Article

Methotrexate recognition by the human reduced folate carrier SLC19A1

https://doi.org/10.1038/s41586-022-05168-0

Received: 7 February 2022

Accepted: 1 August 2022

Published online: 7 September 2022

Check for updates

Nicholas J. Wright[1,5], Justin G. Fedor[1,5], Han Zhang[2], Pyeonghwa Jeong[3], Yang Suo[1], Jiho Yoo[1,4], Jiyong Hong[3], Wonpil Im[2] & Seok-Yong Lee[1]✉

Folates are essential nutrients with important roles as cofactors in one-carbon transfer reactions, being heavily utilized in the synthesis of nucleic acids and the metabolism of amino acids during cell division[1,2]. Mammals lack de novo folate synthesis pathways and thus rely on folate uptake from the extracellular milieu[3]. The human reduced folate carrier (hRFC, also known as SLC19A1) is the major importer of folates into the cell[1,3], as well as chemotherapeutic agents such as methotrexate[4–6]. As an anion exchanger, RFC couples the import of folates and antifolates to anion export across the cell membrane and it is a major determinant in methotrexate (antifolate) sensitivity, as genetic variants and its depletion result in drug resistance[4–8]. Despite its importance, the molecular basis of substrate specificity by hRFC remains unclear. Here we present cryo-electron microscopy structures of hRFC in the apo state and captured in complex with methotrexate. Combined with molecular dynamics simulations and functional experiments, our study uncovers key determinants of hRFC transport selectivity among folates and antifolate drugs while shedding light on important features of anion recognition by hRFC.

Knockout of the gene encoding the reduced folate carrier (RFC) is embryonic lethal in mice[9], and dysfunction of RFC contributes to many disorders associated with folate deficiency, such as megaloblastic anaemia[10], foetal abnormalities, cardiovascular disorders and cancer[11,12]. Human RFC (hRFC) has been implicated as a critical target for chemotherapeutics as well as being a prognostic indicator owing to the increased demands for folates exhibited by rapidly dividing tumour cells[4]. Studies have unequivocally shown that RFC is the determinant for sensitivity to methotrexate (MTX) and related antifolates in tumour cells[4,6–8]. Further, as antifolate chemotherapies such as MTX, pemetrexed (PMX), pralatrexate (PDX) and raltitrexed (TDX) are readily transported by RFC, adaptations through mutations to the gene encoding RFC (*SLC19A1*) have led to several cases of drug resistance[4,5,7,12,13], while single-nucleotide polymorphisms (SNPs) or reduced expression of *SLC19A1* alters drug responses[13,14]. More recently, RFC has also been implicated as an important transporter of immunomodulatory cyclic dinucleotides, such as 2′,3′-cGAMP, which activate the STING pathway[15,16]. Augmenting RFC expression has also been identified as a potential measure to increase folate levels in the brain in cases of folate deficiency[17]. This expands the involvement of RFC to cancer immunotherapy, the host–pathogen response and treatment of neurometabolic disorders[15–17].

RFC is a member of the solute carrier 19 (SLC19) family of transporters that conform to the major facilitator superfamily (MFS) fold[3,18]. The SLC19 family has three members (SLC19A1–SLC19A3); SLC19A1 (RFC) exchanges anions, whereas SLC19A2 and SLC19A3 (ThTr1 and ThTr2, respectively) are organic cation carriers for thiamine[19]. SLC19A1 is a bidirectional folate exchanger with similar efflux and influx Michaelis constants for anion transport (K_t)[20]. Import of folates by RFC is powered by the counter-transport of organic anions, including thiamine mono- and pyrophosphate (TMP and TPP), for which there is a high transmembrane potential[21]. Although SLC19A1 exhibits a strong preference for folates and antifolates, it is broadly specific for a variety of anions, both organic (nucleotides and thiamine phosphates) and inorganic (chloride and phosphate), that act as lower-affinity counter-substrates[20,22]. While the function and importance of RFC have been explored since the 1960s, its structural basis for folate and antifolate specificity as well as anion exchange has not been elucidated[12]. On the other hand, the proton-coupled folate transporter (PCFT), the second route by which folates are taken up by the cell, is also considered a target for antifolate chemotherapeutics[23,24]. The recent PMX-bound PCFT structure provides the molecular basis of antifolate recognition by this transporter[25]. Identifying the nature of the ligand-binding site in hRFC through structural studies and comparing with PCFT would help immensely in the development of optimized therapeutics and overcoming drug-resistant cancers.

