the FEBS Journal

Evolutionary and mechanistic insights into substrate and product accommodation of CTP:phosphocholine cytidylyltransferase from *Plasmodium falciparum*

Gergely N. Nagy[1], Lívia Marton[1], Balázs Krámos[2], Julianna Oláh[2], Ágnes Révész[3], Károly Vékey[3], Frédéric Delsuc[4], Éva Hunyadi-Gulyás[5], Katalin F. Medzihradszky[5,6], Marina Lavigne[7], Henri Vial[7], Rachel Cerdan[7] and Beáta G. Vértessy[1,8]

1 Institute of Enzymology, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Budapest, Hungary
2 Department of Inorganic and Analytical Chemistry, Budapest University of Technology and Economics, Hungary
3 Institute of Organic Chemistry, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Budapest, Hungary
4 Institut des Sciences de l'Evolution, Université Montpellier 2, France
5 Laboratory of Proteomics Research, Biological Research Centre of Hungarian Academy of Sciences, Szeged, Hungary
6 Mass Spectrometry Facility, Department of Pharmaceutical Chemistry, University of California San Francisco, CA, USA
7 Laboratory Dynamique des Interactions Membranaires Normales et Pathologiques, Université Montpellier 2, France
8 Department of Applied Biotechnology and Food Science, Budapest University of Technology and Economics, Hungary

Keywords
CTP:phosphocholine cytidylyltransferase; gene duplication; lipid biosynthesis; malaria; thermodynamics of ligand binding

Correspondence
G. N. Nagy and B. G. Vértessy, Institute of Enzymology, Research Centre for Natural Sciences, Hungarian Academy of Sciences, 29 Karolina street 1113 Budapest, Hungary
Fax: +361 4665465
Tel: +361 2793116
E-mail: gnagy@enzim.hu, vertessy.beata@ttk.mta.hu

(Received 25 December 2012, revised 8 March 2013, accepted 26 March 2013)

doi:10.1111/febs.12282

The enzyme CTP:phosphocholine cytidylyltransferase (CCT) is essential in the lipid biosynthesis of Plasmodia (Haemosporida), presenting a promising antimalarial target. Here, we identified two independent gene duplication events of CCT within Apicomplexa and characterized a truncated construct of *Plasmodium falciparum* CCT that forms a dimer resembling the molecular architecture of CCT enzymes from other sources. Based on biophysical and enzyme kinetics methods, our data show that the CDP-choline product of the CCT enzymatic reaction binds to the enzyme considerably stronger than either substrate (CTP or choline phosphate). Interestingly, in the presence of Mg^{2+}, considered to be a cofactor of the enzyme, the binding of the CTP substrate is attenuated by a factor of 5. The weaker binding of CTP:Mg^{2+}, similarly to the related enzyme family of aminoacyl tRNA synthetases, suggests that, with lack of Mg^{2+}, positively charged side chain(s) of CCT may contribute to CTP accommodation. Thermodynamic investigations by isothermal titration calorimetry and fluorescent spectroscopy studies indicate that accommodation of the choline phosphate moiety in the CCT active site is different when it appears on its own as one of the substrates or when it is linked to the CDP-choline product. A tryptophan residue within the active site is identified as a useful internal fluorescence sensor of enzyme–ligand binding. Results indicate that the catalytic mechanism of *Plasmodium falciparum* CCT may involve conformational changes affecting the choline subsite of the enzyme.

Database
Model data are available in the Protein Model DataBase (PMDB) under the accession number PM0078718 (*Pf*CCT(528–795)) and PM0078719 (*Pf*CCT MΔK)

Abbreviations
CDPCho, CDP-choline, cytidine 5′-diphosphocholine; ChoP, choline phosphate or phosphocholine; CMP segment, functional domains of CCT encompassing catalytic (C), membrane/lipid-binding (M) and putative phosphorylation (P) domains; CTP, cytidine triphosphate; ITC, isothermal titration calorimetry; MESG, 2-amino-6-mercapto-7-methylpurine ribonucleoside; ML, maximum likelihood; PC, phosphatidylcholine; *Pf*CCT MΔK, a truncated construct of *Pf*CCT (*Pf*CCT(528–795)$^{\Delta 720-737}$ encompassing residues 528–795 without the lysine-rich segment 720–737); *Pf*CCT, *Plasmodium falciparum* CTP:phosphocholine cytidylyltransferase; PP_i, pyrophosphate; TCEP, (tris (2-carboxyethyl)phosphine).

Structured digital abstract

- PfCCT MΔK and PfCCT MΔK bind by mass spectrometry studies of complexes (View interaction)
- PfCCT MΔK and PfCCT MΔK bind by comigration in gel electrophoresis (View interaction)
- PfCCT MΔK and PfCCT MΔK bind by molecular sieving (View interaction)

Introduction

Malaria continues to be a major world health problem affecting a hundred million people annually, mostly in African sub-Saharan countries. The disease is caused by the infection and destruction of red blood cells by protozoan parasites belonging to the genus *Plasmodium*, *Plasmodium falciparum* and *Plasmodium vivax* being the most widespread [1,2]. There is thus an urgent need to identify new drug targets and develop new pharmacophores with unique structures and mode of action [3,4].

Successful growth and multiplication of the parasite requires temporally controlled metabolic programmes leading to duplication of the structural components. The main *Plasmodium* membrane constituents are glycerophospholipids produced by *de novo* pathways using precursors that are actively transported from the host cytoplasm [5,6]. Phosphatidylcholine (PC) is the major phospholipid in *P. falciparum* membranes [7,8] and its synthesis relies on numerous pathways, seldom found together in a single organism [5,6]. Choline-mimicking compounds arrest parasite growth with potent *in vivo* antimalarial properties [9]. The clinical candidate albitiazolium [10] that specifically inhibits *de novo* PC biosynthesis is structurally unrelated to existing antimalarial agents and is currently being developed in human phase 2 clinical trials for parenteral cures of severe malaria [11].

De novo PC biosynthesis through the CDP-choline (CDPCho) dependent Kennedy pathway is of crucial importance in *Plasmodium berghei* [12] and *P. falciparum*. Within this pathway, the CTP:phosphocholine cytidylyltransferase (2.7.7.15) enzyme [13–16] that catalyses the second step of the CDPCho pathway by condensing cytidine triphosphate (CTP) to phosphocholine (ChoP) to obtain the energized CDPCho is rate limiting and thus can be considered as a potential molecular drug target.

The CCT genes are annotated in different *Plasmodium* genomes as unique copies. Surprisingly, the plasmodial CCT genes encode two putatively duplicated CCT segments (abbreviated as CMP segments), each including a cytidylyltransferase catalytic domain (C1/C2), a membrane/lipid binding domain (M1/M2) and a putative phosphorylation domain (P1/P2). Between the two CMP segments, there is a long linker (240–278 amino acids), which is partially conserved. Other known CCTs possess a single catalytic domain and also one copy of the membrane binding and phosphorylation domains, and are reported to form homodimers in solution [17]. Hence, the two duplicated segments of plasmodial CCT proteins might associate to form an intramolecular 'pseudo-dimer', although this suggestion has not yet been investigated experimentally. It has been shown that a recombinant *P. falciparum* CCT protein construct (*Pf*CCT$^{528–896}$) containing the second catalytic domain (C2) followed by the membrane interaction domain (M2) is enzymatically active [18].

CCT expression is apparent in all life stages of *Plasmodium* [11,19]. Microarray data (PlasmoDB) [20] show a transcription increase in late schizonts and significant transcription in gametocytes and sporozoites as well. Membrane interaction of an amphipathic segment of the membrane binding domain (M) regulate CCT enzymatic activity while removal of the membrane binding segment from the CCT protein results in a constitutively active enzyme [21]. For the rat CCT enzyme, an *in vitro* kinetic study suggested random mechanism for binding and release of the substrates/products, with Michaelis constant K_M values of 10 mM and 0.3 mM for CTP and ChoP, respectively [22].

Although CCT activity is crucial for Plasmodia at their blood stage, the enzyme:substrate binary complexes have not yet been investigated in detail. To date, no three-dimensional structure of any CCT:substrate complex is available. Current structural concepts rely on the three-dimensional CCT:CDPCho structure of the rat catalytic domain (PDB ID 3HL4) [23] and structures of other cytidylyltransferase:nucleotide complexes (ECT:CMP and GCT:CTP, PDB IDs 3ELB and 1COZ, respectively) [24]. It is also not known whether binary CCT:substrate complexes already adopt a catalytically competent conformation or such conformation is reached only in the ternary complex of the enzyme with the two substrates.

In the present study, we investigated the evolutionary origin of the two domains and the catalytic mechanism and ligand binding properties of *Pf*CCT. We

reconstructed the most likely evolutionary scenario that led to the duplication of the CMP segments in CCT within Apicomplexa. We then designed and characterized a catalytic domain construct of *Pf*CCT using the C2 catalytic domain, and showed that this construct is enzymatically functional and forms a dimer, similarly to other CCTs. We constructed a three-dimensional structural model of this enzyme and used a continuous spectrophotometric method to monitor its steady-state kinetic behaviour. Binding of substrates and products was followed by fluorescence spectroscopy and isothermal titration calorimetry (ITC). Relying on a Trp fluorophore located in the active site, we could characterize alterations in the binding modes of substrate and product containing the choline moiety. Thermodynamic analysis revealed that the presence of Mg^{2+}, the cofactor of CCT catalysis, attenuates the binding of CTP. Taken together, the present ligand binding and kinetic studies of the enzyme contribute to the understanding of the catalytic mechanism of *Pf*CCT.

Results

Phylogenetic analyses revealed two independent duplication events of CMP segments in CCTs within Apicomplexa

The CCT catalytic domain (C) is widely conserved from plants to animals. Putative CCT enzymes harbouring duplicated C domains are found only in three Apicomplexan parasites, *Plasmodium*, *Babesia* and *Theileria*, within the class Aconoidasida. Sequence alignment revealed that the duplicated C domains of plasmodial CCT are highly identical (Fig. 1). This can be exemplified with *Pf*CCT, showing 90% sequence identity and 97% similarity between the N- and the C-terminal catalytic domains (C1 and C2, respectively). To reconstruct the history of duplication events that led to the existence of duplicated CCT genes in Apicomplexa, we performed a phylogenetic analysis. The alignment of the full CCT protein sequences is presented in Fig. S1. It shows that the single CMP segment observed in Coccidia (*Cryptosporidium*, *Neospora* and *Toxoplasma*) is homologous to the N-terminal segment of Aconoidasida (other genera). The maximum likelihood (ML) tree obtained for the 16 full Apicomplexan CCT proteins is shown in Fig. 2A. The phylogenetic relationships among the Apicomplexan species included are fully compatible with the tree recently inferred from genome-scale data [25]. Indeed, *Babesia microti* also appears as a distinct lineage that diverged early within Piroplasmida, rendering the genus *Babesia* paraphyletic. It should also be noted that Piroplasmida members are particularly fast evolving compared with all other Apicomplexans.

