The Malaria Parasite’s Chloroquine Resistance Transporter is a Member of the Drug/Metabolite Transporter Superfamily

Rowena E. Martin and Kiaran Kirk
School of Biochemistry and Molecular Biology, Faculty of Science, The Australian National University, Canberra, Australia

The malaria parasite’s chloroquine resistance transporter (CRT) is an integral membrane protein localized to the parasite’s acidic digestive vacuole. The function of CRT is not known and the protein was originally described as a transporter simply because it possesses 10 transmembrane domains. In wild-type (chloroquine-sensitive) parasites, chloroquine accumulates to high concentrations within the digestive vacuole and it is through interactions in this compartment that it exerts its antimalarial effect. Mutations in CRT can cause a decreased intravacuolar concentration of chloroquine and thereby confer chloroquine resistance. However, the mechanism by which they do so is not understood.

In this paper we present the results of a detailed bioinformatic analysis that reveals that CRT is a member of a previously undefined family of proteins, falling within the drug/metabolite transporter superfamily. Comparisons between CRT and other members of the superfamily provide insight into the possible role of the protein and into the significance of the mutations associated with the chloroquine resistance phenotype. The protein is predicted to function as a dimer and to be oriented with its termini in the parasite cytosol. The key chloroquine-resistance-conferring mutation (K76T) is localized in a region of the protein implicated in substrate selectivity. The mutation is predicted to alter the selectivity of the protein such that it is able to transport the cationic (protonated) form of chloroquine down its steep concentration gradient, out of the acidic vacuole, and therefore away from its site of action.

Introduction

The emergence and spread of malaria parasites that are resistant to the widely-used antimalarial drug chloroquine (CQ) has been a disaster for world health. CQ is a weak base that accumulates in the parasite’s digestive vacuole, a lysosomal compartment in which haemoglobin, taken up from the host cell cytosol via an endocytotic feeding mechanism, is degraded to its component peptides and haem. Within the vacuole CQ interferes with the mechanism by which the potentially toxic haem monomers are converted to the inert crystalline substance haemozoin, causing monomeric haem to accumulate to levels that kill the parasite.

Two proteins—Pgh1 (P-glycoprotein homolog 1; Reed et al. 2000) and CRT (CQ resistance transporter; Fidock et al. 2000)—have been implicated as playing a role in CQ resistance. Both are integral membrane proteins, localized to the parasite’s digestive vacuole membrane. In *Plasmodium falciparum*, the most virulent of the human malaria parasites, mutations in the CRT protein (PICRT) confer CQ resistance on otherwise sensitive parasite strains (Sidhu, Verdis Petard, and Fidock 2002). CQ-resistant parasites have a markedly reduced concentration of CQ in their digestive vacuole (Fitch 1970; Saliba, Folb, and Smith 1998); however, neither the mechanism by which PICRT influences the intravacuolar concentration of the drug nor the normal physiological role of this protein are understood.

In their original description of the protein, Fidock et al. (2000) described PICRT, together with orthologs from other *Plasmodium* species and a more distant homolog from the slime mould *Dictyostelium discoideum*, as being putative channels or transporters containing 10 transmembrane domains (TMDs). Very recently, two preliminary reports have assigned the PfCRT protein to the drug/metabolite transporter superfamily (Martin, Trueman, and Kirk 2003; Tran and Saier 2004). Here we present a detailed bioinformatic analysis of the protein and of the family and superfamily to which it belongs. Comparisons between PICRT and members of the superfamily provide insight into the possible role of the protein and into the significance of the mutations associated with the CQ resistance phenotype.

Materials and Methods

Identifying Homologs of PfCRT

The amino acid sequence of PICRT was queried against the NCBI nonredundant protein database using a BlastP (Altschul et al. 1990) search in which the query sequence was not masked in areas of low compositional complexity. This option was chosen because the filter sometimes excludes from the analysis stretches of hydrophobic residues that correspond to the membrane-spanning regions of the protein and which are therefore of biological interest. Proteins retrieved by this search but excluded from the subsequent phylogenetic study included duplicate sequences and truncated proteins. Several proteins displayed areas of reasonable alignment with PICRT (as appraised by eye), even though the Blast output suggested that the probability of homology was low (P > 10⁻⁴). A detailed bioinformatic analysis was performed on such proteins; for each, the number and position of putative TMDs was used to assess whether the regions of sequence similarity corresponded to regions of alignment between the predicted secondary structures of the retrieved protein and PICRT. Proteins that satisfied this criteria were then queried against the NCBI nonredundant protein database (using BlastP) and the Entrez Conserved Domain Database (using Reverse Position-Specific Blast; Marchler-Bauer et al. 2002) to determine their relationships to other proteins. PICRT was also queried against the Entrez Conserved Domain Database as well as the Conserved Domain Architecture Retrieval Tool database (Geer et al. 2002) and was the subject of a Position-Specific Iterated Blast
Alignment of Hydropathy Profiles

Hydropathy profile alignments were generated at http://bioinformatics.weizmann.ac.il/hydroph/ using Kyte-Doolittle (x - 1) values, a window size of 17, and an algorithm that introduces gaps into the alignment to find the best match between two profiles. The final alignment of the hydropathy plots was compiled and edited in Adobe® Photoshop® 6.0.1.

