44
Water	52	49	36	43	41	48
Ramachandran plot						
Most favored (%)	76.7	97.4	79.5	96.7	82.99	96.85
Allowed (%)	91.07	99.84	94.37	100	93.58	99.84
All-atom clashscore† (percentile)	50.11 (12)	7.6 (99)	37.33 (40)	10.22 (97)	32.66 (40)	6.46 (99)
Overall *MolProbity* score (percentile)	3.72 (10)	1.53 (100)	3.22 (23)	1.74 (99)	3.37 (24)	1.55 (100)

† Clashscore is the number of serious steric overlaps (>0.4 Å) per 1000 atoms.

dramatic improvements to problematic structures in an auto-mated fashion (Joosten *et al.*, 2014). However, in the case of these three structures, *PDB-REDO* was unable to improve the models to current standards of practice. Here, we describe the re-refinement of PDB entries 2za1, 2za2 and 2za3, resulting in a dramatic improvement of the models as suggested by the geometric parameters, density-fit analysis and *R* factors (Table 1). Our most consequential finding is that re-refinement of 2za1 shows that the bound ligand is actually the product UMP, rather than the substrate OMP as reported (Tokuoka *et al.*, 2008). These re-refined structures have been deposited in the PDB as entries 6dsq, 6dsr and 6dss, respectively.

2. Materials and methods

2.1. Structure solution and refinement

All three structures were re-refined from the deposited structure factors using the *PHENIX* suite (Adams *et al.*, 2010). A starting model of the apoenzyme PfOMPDC (PDB entry 3n3m; Y. Liu, L. P. Kotra & E. F. Pai, unpublished work) was used to generate a molecular-replacement solution. In order to minimize model bias, refinement was carried out after randomization of the coordinates of the input model in all cases using the 'shake' feature in *phenix.pdbtools* (Adams *et al.*, 2010) with a maximum displacement of 0.2 Å. Coordinate randomization was accompanied by the removal of all ligands and bound water molecules to produce OMIT maps of bound ligands. Structures were validated using *MolProbity* and, during deposition, by the wwPDB validation service (Williams *et al.*, 2018; Chen *et al.*, 2010). The final re-refined structure

with UMP bound in the active site of PfOMPDC was refined using *REFMAC* and *BUSTER-TNT* (Murshudov *et al.*, 2011; Blanc *et al.*, 2004). For comparison, electron-density maps of the deposited 2za1 structure were also generated using *PHENIX*, *REFMAC* and *BUSTER-TNT*. In both cases, all three refinement programs gave consistent results. Figs. 1, 2(*b*), 2(*c*) and 3 were generated from the electron-density maps produced by refinement in *PHENIX* using *UCSF Chimera* (Pettersen *et al.*, 2004).

3. Results and discussion

Re-assessment of the structure of the orotidine 5′-mono-phosphate decarboxylase (OMPDC) from *P. falciparum* with substrate bound has revealed that this structure is not the Michaelis complex, as reported, but rather a product complex. The structural enzymology of OMPDC has depended on high-quality structures of intermediates in the active site (Richard *et al.*, 2018; Lewis *et al.*, 2017; Reyes *et al.*, 2015; Vardi-Kilshtain *et al.*, 2013; Lee & Houk, 1997). Additionally, the structure of OMPDC from *P. falciparum* is important because of its potential as drug target (Tokuoka *et al.*, 2008; Fujihashi *et al.*, 2015; Cassera *et al.*, 2011). The presence of bound substrates in the study of Tokuoka and coworkers enhances the value of these structures to structure-based drug design (Tokuoka *et al.*, 2008). Indeed, *in silico* drug-design efforts based on these structures have been reported (Takashima *et al.*, 2012; Drinkwater & McGowan, 2014). For both of these reasons, it is important that these structures are well determined and that the ligands are correctly identified (Touw *et al.*, 2016).

The enzyme OMPDC is a member of the $(\beta/\alpha)_8$-barrel superfamily. The *P. falciparum* OMPDC structures reported by Tokuoka and coworkers maintain this fold (Fig. 1).