The ML tree inferred for the 27 CMP protein segments of Apicomplexan CCTs is presented in Fig. 2B. By reconstructing the CMP segment duplication history during the evolution of Apicomplexa, this phylogeny allows the hypothesis of a single ancestral duplication of a CMP segment in the CCT gene predating the separation between Piroplasmida (*Babesia* and *Theileria*) and Haemosporida (*Plasmodium*) to be tested [5]. In this hypothetical scenario, N-terminal and C-terminal segments are predicted to form distinct monophyletic subtrees whose internal relationships should exactly mirror the species tree of Fig. 2A. The ML tree in Fig. 2B clearly shows that this is not the case. Indeed, the sequences of the duplicated C-terminal segments fall into two distinct monophyletic groups, each being closely related to a cluster of N-terminal sequences within both Piroplasmida and Haemosporida. The most likely scenario to explain such a phylogenetic pattern is that two independent domain duplication events occurred in the last common ancestor of Piroplasmida on one side and Haemosporida on the other side. Moreover, the large difference observed in sequence divergence between the N-terminal and C-terminal segments in Piroplasmida and Haemosporida suggests that these two independent duplications have occurred at different points in time. Indeed, the high degree of sequence divergence observed among the duplicated segments in *Babesia* and *Theileria* species might be explained by a duplication event early in the evolutionary history of Piroplasmida. Conversely, the high similarity of duplicated CMP segments observed between *Plasmodium* species may be the result of a more recent duplication event in the Haemosporida lineage. Based on this evolutionary scenario resulting in two very similar catalytic domains of *Pf*CCT and due to the fact that full-length *Pf*CCT could not be expressed and purified for biochemical analysis so far, we decided to study one of the *Pf*CCT catalytic domains as a model.

Homology model and characterization of the *Pf*CCT catalytic domain

A homology model of the two catalytic domains of *Pf*CCT was built based on the published structure of the homodimer catalytic domain of rat CCT (see Supporting information) [23]. Analysis of these models shows that neither of the very few residues that differ in C1 and C2 is bordering the conserved active site of the enzyme. The high degree of sequence identity

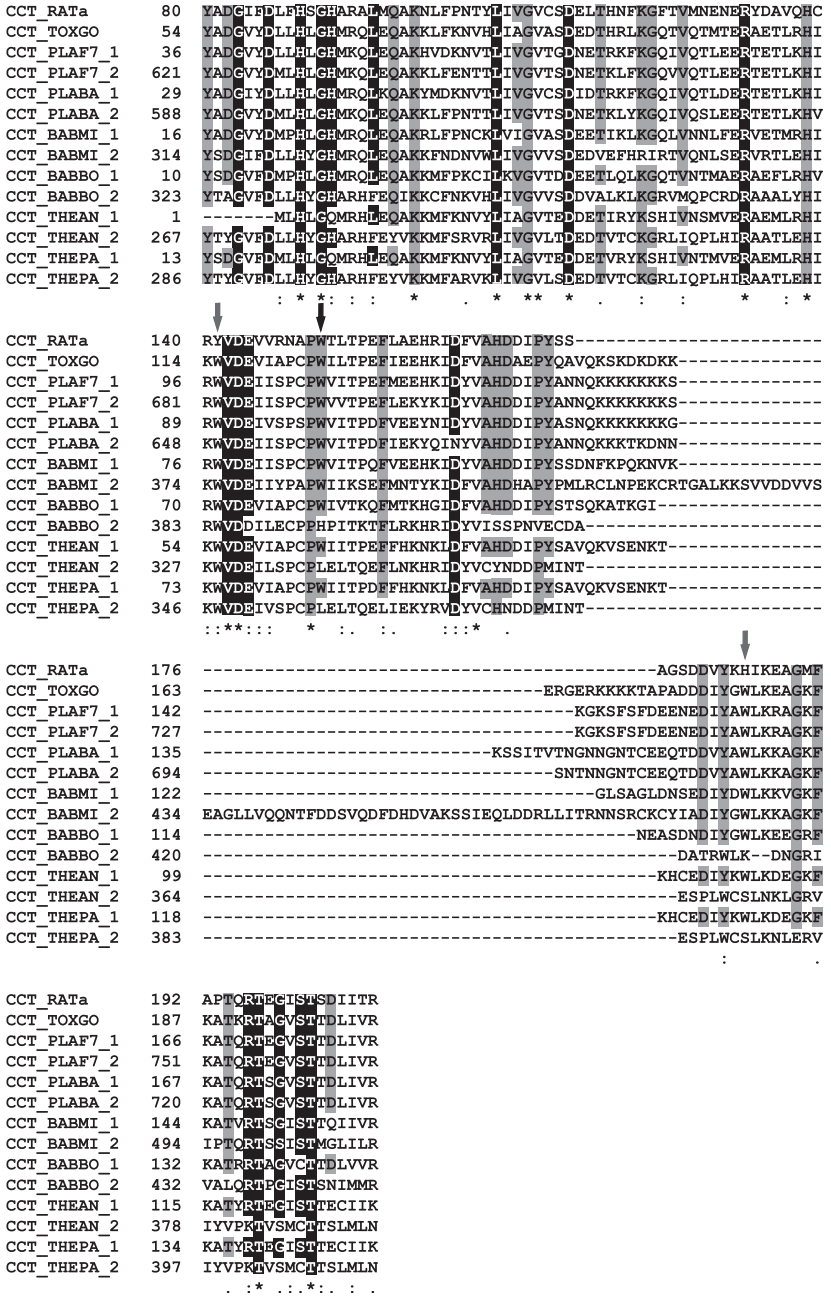

Fig. 1. Residue conservation pattern within duplicated CCT catalytic domains in Apicomplexa. Alignment of cytidylyltransferase catalytic domain sequences, whose boundaries are identified by the PFAM server [56]. RAT, *Rattus norvegicus* CCT (Uniprot entry P19836), (residues 80–208); TOXGO, *Toxoplasma gondii* (B6KSU6) (54–203); PLAF7, *Plasmodium falciparum* 3D7 strain (Q8IEE9) (36–182 and 621–767); PLABA, *Plasmodium berghei* (Q4YBQ9) (29–183 and 588–736); BABMI, *Babesia microti* (I7I8D8) (16–160 and 314–510); BABBO, *Babesia bovis* (A7AU63) (10–158 and 323–448); THEAN, *Theileria annulata* (Q4UFG5) (1–131 and 267–394); THEPA, *Theileria parva* (Q4N412) (13–150 and 286–413). The alignment was initially performed by CLUSTALW and then manually adjusted. Residues conserved among the cytidylyltransferase superfamily (GCT, ECT, CCT enzymes) are highlighted with white letters on a dark grey box; residues conserved among CCT enzymes are highlighted with a light grey box. The positions of the W residues within the *Pf*CCT sequences are shown by grey and black arrows, the latter indicating the conserved W residue bordering the active site.

between the two catalytic domains of *Pf*CCT probably allows one single domain to dimerize as in other studied CCTs. We engineered a truncation construct (*Pf*CCT$^{528–795}$) that contains the C2 catalytic domain but does not contain the membrane-binding M2 domain and is therefore expected to be constitutively active. A lysine-rich region (residues 720–737, termed as K-loop in Fig. 3A), constituting a species-specific insertion in Plasmodia sequences, was removed to aid purification of the recombinant construct (see Materials and methods). Comparing homology models either containing or lacking this lysine-rich segment revealed no structural differences affecting the folding of the catalytic domain (compare models in wheat and blue-white colours in Fig. 3A). The construct *Pf*CCT (528–795)$^{\Delta 720–737}$ lacking the lysine-rich segment (abbreviated as *Pf*CCT MΔK) was expressed in *Escherichia coli*. Removal of the lysine-rich segment improved the thermal stability of the *Pf*CCT construct as judged by an ~ 9 °C increase in melting temperature (T_m), measured by differential scanning fluorimetry (Fig. S2).

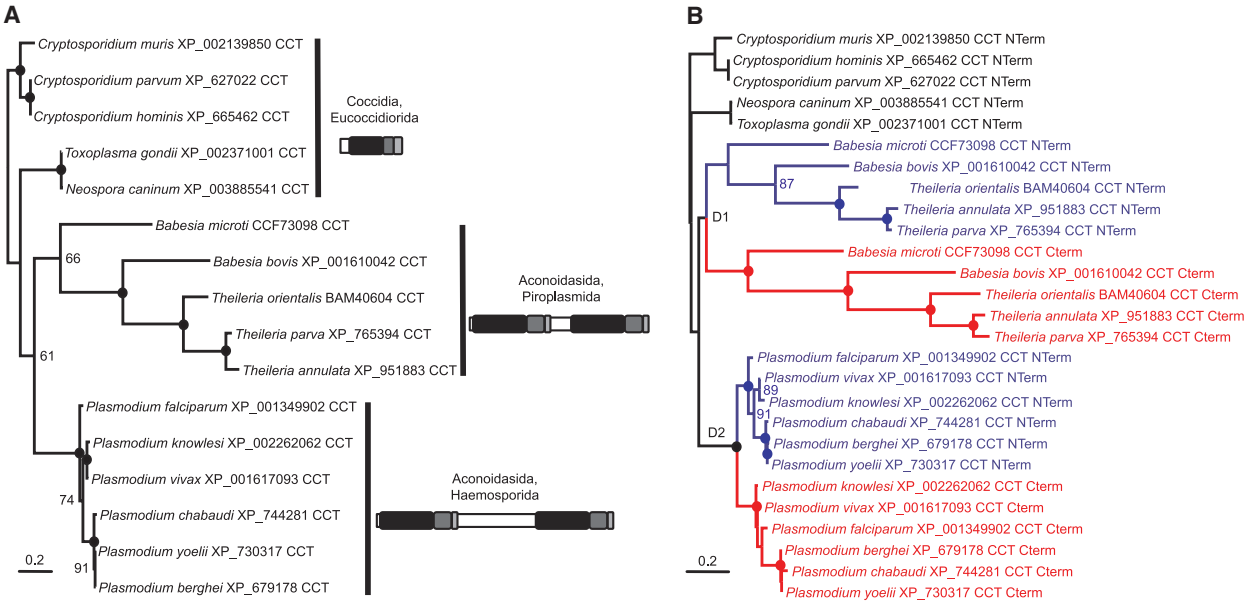

Fig. 2. Phylogenetic analyses of domain duplication within Apicomplexa parasites. (A) ML phylogenetic tree of the 16 full-length CCT Apicomplexan proteins. Numbers at nodes are bootstrap percentages with bullets indicating a value of more than 95. The scale is in mean number of substitutions per site. Structures of the CCT gene with location of CMP segments in the main Ampicoplexan groups are represented following Dechamps *et al.* [5]. (B) ML phylogenetic tree of the 27 CMP protein segments of the CCT gene in Apicomplexa. Numbers at nodes are bootstrap percentages, bullets indicating a value of more than 95. The scale is in mean number of substitutions per site. Nodes without numbers have bootstrap percentages < 70. D1 and D2 indicate the two independent duplication events required to explain the inferred phylogenetic pattern.

