

# SCIENTIFIC REPORTS



OPEN

## Structural determinants of the catalytic mechanism of *Plasmodium* CCT, a key enzyme of malaria lipid biosynthesis

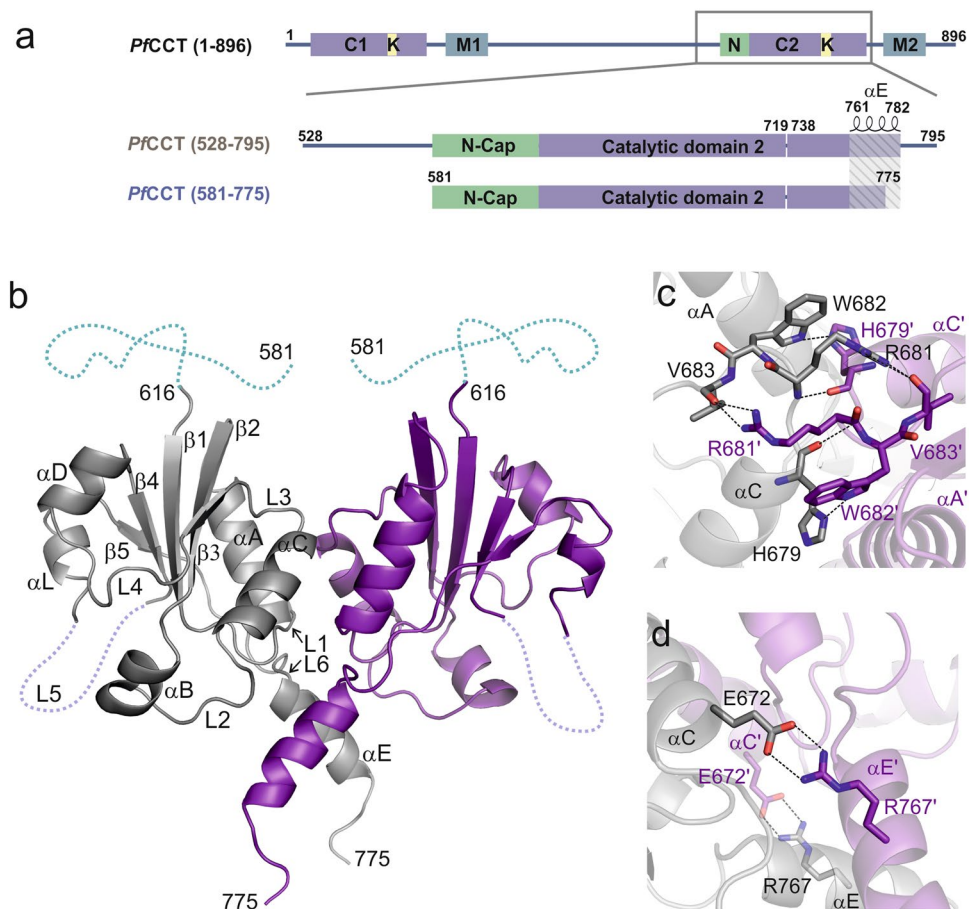
Ewelina Guca<sup>1,7</sup>, Gergely N. Nagy<sup>2,3,8</sup>, Fanni Hajdú<sup>2,3</sup>, Lívia Marton<sup>3,4</sup>, Richard Izrael<sup>2,3</sup>, François Hoh<sup>5,6</sup>, Yinshan Yang<sup>5,6</sup>, Henri Vial<sup>1</sup>, Beata G. Vértessy<sup>2,3</sup>, Jean-François Guichou<sup>5,6</sup> & Rachel Cerdan<sup>1</sup>

The development of the malaria parasite, *Plasmodium falciparum*, in the human erythrocyte, relies on phospholipid metabolism to fulfil the massive need for membrane biogenesis. Phosphatidylcholine (PC) is the most abundant phospholipid in *Plasmodium* membranes. PC biosynthesis is mainly ensured by the *de novo* Kennedy pathway that is considered as an antimalarial drug target. The CTP:phosphocholine cytidyltransferase (CCT) catalyses the rate-limiting step of the Kennedy pathway. Here we report a series of structural snapshots of the PfCCT catalytic domain in its free, substrate- and product-complexed states that demonstrate the conformational changes during the catalytic mechanism. Structural data show the ligand-dependent conformational variations of a flexible lysine. Combined kinetic and ligand-binding analyses confirm the catalytic roles of this lysine and of two threonine residues of the helix  $\alpha$ E. Finally, we assessed the variations in active site residues between *Plasmodium* and mammalian CCT which could be exploited for future antimalarial drug design.

Malaria is a major global health problem and the most life-threatening parasitic disease<sup>1</sup>. In 2016, 216 million cases of malaria and 445 000 deaths were estimated<sup>2</sup>. *Plasmodium falciparum* is the parasite that causes the most fatal form of malaria among all species infecting humans. While the widespread antimalarial efforts featured by the use of insecticide-impregnated bed nets and the artemisinin-based combination therapies (ACT) yielded a marked decrease in malaria incidence until 2015, the rate of decline has recently stalled<sup>2</sup>. Further progress in reducing malaria may be compromised due to the emergence of artemisinin resistant parasites<sup>3,4</sup>. Therefore, it is necessary to search alternative antimalarial therapeutics involving novel targets and mechanisms of action.

During the symptomatic erythrocytic stage of *P. falciparum*, the parasite relies on phospholipids to build the membranes necessary for intense cell division. The major phospholipid is phosphatidylcholine (PC) accounting for 40–50% of the total phospholipid content and the parasite possesses its own enzymatic machinery to produce PC<sup>5,6</sup>. The *de novo* Kennedy pathway is the main route for the PC synthesis in *P. falciparum* and has been identified as a pharmacological target for the treatment of malaria<sup>7–10</sup>. Recently, a link has been established between the level of lysophosphatidylcholine, a major supplier of choline for the Kennedy pathway, and the sexual stage differentiation in *P. falciparum*<sup>11</sup>. This emphasizes the importance of phospholipid biosynthesis at different stages of the parasite development. The second step of the PC biosynthesis process catalyzed by CTP:phosphocholine cytidyltransferase (CCT) [EC: 2.7.7.15] is rate limiting and essential for the survival of the murine parasite at its blood

<sup>1</sup>Dynamique des Interactions Membranaires Normales et Pathologiques, UMR 5235, CNRS, Université de Montpellier, Montpellier, France. <sup>2</sup>Department of Applied Biotechnology and Food Science, Budapest University of Technology and Economics, Budapest, Hungary. <sup>3</sup>Institute of Enzymology, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Budapest, Hungary. <sup>4</sup>Doctoral School of Multidisciplinary Medical Science, University of Szeged, Szeged, Hungary. <sup>5</sup>CNRS UMR5048, Centre de Biochimie Structurale, Université de Montpellier, Montpellier, France. <sup>6</sup>INSERM U1054, Montpellier, France. <sup>7</sup>Present address: Institute for Research in Biomedicine, The Barcelona Institute of Science and Technology, Carrer de Baldiri Reixac 10, 08028, Barcelona, Spain. <sup>8</sup>Present address: Division of Structural Biology, University of Oxford, Roosevelt Drive, Oxford, OX37BN, United Kingdom. Correspondence and requests for materials should be addressed to R.C. (email: [rachel.cerdan@umontpellier.fr](mailto:rachel.cerdan@umontpellier.fr))



**Figure 1.** *PfCCT* domain organization and structure of *PfCCT*<sub>(581-775)</sub>. **(a)** Domain organization of the full length *PfCCT*, *PfCCT*<sub>(528-795)</sub> and *PfCCT*<sub>(581-775)</sub> catalytic domain constructs, lacking a lysine-rich loop between residues 720–737. N-Cap segment (N), Catalytic domains (C), membrane binding domains (M) and lysine-rich sequences (K) are indicated. **(b)** Cartoon representation of *PfCCT*<sub>(581-775)</sub> dimer (PDB:4ZCT). Nomenclature of secondary structures follows that of *RnCCT*<sup>20</sup> and *B. subtilis* GCT<sup>34,36</sup> in this order  $\beta 1$ -L1- $\alpha A$ - $\beta 2$ - $\alpha B$ -L2- $\alpha C$ -L3- $\beta 3$ -L4- $\alpha D$ - $\beta 4$ -L5- $\alpha L$ - $\beta 5$ -L6- $\alpha E$ . The N-terminal disordered part assigned by NMR is depicted as blue dashed line (see also Supplementary Fig. S2). The flexible loop L5 lacking a lineage-specific lysine-rich region (720–737) is indicated by a violet dashed line. **(c,d)** Close-up of dimer interface regions. Residues involved in inter-monomer contacts (dashed line) are shown as sticks. Primes indicate residues and secondary structures of the other monomer.

stage<sup>12,13</sup>. Partial inhibition of *PfCCT* is also part of the mechanism of action of antimalarial compounds targeting lipid biosynthetic pathways<sup>14,15</sup>. CCT transfers a cytidyl group from CTP to phosphocholine (ChoP) via a mechanism involving the sequential binding of both substrates to form a ternary complex<sup>16-19</sup>. This leads to the formation of CDP-choline (CDPCho) by the nucleophilic attack of the ChoP phosphomonoester on the  $\alpha$ -phosphate of CTP and the release of pyrophosphate (PP<sub>i</sub>) as an additional product<sup>19,20</sup>. CCTs are amphitropic enzymes that display a several-fold activity increase upon binding to phosphatidylcholine depleted membrane surfaces by relieving an auto-inhibitory contact between its catalytic and membrane binding domains<sup>21,22</sup>. Most eukaryotic CCTs contain a single catalytic and membrane binding domain and form dimers<sup>22</sup>. In contrast, the gene of *PfCCT* underwent a domain duplication event and thus encodes two catalytic and two membrane binding domains connected by a disordered interdomain segment<sup>16,23</sup> (Fig. 1a). The N-terminal (C1) and C-terminal (C2) catalytic domains share 91% sequence identity over 180 amino-acids. To date, only two high resolution crystal structures of *Rattus norvegicus* CCT (*RnCCT*) catalytic domain were determined, both in complex with CDPCho<sup>20,24</sup>. The first *RnCCT* crystal structure showed an  $\alpha/\beta$  Rossmann fold for the catalytic domain conserved in cytidyltransferase enzymes and allowed the identification of key enzyme interactions with CDPCho<sup>20</sup>. The second crystal structure of the *RnCCT* catalytic domain was solved in the presence of the amphipathic auto-inhibitory (AI) helix of the membrane binding domain<sup>24</sup>. These structural and additional biochemical studies<sup>25,26</sup> provided evidence that the regulatory mechanism of CCT involves a direct interaction between the C-terminal helix  $\alpha E$  of the catalytic domain and the AI helix. Many aspects of the CCT structure and regulatory mechanism have been uncovered recently<sup>21,22,27</sup>. However, 3D structures of the free and substrate-bound enzyme were still to be determined for the detailed understanding of the catalytic mechanism of CCT.