Structural elucidation

Xenopus laevis oocytes expressing wild-type (WT) hRFC exhibited time-dependent accumulation of [^{3}H]MTX (Fig. 1a), with uptake sensitive to competition by extracellular organic anions such as cGAMP (Fig. 1b). Further, this uptake was completely inhibited by the anti-rheumatic drug sulfasalazine, a known hRFC inhibitor[26] (Fig. 1b). hRFC is approximately 60 kDa in size and lacks any rigid extramembrane domains, so a fiducial marker is required for successful

[1]Department of Biochemistry, Duke University School of Medicine, Durham, NC, USA. [2]Departments of Biological Sciences, Chemistry and Bioengineering, Lehigh University, Bethlehem, PA, USA. [3]Department of Chemistry, Duke University, Durham, NC, USA. [4]Present address: College of Pharmacy, Chung-Ang University, Seoul, Republic of Korea. [5]These authors contributed equally: Nicholas J. Wright, Justin G. Fedor. ✉e-mail: seok-yong.lee@duke.edu

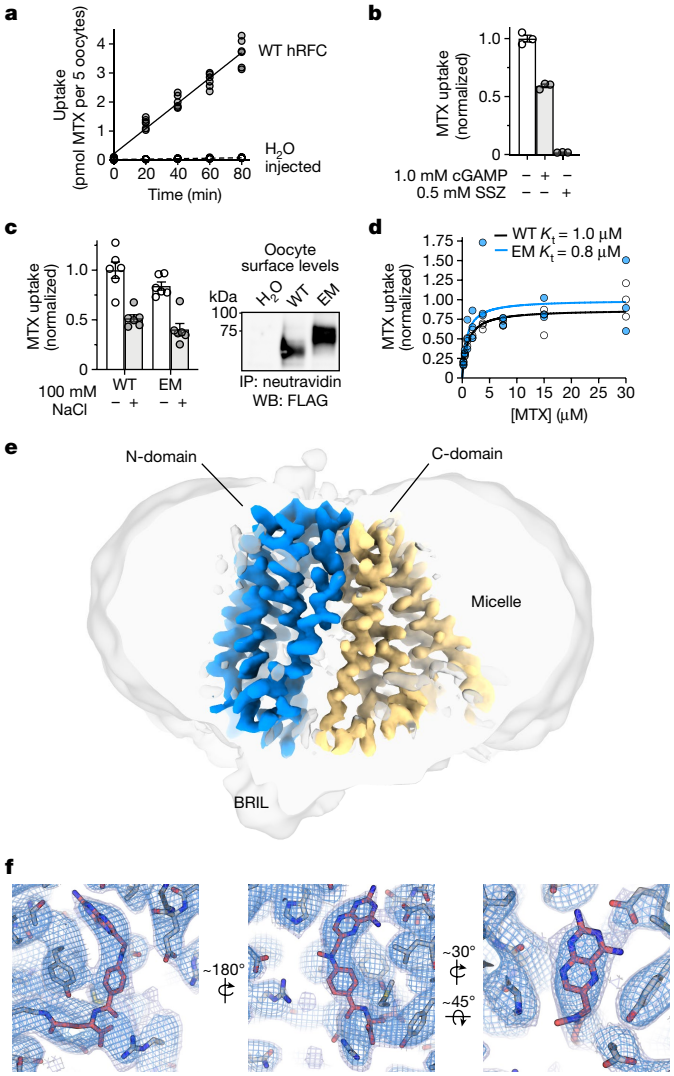


Fig. 1 | Structure of hRFC in complex with MTX. a, Time-dependent accumulation of 1.0 µM [^{3}H]MTX in *X. laevis* oocytes injected with cRNA encoding WT hRFC or water (n = 6 biological replicates, individual measurements shown). **b**, Cold competition of WT hRFC-mediated [^{3}H]MTX uptake (50 nM in 30 min) by cGAMP or sulfasalazine (SSZ; n = 3 biological replicates, shown as individual measurements and mean ± s.e.m.). **c**, Uptake of 1.0 µM [^{3}H]MTX into oocytes expressing WT hRFC or hRFC$_{EM}$ (EM) in 30 min, with relative oocyte membrane expression levels from surface biotinylation, immunoprecipitation (IP) and western blot (WB) analysis (representative blot shown; n = 6 biological replicates for the uptake assay with individual measurements and mean ± s.e.m. shown). Average signal in water-injected controls was used for background correction; average signal in the WT, no-NaCl condition was used for normalization. Uncropped blots shown in Supplementary Fig. 2. **d**, K_t curve for [^{3}H]MTX uptake into oocytes expressing WT hRFC in 30 min (n = 3 biological replicates, individual values shown). Uptake measurements were background corrected using water-injected controls and normalized to the average measurement in the 30 µM condition for each construct. **e**, Cryo-EM reconstruction