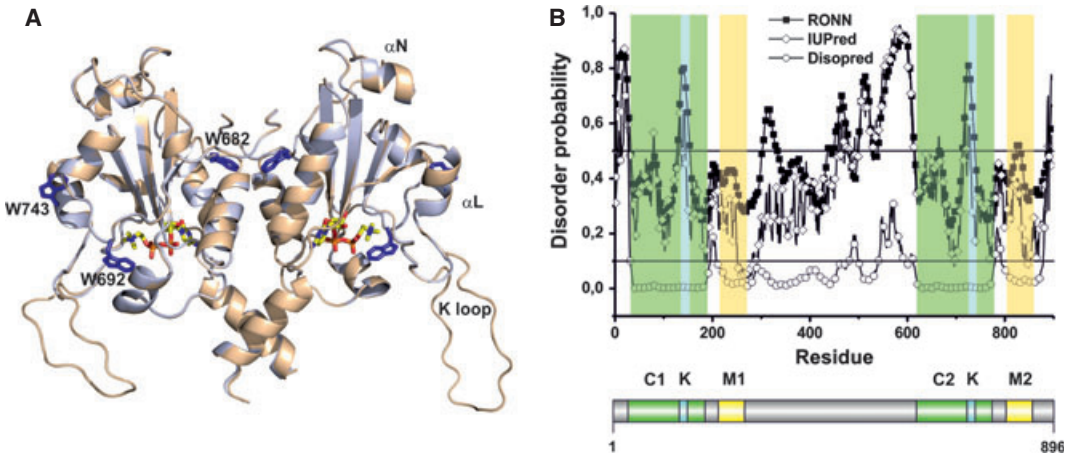

Fig. 3. Structural model of the *Pf*CCT MΔK construct. (A) Superimposed homology models of the homodimer forms of the *Pf*CCT MΔK (528–795)$^{\Delta 720-737}$ construct (shown in cartoon model, coloured blue-white) and the *Pf*CCT(528–795) construct (shown in cartoon model, coloured wheat) containing the C2 domain are superimposed. The models were built with the MODELLER program using the rat CCT crystal structure (PDB ID 3HL4) as template. The CDPCho ligand is shown as sticks, using atomic colouring (carbon, yellow; oxygen, red; phosphorus, orange; nitrogen, blue). The tryptophans W682, W692, W743 are shown as sticks, coloured blue. (B) Disorder profile of full-length *Pf*CCT. The plot shows the probability of disorder predicted by RONN [57], IUPRED [58] and DISOPRED [59]. C1 and C2 domains, coloured green, were classified as ordered by all three predictors. The lysine-rich segments (K, coloured cyan) within the ordered conserved catalytic domains are predicted to be highly flexible.

We used *in silico* analysis of structural disorder (Fig. 3B) to assess further structural characteristics. Within the full-length CCT sequence, there are numerous highly flexible/disordered segments in the linker region connecting the two duplicated segments. The lysine-rich segment is present in both C1 and C2

domains and is predicted to possess high disorder. The region encompassing residues 540–610, present within our engineered *Pf*CCT MΔK construct, is also predicted to possess high flexibility/disorder.

Gel filtration of *Pf*CCT MΔK under native conditions resulted in a single peak, at an elution volume corresponding to a molecular mass of approximately 268 kDa (Fig. S3), which is unexpectedly high (for comparison, the molecular mass of the *Pf*CCT MΔK polypeptide calculated from the amino acid sequence is 30.9 kDa). Denaturing SDS/PAGE analysis revealed that this peak contains one polypeptide component for which the apparent molecular mass based on SDS/PAGE analysis is estimated to be 37 kDa (Fig. 4A), again considerably higher than the calculated molecular mass of 30.9 kDa. SDS/PAGE and gel filtration data therefore indicate anomalous migration of the protein in gel electrophoresis and possible formation of higher oligomers under native conditions. The presence of flexible segments as suggested by the disorder predictors (Fig. 3B) may explain this behaviour and also suggests that estimation of the oligomerization status of the native molecular mass of this protein based on size exclusion chromatography is not reliable, since deviations from the ideally expected hydrodynamic behaviour of globular proteins are expected to greatly perturb elution characteristics.

To investigate the oligomerization state of *Pf*CCT MΔK, we therefore performed electrospray ionization mass spectrometry, which has been shown to provide valuable data on protein–protein complexation [26]. The electrospray mass spectrum of *Pf*CCT MΔK under native conditions (Fig. 4B) reveals two abundant ion series in the mass range m/z 1900–4000. In the upper mass range (m/z 3000–4000), peaks corresponding to a dimer structure can be observed. The charge state distribution is rather narrow, encompassing mainly $z = +18$ to $z = +20$ ions, indicating that the native protein conformation survives the transfer from solution to the mass spectrometer [27]. The molecular mass of the dimer was determined to be 61 840 ± 200 Da (cf. the calculated dimer molecular mass of 61.8 kDa). At the lower mass range (m/z 1900–3000), the ion series corresponding to the monomer form is present. The charge distribution is centred on the +13 ions, and the mass calculated from the distribution is 30 870 ± 200 Da, in very good agreement with the calculated molecular mass. Based on the integrated peak areas, a monomer : dimer ratio of 10 : 13 was estimated. Signals due to higher oligomers – trimer, tetramer etc. – are practically unobservable, setting an upper limit of 3% for their relative abundance on the basis of the signal to noise ratio. The preference for dimer formation suggests that the dimer structure reflects the association existing in solution, i.e. it is not due to non-specific aggregation upon ionization. Dimerization of the *Pf*CCT MΔK construct is also suggested by the three-dimensional structural model and is in agreement with the oligomer properties of other CCT enzymes.

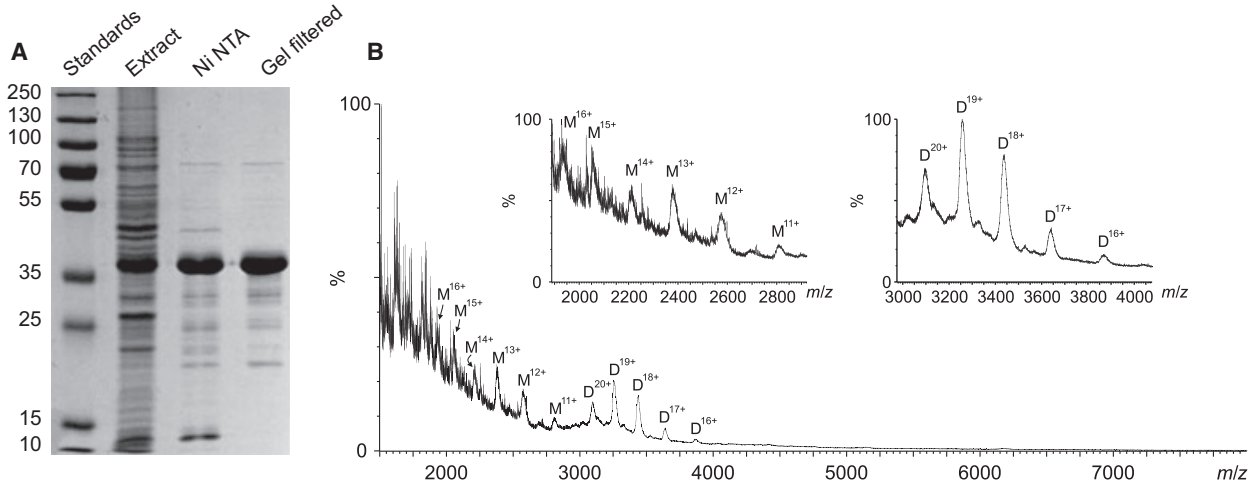

Fig. 4. *Pf*CCT MΔK is predominantly present in homodimer form in solution. (A) Purification steps of *Pf*CCT MΔK followed by 12% SDS/PAGE: lane 1, molecular mass standards; lane 2, *E. coli* extract; lane 3, sample eluted from Ni affinity purification; lane 4, sample after gel filtration. (B) Mass spectrum of *Pf*CCT MΔK protein under native electrospray conditions. M and D denote monomer and dimer signals, respectively, whereas numbers indicate the charge states. The two insets are enlarged graphs for the monomer and dimer ion series regions, with determined molar masses 30 870 (±200) Da and 61 840 (±200) Da, respectively.

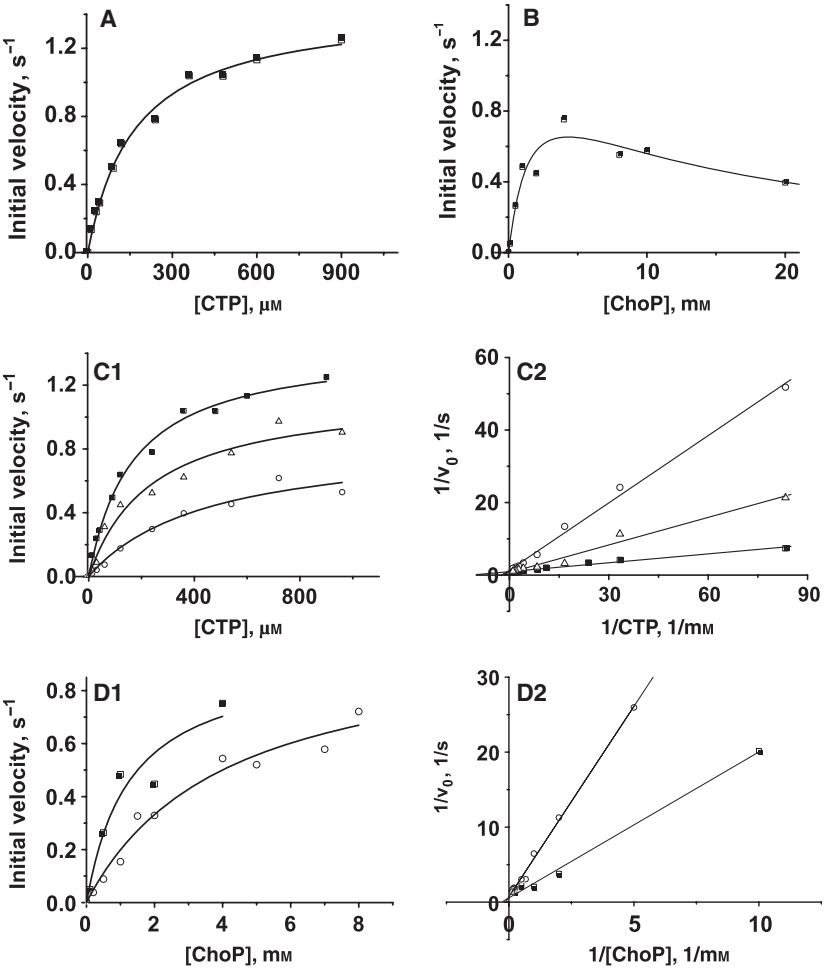

Fig. 5. Steady-state kinetic analysis of the *Pf*CCT MΔK catalysed reaction in the absence and presence of CDPCho product. (A) CTP titration of *Pf*CCT MΔK at a fixed ChoP concentration of 5 mM. The plot shows one representative experiment. Titration data are fitted with the Michaelis–Menten kinetic model assuming no cooperativity. (B) ChoP titration of *Pf*CCT MΔK at a fixed CTP concentration of 900 μM. The plot shows one representative experiment. Titration data are fitted with a kinetic model assuming substrate inhibition without cooperativity. Note the substrate inhibition effect of ChoP as an initial rate decrease is observed at higher substrate concentrations. (C1) CTP titration curves for *Pf*CCT MΔK with addition of CDPCho at different concentrations. The concentration of ChoP was fixed at 5 mM. Global fit of the titration data is shown with the full competitive inhibition model (Eqn 5), yielding K_i, as indicated in the text. (D1) ChoP titration curves for *Pf*CCT MΔK with addition of CDPCho at different concentrations. The concentration of CTP was fixed at 900 μM. For C1 and D1 Lineweaver–Burk plots are also shown (C2 and D2).