<i>PfCCT</i>	$k_{cat}$ (s <sup>-1</sup> )	$K_{M,CTP}$ (μM)	$K_{M,ChoP}$ (μM)
581–775	0.010 ± 0.001	1,100 ± 200	N.D.
528–795 <sup>a</sup>	1.45 ± 0.05	170 ± 20	1,950 ± 100
528–795 (K663A) <sup>b</sup>	0.0008 ± 0.0002	70 ± 20	N.D.
528–795 (Y626F/Q636A)	0.24 ± 0.01	170 ± 10	1,600 ± 100
528–795 (T761A)	0.0010 ± 0.0001	1,270 ± 90	3,700 ± 400
528–795 (T762A)	0.0010 ± 0.0001	610 ± 70	2,100 ± 200

**Table 1.** Kinetic parameters of *PfCCT* enzyme constructs. <sup>a</sup>Experimental data are from Nagy *et al.*<sup>16</sup>. <sup>b</sup>The kinetic constant of *PfCCT*<sub>(528–795)</sub> K663A refer to a phosphohydrolase activity (see Supplementary Fig. S4). Values are the mean (±SD) of at least 2 independent experiments. N.D.: not determined.

<i>PfCCT</i>	$K_d,CDP,Cho$ (μM)	$\Delta H_{CDP,Cho}$ (kcal/mol)	$-T\Delta S_{CDP,Cho}$ (kcal/mol)	$\Delta G_{CDP,Cho}$ (kcal/mol)	$K_d,CTP$ (μM)	$\Delta H_{CTP}$ (kcal/mol)	$-T\Delta S_{CTP}$ (kcal/mol)	$\Delta G_{CTP}$ (kcal/mol)
581–775	47 ± 5	-10.5 ± 0.3	4.7 ± 0.4	-5.8 ± 0.1	N.D.	N.D.	N.D.	N.D.
528–795	34 ± 1	-16.1 ± 0.1	10.1 ± 0.1	-6.0 ± 0.2	60 ± 2	-8.6 ± 0.3	3.0 ± 0.4	-5.7 ± 0.1
528–795 (K663A)	210 ± 60	-11.4 ± 1.5	6.5 ± 1.6	-5.0 ± 1.4	520 ± 50	6.1 ± 9.3	-10.5 ± 0.6	-4.4 ± 0.1
528–795 (Y626F/Q636A)	22.2 ± 0.2	-12.0 ± 0.2	5.8 ± 0.1	-6.2 ± 0.1	60 ± 2	-10.8 ± 0.1	5.1 ± 0.2	-5.7 ± 0.1
528–795 (T761A)	N.D.	N.D.	N.D.	N.D.	170 ± 35	-7.0 ± 0.8	1.9 ± 1.0	-5.1 ± 0.1
528–795 (T762A)	N.D.	N.D.	N.D.	N.D.	170 ± 12	-10.5 ± 0.7	5.5 ± 0.8	-5.1 ± 0.1

**Table 2.** Thermodynamic parameters of ligand binding to *PfCCT* enzyme constructs. Values are the mean (±SD) of at least 2 independent experiments, except for *PfCCT*<sub>(528–795)</sub> Y626F/Q636A (see Methods). N.D.: not determined.

Here, we present the crystal structures of the catalytic domain of the *P. falciparum* enzyme (*PfCCT*) in the free form, as well as in complex with its substrate ChoP or fragments (CMP and choline) and with the product CDPCho, thereby visualizing the free and the ChoP-complexed states of CCT for the first time. The set of high resolution structures allowed the identification of active site conformational changes along the catalytic reaction. Structural insights complemented with kinetic and ligand binding analyses elucidate the role of key active site residues in the *PfCCT* mechanism of action. The presented results allow detailed structural characterization of a potential antimalarial target of the phospholipid metabolism pathway of *Plasmodium* and may help for the design of specific inhibitors.

## Results

**Biochemical characterization of *PfCCT* catalytic domain construct.** In order to provide structural insights into the mechanism of action of *PfCCT*, we set out to crystallize the constitutively active C-terminal catalytic domain of *PfCCT*. Several constructs were tested and our attempts were successful only in case of *PfCCT*<sub>(581–775)</sub> that corresponds to the exact region visible in the crystal structure of *RnCCT* (PDB: 3HL4)<sup>20</sup> (Fig. 1a). While this construct proved to be highly crystallogenic, activity assays revealed its reduced catalytic rate together with an increased  $K_M$  for CTP as compared to the longer construct *PfCCT*<sub>(528–795)</sub> used in our previous studies<sup>16</sup> (Table 1 and Supplementary Fig. S1). We further assessed the functionality of *PfCCT*<sub>(581–775)</sub> by characterizing its substrate and product binding ability. Its affinity for CDPCho ( $K_d = 47 \mu\text{M}$ ) reported by isothermal titration calorimetry (ITC) displayed a moderate attenuation as compared to the one of *PfCCT*<sub>(528–795)</sub> ( $K_d = 34 \mu\text{M}$ ) (Table 2 and Supplementary Fig. S1). Saturation Transfer Difference Nuclear Magnetic Resonance (STD-NMR) experiments indicated STD signal of CTP, ChoP and CDPCho bound to *PfCCT*<sub>(581–775)</sub> at 1 mM, 4 mM and 1 mM concentrations, respectively (Supplementary Fig. S1). This latter result is consistent with previously observed unusually low ChoP affinity of *PfCCT*<sup>16</sup>. We concluded that *PfCCT*<sub>(581–775)</sub> is compatible with substrate and product recognition and may be amenable to structural studies provided that its compromised enzymatic function is taken into account in structure-based interpretations. It is likely that the drop of *PfCCT*<sub>(581–775)</sub> enzyme activity is due to the elimination of the C-terminal flanking part as similarly located truncations in the *RnCCT* catalytic domain also led to severe loss of activity<sup>28,29</sup>. After unsuccessful efforts to crystallize longer *PfCCT* catalytic domain constructs with an extended C-terminal end, we continued crystallization trials with *PfCCT*<sub>(581–775)</sub> and obtained single crystals of the free enzyme and in presence of the enzymatically relevant ligands ChoP, CMP and CDPCho as well as Cho.

**Structure of the free *PfCCT* C-terminal catalytic domain.** The structure of *PfCCT*<sub>(581–775)</sub> was determined at 2.2 Å resolution by molecular replacement using *RnCCT* structure<sup>20</sup> (PDB: 3HL4) as a search model (Table 3). The enzyme forms a globular dimer with an  $\alpha/\beta$  Rossmann fold that is well conserved in CCTs and other members of the cytidyltransferase enzyme family (Fig. 1b). Here, the homodimer state likely mimics the full-length *PfCCT* conformation with two almost identical catalytic sites. The dimerization interface of *PfCCT*