Secondary Structure Predictions

TMpred (www.ch.embnet.org/software/TMPRED_form.html) and TMHMM (www.cbs.dtu.dk/services/TMHMM-2.0/) were used to detect putative membrane-spanning domains in the sequences of the proteins of interest. Predictions of protein orientation were made on the basis of the "positive inside" rule (von Heijne 1986; van Klompenburg et al. 1997) as well as by TMMHM (which incorporates the positive inside rule in its prediction of membrane protein topology [Sonhammer, von Heijne, and Krogh 1998]). The predicted secondary structure of loop 7 of the CRT family proteins was obtained using the PredictProtein server (http://cubic.bioc.columbia.edu/predictprotein/).

Construction of Alignments

The ClustalW program (Thompson, Higgins, and Gibson 1994) in MacVector™ 7.1 was used to generate and edit alignments. Sequences for various DMT proteins have been described (e.g., Jack, Yang, and Saier 2001), and these were used to retrieve many DMT members from the NCBI database. Proteins of different DMT families have diverged considerably at the amino acid level and this can make a one-step alignment method error-prone. We therefore first aligned proteins within a family and then used the ClustalW profile-alignment tool to assemble the families into one large alignment. This alignment corresponded very well with the alignment of the predicted TMDs in these proteins. The first half of the DMT superfamily alignment was aligned to the second using the ClustalW profile alignment tool.

Phylogenetic Analyses

Regions of the alignment that could not be aligned unambiguously were excluded prior to analysis. A phylogenetic tree was estimated using the Neighbor-Joining method (Saitou and Nei 1987) and uncorrected ("p") amino acid distances in MacVector™ 7.1. Ties in the tree were resolved randomly and a bootstrap analysis (Felsenstein 1985) was performed with 1,000 replicates.

Results and Discussion

PfCRT Belongs to a Family of Proteins

The proteins with the greatest similarity to PfCRT are homologs from other Plasmodia species (P. vivax, P. knowlesi, P. yoelii yoelii, P. chabaudi, and P. berghei) and these are retrieved by a BlastP search with P values in the range of 10⁻¹⁶² to 10⁻¹³⁷. The P value is the probability that the sequence similarity shared by the query protein and the retrieved protein arose by chance. A P value < 10⁻⁴ is considered to indicate a significant sequence similarity between the two proteins, consistent with having a related biological function. Following the Plasmodia proteins, the next best hits against PfCRT include a protein from Cryptosporidium parvum (1 × 10⁻²⁵), the D. discoideum protein (4 × 10⁻²⁰), and several proteins from Arabidopsis thaliana (≥ 2 × 10⁻⁵). A phylogenetic analysis performed on the sequence alignment of the Plasmodia, C. parvum, D. discoideum, and A. thaliana proteins (fig. 1) provides good evidence in support of the hypothesis that these proteins form a family.

The CRT Proteins Are Related to Known Transporters

The search for relatives of PfCRT in the NCBI database retrieved many proteins that had P values > 10⁻⁴, indicating low, if any, homology to PfCRT. Nevertheless, several of these proteins showed a reasonable similarity in sequence to PfCRT over specific regions of the alignment. All such proteins are members of the same group of transport proteins, the drug/metabolite transporter (DMT) superfamily. When the D. discoideum, C. parvum, and A. thaliana proteins were queried against the NCBI database, they also retrieved many proven or putative transporters of the DMT superfamily. Three iterations of a Position-Specific Iterated Blast search of the NCBI database using PfCRT as the query sequence retrieved, with good significance, several characterized transporters of the DMT superfamily. Proteins of the same family typically share distinct modules or domains that have a common evolutionary origin and function. The PfCRT protein was queried against the Entrez Conserved Domain Database using Reverse Position-Specific Blast and was found to have weak, but significant, hits to DMT superfamily conserved domains. These observations are consistent with the hypothesis that the proteins of the CRT family form part of, or are related to, the DMT superfamily.

Members of the CRT and DMT Families Have Similar Hydropathy Plots

A simple and commonly used predictor of the structure of a membrane protein is the hydropathy plot,
in which putative TMDs and the connecting hydrophilic, extra-membrane loops are detected as ‘peaks’ and ‘troughs,’ respectively, in a plot of the hydrophobicity index of the polypeptide. Transporters of the same family usually share a very similar hydrophathy plot, even when the underlying polypeptide sequences have diverged considerably (Lolkema and Slotboom 1998). Members of the DMT superfamily, like those of the CRT family, are predicted to contain 10 TMDs. Figure 2 shows an alignment of the hydropathy plots for PfCRT (yellow line), the *D. discoideum* protein (red line), one of the *A. thaliana* proteins (green line), and a protein from the DMT superfamily (the *E. coli* YdeD amino acid effluxer; blue line), consistent with these proteins having similar structure. In pairwise sequence alignments between members of the CRT family and the DMT protein, the regions of sequence similarity were found to correspond to the regions of alignment of predicted TMDs in the two proteins.