Random order kinetic mechanism shown by steady-state analysis

The enzymatic properties of the *Pf*CCT MΔK construct were determined by using a continuous spectrophotometric assay [28] that was adopted for other pyrophosphate producing enzymes [29,30]. Substrate titrations with CTP and ChoP were performed while the other substrate was added at saturating concentration. A plot of the initial rate as a function of CTP concentration provided a saturation curve (see Fig. 5A), which fits the Michaelis–Menten equation providing the apparent kinetic parameters 1.45 ± 0.05 s^{-1} for the turnover number k_{cat} and 168 ± 17 μM for $K_{M,CTP}$ (Table 1). In the ChoP titration, we observed substrate inhibition at [ChoP] > 5 mM. The ChoP titration data were fitted with a substrate inhibition equation (shown in Materials and methods) providing an apparent $K_{M,ChoP}$ of 1.8 mM (Fig. 5B). As a consequence, in the CTP titration experiment, ChoP was used at a sub-inhibitory concentration of 5 mM. None of the CTP and ChoP substrate saturation data showed significant cooperativity, as judged by the Akaike information criterion test comparing the Michaelis and Hill fitted models. Deletion of the

Table 1. Kinetic parameters of *Pf*CCT MΔK catalysis. Kinetic parameters were obtained as described in Materials and methods. One enzyme unit (U) is the number of nanomoles of CDPCho produced per minute.

CTP titration			ChoP titration		
V_{max} (U·mg^{-1})	k_{cat} (s^{-1})	$K_{M,CTP}$ (μM)	k_{cat} (s^{-1})	$K_{M,ChoP}$ (mM)	$K_{I,ChoP}$ (mM)
2810	1.45 ± 0.05	168 ± 17	1.2 ± 0.4	1.8 ± 1.1	10.5 ± 7.5

lysine-rich segment did not modify enzyme activity (data not shown).

In order to assess the binding order of the substrates CTP and ChoP, we performed a kinetic analysis based on CDPCho product inhibition. The initial rate of the CCT reaction was measured as a function of either substrate while the concentrations of the other substrate and the CDPCho product were kept constant (Fig. 5C,D). We found that the inhibitory pattern of CDPCho showed a major competitive character with respect to both CTP and ChoP substrates, which can be interpreted within the framework of a random bi–bi mechanism [31]. The apparent inhibition constant K_i of CDPCho is estimated from the fit to be 65 μM.

Comparison of the heat effects of substrate and product binding at the choline subsite

Since ITC is an especially useful technique for quantitative characterization of biological interactions, we used it to obtain thermodynamic parameters of the binding of substrates and products to *Pf*CCT MΔK. Titration calorimetry experiments were performed in the same buffer as was used in the activity assay, including 5 mM Mg^{2+} (Fig. 6A–D, Table 2).

For both CTP and CDPCho titration of *Pf*CCT MΔK (Fig. 6A,D), the best fit to the experimental data was obtained by fitting a one-site independent binding model. This, together with the observed stoichiometry close to the ratio 1 : 1, is consistent with the kinetic titration data (Table 2). The binding of these ligands to the enzyme did not show any significant cooperativity under our experimental conditions. In the case of the CDPCho product, the dissociation constant K_d of 89 μM is in good agreement with the inhibition constant K_i of 65 μM determined in the steady-state experiments. The CTP equilibrium ligand binding experiment reports similar affinity (K_d = 294 μM) to the value of the respective Michaelis constant (K_M = 168 μM). Both CDPCho and CTP ligands exhibited negative binding enthalpy and entropy values, which indicate that binding of the ligands is enthalpically favoured and entropically disfavoured. The latter observation for CTP binding is consistent with previous NMR data of CTP binding to GCT, a related cytidylyltransferase enzyme [32]. Nevertheless, CDPCho binding to the enzyme occurs with more than twofold higher enthalpy compared with CTP binding, shedding light on the potentially different interactions with these two ligands. These could be due to either different accommodation of the pyrophosphate group of CTP and the β phosphate group of CDPCho and/or the contribution of the interactions of the trimethylammonium moiety of CDPCho at the choline binding subsite in the active site of CCT.

Titration of the enzyme with pyrophosphate (PP_i) product yielded only minor heat effects (Fig. 6C). These calorimetric signals may be interpreted by estimating that the dissociation constant of the CCT:PP_i complex is in the millimolar range. On the other hand, we could not measure any significant calorimetry signal of apo *Pf*CCT MΔK enzyme for ChoP titration (up to 2 mM, already above the K_M determined in the kinetic assay) (Fig. 6B). This result indicates that either ChoP binding to the apoenzyme does not occur under these conditions or the enthalpic character of this binding is negligible when ChoP is present on its own.

In order to reveal any indirect effect of ChoP binding, we applied the displacement ligand titration method that has been shown to be effective in detecting ligand binding events with low millimolar affinities [33]. Immediately following the ChoP titrations, the *Pf*CCT MΔK solutions containing ChoP were further titrated with CDPCho to assess an equilibrium ligand displacement titration of ChoP to CDPCho. These subsequent product titrations yielded thermodynamic parameters very similar to those obtained from apo *Pf*CCT MΔK CDPCho titrations in the absence of ChoP (Fig. 6I,J, Table 2). Hence, the results did not show any detectable binding of ChoP suggesting that under our experimental conditions ChoP binding to the apoenzyme cannot be observed and arguing for a lower limit of the dissociation constant $K_{d,ChoP}$ > 5 mM (to be compared with a K_M of 1.8 mM).

Mg^{2+} attenuates enzyme affinity towards CTP

Mg^{2+} is reported to be required for CCT catalysis [18,34,35]; therefore Mg^{2+} is a generally used additive in CCT assay conditions. However, no data are available on its effect on ligand binding to the enzyme. We therefore carried out a set of ITC experiments in the absence of added Mg^{2+} (Fig. 6E–H). Titration of the enzyme with CDPCho showed similar thermodynamic parameters in the absence and presence of Mg^{2+} cofactor (with dissociation constants 44 and 89 μM, Fig. 6D,H, respectively). CTP affinity, in contrast, seems to be strongly dependent on the presence of the metal cofactor. Addition of 0.5 mM EDTA resulted in significantly higher affinity of the enzyme towards CTP than in the presence of Mg^{2+} (the observed dissociation constant being 59.5 μM compared with 294 μM, respectively) (Fig. 6A,E, Table 2). ChoP titration and PP_i titration provided practically the same results in the absence as well as in the presence of Mg^{2+} (Fig. 6B,C,G,H).

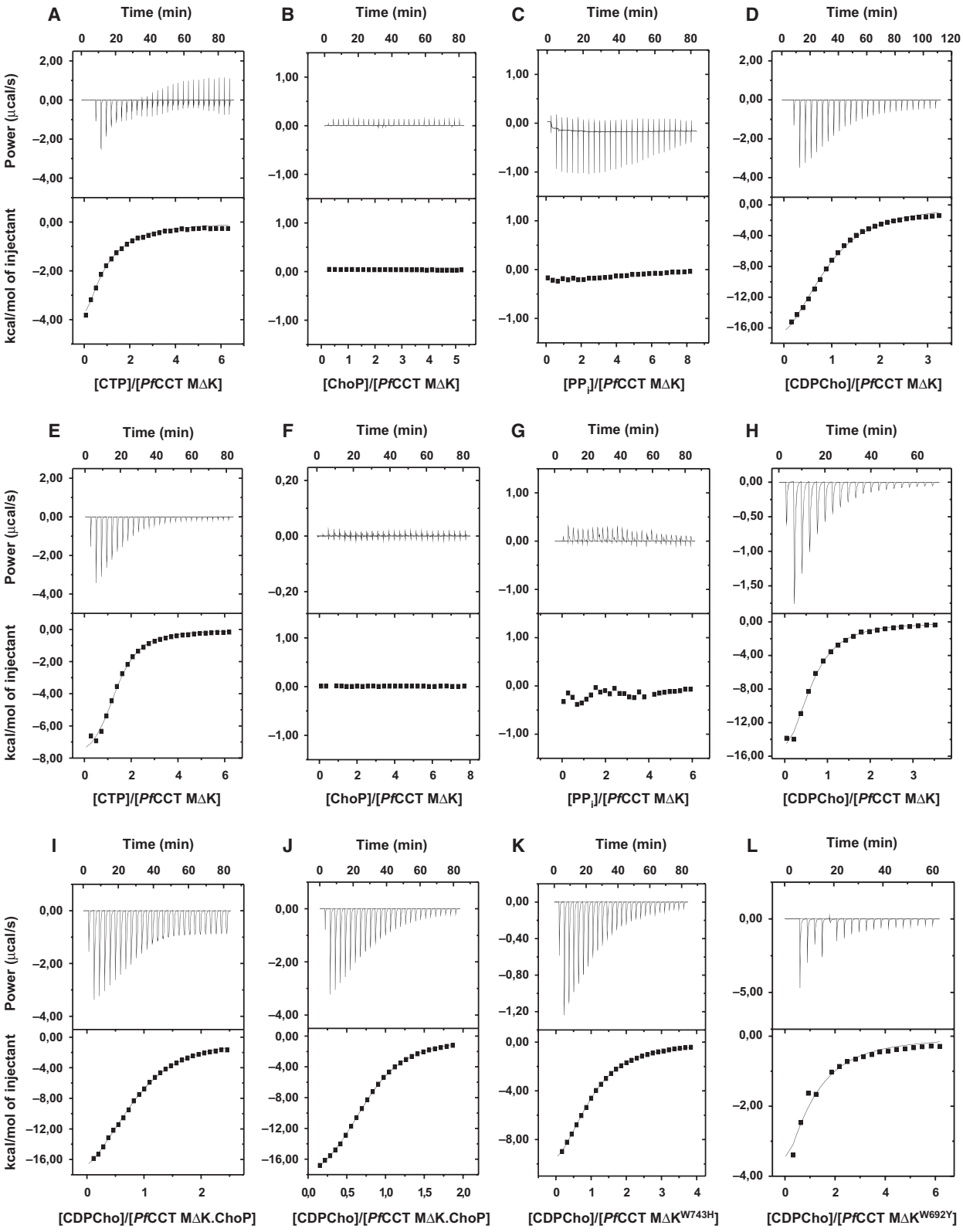
A
Time (min)
Power (μcal/s)
kcal/mol of injectant
[CTP]/[*Pf*CCT MΔK]
B
[ChoP]/[*Pf*CCT MΔK]
C
[PPi]/[*Pf*CCT MΔK]
D
[CDPCho]/[*Pf*CCT MΔK]
E
[CTP]/[*Pf*CCT MΔK]
F
[ChoP]/[*Pf*CCT MΔK]
G
[PPi]/[*Pf*CCT MΔK]
H
[CDPCho]/[*Pf*CCT MΔK]
I
[CDPCho]/[*Pf*CCT MΔK.ChoP]
J
[CDPCho]/[*Pf*CCT MΔK.ChoP]
K
[CDPCho]/[*Pf*CCT MΔK^{W743H}]
L
[CDPCho]/[*Pf*CCT MΔK^{W692Y}]