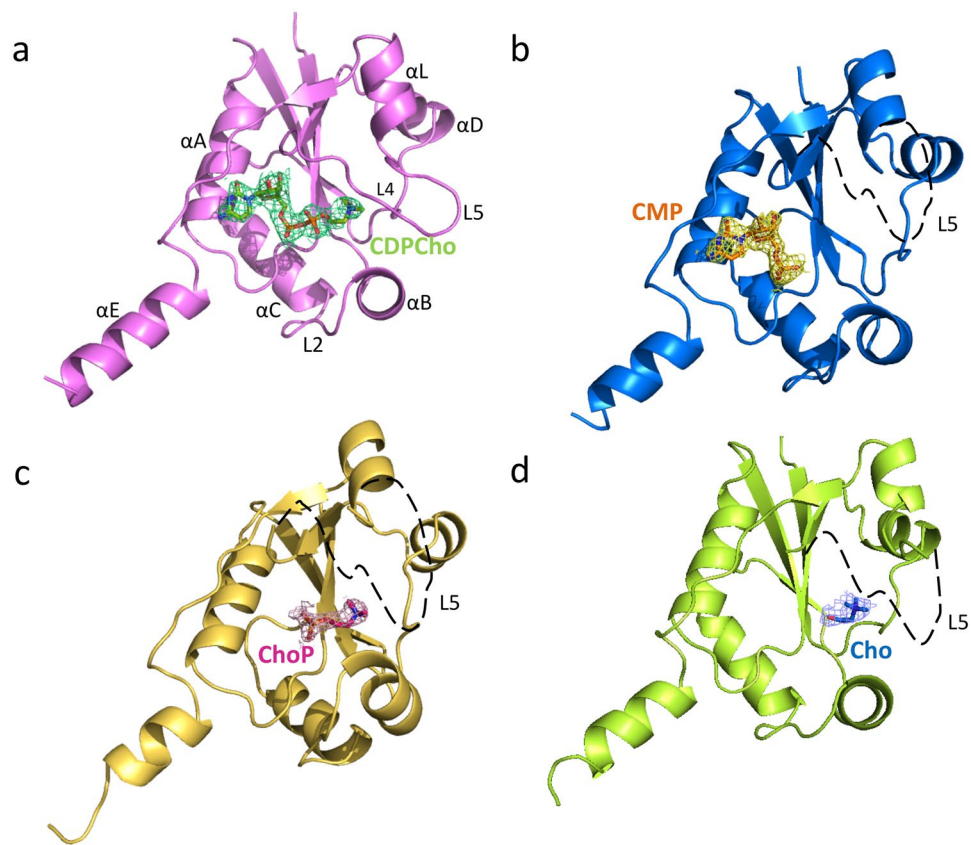
	<i>Pf</i> CCT (PDB: 4ZCT)	<i>Pf</i> CCT-CMP (PDB: 4ZCP)	<i>Pf</i> CCT-Cho (PDB: 4ZCQ)	<i>Pf</i> CCT-ChoP (PDB: 4ZCR)	<i>Pf</i> CCT-CDPCho (PDB: 4ZCS)
<b>Data collection</b>					
Space group	I222	I222	I222	I222	I222
Cell dimensions					
<i>a</i> , <i>b</i> , <i>c</i> (Å)	48.5, 74.4, 119.0	50.5, 69.3, 116.4	50.6, 69.4, 119.0	50.6, 69.6, 117.9	115.5, 149.8, 176.6
$\alpha$ , $\beta$ , $\gamma$ (°)	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90	90, 90, 90
Resolution (Å) <sup>a</sup>	63.02–2.22 (2.44–2.22)	58.22–1.98 (2.03–1.98)	59.92–1.92 (1.99–1.92)	46.52–1.80 (1.85–1.80)	113.7–2.45 (2.51–2.45)
$R_{\text{sym}}$ or $R_{\text{merge}}$	0.087 (0.497)	0.059 (0.307)	0.038 (0.511)	0.054 (0.497)	0.032 (0.377)
$I/\sigma I$	10.7 (3.0)	10.5 (2.8)	15.0 (1.8)	14.3 (1.8)	10.1 (2.5)
CC1/2	0.995 (0.879)	0.999 (0.896)	0.999 (0.737)	0.999 (0.718)	0.997 (0.859)
Completeness (%)	97.9 (97.9)	99.7 (99.7)	99.2 (99.2)	99.6 (99.6)	99.9 (99.6)
Redundancy	4.0 (3.9)	4.6 (4.6)	4.0 (4.0)	6.4 (6.4)	9.3 (9.4)
B-Wilson	17.84	34.40	44.29	39.00	40.80
<b>Refinement</b>					
Resolution (Å)	2.22	1.98	1.92	1.80	2.45
No. reflections	10780	13840	16280	18719	52521
$R_{\text{work}}/R_{\text{free}}$	20.21/24.26	20.92/24.71	20.40/22.60	17.89/23.22	18.06/23.41
No. atoms					
Protein	1120	1023	1042	1050	6810
Ligands		21	7	11	186
PEG					38
Water	98	43	63	91	347
B-factors					
Protein	24.3	41.6	50.4	47.1	45.0
Ligands		38.9	69.3	62.4 <sup>b</sup>	34.1
PEG					77.1
Water	28.3	43.7	51.2	53.9	40.8
R.m.s. deviations					
Bond lengths (Å)	0.007	0.008	0.008	0.009	0.018
Bond angles (°)	1.003	1.104	0.965	1.218	1.970
Ramachandran plot					
Most favored regions (%)	97.58	96.80	98.41	99.18	96.37
Allowed regions (%)	2.42	3.20	1.59	0.82	2.91

**Table 3.** Crystallographic data and refinement statistics. <sup>a</sup>Values in parentheses are for the highest resolution shell. <sup>b</sup>The occupancy of ChoP is 50%.

is mediated by inter-molecular contacts between several structural elements of each chain. Residues of helix  $\alpha C$  of one chain are in contact with residues of loop L3', helix  $\alpha A'$  and helix  $\alpha E'$  of the other chain (Fig. 1c,d). In addition, residues of helix  $\alpha E$  contact residues of helix  $\alpha E'$  and loop L2'. Intermolecular contacts between L3 loop residues of both chains include the <sup>681</sup>RWVDE<sup>685</sup> dimerization motif that is conserved in all cytidylyltransferases. We observed a lack of density in loop L5 for nine residues (<sup>712</sup>IPYANNQK<sup>719E</sup><sup>738</sup>) in a region where a lysine-rich insertion (720–737) had been removed from the construct<sup>16</sup>. The 35 N-terminal amino acids of *Pf*CCT<sub>(581–775)</sub> were also not visible on electron density maps. This part corresponds to the N-terminal pre-catalytic region of *Rn*CCT containing two short helices that constitute an 'N-cap' by folding on the subsequent conserved catalytic core of the enzyme<sup>20,24</sup> (Fig. 1a). To obtain structural information on this N-terminal region, we performed <sup>1</sup>H-<sup>1</sup>H 2D NMR experiments on *Pf*CCT<sub>(581–775)</sub>. The resonances of 33 N-terminal residues were clearly detectable and could be assigned on the <sup>1</sup>H-<sup>1</sup>H 2D NOESY spectrum (Supplementary Fig. S2), arguing for their likely disorder. Few  $\alpha$ -helix specific crosspeaks were found for the <sup>604</sup>DSEKNKG<sup>610</sup> segment (Supplementary Fig. S2).

**Co-crystal structures of CMP-, Cho-, ChoP- and CDPCho-bound *Pf*CCT.** The 3D structures of *Pf*CCT<sub>(581–775)</sub> in complex with CMP, Cho, ChoP or CDPCho (Table 3) possess almost identical overall folds as the free protein (Fig. 2 and Supplementary Table S1). In all structures, the ligands were found in the active site. The occupancy of ChoP was only 50% even when the crystal was obtained after soaking in very high (50 mM) final concentration of ChoP, which is in accordance to the very low affinity described above. Minor conformational changes were observed in loops L2 and L4 between strand  $\beta 3$  and helix  $\alpha D$ . The major change concerned loop L5 (residues 712–738) that is only visible in the presence of CDPCho (Fig. 2a). In this case, ordered conformation of the loop is assisted by the interaction of Y714 with the ligand which is apparently not strong enough with ChoP or choline to render loop L5 ordered (Fig. 3). The nucleotide ligands CMP and CDPCho are anchored in the active site by several polar contacts, hydrogen bonds and a cation- $\pi$  interaction of R755 with the cytosine ring (Fig. 3



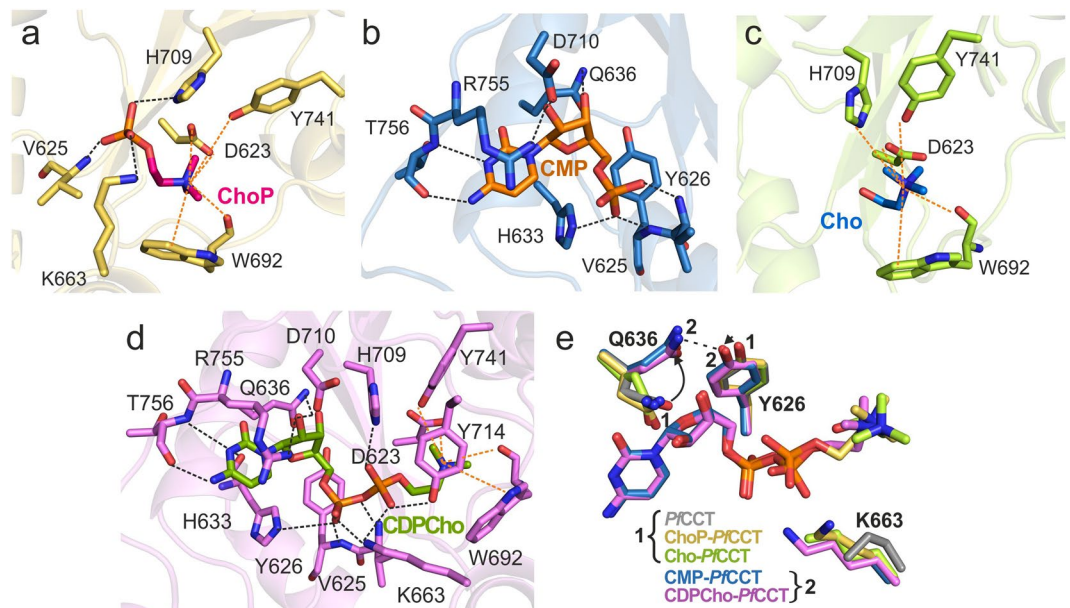


**Figure 2.** X-ray co-structures of CDPCho-*PfCCT*, CMP-*PfCCT*, ChoP-*PfCCT* and Cho-*PfCCT* monomers. Complexes with (a) CDPCho (PDB: 4ZCS) at 2.45 Å resolution, (b) CMP (PDB: 4ZCP) at 1.98 Å resolution, (c) ChoP (PDB: 4ZCR) at 1.80 Å resolution and (d) Cho (PDB: 4ZCQ) at 1.92 Å resolution. In all cases only one monomer of a dimer is presented. The ligands (CDPCho, CMP, ChoP and Cho) are shown in sticks with electron density around them. The  $2F_o - F_c$  electron density omit maps were contoured at 1.0 sigma around the respective ligand within 1.6 Å of the selected atoms. The maps were created using *fft* function from CCP4 software<sup>49</sup>. The dashed black line in (b–d) represents loop L5 which is not visible for these co-structures.

and Supplementary Table S2). In contrast, the choline subsite that accommodates Cho, ChoP and the choline moiety of CDPCho provides ligand recognition with charged as well as cation- $\pi$  interactions forming a composite aromatic box<sup>30</sup>. The coordination of the nucleophile phosphate moiety in ChoP- and CDPCho-liganded structures is ensured by V625, K663 and H709 residues. Analysis of these different liganded *PfCCT* structures demonstrates conformational changes for a set of active site residues responsible for ligand accommodation, including D623, Y626, Q636, K663, W692, H709, Y714, I740, Y741 and R755.