The CRT Family Falls Within the DMT Superfamily

The DMT superfamily consists of a number of discrete membrane protein families including the drug/metabolite effluxer (DME) family, the L-Rhamnose symporter (RhaT) family, the glucose/ribose permease (GRP) family, the *C. elegans* ORF (CEO) family (no members of which have been characterized as yet), the nucleotide-sugar transporter (NST) families, the triosephosphate transporter (TPT) family and the plant organocation permease (POP) family (Jack, Yang, and Saier 2001). An alignment of 368 proteins belonging to the DMT superfamily and CRT family was constructed from 10 CRT, 206 DME, 6 RhaT, 12 GRP, 7 CEO, 76 NST, 41 TPT, and 10 POP proteins. Four transporters from each of three unrelated 10-TMD transporter families—the Ca\(^{2+}\)-cation antiporters (CaCA), the C\(_4\)-dicarboxylate importers (Dcu), and the glutamate:Na\(^{+}\) symporters (ESS)—were included as outgroups.

The ‘best tree’ was estimated using the Neighbor-Joining method and the relationships within the superfamily were found to be similar to those described previously (Jack, Yang, and Saier 2001; Ward 2001; Knappe, Flugge, and Fischer 2003; Livshits et al. 2003; Martinez-Duncker et al. 2003). The NST and TPT families clustered together on a branch separate from the DME proteins and the DME family was comprised of many nodes, one of the outermost being the plant DME proteins. The RhaT, GRP, and CEO families clustered together and are distantly related to the DME proteins. The CRT family placed within the DMT superfamily, in which it branched from the DME proteins, after the RhaT, GRP, and CEO cluster had diverged.

Bootstrap analysis was performed on a reduced data set of 53 proteins that included representatives from the major groups within the DMT superfamily. As shown in figure 3, the bootstrap values support the placement of the CRT proteins within the DMT superfamily, where they cluster as a distinct family that branches between the DME and NST families. The analysis was repeated on another two subsets of ~50 proteins drawn from the full data set of 368 proteins and both yielded results similar to the tree presented in figure 3.

The CRT Proteins Arose from a Gene Duplication Event

Within members of the DMT superfamily there is significant homology between the first half of the protein (encompassing TMDs 1–5) and the second half (encompassing TMDs 6–10), consistent with these transporters having arisen from an ancestral gene that underwent an internal duplication event (Jack, Yang, and Saier 2001). Figure 4 shows alignments of the two halves of a representative selection of DMT proteins together with three members of the CRT family (including PfCRT). The first TMD is aligned with the sixth TMD, the second TMD with the seventh, and so on. As well as there being significant similarities along the whole lengths of the different proteins, there are similarities between the two halves of the proteins, including in the CRT proteins.

The Membrane Orientation of the CRT Proteins

As is evident from figure 4, there is a general trend within the DMT superfamily (including the CRT proteins) for the even-numbered extra-membrane loops to have a greater proportion of positively charged residues than the odd-numbered loops. The N- and C-termini also have
FIG. 4.—The alignment of the first half (encompassing TMDs 1–5) with the second half (encompassing TMDs 6–10) of a representative selection of DMT proteins, including three CRT family members. Residues are shaded as follows: positively charged, blue; negatively charged, red; tryptophan and tyrosine, white; polar, proline, and glycine, grey; remaining nonpolar, yellow. The conserved proline in TMDs 4 and 9 is highlighted in green. The variant N- and C-termini have been omitted. In some proteins loop 2 and/or 7 has been truncated and this is indicated by a solid black line. ‘(p)’ indicates that the assignment of function is only putative. The proteins are as follows: 1, *M. musculus* gi 249992 (Eckhardt et al. 1996); 2, *D. melanogaster* gi 21355345 (Martinez-Duncker et al. 2003); 3, *K. lactis* gi 6016590 (Abeijon, Robbins, and Hirschberg 1996); 4, *C. elegans* gi 20140026 (Berninsone et al. 2001); 5, *C. albicans* gi 1497106 (Nishikawa et al. 2002); 6, *H. sapiens* gi 14009667 (Lubke et al. 2001; Luhn et al. 2001); 7, *Z. mays* gi 1352200 (Fischer et al. 1994); 8, *A. thaliana* gi 25367911 (Santiviago et al. 2002); 9, *B. subtilis* gi 18977747; 10, *E. coli* gi 26343466 (Synchocystis sp. gi 61531731; 11, *M. magnetoelatum* gi 20215041; 12, *P. fluorescens* gi 30059755; 13, *V. cholerae* gi 1506061; 14, *A. tumefaciens* gi 8888063; 15, *P. aeruginosa* gi 15697297; 16, *B. fungorum* gi 22988226; 17, *S. japonicum* gi 27379386; 18, *B. subtilis* gi 19080453; 19, *B. cereus* gi 30319982; 20, *R. metallidurans* gi 22879738; 21, *E. coli* gi 121045; 22, *B. subtilis* gi 1977323; 23, *A. tumefaciens* gi 28211465; 24, *F. nucleatum* gi 19705391; 25, *L. plantarum* gi 28379273; 26, *G. max* gi 10768774; 27, *A. thaliana* gi 14415262; 28, *A. thaliana* gi 25408055; 29, *O. sativa* gi 34913778; 30, *A. thaliana* gi 21383591; 31, *P. falculorum* gi 23612473; 32, *D. discoideum* gi 1139714; 33, *M. musculus* gi 765558; 34, *H. sapiens* gi 30128268; 35, *C. albicans* gi 14971021; 36, *C. elegans* gi 17535737.