Table 2. Thermodynamic parameters of ligand binding to *Pf*CCT MΔK constructs determined by direct and displacement titrations at 20 °C.

*Pf*CCT construct		Mg^{2+}	*N*	K_d (μM)	ΔH (kcal·mol^{-1})	ΔS (cal·mol^{-1}·K^{-1})
MΔK	CDPCho	5 mM	0.88 ± 0.03	88.5 ± 7.2	−24.0 ± 1.0	−63.3 ± 19.0
MΔK	CTP	5 mM	0.62 ± 0.01	294 ± 4.0	−9.6 ± 1.7	−16.6 ± 3.6
MΔK	CDPCho (5 mM ChoP present)[a]	5 mM	0.85 ± 0.01	88.5 ± 3.4	−23.2 ± 0.4	−60.5 ± 2.5
MΔK	CDPCho	None	0.57 ± 0.02	44.4 ± 3.3	−22.9 ± 1.0	−58.1 ± 5.0
MΔK	CTP	None	0.96 ± 0.02	59.5 ± 1.8	−8.6 ± 0.3	−10.1 ± 0.5
MΔK	CDPCho (2 mM ChoP present)[a]	None	0.75 ± 0.01	42.9 ± 0.8	−20.7 ± 0.1	−50.7 ± 1.0
MΔK^{W743H}	CDPCho	None	0.96 ± 0.01	67.7 ± 0.9	−13.5 ± 0.1	−27.0 ± 0.4
MΔK^{W692Y}	CDPCho	None	0.83 ± 0.24	218 ± 50	−9.8 ± 4.5	−16.8 ± 8.6

[a] Displacement titrations.

Intrinsic fluorescent signal points to differences between ChoP and CDPCho bound states of the enzyme

In order to gain information on the microenvironment of the active site, we investigated the change in the intrinsic Trp fluorescence signal. *Pf*CCT MΔK has three tryptophan residues: W682, W692 and W743 (Fig. 1). W682 is buried at the dimer interface while the other two residues are more solvent-accessible; W743 is part of the αL-helix, located on the construct surface (Fig. 3A), while W692 establishes a cation–π interaction with the trimethylammonium moiety of the CDPCho product at the active site (Figs 3A and 7) [23]. To identify the Trp residue that is responsible for the observed signal changes, we created two constructs by site-directed mutagenesis, performing a W/H exchange at position 743 and a W/Y exchange at position 692, respectively. The *Pf*CCT MΔK^{W743H} showed similar ligand binding (Fig 6K, Table 2) and catalytic properties, whereas for the *Pf*CCT MΔK^{W692Y} mutant the interaction with CDPCho is still present but becomes somewhat perturbed (Fig 6L, Table 2, K_d increases threefold).

Addition of CDPCho led to an increase in the intensity of the intrinsic Trp fluorescence signal of *Pf*CCT MΔK (Fig. 8A). Moreover, the fluorescence emission maximum shifted from 347.7 to 344.2 nm. Plotting the emission maximum shifts (λ_{max}) against CDPCho concentration resulted in a saturation curve, indicating saturation of enzyme active sites with CDPCho (Fig. 8B). The dissociation constant of *Pf*CCT MΔK

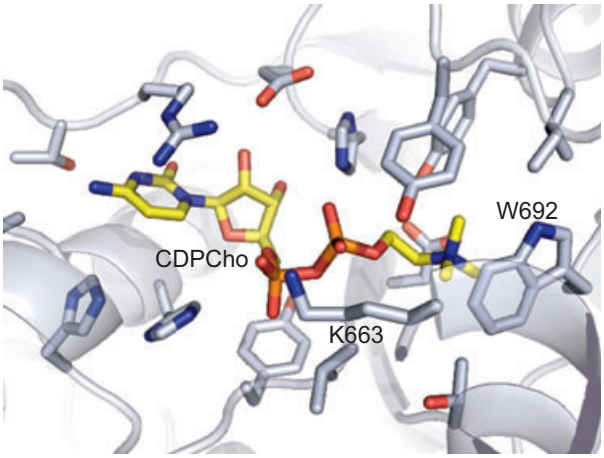

Fig. 7. Active site close-up of the *Pf*CCT MΔK model with bound CDPCho product. CDPCho is shown as sticks with atom colouring (carbon, yellow; oxygen, red; phosphorus, orange; nitrogen, blue). Residue side chains coordinating the CDPCho are shown as sticks coloured by element (carbon, grey; oxygen, red; phosphorus, orange; nitrogen, blue). Note W692 establishing a cation–π interaction with the trimethylammonium moiety of CDPCho.

Fig. 6. Equilibrium ligand binding to *Pf*CCT constructs followed by ITC. (A) Titration of 301 μM *Pf*CCT MΔK (monomer concentration) with 9 mM CTP in the presence of 5 mM Mg^{2+}. (B) Titration of 257 μM *Pf*CCT MΔK with 30 mM ChoP in the presence of 5 mM Mg^{2+}. (C) Titration of 257 μM *Pf*CCT MΔK with 10 mM PP_i in the presence of 5 mM Mg^{2+}. (D) Titration of 219 μM *Pf*CCT MΔK with 3.5 mM CDPCho in the presence of 5 mM Mg^{2+}. (E) Titration of 360 μM *Pf*CCT MΔK with 7.5 mM of CTP in the presence of 0.5 mM EDTA. (F) Titration of 329 μM *Pf*CCT MΔK with 12 mM ChoP. (G) Titration of 356 μM *Pf*CCT MΔK with 10 mM PP_i. (H) Titration of 146 μM *Pf*CCT MΔK with 2.0 mM CDPCho. (I) Titration of 257 μM *Pf*CCT MΔK with 3 mM CDPCho in the presence of 5.0 mM ChoP and 5 mM Mg^{2+}. Note that the presence of 5.0 mM ChoP has no effect on the binding characteristics of CDPCho (cf. D). (J) Titration of 329 μM *Pf*CCT MΔK with 3.0 mM CDPCho in the presence of 2.0 mM ChoP. Note that the presence of 2.0 mM ChoP has no effect on the binding characteristics of CDPCho (cf. H). (K) Titration of 165 μM *Pf*CCT MΔK^{W743H} with 3.0 mM CDPCho. (L) Titration of 179 μM *Pf*CCT MΔK^{W692Y} with 2.5 mM CDPCho. Integrated heat curves were corrected by subtracting the heat of dilution.

with the CDPCho complex was found to be 19.2 ± 2.4 μM, which is in adequate agreement with the value of 44 ± 2 μM determined by ITC (Table 2). Titration of *Pf*CCT MΔK with the ChoP substrate is not accompanied by either an increase of fluorescence intensity or a blue shift of the fluorescence emission maximum (Fig. 8B). For the *Pf*CCT MΔK^{W743H} mutant, titration with CDPCho resulted in a similar intensity increase and a shift of the emission maximum compared with the wild type, with an observed binding constant of 32.4 ± 5.8 μM (Fig. 8B). In contrast, no such fluorescence signal change was observed upon titration of *Pf*CCT MΔK^{W692Y} with CDPCho (Fig. 8B). We conclude that the changes of fluorescent emission spectral parameters induced by binding of CDPCho to the enzyme are due to the W692 residue present within the active site.

Discussion

Here we report a detailed ligand binding and kinetic study of a truncated construct of *Pf*CCT (*Pf*CCT MΔK) that contains a single catalytic domain but lacks the membrane binding domain. To evaluate compatibility of the construct with the full-length *Pf*CCT sequence, we investigated the conservation pattern of the duplicated CMP segments in Apicomplexa. Results of the phylogenetic analyses indicate two independent ancestral gene duplication events in Haemosporida (Plasmodia) and Piroplasmida, the former having occurred more recently. This indicates a reduced possibility for the functional divergence of the C1 and C2 domains. Widespread sequence conservation and the rat CCT crystal structure [23] enabled the creation of a homology model of the investigated *Pf*CCT construct containing the C2 domain. These considerations together with the present results indicating dimerization and the steady-state kinetic characteristics of *Pf*CCT MΔK verify that this construct constitutes a reliable catalytically functional model for the physiologically expected C1/C2 pseudo-heterodimer catalytic domain assembly of full-length *Pf*CCT.