### Deciphering the role of K663 in catalysis by structural, kinetic and thermodynamic analyses.

K663 is a special active site residue that displays conformational variety in different liganded complexes (Fig. 3e). A catalytically essential role of its equivalent residue in *RnCCT* (K122) had already been demonstrated using molecular dynamics (MD) simulations<sup>24</sup> and mutagenesis studies<sup>31</sup>. Moreover, it was suggested that impeding its conformational dynamics has a major role in the enzyme regulatory mechanism<sup>24</sup>. Here, we found that in the free and the CMP-bound *PfCCT* structures, the reliable positioning of the K663 sidechain is precluded due to the weak electron density, suggesting its high flexibility or disorder (Supplementary Fig. S3). The presence of Cho in the active site at least partially stabilizes the K663 side chain, while ChoP or CDPCho accommodation renders this residue fully visualized. The K663 terminal amino moiety points towards the phosphate group of ChoP or both phosphate moieties of CDPCho, in somewhat different conformations (Fig. 3e and Supplementary Fig. S3). These structural insights showing variable detection and interactions of K663 prompted us to an in-depth definition of its role throughout the course of the catalytic mechanism. We interrogated the functionality of this residue by site directed mutagenesis followed by biochemical analysis. Exchange of K663 to alanine rendered *PfCCT*<sub>(528–795)</sub> practically inactive (Table 1). The enzymatic activity was monitored by a coupled pyrophosphatase (PPase) assay and consisted of two consecutive steps: the cleavage of the byproduct pyrophosphate into two phosphates and the addition of phosphates to a nucleotide derivative, resulting a chromophore substance. Surprisingly, we observed a remnant enzymatic activity upon *PfCCT* K663A enzyme addition, even in absence of both the ChoP substrate and the PPase auxiliary enzyme (Supplementary Fig. S4). The accumulation of free phosphate in the absence of the substrate, ChoP, argued for a CTP phosphohydrolase (CTP  $\rightarrow$  CDP + P<sub>i</sub>) activity of the K663A mutant (Supplementary Fig. S4). Thus in a straightforward manner, we could conveniently distinguish

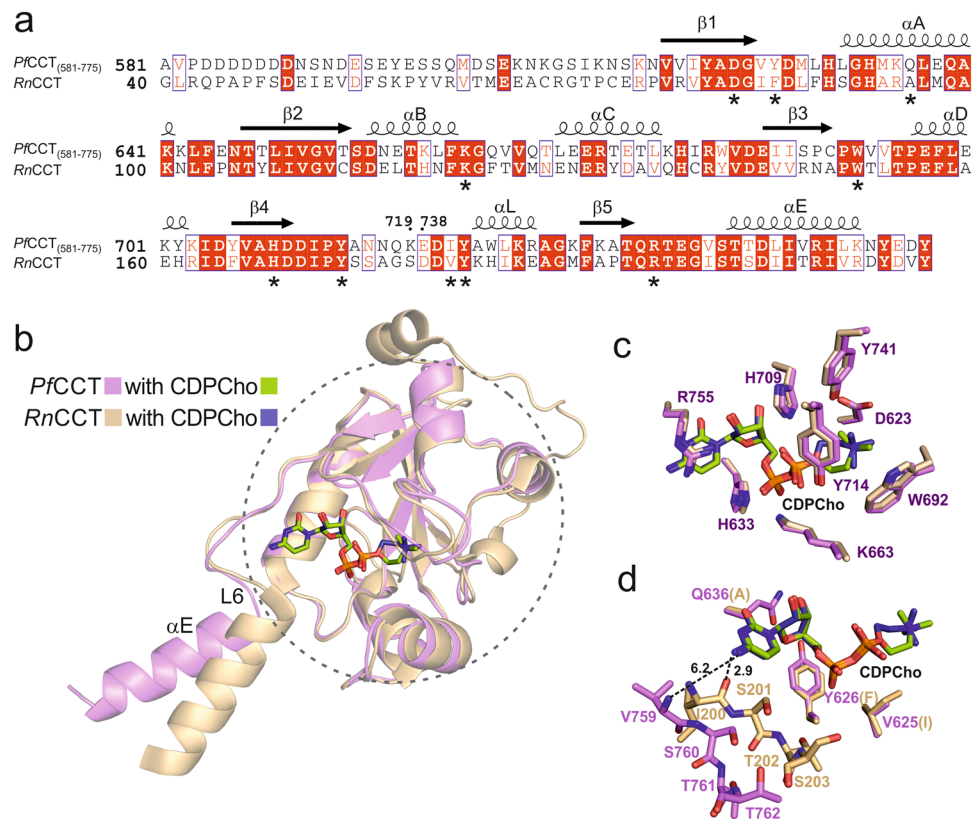


**Figure 3.** Ligand interactions and conformational changes of *PfCCT*<sub>(581–775)</sub> upon ligand binding. Close-up views of the substrate/product binding sites of *PfCCT*. (a) CDPCho-*PfCCT* structure, (b) CMP-*PfCCT* structure, (c) ChoP-*PfCCT* structure and (d) Cho-*PfCCT* structure. Ligand interaction residues are labelled and shown in stick representation. Interactions with the nucleotide and choline moieties of ligands are represented by black and orange dashed lines, respectively. (e) Conformational changes of active site residues Y626, Q636 and K663 in different ligand complexes of *PfCCT*<sub>(581–775)</sub>. Binding of nucleotide ligands induces a shift of Q636 and Y626 from state 1 to state 2. The K663 residue sidechain is not rendered fully visible in the free enzyme as well as in Cho and CMP complexes.

the cytidyltransferase from the CTP phosphohydrolase activities in our coupled enzyme assay for *PfCCT*<sub>(528–795)</sub> and its K663A mutant, respectively. For reliable characterization of ligand interactions of this lysine, we measured the equilibrium ligand binding of CDPCho and CTP to *PfCCT*<sub>(528–795)</sub> K663A (Table 2 and Supplementary Fig. S4). CDPCho binding affinity of the mutant suffered a four- to six-fold attenuation as evidenced by ITC measurements. An even more drastic effect was observed for CTP binding to K663A mutant where the reduced binding affinity was coupled with a fundamental change in the energetics of binding as compared to wild-type *PfCCT*<sub>(528–795)</sub>. The observed endothermic character of binding ( $\Delta H_{CTP} > 0$ ) caused an entropy-favored binding mechanism that opposes the definitely exothermic interaction with the wild-type enzyme. This change can reflect either the increased conformational entropy due to a poorly defined binding conformation of the triphosphate of CTP or to a lesser extent, the solvent release upon binding to the mutated active site. Collectively, K663 appears to facilitate multiple steps in the catalytic mechanism. It has a decisive role in ChoP coordination as evidenced by its structural contacts (Supplementary Fig. S3 and Table S2) and by its absolute requirement for the classical CCT activity with ChoP as substrate. It also substantially contributes to the efficiency of CTP binding to the free enzyme.

### Comparison of *PfCCT* and *RnCCT* structures and functional characterization of residue alterations.

The *PfCCT* and *RnCCT* catalytic domains share 44% sequence identity<sup>23</sup> and high structural similarity (Fig. 4a,b). This especially holds for the core region of the conserved  $\alpha/\beta$  Rossmann fold indicated by a dashed circle in Fig. 4b, that hosts the inner part of the active site accommodating CDPCho. This region, defined by residues 618–760 in *PfCCT* displays an all-atom RMSD value of 0.73 Å with the respective region of the *RnCCT* structure complexed with CDPCho<sup>20</sup>. CDPCho is found in the same zig-zag position in both *PfCCT* and *RnCCT* co-structures with only the two methylene groups of this ligand adopting different conformations. The interaction network of CDPCho is also largely similar (Fig. 4c). The differences in the catalytic site concern the residues V625, Y626, Q636 and V759 of *PfCCT* corresponding to hydrophobic residues I84, F85, A95 and I200 of *RnCCT*, respectively (Fig. 4d). A notable structural difference involves Q636 sidechain that enables a direct interaction with the ribose 3'OH moiety in *PfCCT*, whereas a similar contact in *RnCCT* is established by an ordered water molecule next to A95. Y626 and Q636 constitute the wall of the active site cavity around the hole for the ribose moiety of the ligands. Both residues display conformational changes during the course of catalysis (Fig. 3e). The position corresponding to Y626 is strictly reserved for aromatic (Y/F) residues in the HxGH nucleotidyltransferase enzyme superfamily because the edge-on orientation of the aromatic sidechain secures the parallel conformation and hydrogen bond contacts of the two histidines from the HxGH motif, critical for their proper catalytic function<sup>32</sup> (Supplementary Fig. S5). In a nucleotide-bound form, Q636 forms a hydrogen bridge to Y626, enabled by the phenolic OH group of the latter that is additionally present as compared to its *RnCCT* surrogate F85. In order to evaluate the impact of these residue changes, we engineered a Y626F/Q636A double mutant in *PfCCT*<sub>(528–795)</sub> to