Fig. 3.—Neighbour-Joining tree of the DMT superfamily based on uncorrected distances. The analysis included three members of the CRT family and 38 DMT sequences. Four of the DMT proteins are characterized transporters of the nucleotide-sugar transporter (NST) family and 34 are known or putative members of the drug/metabolite efflux (DME) family. Included for comparison were four sequences from each of three unrelated 10 TMD transporter families: the Ca\(^{2+}\) cation antiporters (CaCA), the C\(_4\)-dicarboxylate importers (Dcu), and the Glutamate:Na\(^{+}\) symporters (ESS). Numbers indicate the percentage of 1,000 bootstrap replicates which support the topology shown. The scale bar represents the number of substitutions per site for a unit branch length.

**FIG. 4.**—The alignment of the first half (encompassing TMDs 1–5) with the second half (encompassing TMDs 6–10) of a representative selection of DMT proteins, including three CRT family members. Residues are shaded as follows: positively charged, blue; negatively charged, red; tryptophan and tyrosine, white; polar, proline, and glycine, grey; remaining nonpolar, yellow. The conserved proline in TMDs 4 and 9 is highlighted in green. The variant N- and C-termini have been omitted. In some proteins loop 2 and/or 7 has been truncated and this is indicated by a solid black line. ‘(p)’ indicates that the assignment of function is only putative. The proteins are as follows: 1, *M. musculus* gi 24999277 (Eckhardt et al. 1996); 2, *D. melanogaster* gi 21355345 (Martinez-Duncker et al. 2003); 3, *K. lactis* gi 6016590 (Abeijon, Robbins, and Hirschberg 1996); 4, *C. elegans* gi 20140026 (Berninsone et al. 2001); 5, *C. albicans* gi 1497106 (Nishikawa et al. 2002); 6, *H. sapiens* gi 14009667 (Lubke et al. 2001; Luhn et al. 2001); 7, *Z. mays* gi 1352200 (Fischer et al. 1994); 8, *A. thaliana* gi 25367911 (Santiviago et al. 2002); 9, *B. subtilis* gi 18977747; 10, *E. coli* gi 26343466 (Synchocystis sp. gi 61531731; 11, *M. magnetoelatum* gi 20215041; 12, *P. fluorescens* gi 30059755; 13, *V. cholerae* gi 1506061; 14, *A. tumefaciens* gi 8888063; 15, *P. aeruginosa* gi 15697297; 16, *B. fungorum* gi 22988226; 17, *S. japonicum* gi 27379386; 18, *B. subtilis* gi 19080453; 19, *B. cereus* gi 30319982; 20, *R. metallidurans* gi 22879738; 21, *E. coli* gi 121045; 22, *B. subtilis* gi 1977323; 23, *A. tumefaciens* gi 28211465; 24, *F. nucleatum* gi 19705391; 15, *M. lotii* gi 13473318; 16, *M. rubra* gi 20617188; 17, *G. max* gi 10768774; 18, *A. thaliana* gi 14415262; 19, *A. thaliana* gi 25408055; 20, *O. sativa* gi 34913778; 21, *A. thaliana* gi 21383591; 22, *P. falculorum* gi 23612473; 23, *D. discoideum* gi 1139714; 24, *M. musculus* gi 765558; 25, *H. sapiens* gi 30128268; 26, *C. albicans* gi 14971021; 27, *C. elegans* gi 17535737.
a preponderance of positive charge. This makes it likely that the even-numbered loops, together with the N- and C-termini, are located at the cytoplasmic face of the membrane (predicted by the positive inside rule [von Heijne 1986; van Klompenburg et al. 1997] and by TMHMM [Sonnhammer, von Heijne, and Krogh 1998]). Such an orientation (i.e., termini in the cytosol) has been proven experimentally for two DMT proteins, the mouse CMP-sialic acid transporter (NST family; Eckhardt, Gotza, and Gerardy-Schahn 1999) and the PecM protein of Erwinia chrysanthemi (DME family Rouanet and Nasser 2001).

For the RhaT, GRP, and CEO proteins it is the odd-numbered loops that have the higher proportion of positive charge, suggestive of the opposite topology, and this has been demonstrated experimentally for the S. typhimurium RhaT protein (for which the termini were found to be noncytosolic [Tate and Henderson 1993]).

The Significance of the Extra-Membrane Loops

Having established that PfCRT is a member of the DMT superfamily it is possible to assign putative functions to different regions of the PfCRT protein on the basis of previous studies of other members of the superfamily (fig. 4).

A striking feature of figure 4 is the conservation of structural elements throughout the superfamily, despite the considerable divergence in sequence, mode of transport, and substrate specificity. For instance, the length and composition of the loops are conserved between proteins from different families. Furthermore, the loop regions in the second half of the transporter bear significant similarity to the corresponding loops in the first half. Loops 3 and 8 and 4 and 9, are particularly well conserved. Studies with proteins from both the NST and DME families have revealed that the insertion of an epitope tag or reporter molecule in loops 3, 4, 8, and 9 can inactivate the transporter and in some instances cause the protein to be localized incorrectly within the cell and/or degraded (Tate and Henderson 1993; Eckhardt, Gotza, and Gerardy-Schahn 1999; Rouanet and Nasser 2001). By contrast, the same sequences introduced into loops 1, 2, 5, and 6 have no effect on transporter activity or localization.