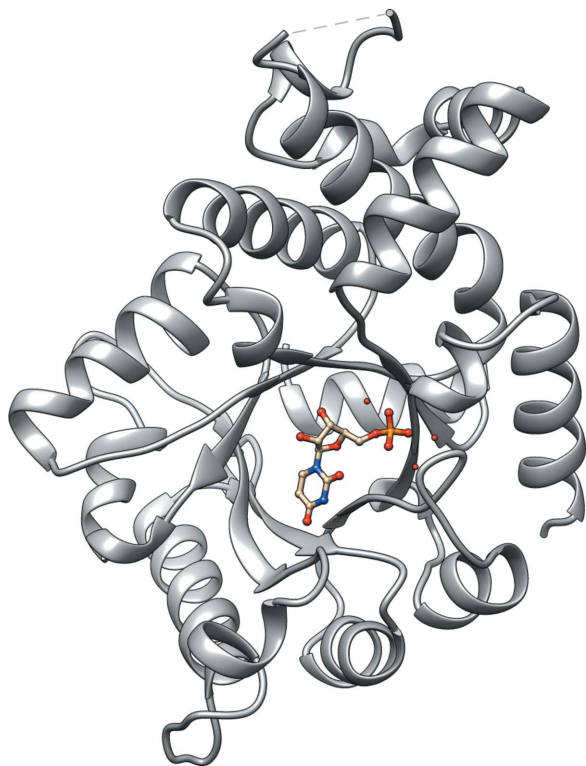However, the three structures of the set are not well validated overall (Wlodawer *et al.*, 2018; Dauter *et al.*, 2014; Kleywegt & Jones, 1996). By both clashscore and number of Ramachandran outliers, the OMP-bound structure with PDB code 2za1, in particular, is seen to be among the lowest scoring in the PDB (12th percentile). The re-refinement reported here solves these overall problems (Table 1, Fig. 2*a*). The number of residues in disfavored regions of the Ramachandran plot is reduced to an appropriate level, and the overall clashscores improve to the very topmost percentile of structures in the PDB (Fig. 2*a*). A specific example of the results of these improvements can be seen in Fig. 2(*b*), where a region of the protein that was formerly fitted as an extended loop because of a residue in a disallowed conformation is shown to form a proper α-helix once the residue is modeled with allowed Ramachandran angles. The potential importance of the re-fitting to the active site is shown in Fig. 2(*c*), which depicts the backbone atoms in the apoenzyme that are within 5 Å of where the substrate would bind. Although the original and refitted structures agree overall, regions of difference that affect the shape of the substrate-binding pocket can be observed.