**Figure 4.** Structural comparison of *PfCCT* and *RnCCT* catalytic domain in complex with CDPCho. **(a)** Sequence alignment of *PfCCT*<sub>(581–775)</sub> and the corresponding region of *RnCCT*. Strictly conserved residues are indicated as white letters on red background and similar residues as red letters on white background. The secondary structure elements are added above the sequence. We also noted by a star all residues mentioned in the text, including active site residues and non-conserved residues. Residue numbers are noted where the lysine-rich specific loop (720–737) has been removed. The layout was created by ESPript 3.0<sup>57</sup>. **(b)** Structural alignment of *PfCCT* (violet) and *RnCCT* (PDB: 3HL4) (light orange) monomers complexed with CDPCho. CDPCho is coloured by atoms with green and deep blue carbons in *PfCCT* and *RnCCT* complex structures, respectively. Dashed circle designates the core region which is highly similar in *PfCCT* and *RnCCT* structures. **(c)** Close-up view of the CDPCho interaction network. **(d)** Superposition of 3 non-conserved active site residues between *PfCCT* and *RnCCT* and displacement of the signature sequence fragment (*PfCCT*:<sup>759</sup>VSTT<sup>762</sup>; *RnCCT*:<sup>200</sup>ISTS<sup>203</sup>) located at the border of loop L6 and helix  $\alpha$ E. Residues V625, Y626 and Q636 in *PfCCT* correspond to I84, F85 and A95 in *RnCCT*, respectively. Distances between atoms showed by dashed line are in Å.

mimic the *RnCCT* enzyme. This construct displayed a six-fold attenuated catalytic rate together with an unaltered  $K_{M,CTP}$  value compared to *PfCCT*<sub>(528–795)</sub> (Table 1). Thermodynamic analysis revealed unaltered CTP affinity of the mutant with a slightly increased favorable binding enthalpy counterbalanced by a larger entropic penalty (Table 2). A modest effect of the residue substitutions was uncovered by the analysis of CDPCho binding. A somewhat tighter binding of CDPCho was observed in the case of the *RnCCT*-mimicking construct *PfCCT*<sub>(528–795)</sub> Y626F/Q636A, accompanied by a decrease of both the favorable enthalpy and the unfavorable entropy components (Table 2). The differences observed for CDPCho binding to *PfCCT* or to the *RnCCT*-mimicking construct may be interpreted by a slightly more constrained CDPCho interaction in the wild-type *PfCCT* enzyme. This is also supported by the reduced flexibility of the nucleotide binding pocket in the CDPCho-*PfCCT* structure in comparison to the CDPCho-*RnCCT* structure (PDB: 3HL4), as suggested by *B*-factor values.

**Conformational displacement at the PP<sub>i</sub> binding region contributes to low activity of *PfCCT*<sub>(581–775)</sub>.** In contrast to the core region, the peripheral segments of the *PfCCT* catalytic domain display considerable structural changes compared to *RnCCT*. These include the N-terminal flexible segment of *PfCCT* and the C-terminal protruding part consisting of loop L6 and helix  $\alpha$ E (Fig. 4b). This C-terminal region contributes to the active site assembly by providing contacts between CTP and the strictly conserved CCT signature sequence RTEG(I/V)ST(S/T) (Supplementary Fig. S6). The terminal two residues of the signature sequence constitute the N-terminus of the helix  $\alpha$ E and are expected to coordinate PP<sub>i</sub> at the bottom of the active site<sup>20</sup>. The displacement of this functionally relevant part leads to the loss of contact between V759 main chain and the cytosine ring and to a ~5 Å displacement of T761 and T762 compared to what is observed in *RnCCT* (Fig. 4d). To probe the functional role of these threonine residues, we designed two T/A single point mutants in *PfCCT*<sub>(528–795)</sub> and characterized changes



in enzyme kinetics and CTP binding. We found that these alterations rendered the enzyme practically inactive, coupled with a substantial increase of  $K_{M,CTP}$  values from 170  $\mu\text{M}$  for *PfCCT*<sub>(528–795)</sub> to 1,270  $\mu\text{M}$  and 610  $\mu\text{M}$  for *PfCCT*<sub>(528–795)</sub> T761A and *PfCCT*<sub>(528–795)</sub> T762A, respectively (Table 1). Nevertheless, analysis of CTP binding to these mutants followed by ITC reported a less substantial, three-fold attenuation effect on binding affinity ( $K_{d,CTP}$  values in Table 2). Collectively, kinetic analysis and ITC results suggest that while these threonine residues possibly contact CTP and  $\text{PP}_i$ , they have a more prominent contribution to catalysis than to ligand binding.

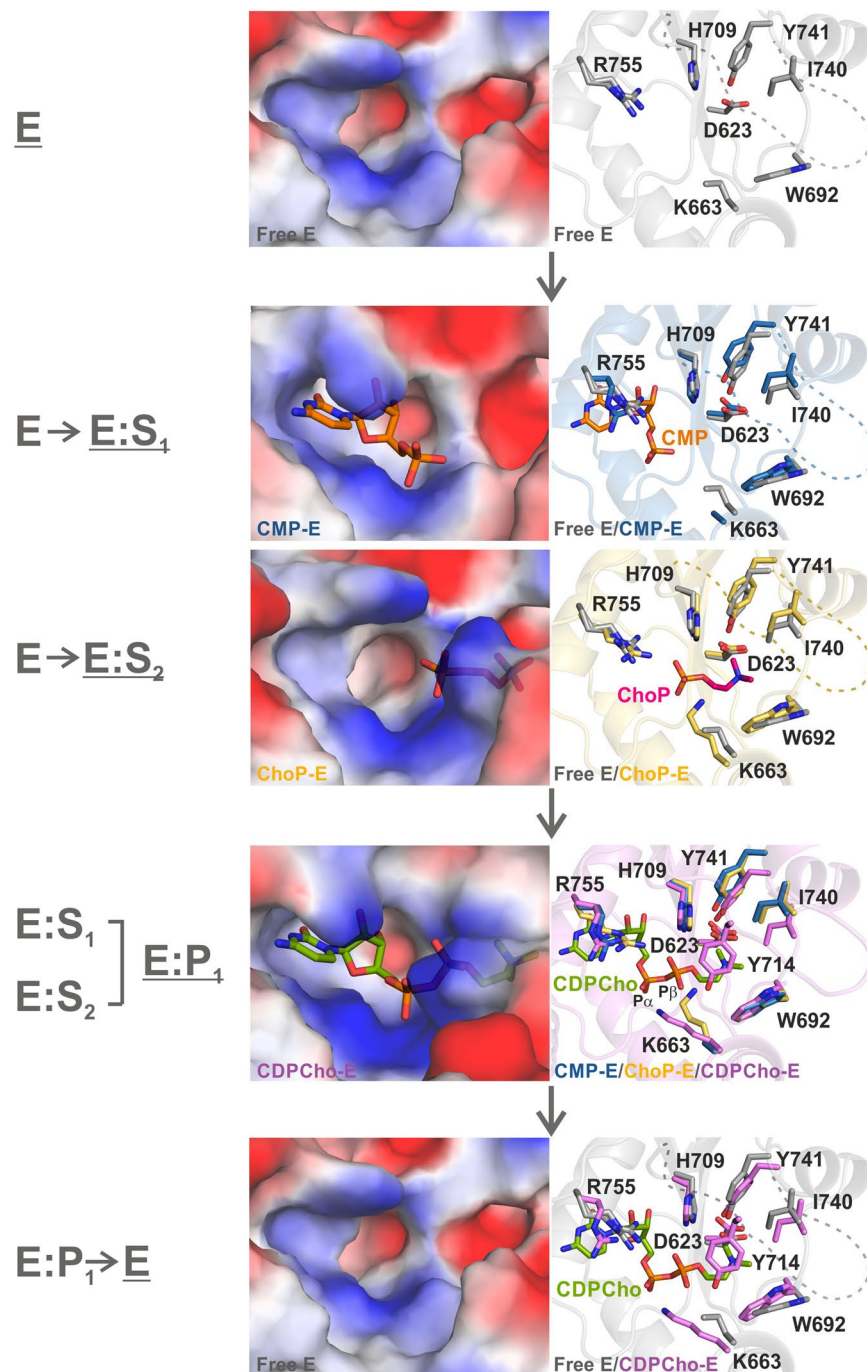
## Discussion

Prior to our study, only two high resolution structures of CCT catalytic domain and six structures of members of the Rossmann-folded cytidylyltransferase enzyme family were available from the Protein Data Bank. These include ECT (CTP:phosphoethanolamine cytidylyltransferase) and GCT (CTP:glycerol-3-phosphate cytidylyltransferase) catalytic domain structures that together with CCT share the conserved fold architecture as well as some key active site residues comprising the RTEG(I/V)ST(S/T) and HxGH signature sequences. All these enzymes are specialized in the cytidylyl transfer reaction with a hallmark of exploiting basic residues for substrate coordination and catalysis<sup>31,33–35</sup>. The herein presented structural data complemented with biochemical characterization provide an unprecedented step-by-step view of the enzymatic cycle of *PfCCT*. The structural snapshots of subsequent steps of the catalytic mechanism readily identify steric fluctuations and conformational changes of individual residues at the choline and nucleotide subsites (Fig. 5). A limitation of the structural set originates from the C-terminal truncation of the catalytic domain with the removal of a part of the helix  $\alpha\text{E}$  and the subsequent linker segment. It is very likely that this engineered modification is responsible for the altered orientation of the helix  $\alpha\text{E}$  compared to CCT, ECT and GCT structures which leads to the detachment of its N-terminal T761 and T762 residues from the vicinity of the active site. We hypothesize that the apparent displacement of these catalytically important threonines may contribute to the low activity of the shortened *PfCCT*<sub>(581–775)</sub> construct used for crystallographic studies. We suspect that the helix  $\alpha\text{E}$  in *PfCCT*<sub>(581–775)</sub> possesses significant conformational freedom in solution and that its observed conformation in the crystal structures is stabilized by crystal packing with local  $\alpha\text{E}$  lattice contacts forming  $\alpha\text{E}$  homotopic tetramers from different monomers (not shown).