Compared to the other loops, 2 and 7 show less sequence conservation and show significant variation in length between DMT proteins. In most DMT proteins loop 7 is a relatively long hydrophilic domain, changes in which have been shown to influence transporter activity. The activity of the mouse CMP-sialic acid transporter is reduced when an epitope is inserted into loop 7 (but not loop 2), and increasing the length of the tag causes a further reduction in transporter activity (Eckhardt, Gotza, and Gerardy-Schahn 1999). As illustrated in figure 5, loop 7 is especially long and well conserved in composition in the proteins of the CRT family, and the predicted structure of this region (given by the PredictProtein sever) was similar for each protein: a ‘compact globular domain’ that is formed by a nine amino acid alpha helix followed by two short beta sheets (although in some proteins the first beta sheet may instead be an alpha-helix).

The predicted orientation of PfCRT in the membrane (i.e., with the protein termini in the cytosol) places loop 7 in the digestive vacuole, where it may have a role in modulating the activity of the transporter.

The Presence of Helix Packing Motifs in Transmembrane Domains 5 and 10 Indicates that PfCRT May Function as a Dimer

While the structure of the DMT transporters has been retained over time and evolution, the underlying amino acid sequence has proven more plastic. Nevertheless, some regions of sequence have been strongly conserved. In TMDs 5 and 10 there are two conserved glycines that are separated by six hydrophobic residues (fig. 4; Knappe, Flugge, and Fischer 2003). This motif (GxxxxxxG) is a common feature of membrane proteins, in which it is thought to facilitate the packing of membrane-spanning helices, leading to the association of TMDs to form oligomers (Liu, Engelman, and Gerstein 2002). Other small residues (alanine, serine, and threonine) can replace one of the glycines in a glycine-packing motif (Russ and Engelman 2000; Eilers et al. 2002) and such a substitution has occurred in a number of the DMT proteins, including PfCRT (fig. 4). TMDs 5 and 10 are known to have a role in mediating the formation of homo-dimers by the NST and TPT transporters (Abe, Hashimoto, and Yoda 1999; Ishida et al. 1999; Streafield et al. 1999; Gao and Dean 2000) and nonconservative mutations within the putative packing motifs in TMDs 5 and 10 abrogate transporter activity as well as stability (Ishida et al. 1999; Streafield et al. 1999). Although it is not known whether members of the DME
family function as dimers, the conservation of the GxxxxxxG motif in DME (as well as CRT proteins) are consistent with their doing so.

Different Transmembrane Domains Are Implicated in Substrate Recognition, Binding, and Translocation

Many studies with NST or TPT proteins have shown that substitutions at conserved positions in TMDs 3, 4, 8, and 9 interfere with the binding and translocation of the substrate (a diverant alliance). The mutations either impair the rate of transport or abolish transport altogether, while not affecting either the localization of the protein within the cell or its ability to form homo-dimers (Abe, Hashimoto, and Yoda 1999; Beminson et al. 2001; Gao, Nishikawa, and Dean 2001; Luhre et al. 2001; Luhn et al. 2001; Oelmann, Stanley, and Gerardy-Schahn 2001; Ettioni et al. 2002; Knappe, Flugge, and Fischer 2003). A large proportion of these functionally important residues are located in TMDs 4 or 9; in both cases they are concentrated in the center of the helix where they form a strongly conserved region of polar and positively charged residues that is thought to constitute a substrate binding motif (Fischer et al. 1994; Gao, Nishikawa, and Dean 2001; Luhre et al. 2001).

The binding motif regions in TMDs 4 and 9 are also well conserved in DME and CRT proteins, but they do not contain positively charged residues. Instead, there is a conserved proline, usually in both TMDs 4 and 9, although a few proteins (including the Arabidopsis CRT protein) have a proline in only one of the binding motifs (fig. 4). The role of these conserved prolines in the function of DME transporters has not been studied, but proline residues located in the central part of a transmembrane helix are known to be essential for the biological activity of a number of membrane transport proteins (Webb, Rosenberg, and Cox 1992; Lin, Itokawa, and Uhl 2000; Shelden et al. 2001; Koike et al. 2004). The proline ring distorts the normal structure of a membrane-spanning helix via two mechanisms: a kink is introduced in the helix backbone to avoid a steric clash with the proline ring, and the hydrogen bonds that would normally stabilize this region of the helix are unable to form (Woolfson and Williams 1990; Visiers, Braunheim, and Weinstein 2000; Cordes, Bright, and Sansom 2002). This results in a flexible ‘hinge’ point in the helix that has the capacity to form hydrogen bonds (perhaps with a substrate). The presence of a proline hinge in the putative binding motif of DME and CRT proteins suggests a role for this residue in the binding and translocation of substrates.

The codons encoding both proline and glycine are GC-rich and, as such, these codons are the least likely to be retained by chance in the AT-rich genomes of Plasmodium and Dictyostelium (Stevens and Arkin 2000). This is consistent with there being a critical role for the putative helix-packing glycine motifs and the intramembrane proline residues in the function of CRT proteins.