at a resolution of 3.3 Å of MTX-modified hRFC, with the transporter N-domain shown in blue, the transporter C-domain shown in gold and the detergent micelle shown at a lower map threshold in transparent grey for reference (map threshold shown at 0.4 for transporter and 0.1 for detergent micelle). **f**, Cryo-EM volume corresponding to MTX in the transporter central cavity (two map thresholds shown; dark blue, 0.1; light blue, 0.05).

single-particle three-dimensional (3D) reconstruction. Instead of using a monoclonal antibody or nanobody[25,27], we replaced a short segment of the disordered loop connecting transmembrane helices (TMs) 6 and 7 (residues 215–241) with the engineered apocytochrome b_{562} variant BRIL[28] to enable cryo-electron microscopy (cryo-EM) analysis (Supplementary Fig. 1 and Extended Data Fig. 1a). In *X. laevis* oocytes, the resulting construct, hRFC$_{EM}$, exhibited surface expression and mediated chloride-sensitive uptake of MTX to levels comparable with WT protein (Fig. 1c). Both WT hRFC and hRFC$_{EM}$ exhibited MTX K_t values of ~1 µM (Fig. 1d), in line with previous reports for the WT carrier[20,22]. We first obtained a cryo-EM reconstruction of hRFC$_{EM}$ prepared in the presence of MTX to a resolution of 3.8 Å (Extended Data Fig. 2 and Extended Data Table 1). We term this structure hRFC$_{EM}$. The final reconstruction features weak signal for the apparently flexible BRIL domain (Extended Data Fig. 2c); therefore, the utility of BRIL as a fiducial during particle alignment is unclear. Unfortunately, we were unable to observe cryo-EM density within the central cavity corresponding to MTX. We therefore solved the true apo structure of hRFC (termed apo hRFC$_{EM}$) to a resolution of 3.6 Å (Extended Data Fig. 3a–d and Extended Data Table 1). Comparing the maps and models for hRFC$_{EM}$ with or without MTX (Cα root-mean-squared deviation of 0.1 Å) indicated that the structures are essentially identical (Extended Data Fig. 3e,f). While there are weak densities in roughly the same position within the cavity in both of these 3D reconstructions (Extended Data Fig. 3g), whether these peaks represent signal or noise is unclear. Considering their presence in both ligand-added and apo conditions, assignment is currently not possible.