The present structural set together with previously obtained data offers a stepwise understanding of the CCT catalytic mechanism (Fig. 5). Binding of CMP, a fragment of CTP substrate, to the free enzyme active site induces a major conformational change of R755 with a  $\sim 4 \text{ \AA}$  displacement to provide a cation- $\pi$  stacking interaction with the cytosine group. A similar nucleotide ligand induced movement of the corresponding arginine has previously been observed by MD simulations for *RnCCT*<sup>24</sup> and is also seen in ECT and GCT co-structures. In case of *PfCCT*, a flip of the non-conserved Q636 sidechain additionally provides space for the accommodation of the ribose. Notably, its alanine counterpart in *RnCCT* lacks the ability for this sidechain flip. From a drug design perspective, targeting Q636 might enhance the selectivity to *PfCCT* versus human CCT of a future chemical hit compound. The burying of the cytosine part is observed in all available cytidylyltransferase structures complexed with nucleotide ligands and is expected to contribute to the favorable binding enthalpy. ChoP binding to the choline subsite is accomplished by a slight inward movement of I740 and Y741 sidechains. Key residue contacts to ChoP that emerge from these movements involve the orientation of K663 towards the phosphate moiety and of W692 to provide cation- $\pi$  interaction to the trimethylammonium moiety (Fig. 5). Loop L5 is not visible at this stage and is seen ordered only in the CDPCho product-bound form. The conformational fluctuations of this segment may constitute a side-entry of the active site that enables access of ChoP. While the disorder to order transition of loop L5 is first observed in the presented *PfCCT* structures, the flexibility of this loop was also evidenced in *RnCCT* by the comparatively high local *B*-factors as well as by fluctuations observed in MD simulations with or without CDPCho<sup>24</sup>. Intriguingly, loop L5 hardly establishes contacts to other structural elements, thus its anchoring to the active site is presumably fostered by the cation- $\pi$  and polar interactions of Y714 with CDPCho (Fig. 5). Closure of the choline subsite in the CDPCho-bound state is also accompanied by the further inward movement of I740 and Y741. Meanwhile, K663 moves to coordinate the  $\text{P}_\alpha$  group in addition to the  $\text{P}_\beta$  group of CDPCho that originated from ChoP, allowing the phenolic OH of Y714 to coordinate  $\text{P}_\beta$  moiety (Fig. 5). K663 may also be involved in the formation of a productive  $\text{ES}_1\text{S}_2$  ternary complex and in the catalytic step. Its position and function during the transition state remains to be resolved. It may either coordinate nucleophile phosphate of ChoP or provide a bridge between the two phosphate moieties directly involved in catalysis, a conformation apparent in CDPCho liganded *PfCCT* and *RnCCT* structures and reminiscent of the position of K46 in GCT<sup>34</sup>. So far no CCT structure complexed to CTP is available. Structural insights of related GCT:CTP<sup>36</sup> and ECT:CTP (PDB: 4XSV) complexes reveal that the serine and threonine residues corresponding to T761 and T762 of *PfCCT* participate in the coordination of CTP phosphate moieties. The mutation into alanine of the threonine residue of *BsGCT*, equivalent to *PfCCT* T761, results in a pronounced decrease of the enzyme activity<sup>37</sup>. Our biochemical results indicate the importance of these threonine residues in CTP binding and catalysis. Interestingly, MD simulations show that the binding of the AI helix represses the dynamics of the Lys122 of *RnCCT* (Lys663 of *PfCCT*) but also of helix  $\alpha\text{E}$  which includes serine/threonine counterparts of T761 and T762<sup>24,27</sup>.

In conclusion, we provide novel structural insights into the *PfCCT*-catalyzed enzymatic steps that are supported by biochemical characterization of key elements in CCT action. A detailed knowledge at the molecular level of the catalytic mechanism of *PfCCT* constitutes the first step to the rational identification of efficient and specific inhibitors. The PC biosynthesis pathway is already validated as an antimalarial target at the parasite blood-stage<sup>15,38</sup>. Recently, the activity of this pathway has also been directly linked to *P. falciparum* transmission-stage formation<sup>11</sup>. Inhibiting *PfCCT*, the key enzyme of PC pathway, should therefore have a strong impact at multiple stages of the parasite development.





**Figure 5.** Structural views of *PfCCT* active site at the different enzymatic steps. Electronic surface representations of the binding pockets (left) reveal polarity and cavity size changes along the enzymatic cycle. Close-up views of the active site (right) show key residues superimposed from two consecutive steps of the catalytic mechanism revealing the conformational changes upon substrate/product binding. Protein fold is shown in the background as cartoon for the later step. Residues from the free enzyme (E) are in grey, from the CMP-bound enzyme in blue, from the ChoP-bound enzyme in yellow and from the CDPCho-bound enzyme in violet. The ligands are colored as followed: CMP, orange (S<sub>1</sub>); ChoP, dark pink (S<sub>2</sub>) and CDPCho green (P<sub>1</sub>). In the free enzyme, R755 is found in two alternative conformations. Note that instead of the proper substrate CTP, CMP is present in the nucleotide subsite in E:S<sub>1</sub> state. The  $\alpha$  and  $\beta$  phosphate moieties of CDPCho are labelled on the relevant panel.

## Methods

**Cloning, bacterial expression and purification of CCT constructs.** The *PfCCT* cDNA sequence (PlasmoDB: PF3D7\_1316600) was codon-optimized for expression in *E. coli* (GenScript). The *PfCCT*<sub>(581–775)</sub> construct was obtained using previously described *PfCCT*<sub>(528–795)</sub> lacking the lysine-rich *Plasmodium* specific loop (720–737)<sup>16</sup>. The DNA fragment was cloned into the expression vector pET15bTEV (modified from pET15b,

Novagen) using the NdeI/BamHI sites in order to produce N-terminal  $6 \times$  His-tagged protein where the  $6 \times$  His-tag is cleavable by tobacco etch virus (TEV) protease. The Y626F/Q636A double mutant, the T761A, T762A and the K663A mutant constructs of His-tagged *PfCCT*<sub>(528–795)</sub> were produced by site-directed mutagenesis using the QuikChange method (Agilent). Primers were synthesized by Eurofins MWG GmbH. Primer sequences are given in Supplementary Table S3. *E. coli* BL21 (DE3) cells transformed with the plasmid were grown in Luria-Bertani (LB) broth containing  $100 \text{ mg} \cdot \text{ml}^{-1}$  ampicillin at  $37^\circ\text{C}$  to  $\text{OD} = 0.6$ . Protein expression was induced by  $0.5 \text{ mM}$  IPTG and bacteria were continuously cultured for 24 h at  $16^\circ\text{C}$ . Cells were harvested and re-suspended in lysis buffer containing  $20 \text{ mM}$  Tris/HCl pH 7.5,  $0.15 \text{ M}$  NaCl and  $2 \text{ mM}$  ethanethiol (EtSH). Supernatant containing protein was obtained by lysing cells using a French press, followed by centrifugation at  $40\,000 \text{ g}$  for 1 h. The supernatant was loaded onto a HisTrap HP column (GE Healthcare) pre-equilibrated with lysis buffer. His-*PfCCT*<sub>(581–775)</sub> was eluted with  $150 \text{ mM}$  and  $250 \text{ mM}$  imidazole in lysis buffer. Removal of the His<sub>6</sub>-tag was performed after the affinity column by incubation with recombinant His<sub>6</sub>-TEV protease for one night at room temperature (protease:protein ratio of 1:100 w-w). After concentration, 1 ml of protein sample was further purified by gel filtration using a HiLoad 16/60 Superdex 200 pg (GE Healthcare) in lysis buffer. Purification of His-tagged *PfCCT*<sub>(528–795)</sub> mutant constructs is described in Nagy *et al.*<sup>16</sup>.

**Enzyme kinetics.** Enzyme activity measurements were performed as described previously, using a continuous coupled colorimetric assay with pyrophosphatase (PPase), purine nucleoside phosphorylase (PNP) auxiliary enzymes and the auxiliary substrate 7-methyl-6-thioguanosine (MESG)<sup>16,30,39</sup>. For CTP substrate titrations, CTP concentration was varied between  $100 \mu\text{M}$  and  $2 \text{ mM}$  while ChoP concentration was kept constant at  $5 \text{ mM}$ . For ChoP titrations, ChoP concentration was varied between  $0.5 \text{ mM}$  and  $20 \text{ mM}$  while CTP concentration was kept constant at  $1 \text{ mM}$ . Initial velocities were calculated from the first 10% of the progress curves using phosphate calibration as reference<sup>16</sup>. Due to the low activity of *PfCCT*<sub>(581–775)</sub> and *PfCCT*<sub>(528–795)</sub> T761A, T762A variants, the enzyme was used at  $10 \mu\text{M}$  concentration and the slope of the absorbance change was determined after monitoring the reaction for 20 minutes. For measurements with *PfCCT*<sub>(528–795)</sub> K663A the following reagent concentrations were used in the enzyme assay:  $[\text{PPase}] = 0.17 \text{ U} \cdot \text{ml}^{-1}$ ,  $[\text{PNP}] = 1.25 \text{ U} \cdot \text{ml}^{-1}$ ,  $[\text{MESG}] = 0.1 \text{ mM}$ ,  $[\text{Mg}^{2+}] = 5 \text{ mM}$  and  $[\text{P}f\text{CCT}_{(528–795)} \text{ K663A}] = 10 \mu\text{M}$ , to yield adequate signal-to-noise ratio ( $\text{S/N} > 10$ ). Three independent experiments were performed for CTP and ChoP kinetic titrations of *PfCCT*<sub>(528–795)</sub> K663A. Global fit of CTP and ChoP kinetic titrations for each enzyme variant was performed in Origin 8.6 using Michaelis–Menten equation, to yield a single  $k_{\text{cat}}$  parameter.