The observed competitive inhibition pattern of CDPCho versus the two substrates indicates that CDPCho can bind to the same enzyme conformation as either substrate. Interpretation of our product inhibition measurements in the light of the above considerations suggests a random binding mechanism for *Pf*CCT MΔK, in agreement with previous data for rat CCT [22]. In such a kinetic mechanism, either substrate (CTP or ChoP) can bind to the apoenzyme on its own. However, in our ligand binding experiments based on either fluorescence spectroscopy or ITC, no binding of ChoP to the apoenzyme could be observed, in contrast to the suggested kinetic mechanism. It has to be emphasized that both techniques could be used with success to follow binding of CDPCho to the apoenzyme, and although fluorescence spectroscopy is an indirect method ITC is considered to be a generally applicable direct technique to evaluate molecular inter-

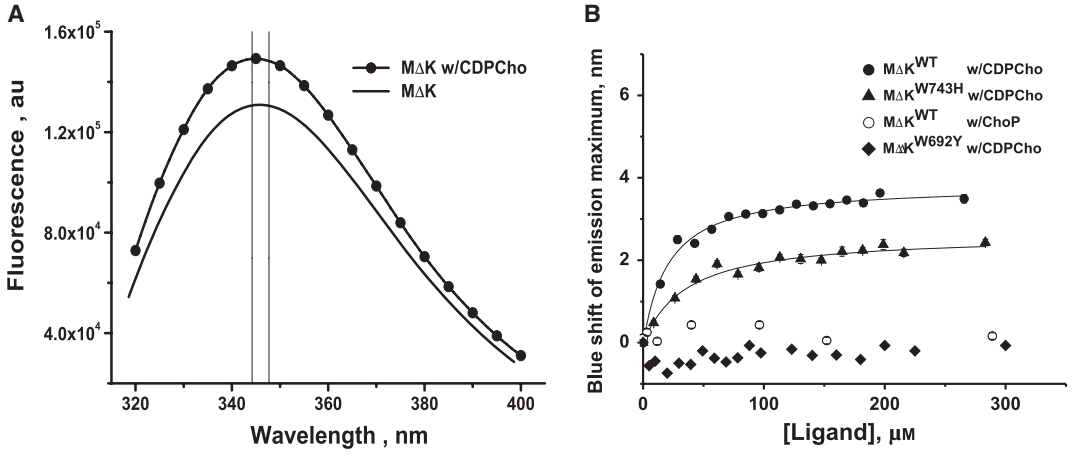

Fig. 8. Equilibrium CDPCho binding to *Pf*CCT MΔK constructs by Trp fluorescence. (A) Emission spectrum of *Pf*CCT MΔK in the absence of ligands (line without symbols) and in the presence of 100 μM CDPCho (line with full circles). Note the intensity increase and the blue shift upon CDPCho addition. (B) Equilibrium binding titration experiments of *Pf*CCT MΔK and its point mutants with CDPCho. The shift of the emission maximum in the presence of different amounts of ligand was analysed by fitting the Gauss peak function to the emission spectrum (see Materials and methods). *Pf*CCT MΔK titrations with CDPCho and ChoP are shown as full and open circles, respectively. Titrations of *Pf*CCT MΔK^{W743H} and *Pf*CCT MΔK^{W692Y} with CDPCho are shown as full triangles and full diamonds, respectively. Note the absence of the emission maxima shift upon ChoP titration.

actions. To reconcile the data from the kinetic and ligand binding experiments, it is important to realize that kinetic studies have their own limitations as well. Accordingly, it was shown that an observed kinetic pattern for the random mechanism may not identify the relevant set of elementary steps [36].

The fluorimetric titrations in our study rely on the signal of W692, located in the choline subsite of the active site. The homology model reveals that this residue possibly monitors CDPCho binding to the active site by juxtaposing the trimethylammonium moiety of CDPCho to the indole ring of W692 (Fig. 7). The lack of any fluorescent signal upon ChoP addition is probably due to a different microenvironment of the choline subsite. ITC experiments also including displacement ligand titrations reported considerable heat effects for CDPCho binding, whereas no such heat effects were found for ChoP, strengthening the suggestion that the interaction of the ChoP moiety with the choline binding subsite in the enzyme depends on whether ChoP is present on its own or within CDPCho.

Mg^{2+} was reported to be a cofactor of cytidylyltransferases [35]; however, we found that CTP and CDPCho bind to the CCT enzyme even in the absence of Mg^{2+} (cf. Fig. 6C,F). This correlates well with successful co-crystallization of enzymes belonging to the cytidylyltransferase family with CDPCho and CTP in the absence of Mg^{2+} [23,24]. Moreover, we also observed that addition of Mg^{2+} significantly decreased the binding affinity of CCT towards CTP. In the absence of the divalent cation, binding of the nucleotide needs to be accomplished by adequate charge compensation of the phosphate chain, presumably performed by positively charged and polar residues. Considering the sequence conservation within the catalytic domain of CCTs as well as the structural models, residues R755 and K663 as well as T761 and T762 [37] may be proposed to fulfil such coordinating roles for the phosphate chain. We propose that Mg^{2+} may be important for efficient catalysis but not for nucleotide binding for CCT enzymes belonging to the HxGH enzyme superfamily [38]. Lack of requirement for Mg^{2+} during nucleotide binding to the enzyme may also present a possible mode of action in the HxGH enzyme superfamily. Such a situation was experimentally verified for a class I aminoacyl tRNA synthetase that binds its cognate nucleotide ATP with considerably higher affinity in the absence rather than in the presence of Mg^{2+} [39].

In our study, we delineated the possible evolutionary steps leading towards Plasmodia CCT proteins, identified a useful fluorophore within the active site of CCT and provided insights into kinetic and ligand binding characteristics of a *Pf*CCT construct that can be expressed and purified for detailed studies. Based on these results, further research directions are suggested to focus on a detailed in-depth kinetic study of the elementary steps of CCT catalysis and on generating experimental three-dimensional structures to provide a full coherent knowledge of the mechanism of action of this important enzyme.

Materials and methods

Chemicals

Restriction enzymes, T4 DNA ligase, DNA polymerases and DNA purification kit were obtained from Machenary-Nagel (Düren, Germany). CTP, CDPCho, Sypro Orange, inorganic pyrophosphatase, purine nucleoside phosphorylase and antibiotics were purchased from Sigma-Aldrich (St Louis, MO, USA). ChoP was from TCI Europe N.V. (Antwerp, Belgium).

Isopropyl thio-β-D-galactoside was obtained from Fisher Scientific GmbH (Schwerte, Germany). MESG (2-amino-6-mercapto-7-methylpurine ribonucleoside) was obtained from Berry and Associates (Dexter, MI, USA). Nickel-nitrilotriacetic acid was from Qiagen (Düsseldorf, Germany), protease inhibitor cocktail tablets were purchased from Roche (Basel, Switzerland). All other chemicals were of analytical grade of the highest purity available.

Phylogenetic analyses

A first data set containing the amino acid sequences of CCT proteins was assembled from GenBank for 16 Apicomplexan taxa including species belonging to the genera *Crytosporidium*, *Neospora*, *Toxoplasma*, *Babesia*, *Theileria* and *Plasmodium*. These full-length protein sequences were aligned using the probabilistic consistency-based program PROBCONS [40] with default parameters on the Phylogeny.fr portal [41]. This resulted in an alignment with 1122 amino acid sites. In order to reconstruct the history of domain duplication within these proteins, a second data set was constructed by extracting the regions of the full alignment corresponding to the N-terminal and C-terminal CMP segments, leading to 16 and 11 sequences, respectively. Indeed, only members of the genera *Babesia*, *Theileria* and *Plasmodium* possess a duplicated C-terminal domain, whereas *Crytosporidium*, *Toxoplasma* and *Neospora* exhibit a CCT gene with a single CMP segment. The amino acid sequences of the 27 extracted domains were realigned using the phylogeny-aware PRANK program on the webPRANK server [42] with default parameters. This resulted in a final alignment with 362 sites.

Ambiguously aligned sites were subsequently removed from both data sets using the GBLOCKS program [43] with

default relaxed parameters. This led to final alignments with 456 and 217 amino acid sites for the full gene and domain data sets, respectively. ML phylogenetic reconstruction was conducted on these two data sets using PHYML 3.0 [44] with NNI branch swapping on a BIONJ starting tree. For both data sets, the LG+G4 model of protein sequence evolution was selected as the best fitting model by PROTTEST 3 [45]. The statistical robustness of phylogenetic inference was assessed by performing 100 bootstrap replicates using the same heuristic search strategy as for the initial analyses.

Homology modelling

Homodimer homology models have been built for constructs *Pf*CCT(528–795) and *Pf*CCT MΔK using the catalytic domain of the rat CCT which included the CDPCho product (PDB ID 3HL4) [23]. Both homology models were built using the same computational procedure. The alignment previously published in the literature [23] was slightly modified in the neighbourhood of the lysine-rich segment. The identity of the alignment was 42.21%. For homology modelling only residues 581–775 were considered, as in the rat CCT X-ray structure no template was found for the starting and end sequences of the construct. The MODELLER 9v8 software [46] was used to create 100 homology models for both constructs using the same alignment. The low energy conformations with low objective function values were visually inspected and one with appropriately coordinated CDPCho ligand was selected for each construct. The selected structures were optimized with CHARMM27 force field [47] and CHARMM software [48] using the GBSW implicit solvent model. The topology file and non-standard parameters added for the CDPCho ligand can be found in Supporting information.

The selected homology models were validated using the PROCHECK [49], WHAT_CHECK [50] and ERRAT [51] programs (Doc. S2, Figs S4 and S5, Tables S1 and S2). The two homology models exhibited very similar characteristics and differed significantly only in the presence/absence of the *Plasmodia* lineage-specific, lysine-rich segment. The models indicated that this insert lay reasonably far away from the protein domain that harbours the active site of the constructs; thus it is unlikely to have a major influence on the binding of the CDPCho ligand. Based on this and on the results of disorder predictions the homology model of *Pf*CCT MΔK protein was chosen for modelling the active site of *Pf*CCT with ligands.

Cloning and mutagenesis

The *Pf*CCT cDNA sequence (PF3D7_1316600) was codon-optimized for expression in *E. coli* (GenScript). The DNA fragment corresponding to residues 528–795 of *Pf*CCT was cloned into the expression vector pET15b (Novagen) using the *Nhe*I/*Bam*HI sites for production of N-terminal 6× His-tagged protein. Site-directed mutagenesis to produce the MΔK (528–795, Δ720–737) construct of *Pf*CCT lacking the lysine-rich *Plasmodium*-specific segment (720–737) and its W692Y and W743H point mutants was performed by the Quickchange method (Agilent) (mutagenesis primers are shown in Table S3). All constructs were verified by DNA sequencing.

Protein expression and purification

The recombinant His-tagged *Pf*CCT MΔK (528–795, Δ720–737) protein was expressed using the BL21 (DE3) Rosetta *E. coli* expression system. Briefly, 0.5 L of LB medium was inoculated with Rosetta cells from a 5 mL overnight culture and grown at 37 °C until A_{600} reached 0.4. After 30 min cooling at 20 °C, protein expression was induced by 0.5 mM isopropyl thiogalactoside for 20 h at 20 °C. The cells were harvested by centrifugation at 4000 ***g*** for 20 min and stored at −80 °C. Thawed cells were resuspended in 30 mL of 20 mM Hepes pH 7.5, 100 mM NaCl, 1 mM phenylmethylsulfonyl fluoride, 1% Triton X-100, 3 $\mu g \cdot mL^{-1}$ RNase, 3 $\mu g \cdot mL^{-1}$ DNase, EDTA containing protease inhibitor cocktail tablet (Roche), 1 mM benzamidine and 10 mM β-mercaptoethanol. The solution was sonicated and centrifuged, and then the supernatant was supplemented with 1.5 mM Mg^{2+} to form a chelate complex with EDTA present and was applied onto a pre-equilibrated benchtop nickel-nitrilotriacetic acid-agarose affinity-chromatography column (Qiagen). The protein was eluted with Hepes 20 mM, NaCl 100 mM (pH 7.5) (buffer A) containing 250 mM imidazole. The eluted samples were dialysed overnight in buffer A and the next day the sample was concentrated and further purified on a HiLoad XK-16 Superdex 200 column (GE Healthcare) previously equilibrated with buffer A. The purified sample was shown to be over 90% homogeneous based on SDS/PAGE analysis, concentrated to 10 $mg \cdot mL^{-1}$ by using an ultrafiltration membrane (Sartorius), then supplemented with 10% glycerol, and finally flash-frozen in liquid N_2 in aliquots and stored at −80 °C.