We reasoned that there are three possibilities explaining the absence of a defined MTX density in hRFC$_{EM}$. First, MTX uptake activity is reduced by extracellular chloride (Fig. 1c), so chloride in the purification buffers probably hinders MTX binding to purified hRFC$_{EM}$ in vitro. Second, being able to accommodate a range of substrates, RFC may bind MTX dynamically within the cavity, broadening and weakening substrate density in the cryo-EM maps. Finally, as RFC is an exchanger, substrates may exhibit high off-rates in the conformation captured by cryo-EM. As a case in point, the crystal structure of the arginine transporter AdiC in complex with arginine required a mutation that stabilized ligand binding[29].

For the above reasons, an alternative approach was required, especially because the purified protein exhibited a lack of tolerance for reduced salt concentrations. *N*-hydroxysuccinimide-conjugated MTX (NHS–MTX) is a reagent reported to inhibit hRFC specifically and irreversibly through covalent modification of K411 (refs. [30–32]). Cell membranes containing overexpressed hRFC$_{EM}$ were thoroughly washed in a low-anion buffer and then treated with NHS–MTX, after which typical ionic conditions (150 mM NaCl) were restored for detergent extraction and subsequent purification. Spectral analysis of the resulting purified hRFC$_{EM}$ treated with NHS–MTX indicated a labelling ratio of 1:1.1 for hRFC$_{EM}$:MTX (Extended Data Fig. 1c). We then solved the structure of NHS–MTX-treated hRFC$_{EM}$ to an overall resolution of 3.3 Å (Fig. 1e, Extended Data Fig. 4a–d and Extended Data Table 1). The cryo-EM density corresponding to MTX in the focused maps is of good quality and facilitated unambiguous ligand placement (Fig. 1f). We term this structure hRFC$_{EM}$–MTX.

Environment of the central cavity

hRFC exhibits a canonical MFS transporter fold, where all three hRFC$_{EM}$ structures adopt an inward-facing conformation (Fig. 2a). Features unique to the hRFC fold include broken helices at TM1, TM4 and TM7, which line the central cavity in which MTX binds. Notably, the middle of TM1 features an unstructured region of approximately eight residues in length (Fig. 2a). The central cavity is highly conserved and consists of two regions of distinct surface electrostatics (Fig. 2b). More proximal to the intracellular matrix are charged residues R42, R133, R157, R373 and K411, which contribute to a highly electropositive surface potential.