While discrete regions of the DMT proteins are implicated in the binding and translocation of the substrate, other elements of the transporters have been implicated in the recognition of, and discrimination between, substrates. The participation of residues from a number of TMDs in substrate recognition is not uncommon among transporters. For example, in members of the major-facilitator superfamily, eight out of the 12 TMDs are predicted to be involved in determining the substrate specificity of the transporter (Hirai et al. 2003). The functional analyses of chimeras constructed from two human NSTs, the UDP-galactose transporter, and the CMP-sialic acid transporter, revealed the participation of TMDs 1, 2, 3, 7, and 8 in determining substrate specificity (Aoki, Ishida, and Kawakita 2001, 2003). The involvement of TMDs 3 and 8 is not surprising, as these domains are also thought to influence the binding and translocation of the substrate; residues in these helices are likely to face the translocation pore where they may interact with the substrate.

Amino acid residues are considered to be potentially ‘substrate-specific’ when they are conserved among transporters with identical substrates, but are different between transporters of differing substrate specificity. In DMT proteins there are many such residues present in TMDs 2 and 7 (fig. 4). Indeed, TMD 7 is one of the helices that varies the most in sequence between DMT proteins of different subgroups. Consistent with this observation, the substrate specificity of the UDP-galactose transporter is broadened to include CMP-sialic acid simply by replacing TMD 7 with the corresponding helix from the CMP-sialic acid transporter. TMD 1 is also crucial for determining the substrate specificity of hybrids between the UDP-galactose and CMP-sialic acid transporters, and it is thought to have a similar involvement in the substrate specificity of the UDP-galactose transporter from fruit fly (Segawa, Kawakita, and Ishida 2002). Similarly, in TPT proteins, residues in TMD 1 are proposed to line the translocation pore of the transporter (Knappe, Flugge, and Fischer 2003).

TMDs 7, 8, 9, and 10 all fulfill the same role in transporter function as the corresponding domains in the first half of the protein (fig. 6). It might therefore be expected that TMD 6, as the counterpart of TMD 1 in the second half of the protein, should also play a role in determining substrate specificity. Experimental support for this comes from a study with the hamster CMP-sialic acid transporter (Eckhardt, Gotza, and Gerardy-Schahn 1998). Mutation of a glycine residue in TMD 6 to glutamate, glutaminine, or isoleucine severely reduces transporter activity without affecting the expression or trafficking of the protein. Overexpression of the mutant proteins restores a low level of CMP-sialic acid transport, consistent with the mutations having affected the affinity of the transporter for its substrate.

The Chloroquine Resistance-Conferring K76 Mutation Lies in a Region Implicated in Substrate Selectivity

The substitution of the lysine at position 76 for threonine (K76T) has been identified as a crucial determinant of PICRT-mediated CQ resistance (Sidhu, Verdier-Pinard, and Fidock 2002). As depicted in figure 6, this mutation lies towards the C-terminal end of TMD 1, a region of the transporter that we predict to be involved in substrate recognition. Field isolates from both Old and New World strains of CQ-resistant parasites all have the critical
K76T mutation, accompanied by a number of what are thought to be ‘compensatory’ PfCRT mutations that enable the transporter to maintain a semblance of its normal physiological role (Carlton et al. 2001). The A. thaliana CRT proteins, like the PfCRT proteins of CQ-sensitive strains, have a positively charged residue in this position (fig. 5). By contrast, the D. discoideum and C. parvum proteins are more similar to the PfCRT proteins from resistant strains of P. falciparum in that they have a serine (serine and threonine are hydroxy amino acids and a S → T mutation is considered to be a conservative substitution).

A CQ-sensitive isolate from Sudan (106/1) contains all of the PfCRT mutations found in CQ-resistant strains from the Old World, with the exception of the K76T mutation which has reverted to the wild-type form (Fidock et al. 2000). When this strain was subjected to CQ-selection pressure in vitro, two CQ-resistant clones arose, both with novel mutations in the K76 position (K76I and K76N; Fidock et al. 2000; Cooper et al. 2002). Characterization of these mutants revealed that the K76I mutation had the unusual effect of increasing the parasite’s sensitivity to quinine while decreasing its sensitivity to the diastereomer quinidine. This adds further support to the hypothesis that TMD 1 of PfCRT, and in particular position 76, is involved in determining the substrate specificity of the transporter. The absence of the K76I and K76N mutations in CQ-resistant field populations (and the prevalence of the K76T mutation) may reflect a less fit phenotype in vivo for these artificially-derived mutants (Cooper et al. 2002). This is consistent with the observation that only two types of amino acid residues are found in this position among other proteins of the CRT family: positively charged or hydroxy.

Warhurst and colleagues (2002) have suggested previously that PfCRT is related to proteins of the chloride channel (ClC) family, members of which have been determined to possess 18 alpha helices (Dutzler et al. 2002). In their analysis, the region of PfCRT containing the K76T mutation was predicted to correspond to helix C of the Salmonella typhimurium ClC protein and the loss of the positive charge in this position was postulated to alter the specificity of the parasite chloride channel, such that it permitted the transport of chloroquine. However, residues in helix C of the S. typhimurium ClC protein are not known to influence substrate specificity, nor do they line either the channel’s selectivity filter or translocation pore (Dutzler et al. 2002). The mechanism by which the K76T mutation would influence the selectivity of a channel of this type, and thereby confer chloroquine resistance, is therefore unclear.