**ITC measurements.** Calorimetric measurements were performed on an ITC 200 titration calorimeter (Malvern) at  $20^\circ\text{C}$ . The protein samples were dialyzed against  $20 \text{ mM}$  HEPES/NaOH pH 7.5,  $100 \text{ mM}$  NaCl, including  $0.5 \text{ mM}$  Tris(2-carboxyethyl)phosphine (TCEP) as reducing agent. CTP and CDPCho were diluted from a concentrated stock into the dialysis buffer with subsequent pH adjustment. Concentration of the titrating nucleotide ligands were determined using molar extinction coefficients  $9000 \text{ M}^{-1} \text{ cm}^{-1}$  at  $271 \text{ nm}$  for CTP and  $5000 \text{ M}^{-1} \text{ cm}^{-1}$  at  $260 \text{ nm}$  for CDPCho<sup>16</sup>. In the experimental setup, the cell of the instrument was filled with protein and the syringe with the ligand. The titrations typically included 19 steps of injection with  $2 \mu\text{l}$  of ligand per injection spaced 180 s apart from each other, with the injection syringe rotating at  $750 \text{ r.p.m.}$  The data were analyzed using MICROCAL ORIGIN software following the directions of the manufacturer. To account for dilution, the invariant integrated heat value of the last 3–5 data points was subtracted from all data points. The one set of independent sites binding model was applied to data for determination of thermodynamic parameters: dissociation constant ( $K_{\text{d}}$ ), stoichiometry ( $N$ ), enthalpy ( $\Delta H$ ) and entropy ( $\Delta S$ ). Due to the low apparent affinity, the  $N$  value was fixed to a constant value of 1 in CTP titrations to *PfCCT*<sub>(528–795)</sub> K663A according to the low  $c$  value ITC data analysis protocol<sup>40,41</sup>. For CTP and CDPCho titrations of *PfCCT*<sub>(528–795)</sub> K663A, and CTP titrations of T761A and T762A variants of *PfCCT*<sub>(528–795)</sub>, mean and SD from two independent experiments are provided, whereas for CDPCho titrations of *PfCCT*<sub>(581–775)</sub> and *PfCCT*<sub>(528–795)</sub> mean and SD from three independent experiments are provided. For *PfCCT*<sub>(528–795)</sub> (Y626F/Q636A), errors of model fit to data of a single ligand binding experiment are given.

**$2\text{D } ^1\text{H NMR Spectroscopy.}$**  NMR experiments for the assignments of the flexible part of *PfCCT*<sub>(581–775)</sub> were recorded at  $283 \text{ K}$  on a Bruker 700 MHz NMR spectrometer equipped with a cryogenic  $5 \text{ mm}$  Z-gradient probe head. The aqueous samples were containing  $0.5 \text{ mM}$  unlabeled *PfCCT*<sub>(581–775)</sub> in  $20 \text{ mM}$  Tris/HCl pH 7.5,  $150 \text{ mM}$  NaCl and  $2 \text{ mM}$  EtSH in  $10\%$   $\text{D}_2\text{O}$  as an internal lock. Two-dimensional  $^1\text{H}$ - $^1\text{H}$  TOCSY (isotropic mixing:  $60 \text{ ms}$ ) and  $^1\text{H}$ - $^1\text{H}$  NOESY (mixing time:  $200 \text{ ms}$ ) were performed for the assignments of the backbone and aliphatic side chain  $^1\text{H}$  resonances. Water suppression was achieved with the WATERGATE sequence<sup>42</sup> in both experiments. NMR spectra were processed using GIFA software<sup>43</sup>.

**Crystallization and diffraction data collection.**  $10 \text{ mg/ml}$  of *PfCCT*<sub>(581–775)</sub> with  $20 \text{ mM}$  of CDPCho (final protein concentration  $8 \text{ mg/ml}$ ) in the lysis buffer were screened by commercially available Classics screen (NeXtal, Qiagen) by sitting drop method.  $200 \text{ nl}$  sitting drops were set up using a 1:1 ratio between protein and precipitant in 96-well Greiner crystallization plates using a Cartesian dispensing robot<sup>44</sup>. After one month crystals of protein-CDPCho complex appeared in  $0.2 \text{ M}$  NaF,  $0.1 \text{ M}$  HEPES pH 7.5,  $20\%$  PEG 3350 at  $20^\circ\text{C}$ . The ligand-free *PfCCT* crystals were grown by the hanging-drop vapor-diffusion method at  $4^\circ\text{C}$  by mixing  $1 \mu\text{l}$  of protein solution at concentration of  $10 \text{ mg/ml}$  and  $1 \mu\text{l}$  reservoir solution ( $0.08 \text{ M}$  NaF,  $0.1 \text{ M}$  Tris pH 7.5,  $24\%$  PEG 3350) and appeared after 3 weeks. CMP-bound and ChoP-bound protein crystals were obtained by soaking crystals with CMP and ChoP dissolved in lysis buffer ( $20 \text{ mM}$  Tris/HCl pH 7.5,  $0.15 \text{ M}$  NaCl and  $2 \text{ mM}$  EtSH). To obtain these co-crystals, ligand-free crystals were first grown by the hanging-drop vapor-diffusion method at  $20^\circ\text{C}$  by mixing  $1 \mu\text{l}$  of protein solution at concentration  $8 \text{ mg/ml}$  and  $1 \mu\text{l}$  reservoir solution (pH 8.0,  $20\%$  PEG

4000). Crystals appeared after 2 months and were soaked with 100 mM CMP and 100 mM ChoP (50 mM final concentration) for 5 days and 4 weeks, respectively. Cho-bound crystals were obtained by the hanging-drop vapor-diffusion method in 0.08 M NaF, 0.1 M Tris pH 7.5, 24% PEG 3350 at 20 °C by co-crystallization with 50 mM Cho dissolved in the lysis buffer. Crystals appeared after one week and were left to grow for an additional three weeks. Crystals were mounted and soaked in oil as cryosolvent, then frozen in liquid nitrogen. Diffraction data were collected at beamlines ID29, ID23-1 and ID23-2 of the European Synchrotron Radiation Facilities in Grenoble, France. These crystals belong to space group  $I_{222}$  with the asymmetric unit containing one monomer in the case of the free, CMP-bound, ChoP-bound and Cho-bound forms, and six monomers in the case of the CDPCho-bound form.

**Structure determination and refinement.** The indexing and integration of obtained data sets were carried out using XDS (Pilatus)<sup>45,46</sup>. The scaling, merging and calculating of structural factor amplitudes were executed using programs TRUNCATE<sup>47</sup> and SCALA<sup>48</sup> from CCP4 software<sup>49</sup>. Crystal structures were constructed by molecular replacement using Phaser<sup>50</sup> with *Rn*CCT (PDB: 3HL4) and free *Pf*CCT (PDB: 4ZCT) as search models, followed by model building in Coot<sup>51</sup>. The structure refinements were done using REFMAC 5.8.0073<sup>52,53</sup> and PHENIX 1.9.1692<sup>54</sup>. Molecular graphics representations were created using PyMol<sup>55</sup>. All the ligand structures were created using PRODRG server<sup>56</sup>. Crystal diffraction data and final refinement statistics were summarized in Table 3.

**Accession codes.** Atomic coordinates and structure factors for the reported crystal structures have been deposited in the PDB under ID codes 4ZCT, 4ZCP, 4ZCQ, 4ZCR and 4ZCS for the free, CMP-, Cho-, ChoP- and CDPCho-bound *Pf*CCT, respectively.