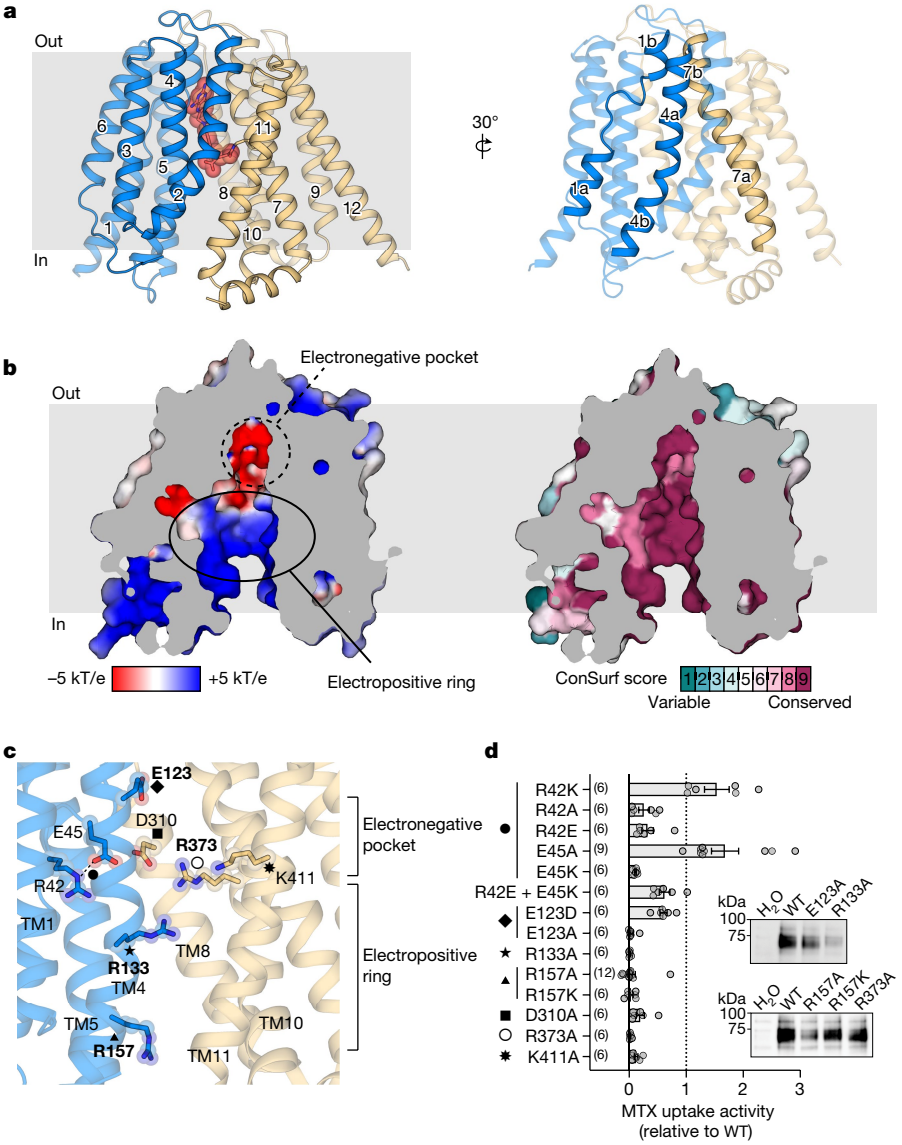


Fig. 2 | Transporter architecture and chemical environment of the hRFC central cavity. a, Overall structure of NHS–MTX-modified hRFC, with bent helices TM1, TM4 and TM7 highlighted at right. **b**, Electrostatic surface of the central cavity shown at left, with conservation analysis of the cavity-lining residues shown at right (electrostatics calculated with APBS[51] and conservation levels determined with ConSurf[52]; detailed description in Methods). **c**, Charged residues within the central cavity, with key residues labelled in bold. **d**, [^{3}H]MTX uptake activity of mutants of interest with mutations mapping to charged residues, background corrected and relative to WT (individual values and mean ± s.e.m. shown, with the number of biological replicates per condition indicated in brackets). Inset, relative oocyte membrane expression levels from surface biotinylation, immunoprecipitation and western blot analysis (representative blots shown). Uncropped blots shown in Supplementary Fig. 2.

Distal to the cavity opening are residues E45, E123 and D310, which contribute to the apparent electronegative surface potential at this site (Fig. 2c). We therefore term these central cavity regions the 'electropositive ring' and 'electronegative pocket', respectively. Mutants of select charged residues were assessed for MTX uptake activity in oocytes, most of which exhibited reduced or abolished activity relative to WT hRFC. In particular, the E123, R133, R157 and R373 residues are highly sensitive to charge perturbation, as their substitution to alanine resulted in little to no detectable MTX uptake activity above background (Fig. 2d). There is an absolute requirement for arginine at position 157, as little to no activity was detected for mutants with either alanine or lysine substitution (Fig. 2d). Oocyte surface expression was confirmed for these particular hRFC mutants (Fig. 2d). Residues R42, E45, D310 and K411 appear to exhibit less strict charge requirements, although charge elimination or substitution at these positions affects MTX uptake. Taken in concert with previous mutagenesis studies[18,32,33], our data highlight the functional importance of the unique chemical environment of the hRFC central cavity.