Apart from the K76T mutation, there is another PfCRT mutation (A220S) conserved in most CQ-resistant isolates and absent from CQ-sensitive strains (with the exception of the 106/1 strain). The A220S mutation is located in TMD 6 and does not confer resistance in the absence of the K76T mutation (e.g., 106/1). However, its presence in most CQ-resistant strains analyzed to date suggests that it acts in synergy with K76T, perhaps by...
aiding the recognition of CQ as a substrate for PfCRT or as a compensatory mutation that stabilizes the interaction of the transporter with its physiological substrate(s). The location of the A220 mutation in a PfCRT domain predicted to participate in substrate recognition is consistent with both of these scenarios.

Substrates Effluxed by DME Transporters

The members of the DMT superfamily bearing the closest similarity to the CRT proteins in the region of the substrate binding motif fall within the DME transporter subfamily. Substrates for DME transporters include amino acids, weak bases, and organic cations. The YdeD protein of E. coli exports cysteine metabolites (Dassler et al. 2000), whereas the E. coli YbiF protein exports a broad range of amino acids including homoserine, threonine, lysine, and histidine (Livshits et al. 2003). In other species of bacteria, DME proteins are implicated in the efflux of methyamine (MttP; Ferguson and Krzycki 1997), the di-cationic herbicide methyl viologen (YddG; Santiviago et al. 2002) and the pigment indigoidine, which is, like chloroquine, a weak base (PecM; Rouanet and Nasser 2001). The fact that DME transporters are known to transport both weak bases and divergent organic cations lends support to the hypothesis that the CQ-resistant form of PfCRT transports the chloroquine in the di-cationic form.

DME systems are postulated to be H\(^+\)-coupled and this has been confirmed experimentally for at least one DME transporter (the E. coli YbiF protein) [Livshits et al. 1993].

The Role of PfCRT in Chloroquine Resistance

Figure 7 shows a model for the mechanism of PfCRT-mediated resistance to CQ. The protein is shown as a dimer, functioning to export ‘metabolites’ (perhaps amino acids or peptides, and perhaps in symport with H\(^+\)) from the parasite’s digestive vacuole (DV). The ‘positive-inside’ rule, and a presumed inwardly-positive electrical potential across the DV membrane, predicts the N- and C-termini to be cytosolic and the ‘compact globular domain’ of loop 7 to be located at the vacuolar face of the membrane. CQ is a diprotic weak base and therefore accumulates in the acidic DV in the protonated (positively charged, CQ\(^{2+}\)) form. In parasites expressing ‘wild type’ PfCRT, the positive charge of K76 prevents the interaction of CQ\(^{2+}\) with the transporter. The CQ resistance-conferring K to T mutation removes the positive charge and alters the substrate selectivity, allowing CQ\(^{2+}\) to interact with the transporter and to be effluxed from the DV, perhaps in symport with H\(^+\). This results in a reduction of the concentration of CQ within the vacuole.

Substrates Effluxed by DME Transporters

The members of the DMT superfamily bearing the closest similarity to the CRT proteins in the region of the substrate binding motif fall within the DME transporter subfamily. Substrates for DME transporters include amino acids, weak bases, and organic cations. The YdeD protein of E. coli exports cysteine metabolites (Dassler et al. 2000), whereas the E. coli YbiF protein exports a broad range of amino acids including homoserine, threonine, lysine, and histidine (Livshits et al. 2003). In other species of bacteria, DME proteins are implicated in the efflux of methyamine (MttP; Ferguson and Krzycki 1997), the di-cationic herbicide methyl viologen (YddG; Santiviago et al. 2002) and the pigment indigoidine, which is, like chloroquine, a weak base (PecM; Rouanet and Nasser 2001). The fact that DME transporters are known to transport both weak bases and divergent organic cations lends support to the hypothesis that the CQ-resistant form of PfCRT transports the chloroquine in the di-cationic form.

DME systems are postulated to be H\(^+\)-coupled and this has been confirmed experimentally for at least one DME transporter (the E. coli YbiF protein) [Livshits et al. 1993].

The Role of PfCRT in Chloroquine Resistance

Figure 7 shows a model for the mechanism of PfCRT-mediated resistance to CQ. The protein is shown as a dimer, functioning to export ‘metabolites’ (perhaps amino acids or peptides, and perhaps in symport with H\(^+\)) from the parasite’s digestive vacuole (DV). The ‘positive-inside’ rule, and a presumed inwardly-positive electrical potential across the DV membrane, predicts the N- and C-termini to be cytosolic and the ‘compact globular domain’ of loop 7 to be located at the vacuolar face of the membrane. CQ is a diprotic weak base and therefore accumulates in the acidic DV in the protonated (positively charged, CQ\(^{2+}\)) form. In parasites expressing ‘wild type’ PfCRT, the positive charge of K76 prevents the interaction of CQ\(^{2+}\) with the transporter. The CQ resistance-conferring K to T mutation removes the positive charge and alters the substrate selectivity, allowing CQ\(^{2+}\) to interact with the transporter and to be effluxed from the DV, perhaps in symport with H\(^+\). This results in a reduction of the concentration of CQ within the vacuole.