SDS/PAGE analysis of the purification procedure is shown in Fig. 4A. Protein concentration is in monomers throughout the whole paper. It was determined spectrophotometrically from the absorbance at 280 nm by using an extinction coefficient of 31 400 $M^{-1} \cdot cm^{-1}$ calculated on the basis of amino acid composition by using the PROTPARAM server [52] and a molecular weight of 30 898 for the His-tagged monomer (throughout the present study, molar concentrations of the CCT protein construct refer to the monomer). Absorbance measurements were carried out using a Nanodrop spectrophotometer. Protein samples were identified by mass spectrometric analysis of the tryptic digest on an LCQ-Fleet Ion Trap LC/MS.

Differential scanning fluorimetry

Solutions of 5× Sypro Orange and 0.5 mg·mL^{-1} protein were prepared in a final volume of 25 μL in 96-well PCR plates. The plates were sealed with adhesive PCR film (Thermo Scientific) and heated in a Stratagene Mx3000P real-time PCR instrument (Agilent) from 25 to 85 °C in increments of 0.5 °C. Three parallels were set up for each condition. To obtain the midpoint of the protein melting transition defined as a melting temperature T_m, a Boltzmann model (Eqn 1) was used to fit the fluorescence data using ORIGINPRO 8 (OriginLab Corp, Northampton, MA, USA).

$$F_T = F_0 + \frac{F_1 - F_0}{1 + \exp(T_m - T)/dT} \quad (1)$$

where F_T is the fluorescence intensity at temperature T, T_m is the melting temperature, and F_0 and F_1 are the pre-transitional and post-transitional fluorescence intensities, respectively. Data points after the fluorescence intensity maximum F_1 were excluded from fitting.

Mass spectometry

In the mass spectrometric study of protein complexes, a commercial Waters QTOF Premier instrument equipped with electrospray ionization source was used in the positive ion mode. Mass spectra were obtained under native conditions: namely, the ions were generated from aqueous 20 mM NH_4HCO_3 buffer solution (pH 7.8) containing the *Pf*CCT MΔK protein at 40 μM monomer concentration. These conditions allow transfer of the native protein complex present in the solution into the gas phase. The capillary voltage was 2800 V, the sampling cone voltage was 128 V and the temperature of the source was kept at 90 °C. Mass spectra were recorded in the mass range 1500–8000 m/z.

Fluorescence spectroscopy

Emission curves were recorded on a Jobin Yvon SpeX Fluoromax-3 spectrofluorometer using 10-mm path length 20 °C thermostatted cuvettes with excitation at 295 nm (slit 1 nm), emission between 320 and 400 nm (slit 5 nm). 3 μM of protein in buffer A was titrated with addition of 1–2 μL aliquots from concentrated stock ligand solutions. After 3 min incubation, three emission scans were recorded and then averaged. The emission maximum was determined by fitting the upper 20% fluorescence intensity of the emission curves with the Gauss distribution equation:

$$y = y_0 + A\exp[-(\lambda - \lambda_{max})^2/2\omega^2] \quad (2)$$

where λ is the wavelength in nanometers and λ_{max} is the calculated emission peak maximum in nanometers.

The dissociation constant of ligand binding was calculated by fitting to the calculated emission shift data with a quadratic equation (Eqn 3) as previously described [53–55]:

$$y = A[(c + K + x) - [(c + K + x)^2 - 4cx]^{0.5}]/2c \quad (3)$$

where x is the ligand concentration (μM), y is the shift in emission on ligand addition (nm), A is the emission shift observed at saturation, c is the enzyme concentration (μM) and K is the dissociation constant (μM).

Steady-state kinetic assay

A continuous coupled enzyme activity assay was employed based on [29]. The assay was performed in 10-mm path length 20 °C thermostatted cuvettes in a dual wavelength Specord 200 spectrophotometer, with a reaction volume of 600 μL. Enzyme and substrate concentrations were set such that CCT catalysis was the rate limiting step: [MΔK] = 0.16 μM, [inorganic pyrophosphatase] = 2.08 U·mL^{-1}, [purine nucleoside phosphorylase] = 2.08 U·mL^{-1}, [Mg^{2+}] = 5 mM, [MESG] = 0.33 mM, and variable concentrations of phosphocholine and CTP substrates in buffer A. The reaction was started by a final quick addition of *Pf*CCT MΔK to the reaction mixture. Initial velocity was determined from the slope of the first 10% of the progress curve. For CTP substrate titrations, CTP concentration was varied between 12 μM and 1.2 mM while ChoP concentration was kept at 5 mM. For ChoP substrate titrations, ChoP concentration was varied between 0.2 and 20 mM while CTP concentration was kept at 1 mM. Kinetic data were fitted with Eqns (4) and (5) (Michaelis–Menten equation and substrate inhibition equation, respectively):

$$v = \frac{V_{max}[S]}{K_M + [S]} \quad (4)$$

$$v = \frac{V_{max}}{1 + K_M/[S] + [S]/K_i} \quad (5)$$

In these equations, v is the reaction rate (s^{-1}), V_{max} is the maximum velocity of the reaction ($s^{-1}\cdot\mu M^{-1}$), [S] is the concentration of the substrate (μM) and K_i describes the binding of a substrate molecule to the enzyme (μM), with the result being a decrease by half in the rate of the reaction.

Product inhibition kinetic measurements

Product inhibition studies were performed using the above assay with CDPCho as inhibitor present in 0, 40 and 100 μM concentrations. Substrate concentration was varied in the presence of a fixed saturating concentration of the other substrate. The inhibition constant K_i was calculated by global fit of the CTP titration at saturated ChoP con-

centration curves with Eqn (6) describing full competitive inhibition:

$$v = V_{max}/[1 + (K_M/[S])(1 + [I]/K_i)] \quad (6)$$

ITC measurements

Calorimetric measurements were performed on a MicroCal-ITC 200 titration calorimeter (GE Healthcare) at 20 °C. The protein samples were dialysed against buffer A, including 10 mM β-mercaptoethanol or 0.5 mM TCEP (tris (2-carboxyethyl)phosphine) as reducing agent and 5 mM Mg^{2+} or 0.5 mM EDTA as indicated. Ligands were freshly dissolved in dialysis buffer and subsequently pH was adjusted, prior to ITC measurements.

CTP and CDPCho concentrations were determined using molar extinction coefficients 9000 $M^{-1}{\cdot}cm^{-1}$ at 271 nm, pH 7.0, and 5000 $M^{-1}{\cdot}cm^{-1}$ at 260 nm, respectively. In the experimental setup, the cell of the instrument was filled with protein and the syringe with the respective ligand. Each titration typically included 27 steps of injection with 1.5 μL of ligand per injection spaced 180 s apart from each other, with the injection syringe rotating at 500 r.p.m.

In the case of displacement ligand titrations, the enzyme previously titrated with ChoP, reaching a final concentration of 2–5 mM, was immediately further titrated with CDPCho. The data were analysed using MICROCAL ORIGIN software, following the directions of the manufacturer, in order to calculate the thermodynamic parameters: dissociation constant (K_d), stoichiometry (N), enthalpy (ΔH) and entropy (ΔS). The heats of dilution obtained from titration of the same ligand solutions into buffer were subtracted from the titration data before the data were fitted to the appropriate model.

Acknowledgements

This work was supported by the French Agence Nationale de la Recherche (ANR-09-BLAN-0397) (ML, HV, RC), the European Community (FP7/EviMalar Network of Excellence No. 242095) (HV, RC), ANR-NKTH ADD-MAL (Agence Nationale de la Recherche – Nemzeti Kutatasi es Fejlesztesi Hivatal, 'Antimalarial drug discovery and development of new *in vitro* assays for the optimization of antimalarial therapy') (BGV, GNN, LM) and the New Széchenyi Plan (TÁMOP-4.2.2/B-10/1-2010-0009) (JO, BK, BGV, GNN). JO acknowledges receipt of an EU Marie Curie ERG Fellowship (Project Oestrometab). This work has been supported by the Hungarian Scientific Research Fund [grant no. OTKA NK83857 (ÁR, KV) and NK84008 (BGV)]. GNN was supported by the Zsuzsa Szabó Foundation and the Richter Gedeon Centenarium Foundation. This study represents the contribution ISEM 2013-031 of the Institut des Sciences de l'Evolution de Montpellier.