MTX recognition by hRFC

MTX occupies the central cavity of $hRFC_{EM}$ and is physically connected to the transporter through an amide covalent linkage, containing δ-carbon and ε-oxygen atoms of the L-glutamate moiety (L-Glu) of MTX and the ζ-nitrogen of transporter residue K411 (Fig. 3a). MTX comprises three groups: a pteridine ring, *p*-aminobenzoate (PABA) and L-Glu (Fig. 3b). Binding within the electropositive ring of hRFC, the MTX L-Glu moiety contacts TM4 by residue R133, through a close interaction with the α-carboxylate (Fig. 3a,b). Indeed, R133A substitution completely abolished uptake activity (Fig. 2d), and previous work has demonstrated the importance of the MTX α-carboxylate for hRFC-mediated uptake[32]. When comparing the $hRFC_{EM}$–MTX and apo

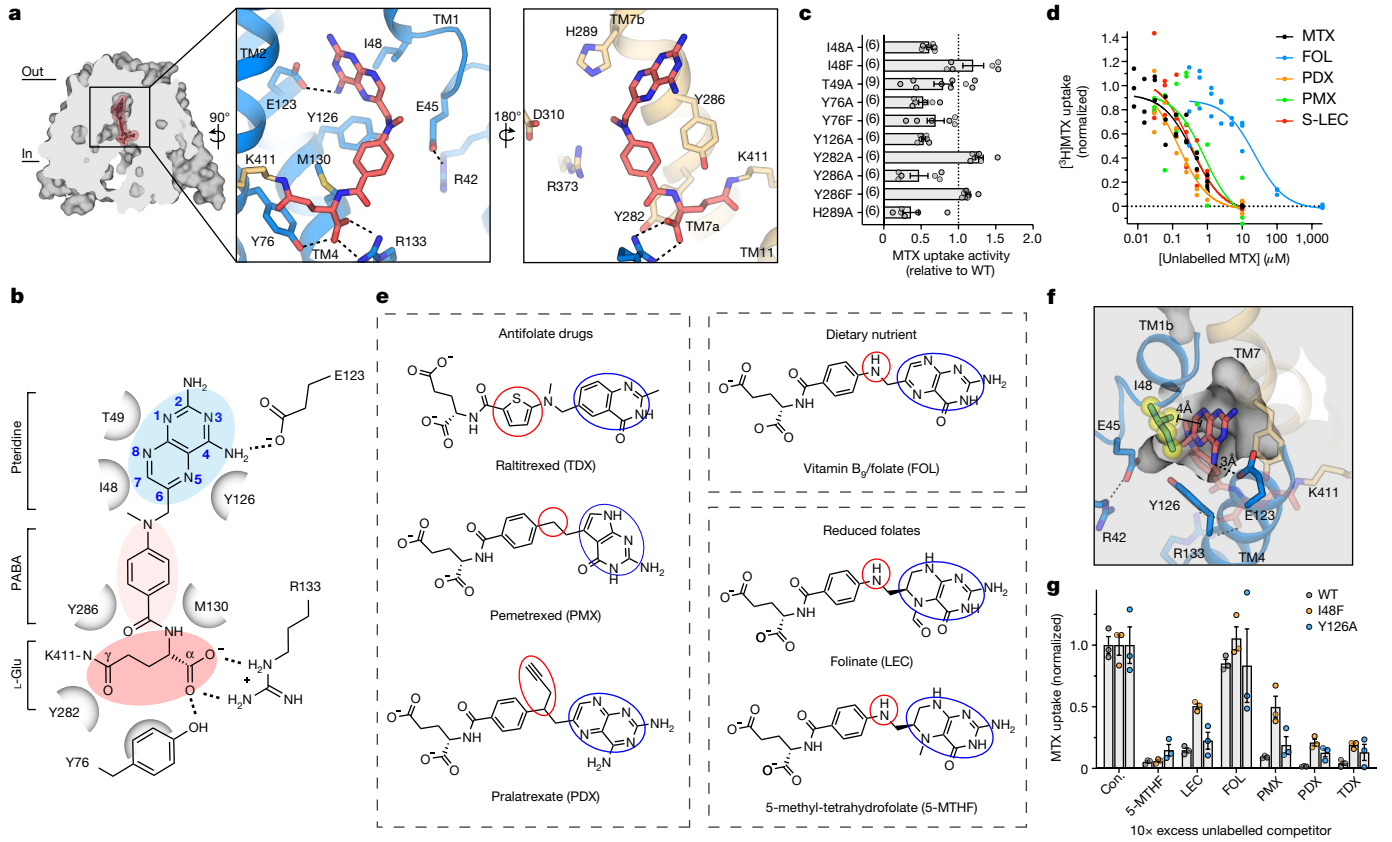


Fig. 3 | Structural determinants of folate and antifolate drug recognition by hRFC. a, Detailed transporter–MTX interactions within the central cavity. **b**, Ligplot schematic of MTX–transporter interactions, with key chemical positions labelled. **c**, [^{3}H]MTX uptake activity for mutants of select MTX-proximal residues, background corrected and relative to WT (individual values and mean ± s.e.m., with the number of biological replicates for each condition shown in brackets). **d**, Cold competition by select folates and antifolate drugs against 100 nM [^{3}H]MTX uptake in oocytes ($n = 3$ biological replicates, individual measurements shown). S-LEC, 6-(*S*)-enantiomer of folinate. **e**, Chemical structures of select antifolate drugs and folates, with key differences circled in blue (heterocyclic head group) or red (linker region). LEC and 5-methyl-tetrahydrofolate (5-MTHF) are shown. **f**, Structural features of the electronegative pocket, with important interactions with MTX highlighted. **g**, Cold competition of [^{3}H]MTX uptake by various folates and antifolate drugs, for WT, I48F and Y126A hRFC, background corrected and normalized to the positive control (Con.) ($n = 3$ biological replicates, with individual values and mean ± s.e.m. shown). See Methods for compound concentrations.

hRFC$_{EM}$ structures, there are subtle conformational changes centred at R133, which appear to be induced by MTX occupancy (Extended Data Fig. 4e–g). Additionally, A132 is located in this broken portion of TM4, a position that is mutated to proline in an MTX-resistant mouse cell line (Extended Data Fig. 5 and Extended Data Table 2), further implying the functional importance of this region[34].