Substrates Effluxed by DME Transporters

The members of the DMT superfamily bearing the closest similarity to the CRT proteins in the region of the substrate binding motif fall within the DME transporter subfamily. Substrates for DME transporters include amino acids, weak bases, and organic cations. The YdeD protein of E. coli exports cysteine metabolites (Dassler et al. 2000), whereas the E. coli YbiF protein exports a broad range of amino acids including homoserine, threonine, lysine, and histidine (Livshits et al. 2003). In other species of bacteria, DME proteins are implicated in the efflux of methyamine (MttP; Ferguson and Krzycki 1997), the di-cationic herbicide methyl viologen (YddG; Santiviago et al. 2002) and the pigment indigoidine, which is, like chloroquine, a weak base (PecM; Rouanet and Nasser 2001). The fact that DME transporters are known to transport both weak bases and divergent organic cations lends support to the hypothesis that the CQ-resistant form of PfCRT transports the chloroquine in the di-cationic form.

DME systems are postulated to be H\(^+\)-coupled and this has been confirmed experimentally for at least one DME transporter (the E. coli YbiF protein) [Livshits et al. 1993].

The Role of PfCRT in Chloroquine Resistance

Figure 7 shows a model for the mechanism of PfCRT-mediated resistance to CQ. The protein is shown as a dimer, functioning to export ‘metabolites’ (perhaps amino acids or peptides, and perhaps in symport with H\(^+\)) from the parasite’s digestive vacuole (DV). The ‘positive-inside’ rule, and a presumed inwardly-positive electrical potential across the DV membrane, predicts the N- and C-termini to be cytosolic and the ‘compact globular domain’ of loop 7 to be located at the vacuolar face of the membrane. CQ is a diprotic weak base and therefore accumulates in the acidic DV in the protonated (positively charged, CQ\(^{2+}\)) form. In parasites expressing ‘wild type’ PfCRT, the positive charge of K76 prevents the interaction of CQ\(^{2+}\) with the transporter. The CQ resistance-conferring K to T mutation removes the positive charge and alters the substrate selectivity, allowing CQ\(^{2+}\) to interact with the transporter and to be effluxed from the DV, perhaps in symport with H\(^+\). This results in a reduction of the concentration of CQ within the vacuole.

Experimental testing of the key aspects of the model presented in figure 7 await the expression of PfCRT in a heterologous system in a form in which the transport properties of the protein can be investigated directly. PfCRT has been successfully expressed in yeast (Zhang, Howard, and Roepe 2002); however, there has not, as yet, been any direct demonstration of its transport function. Efforts are...
presently underway to express the protein in *Xenopus laevis* oocytes and to measure and compare the transport of radiolabeled chloroquine via PfCRT with and without the K76T mutation. It is predicted that oocytes expressing PfCRT from CQ-resistant strains (i.e., having the K76T mutation) in their plasma membrane will transport $[^3H]$CQ, whereas those expressing wild-type PfCRT from CQ-sensitive strains (having K76) will not. The successful expression of PfCRT in *Xenopus* oocytes will also allow: (1) a direct test of the hypothesis that PfCRT proteins from both CQ-resistant and CQ-sensitive strains transport amino acids/peptides; (2) screening of other classes of substrate for their ability to be transported via PfCRT; (3) the determination of whether PfCRT-mediated transport is H$^+$ coupled; and (4) an investigation of whether, as has been proposed (Warhurst 2003), the chloroquine resistance reversal agent verapamil interacts directly with PfCRT (in TMD 1) to inhibit the transport of CQ. Such experiments have the potential to yield important insights into the molecular mechanism underlying chloroquine resistance.

**Acknowledgments**

The authors are grateful to Stephen Allen and Jan Martin for technical assistance and to Rhys Hayward, Kevin Saliba, and John Trueman for helpful discussions. This work was supported by the Australian National Health and Medical Research Council (grant ID 179804).

**Literature Cited**


assembly and export from the endoplasmic reticulum. J. Biol. Chem. 275:17718–17727.


development, and plastid-dependent nuclear gene expression. 
Tate, C. G., and P. J. Henderson. 1993. Membrane topology of 
the L-rhamnose-H⁺ transport protein (RhaT) from enter- 
ClustalW: improving the sensitivity of progressive multiple 
sequence alignment through sequence weighting, position- 
specific gap penalties and weight matrix choice. Nucleic 
Tran, C. V., and M. H. Saier, Jr. 2004. The principal chloroquine 
resistance protein of Plasmodium falciparum is a member of 
the drug/metabolite transporter superfamily. Microbiology 
150:1–3.
van Klompenburg, W., I. Nilsson, G. von Heijne, and B. de 
Kruijff. 1997. Anionic phospholipids are determinants of 
Visiers, I., B. B. Braunheim, and H. Weinstein. 2000. Prokink: 
a protocol for numerical evaluation of helix distortions by 
von Heijne, G. 1986. The distribution of positively charged 
residues in bacterial inner membrane proteins correlates with 
the trans-membrane topology. EMBO J. 5:3021–3027.
proteins from the model plant Arabidopsis thaliana. Bio- 
informatics 17:560–563.
Warhurst, D. C. 2003. Polymorphism in the Plasmodium falci- 
parum chloroquine-resistance transporter protein links 
verapamil enhancement of chloroquine sensitivity with the 
analysis of the Escherichia coli phosphate-specific transport 
system, a member of the traffic ATPase (or ABC) family of 
membrane transporters, a role for proline residues in trans-
Woolfson, D. N., and D. H. Williams. 1990. The influence of proline 
residues on alpha-helical structure. FEBS Lett. 277:185–188.
antimalarial drug resistance protein Pfcrt expressed in yeast. J. 

Michele Vendruscolo, Associate Editor
Accepted June 28, 2004