More so than the global structure, modeling of the ligand in these structures is, of course, critical to their interpretation. From an enzymological perspective, the decarboxylation reaction catalyzed by OMPDC is a very well studied example of the rate acceleration that can arise from control of the electrostatic environment around a substrate (Fried & Boxer, 2017; Amyes *et al.*, 2017; Jordan & Patel, 2013). OMPDC has been shown to convert substrate to product with minimal active-site rearrangement (Richard *et al.*, 2018; Fried & Boxer, 2017). This aspect of the enzyme is important because the experimental method of Tokuoka and coworkers for forming the co-complex of the enzyme and substrate relied on soaking the crystals in OMP. The turnover rate of the wild-type
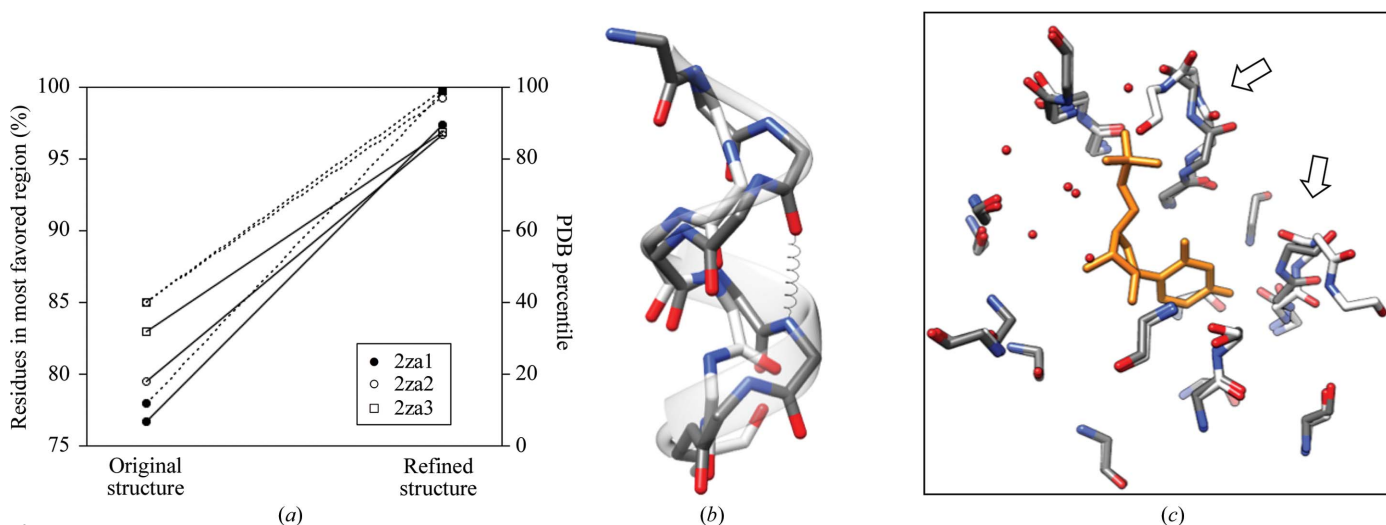

Figure 1
Schematic of the structure of the PfOMPDC monomer with the reaction product uridine 5'-monophosphate (UMP) bound.

Figure 2
Global improvement of the PfOMPDC structure upon refitting. (*a*) The percentage of residues in the most favored region of Ramachandran space (solid lines) is increased in all cases. The all-atom clashscore (dotted lines), as a percentile score among all structures in the PDB, is dramatically increased. (*b*) An apoenzyme region originally modeled in PDB entry 2za2 (white) as an extended loop, owing to a residue with the backbone carbonyl misaligned, is shown by the re-refined structure (gray) to form a typical α-helix with backbone atoms positioned within hydrogen-bonding distance (shown by the gray spring). (*c*) Comparison of the protein backbone near the active site shows potentially important differences (arrows) that change the shape of the binding pocket. Residues in PDB entry 2za2 (white) that would be within 5 Å of the substrate (orange, modeled from PDB entry 2za1) are compared with those of the re-refined apoenzyme structure (gray).
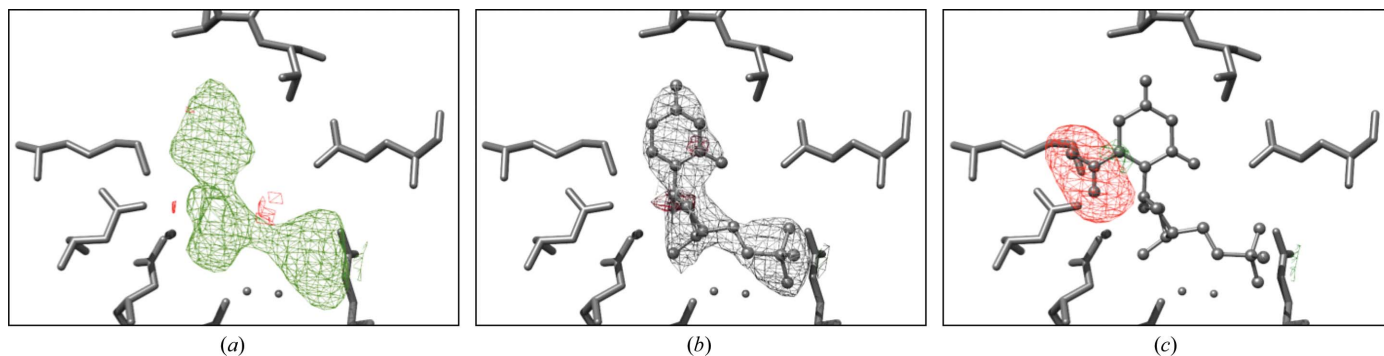
Figure 3
Ligand electron density in the PfOMPDC active site after re-refinement of the deposited structure factors (PDB entry 2za1). (a) A simulated-annealing OMIT map showing the $F_o - F_c$ difference density map in green for positive and red for negative difference density (both contoured at 3.0 r.m.s.d.). (b) Modeling UMP into this unbiased ligand density, showing the $2F_o - F_c$ map in gray (1.5 r.m.s.d.) and difference density in green and red (± 3.0 r.m.s.d.). (c) Modeling OMP into ligand electron density, with the difference density $F_o - F_c$ map in green and red (± 3.0 r.m.s.d.).