## References

- White, N. J. *et al.* Malaria. *Lancet* **383**, 723–735 (2014).
- WHO. World Malaria Report 2016. *World Health Organization*, 1–186 (2016).
- Breman, J. G. Resistance to artemisinin-based combination therapy. *Lancet Infect. Dis.* **12**, 820–822 (2012).
- Dondorp, A. M. *et al.* Artemisinin resistance in *Plasmodium falciparum* malaria. *N. Engl. J. Med.* **361**, 455–467 (2009).
- Botte, C. Y. *et al.* Atypical lipid composition in the purified relic plastid (apicoplast) of malaria parasites. *Proc. Natl. Acad. Sci. USA* **110**, 7506–7511 (2013).
- Dechamps, S., Shastri, S., Wengelnik, K. & Vial, H. J. Glycerophospholipid acquisition in *Plasmodium* - a puzzling assembly of biosynthetic pathways. *Int. J. Parasitol.* **40**, 1347–1365 (2010).
- Vial, H. *et al.* In *Apicomplexan Parasites: Molecular Approaches toward Targeted Drug Development* (ed Katja Becker) 137–162 (Wiley-VCH Press, 2011).
- Vial, H. J. *et al.* Prodrugs of bithiazolium salts are orally potent antimalarials. *Proc. Natl. Acad. Sci. USA* **101**, 15458–15463 (2004).
- Wein, S. *et al.* High Accumulation and *In Vivo* Recycling of the New Antimalarial Albitiazolium Lead to Rapid Parasite Death. *Antimicrob Agents Chemother* **61**, e00352–00317 (2017).
- Wengelnik, K. *et al.* A class of potent antimalarials and their specific accumulation in infected erythrocytes. *Science* **295**, 1311–1314 (2002).
- Brancucci, N. M. B. *et al.* Lysophosphatidylcholine Regulates Sexual Stage Differentiation in the Human Malaria Parasite *Plasmodium falciparum*. *Cell* **171**, 1532–1544 (2017).
- Dechamps, S. *et al.* The Kennedy phospholipid biosynthesis pathways are refractory to genetic disruption in *Plasmodium berghei* and therefore appear essential in blood stages. *Mol. Biochem. Parasitol.* **173**, 69–80 (2010).
- Sen, P., Vial, H. J. & Radulescu, O. Kinetic modelling of phospholipid synthesis in *Plasmodium knowlesi* unravels crucial steps and relative importance of multiple pathways. *BMC systems biology* **7**, 123 (2013).
- Gonzalez-Bulnes, P. *et al.* PG12, a phospholipid analog with potent antimalarial activity, inhibits *Plasmodium falciparum* CTP:phosphocholine cytidyltransferase activity. *J. Biol. Chem.* **286**, 28940–28947 (2011).
- Wein, S. *et al.* Transport and pharmacodynamics of albitiazolium, an antimalarial drug candidate. *Br. J. Pharmacol.* **166**, 2263–2276 (2012).
- Nagy, G. N. *et al.* Evolutionary and mechanistic insights into substrate and product accommodation of CTP:phosphocholine cytidyltransferase from *Plasmodium falciparum*. *FEBS J.* **280**, 3132–3148 (2013).
- Park, Y. S., Sweitzer, T. D., Dixon, J. E. & Kent, C. Expression, purification, and characterization of CTP:glycerol-3-phosphate cytidyltransferase from *Bacillus subtilis*. *J. Biol. Chem.* **268**, 16648–16654 (1993).
- Sundler, R. Ethanolaminephosphate cytidyltransferase. Purification and characterization of the enzyme from rat liver. *J. Biol. Chem.* **250**, 8585–8590 (1975).
- Veitch, D. P., Gilham, D. & Cornell, R. B. The role of histidine residues in the HXGH site of CTP:phosphocholine cytidyltransferase in CTP binding and catalysis. *Eur. J. Biochem.* **255**, 227–234 (1998).
- Lee, J., Johnson, J., Ding, Z., Paetzel, M. & Cornell, R. B. Crystal structure of a mammalian CTP: phosphocholine cytidyltransferase catalytic domain reveals novel active site residues within a highly conserved nucleotidyltransferase fold. *J. Biol. Chem.* **284**, 33535–33548 (2009).
- Cornell, R. B. Membrane lipid compositional sensing by the inducible amphipathic helix of CCT. *Biochim. Biophys. Acta* **1861**, 847–861 (2016).
- Cornell, R. B. & Ridgway, N. D. CTP:phosphocholine cytidyltransferase: Function, regulation, and structure of an amphitropic enzyme required for membrane biogenesis. *Prog. Lipid Res.* **59**, 147–171 (2015).
- Contet, A. *et al.* *Plasmodium falciparum* CTP:phosphocholine cytidyltransferase possesses two functional catalytic domains and is inhibited by a CDP-choline analog selected from a virtual screening. *FEBS Lett.* **589**, 992–1000 (2015).
- Lee, J., Taneva, S. G., Holland, B. W., Tieleman, D. P. & Cornell, R. B. Structural basis for autoinhibition of CTP:phosphocholine cytidyltransferase (CCT), the regulatory enzyme in phosphatidylcholine synthesis, by its membrane-binding amphipathic helix. *J. Biol. Chem.* **289**, 1742–1755 (2014).
- Ding, Z. *et al.* A 22-mer segment in the structurally pliable regulatory domain of metazoan CTP: phosphocholine cytidyltransferase facilitates both silencing and activating functions. *J. Biol. Chem.* **287**, 38980–38991 (2012).
- Huang, H. K. *et al.* The membrane-binding domain of an amphitropic enzyme suppresses catalysis by contact with an amphipathic helix flanking its active site. *J. Mol. Biol.* **425**, 1546–1564 (2013).
- Ramezanzpour, M., Lee, J., Taneva, S. G., Tieleman, D. P. & Cornell, R. B. An auto-inhibitory helix in CTP:phosphocholine cytidyltransferase hijacks the catalytic residue and constrains a pliable, domain-bridging helix pair. *J. Biol. Chem.* **293**, 7070–7084 (2018).



28. Cornell, R. B. *et al.* Functions of the C-terminal domain of CTP: phosphocholine cytidyltransferase. Effects of C-terminal deletions on enzyme activity, intracellular localization and phosphorylation potential. *Biochem. J.* **310**, 699–708 (1995).
29. Yang, W., Boggs, K. P. & Jackowski, S. The association of lipid activators with the amphipathic helical domain of CTP:phosphocholine cytidyltransferase accelerates catalysis by increasing the affinity of the enzyme for CTP. *J. Biol. Chem.* **270**, 23951–23957 (1995).
30. Nagy, G. N. *et al.* Composite aromatic boxes for enzymatic transformations of quaternary ammonium substrates. *Angew. Chem. Int. Ed. Engl.* **53**, 13471–13476 (2014).
31. Helmink, B. A., Braker, J. D., Kent, C. & Friesen, J. A. Identification of lysine 122 and arginine 196 as important functional residues of rat CTP:phosphocholine cytidyltransferase alpha. *Biochemistry* **42**, 5043–5051 (2003).
32. D'Angelo, I. *et al.* Structure of nicotinamide mononucleotide adenyltransferase: a key enzyme in NAD(+) biosynthesis. *Structure* **8**, 993–1004 (2000).
33. Maheshwari, S. *et al.* Biochemical characterization of Plasmodium falciparum CTP:phosphoethanolamine cytidyltransferase shows that only one of the two cytidyltransferase domains is active. *Biochem. J.* **450**, 159–167 (2013).
34. Patridge, K. A. *et al.* Glycerol-3-phosphate cytidyltransferase. Structural changes induced by binding of CDP-glycerol and the role of lysine residues in catalysis. *J. Biol. Chem.* **278**, 51863–51871 (2003).
35. Tian, S. *et al.* Human CTP:phosphoethanolamine cytidyltransferase: enzymatic properties and unequal catalytic roles of CTP-binding motifs in two cytidyltransferase domains. *Biochem. Biophys. Res. Commun.* **449**, 26–31 (2014).
36. Weber, C. H., Park, Y. S., Sanker, S., Kent, C. & Ludwig, M. L. A prototypical cytidyltransferase: CTP:glycerol-3-phosphate cytidyltransferase from bacillus subtilis. *Structure* **7**, 1113–1124 (1999).
37. Park, Y. S. *et al.* Identification of functional conserved residues of CTP:glycerol-3-phosphate cytidyltransferase. Role of histidines in the conserved HXGH in catalysis. *J. Biol. Chem.* **272**, 15161–15166 (1997).
38. Peyrottes, S. *et al.* Choline analogues in malaria chemotherapy. *Curr. Pharm. Des.* **18**, 3454–3466 (2012).
39. Marton, L. *et al.* Molecular Mechanism for the Thermo-Sensitive Phenotype of CHO-MT58 Cell Line Harboring a Mutant CTP:Phosphocholine Cytidyltransferase. *PLoS One* **10**, e0129632 (2015).
40. Tellinghuisen, J. Isothermal titration calorimetry at very low c. *Anal. Biochem.* **373**, 395–397 (2008).
41. Turnbull, W. B. & Daranas, A. H. On the value of c: can low affinity systems be studied by isothermal titration calorimetry? *J. Am. Chem. Soc.* **125**, 14859–14866 (2003).
42. Poggio, M., Saudek, V. & Sklenar, V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR* **2**, 661–665 (1992).
43. Pons, J. L., Malliavin, T. E. & Delsuc, M. A. Gifa V. 4: A complete package for NMR data set processing. *J. Biomol. NMR* **8**, 445–452 (1996).
44. Walter, T. S. *et al.* A procedure for setting up high-throughput nanolitre crystallization experiments. Crystallization workflow for initial screening, automated storage, imaging and optimization. *Acta Crystallogr. D Biol. Crystallogr.* **61**, 651–657 (2005).
45. Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystallogr.* **26**, 795–800 (1993).
46. Kabsch, W. Xds. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
47. French, S. & Wilson, K. On the treatment of negative intensity observations. *Acta Crystallogr. Sect. A* **34**, 517–525 (1978).
48. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 72–82 (2006).
49. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).
50. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).
51. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
52. 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).
53. Murshudov, G. N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 355–367 (2011).
54. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
55. DeLano, W. L. The PyMOL Molecular Graphics System. <http://www.pymol.org>, DeLano Scientific, San Carlos, CA, USA (2002).
56. Schüttelkopf, A. W. & van Aalten, D. M. F. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. Section D* **D60**, 1355–1363 (2004).
57. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320–324 (2014).

## Acknowledgements

This work was supported by the French Agence Nationale de la Recherche (ANR-09-BLAN-0397) and by the National Research, Development and Innovation Office (K109486, K119493), the Hungarian Academy of Sciences Medinprot program and ICGEB CRP/HUN14-01. The authors thank the support of collaboration visits by the French-Hungarian Bilateral S&T Cooperation (TÉT\_14\_FR-1-2015-0024). This work was also supported by the French Infrastructure for Integrated Structural Biology (FRISBI) ANR-10-INBS-05. EG was funded by the European Commission FP7 Marie-Curie Initial Training Network “ParaMet” [grant number 290080]. Financial support from ÜNKP-16-2 to FHa is gratefully acknowledged. The authors also acknowledge the European Synchrotron Radiation Facilities (Grenoble, France) for provision of synchrotron radiation facilities and we are grateful to staff scientists for assistance in using the beamlines ID29, ID23-1 and ID23-2. We thank Dr. Karine De Guillen for help with the STD-NMR experiments and Dr. Kai Wengelnik for helpful discussions and advices.

## Author Contributions

E.G. purified and crystallized the proteins. G.N.N., F. Ha., L.M. and R.I. cloned and purified the mutants, performed and analyzed kinetic and ITC experiments. E.G., F. Ho. and J.-F.G. collected diffraction data, built and refined the structures. E.G. and Y.Y. performed and analyzed NMR experiments. H.V., B.G.V., J.-F.G. and R.C. designed the study and supervised the work. E.G., G.N.N., F. Ha., L.M., R.I., B.G.V. and R.C. wrote the manuscript. All authors reviewed the manuscript.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-018-29500-9>.

**Competing Interests:** The authors declare no competing interests.



**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018