References

1 WHO (2011) World Malaria Report 2011. World Health Organization, Geneva.
2 Murray CJL, Rosenfeld LC, Lim SS, Andrews KG, Foreman KJ, Haring D, Fullman N, Naghavi M, Lozano R & Lopez AD (2012) Global malaria mortality between 1980 and 2010: a systematic analysis. *Lancet* **379**, 413–431.
3 Fairhurst RM, Nayyar GM, Breman JG, Hallett R, Vennerstrom JL, Duong S, Ringwald P, Wellems TE, Plowe CV & Dondorp AM (2012) Artemisinin-resistant malaria: research challenges, opportunities, and public health implications. *Am J Trop Med Hyg* **87**, 231–241.
4 Dondorp AM, Fairhurst RM, Slutsker L, Macarthur JR, Breman JG, Guerin PJ, Wellems TE, Ringwald P, Newman RD & Plowe CV (2011) The threat of artemisinin-resistant malaria. *N Engl J Med* **365**, 1073–1075.
5 Dechamps S, Shastri S, Wengelnik K & Vial HJ (2010) Glycerophospholipid acquisition in Plasmodium – a puzzling assembly of biosynthetic pathways. *Int J Parasitol* **40**, 1347–1365.
6 Vial H & Mamoun C (2005) Plasmodium lipids: metabolism and function. In *Molecular Approach to Malaria* (Sherman I, ed.), pp. 327–352. ASM Press, Washington DC.
7 Vial HJ & Ancelin ML (1992) Malarial lipids. An overview. *Subcell Biochem* **18**, 259–306.
8 Vial HJ, Eldin P, Tielens AG & van Hellemond JJ (2003) Phospholipids in parasitic protozoa. *Mol Biochem Parasitol* **126**, 143–154.
9 Wengelnik K, Vidal V, Ancelin ML, Cathiard AM, Morgat JL, Kocken CH, Calas M, Herrera S, Thomas AW & Vial HJ (2002) A class of potent antimalarials and their specific accumulation in infected erythrocytes. *Science* **295**, 1311–1314.
10 Vial HJ, Wein S, Farenc C, Kocken C, Nicolas O, Ancelin ML, Bressolle F, Thomas A & Calas M (2004) Prodrugs of bisthiazolium salts are orally potent antimalarials. *Proc Natl Acad Sci USA* **101**, 15458–15463.
11 Wein S, Maynadier M, Bordat Y, Perez J, Maheshwari S, Bette-Bobillo P, Tran Van Ba C, Penarete-Vargas D, Fraisse L, Cerdan R *et al.* (2012) Transport and pharmacodynamics of albitazolium, a candidate antimalarial drug. *Br J Pharmacol* **166**, 2263–2276.
12 Dechamps S, Wengelnik K, Berry-Sterkers L, Cerdan R, Vial HJ & Gannoun-Zaki L (2010) 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.
13 Jackowski S (1994) Coordination of membrane phospholipid synthesis with the cell cycle. *J Biol Chem* **269**, 3858–3867.
14 Kent C (1995) Eukaryotic phospholipid biosynthesis. *Annu Rev Biochem* **64**, 315–343.
15 Kent C (2005) Regulatory enzymes of phosphatidylcholine biosynthesis: a personal perspective. *Biochim Biophys Acta* **1733**, 53–66.
16 Fagone P & Jackowski S (2012) Phosphatidylcholine and the CDP-choline cycle. *Biochim Biophys Acta* **1831**, 523–532.
17 Taneva S, Dennis MK, Ding Z, Smith JL & Cornell RB (2008) Contribution of each membrane binding domain of the CTP:phosphocholine cytidylyltransferase-alpha dimer to its activation, membrane binding, and membrane cross-bridging. *J Biol Chem* **283**, 28137–28148.
18 Yeo HJ, Larvor MP, Ancelin ML & Vial HJ (1997) *Plasmodium falciparum* CTP:phosphocholine cytidylyltransferase expressed in *Escherichia coli*: purification, characterization and lipid regulation. *Biochem J* **324**, 903–910.
19 Yeo HJ, Sri Widada J, Mercereau-Puijalon O & Vial HJ (1995) Molecular cloning of CTP:phosphocholine cytidylyltransferase from *Plasmodium falciparum*. *Eur J Biochem* **233**, 62–72.
20 Llinas M, Bozdech Z, Wong ED, Adai AT & DeRisi JL (2006) Comparative whole genome transcriptome analysis of three *Plasmodium falciparum* strains. *Nucleic Acids Res* **34**, 1166–1173.
21 Friesen JA, Campbell HA & Kent C (1999) Enzymatic and cellular characterization of a catalytic fragment of CTP:phosphocholine cytidylyltransferase alpha. *J Biol Chem* **274**, 13384–13389.
22 Veitch DP, Gilham D & Cornell RB (1998) The role of histidine residues in the HXGH site of CTP: phosphocholine cytidylyltransferase in CTP binding and catalysis. *Eur J Biochem* **255**, 227–234.
23 Lee J, Johnson J, Ding Z, Paetzel M & Cornell RB (2009) Crystal structure of a mammalian CTP: phosphocholine cytidylyltransferase catalytic domain reveals novel active site residues within a highly conserved nucleotidyltransferase fold. *J Biol Chem* **284**, 33535–33548.
24 Weber CH, Park YS, Sanker S, Kent C & Ludwig ML (1999) A prototypical cytidylyltransferase: CTP: glycerol-3-phosphate cytidylyltransferase from *Bacillus subtilis*. *Structure* **7**, 1113–1124.
25 Cornillot E, Hadj-Kaddour K, Dassouli A, Noel B, Ranwez V, Vacherie B, Augagneur Y, Bres V, Duclos A, Randazzo S *et al.* (2012) Sequencing of the smallest Apicomplexan genome from the human pathogen *Babesia microti*. *Nucleic Acids Res* **40**, 9102–9114.
26 Loo JA (2000) Electrospray ionization mass spectrometry: a technology for studying noncovalent macromolecular complexes. *Int J Mass Spectrom* **200**, 175–186.
27 Ashcroft AE (2005) Recent developments in electrospray ionisation mass spectrometry: noncovalently bound protein complexes. *Nat Prod Rep* **22**, 452–464.
28 Webb MR (1992) A continuous spectrophotometric assay for inorganic phosphate and for measuring phosphate release kinetics in biological systems. *Proc Natl Acad Sci USA* **89**, 4884–4887.
29 Lloyd AJ, Thomann HU, Ibba M & Soll D (1995) A broadly applicable continuous spectrophotometric assay for measuring aminoacyl-tRNA synthetase activity. *Nucleic Acids Res* **23**, 2886–2892.
30 Varga B, Barabas O, Takacs E, Nagy N, Nagy P & Vertessy BG (2008) Active site of mycobacterial dUTPase: structural characteristics and a built-in sensor. *Biochem Biophys Res Commun* **373**, 8–13.
31 Fromm H (1975) Initial Rate Enzyme Kinetics, 1st edn. Springer, Berlin.
32 Stevens SY, Sanker S, Kent C & Zuiderweg ER (2001) Delineation of the allosteric mechanism of a cytidylyltransferase exhibiting negative cooperativity. *Nat Struct Biol* **8**, 947–952.
33 Christensen T, Gooden DM, Kung JE & Toone EJ (2003) Additivity and the physical basis of multivalency effects: a thermodynamic investigation of the calcium EDTA interaction. *J Am Chem Soc* **125**, 7357–7366.
34 Mages F, Rey C, Fonlupt P & Pacheco H (1988) Kinetic and biochemical properties of CTP:choline-phosphate cytidylyltransferase from the rat brain. *Eur J Biochem* **178**, 367–372.
35 Tilley DM, Evans CR, Larson TM, Edwards KA & Friesen JA (2008) Identification and characterization of the nuclear isoform of *Drosophila melanogaster* CTP: phosphocholine cytidylyltransferase. *Biochemistry* **47**, 11838–11846.
36 Frieden C (1976) On the kinetic distinction of ordered and random bireactant enzyme systems. *Biochem Biophys Res Commun* **68**, 914–917.
37 Huang HK, Taneva SG, Lee J, Silva LP, Schriemer DC & Cornell RB (2012) 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.
38 Bork P, Holm L, Koonin EV & Sander C (1995) The cytidylyltransferase superfamily: identification of the nucleotide-binding site and fold prediction. *Proteins* **22**, 259–266.
39 Kapustina M, Weinreb V, Li L, Kuhlman B & Carter CW Jr (2007) A conformational transition state

accompanies tryptophan activation by *B. stearothermophilus* tryptophanyl-tRNA synthetase. *Structure* **15**, 1272–1284.
40 Do CB, Mahabhashyam MS, Brudno M & Batzoglou S (2005) ProbCons: probabilistic consistency-based multiple sequence alignment. *Genome Res* **15**, 330–340.
41 Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard JF, Guindon S, Lefort V, Lescot M *et al.* (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* **36**, W465–469.
42 Loytynoja A & Goldman N (2010) webPRANK: a phylogeny-aware multiple sequence aligner with interactive alignment browser. *BMC Bioinformatics* **11**, 579.
43 Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Mol Biol Evol* **17**, 540–552.
44 Guindon S, Delsuc F, Dufayard JF & Gascuel O (2009) Estimating maximum likelihood phylogenies with PhyML. *Methods Mol Biol* **537**, 113–137.
45 Darriba D, Taboada GL, Doallo R & Posada D (2011) ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* **27**, 1164–1165.
46 Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, Pieper U & Sali A (2006) Comparative protein structure modeling using Modeller. *Curr Protoc Protein Sci*. doi:10.1002/0471140864.ps0209s50.
47 MacKerell AD Jr, Banavali N & Foloppe N (2000) Development and current status of the CHARMM force field for nucleic acids. *Biopolymers* **56**, 257–265.
48 Brooks BR, Brooks CL III, Mackerell AD Jr, Nilsson L, Petrella RJ, Roux B, Won Y, Archontis G, Bartels C, Boresch S *et al.* (2009) CHARMM: the biomolecular simulation program. *J Comput Chem* **30**, 1545–1614.
49 Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R & Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **8**, 477–486.
50 Hooft RW, Vriend G, Sander C & Abola EE (1996) Errors in protein structures. *Nature* **381**, 272.
51 Colovos C & Yeates TO (1993) Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci* **2**, 1511–1519.
52 Artimo P, Jonnalagedda M, Arnold K, Baratin D, Csardi G, de Castro E, Duvaud S, Flegel V, Fortier A, Gasteiger E *et al.* (2012) ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res* **40**, W597–603.
53 Pecsi I, Szabo JE, Adams SD, Simon I, Sellers JR, Vertessy BG & Toth J (2011) Nucleotide pyrophosphatase employs a P-loop-like motif to enhance catalytic power and NDP/NTP discrimination. *Proc Natl Acad Sci USA* **108**, 14437–14442.
54 Pecsi I, Leveles I, Harmat V, Vertessy BG & Toth J (2010) Aromatic stacking between nucleobase and enzyme promotes phosphate ester hydrolysis in dUTPase. *Nucleic Acids Res* **38**, 7179–7186.
55 Varga B, Barabas O, Kovari J, Toth J, Hunyadi-Gulyas E, Klement E, Medzihradszky KF, Tolgyesi F, Fidy J & Vertessy BG (2007) Active site closure facilitates juxtaposition of reactant atoms for initiation of catalysis by human dUTPase. *FEBS Lett* **581**, 4783–4788.
56 Finn RD, Mistry J, Tate J, Coggill P, Heger A, Pollington JE, Gavin OL, Gunasekaran P, Ceric G, Forslund K *et al.* (2010) The Pfam protein families database. *Nucleic Acids Res* **38**, D211–222.
57 Yang ZR, Thomson R, McNeil P & Esnouf RM (2005) RONN: the bio-basis function neural network technique applied to the detection of natively disordered regions in proteins. *Bioinformatics* **21**, 3369–3376.
58 Dosztanyi Z, Csizmok V, Tompa P & Simon I (2005) The pairwise energy content estimated from amino acid composition discriminates between folded and intrinsically unstructured proteins. *J Mol Biol* **347**, 827–839.
59 Ward JJ, McGuffin LJ, Bryson K, Buxton BF & Jones DT (2004) The DISOPRED server for the prediction of protein disorder. *Bioinformatics* **20**, 2138–2139.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site:

Fig. S1. Alignment of 16 full-length Apicomplexan CCT sequences.
Fig. S2. Effect of the lineage-specific lysine-rich segment on the thermal stability of CCT constructs.
Fig. S3. Gel filtration of *Pf*CCT MΔK.
Fig. S4. Ramachandran plot of the template and the homology models.
Fig. S5. ERRAT results.
Doc. S1. Validation of the homology models.
Doc. S2. Topology file and non-standard parameters added for the CDPCho ligand.
Table S1. Summary of PROCHECK results.
Table S2. Summary of WHAT_CHECK results.
Table S3. Primers used for site-directed mutagenesis.