enzyme is very high, and catalysis is thought to occur with little rearrangement of the active site (Fujihashi *et al.*, 2013; Wu *et al.*, 2000). Little barrier exists to the enzyme simply converting the substrate in these soaking experiments to product. Indeed, the originally deposited 2za1 structure itself gives evidence that the substrate OMP is an imperfect fit to the observed electron density. The *B* factors of the carboxylate atoms are significantly higher than those of the uridine, particularly in the *B* subunit (average carboxylate *B* factor of 109 Å^2 compared with 73 Å^2 for the rest of the OMP cofactor). Another noteworthy aspect of modeling the carboxylate into this electron density is that the interatomic distances between the OMP carboxylate O atom and the nearest Asp136 carboxylate O atom are 2.3 and 2.5 Å in the *A* and *B* subunits, respectively. This proximity was interpreted as evidence of the mechanistic role of electrostatic destabilization of the substrate carboxylate on the part of the enzyme (Tokuoka *et al.*, 2008). A more parsimonious explanation, perhaps, is that the substrate carboxylate is not present in this space.

The validation metrics of the original structures raise the possibility that model bias has not been well controlled in these structures (Hodel *et al.*, 1992). In fact, the relatively large gap observed between R_{cryst} and R_{free} for all three structures is a sign of overfitting (Kleywegt & Jones, 1997). Model bias is a special concern when it comes to ligand fitting (Wlodawer *et al.*, 2018). Unbiased re-refinement of PDB entry 2za1 shows that the OMIT map for the bound ligand contains very little electron density in the region of the carboxylate (Fig. 3a). The placement of UMP into this electron density and subsequent refinement confirms that no positive difference density exists in the carboxylate region (Fig. 3b). Finally, attempting to fit OMP into this electron density reveals the extent of negative difference density around these atoms (Fig. 3c). To be sure that the observed differences in electron density around the orotidine carboxylate were not owing to idiosyncrasies of the refinement software, both the originally deposited structure 2za1 and the re-refined structure 6dsq were evaluated using *REFMAC* and *BUSTER-TNT* to complement the *PHENIX* refinement. All three programs revealed pronounced negative difference density around the carboxylate in PDB entry 2za1, suggesting its absence. At the same time, all three refinement

packages showed that UMP fits well into the observed electron density, without any positive difference density, in the absence of the carboxylate. Therefore, we conclude that it is the reaction product UMP and not the substrate OMP that is bound in PDB entry 2za1.

Indeed, this conclusion is consistent with the thorough reporting of Tokuoka and coworkers of the differences among the three structures, in which the two ligand-bound structures 2za1 and 2za3 are seen to be essentially identical, differing in the same ways from the apo structure 2za2.

Funding information

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
Amyes, T. L., Malabanan, M. M., Zhai, X., Reyes, A. C. & Richard, J. P. (2017). *Protein Eng. Des. Sel.* **30**, 157–165.
Blanc, E., Roversi, P., Vonrhein, C., Flensburg, C., Lea, S. M. & Bricogne, G. (2004). *Acta Cryst.* D**60**, 2210–2221.
Cassera, M. B., Zhang, Y., Hazleton, K. Z. & Schramm, V. L. (2011). *Curr. Top. Med. Chem.* **11**, 2103–2115.
Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. & Richardson, D. C. (2010). *Acta Cryst.* D**66**, 12–21.
Dauter, Z., Wlodawer, A., Minor, W., Jaskolski, M. & Rupp, B. (2014). *IUCrJ*, **1**, 179–193.
Drinkwater, N. & McGowan, S. (2014). *Biochem. J.* **461**, 349–369.
Fried, S. D. & Boxer, S. G. (2017). *Annu. Rev. Biochem.* **86**, 387–415.
Fujihashi, M., Ishida, T., Kuroda, S., Kotra, L. P., Pai, E. F. & Miki, K. (2013). *J. Am. Chem. Soc.* **135**, 17432–17443.
Fujihashi, M., Mnpotra, J. S., Mishra, R. K., Pai, E. F. & Kotra, L. P. (2015). *J. Genet. Genomics*, **42**, 221–234.
Hodel, A., Kim, S.-H. & Brünger, A. T. (1992). *Acta Cryst.* A**48**, 851–858.
Joosten, R. P., Long, F., Murshudov, G. N. & Perrakis, A. (2014). *IUCrJ*, **1**, 213–220.
Jordan, F. & Patel, H. (2013). *ACS Catal.* **3**, 1601–1617.
Kleywegt, G. J. & Jones, T. A. (1996). *Structure*, **4**, 1395–1400.
Kleywegt, G. J. & Jones, T. A. (1997). *Methods Enzymol.* **277**, 208–230.

Lee, J. K. & Houk, K. N. (1997). *Science*, **276**, 942–945.

Lewis, C. A., Shen, L., Yang, W. & Wolfenden, R. (2017). *Biochemistry*, **56**, 1498–1503.

Murshudov, G. N., Skubák, P., Lebedev, A. A., Pannu, N. S., Steiner, R. A., Nicholls, R. A., Winn, M. D., Long, F. & Vagin, A. A. (2011). *Acta Cryst.* D**67**, 355–367.

Petsko, G. A. & Ringe, D. (2004). *Protein Structure and Function.* London: New Science Press.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C. & Ferrin, T. E. (2004). *J. Comput. Chem.* **25**, 1605–1612.

Rathod, P. K. & Reyes, P. (1983). *J. Biol. Chem.* **258**, 2852–2855.

Reyes, A. C., Zhai, X., Morgan, K. T., Reinhardt, C. J., Amyes, T. L. & Richard, J. P. (2015). *J. Am. Chem. Soc.* **137**, 1372–1382.

Richard, J. P., Amyes, T. L. & Reyes, A. C. (2018). *Acc. Chem. Res.* **51**, 960–969.

Takashima, Y., Mizohata, E., Krungkrai, S. R., Fukunishi, Y., Kinoshita, T., Sakata, T., Matsumura, H., Krungkrai, J., Horii, T. & Inoue, T. (2012). *J. Biochem.* **152**, 133–138.

Tokuoka, K., Kusakari, Y., Krungkrai, S. R., Matsumura, H., Kai, Y., Krungkrai, J., Horii, T. & Inoue, T. (2008). *J. Biochem.* **143**, 69–78.

Touw, W. G., Joosten, R. P. & Vriend, G. (2016). *J. Mol. Biol.* **428**, 1375–1393.

Vardi-Kilshtain, A., Doron, D. & Major, D. T. (2013). *Biochemistry*, **52**, 4382–4390.

Williams, C. J. *et al.* (2018). *Protein Sci.* **27**, 293–315.

Wlodawer, A., Dauter, Z., Porebski, P. J., Minor, W., Stanfield, R., Jaskolski, M., Pozharski, E., Weichenberger, C. X. & Rupp, B. (2018). *FEBS J.* **285**, 444–466.

Wu, N., Mo, Y., Gao, J. & Pai, E. F. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 2017–2022.