Proximal to the electronegative pocket, the PABA group of MTX is clasped by residues Y126, M130 and Y286. The pteridine ring of MTX, further towards the extracellular side, is bound within the electronegative pocket, where it interacts closely with elements of the partially unwound TM1, including E45, I48 and T49 (Fig. 3a,b). In particular, E123 forms a tight interaction (~3 Å) with the pteridine ring of MTX. We mutated residues within these regions of the structure and found that many affected drug uptake (Fig. 3c). E123 appears most critical, as substitution to alanine completely abolished uptake activity, while the conservative substitution to aspartate partially restored activity (Fig. 2d).

Selectivity determinants of drug uptake

A hallmark functional feature of hRFC-mediated uptake is the preference of hRFC for reduced folates and antifolate drugs over vitamin B_9 (folate, FOL) and other anionic compounds (Fig. 3d). The many folate substrates of hRFC predominately vary in the identity of the heterocyclic ring. Often, a pterin or pteridine ring is found at this position, as in MTX, with exceptions including the pyrrolopyrimidine ring in PMX (Fig. 3e). While ring position 4 (C4) is a carbonyl in pterins, pteridines feature an amine. Further, reduced folates and FOL differ in the pterin oxidation state at ring positions 5–8 (Fig. 3e).

In the hRFC$_{EM}$–MTX structure, the partially unwound TM1 is stabilized by a salt bridge formed by residues R42 and E45 (Fig. 3f), along with a direct contact with W107 of TM3 (Extended Data Fig. 5). The overall conformation of TM1 positions I48 in direct contact with atom positions 5–8 of the pteridine ring, which are aromatic (Fig. 3f). The pteridine ring is also in close proximity to E123, which has a critical role in MTX recognition and transport (Figs. 2d and 3a,f). On the basis of these findings (Fig. 3a,f) and the fact that one major difference among folates, reduced folates and antifolate drugs is the identity of the chemical moiety analogous to the MTX pteridine (Fig. 3e), we considered that I48 and E123 may be important contributors to the folate and antifolate drug selectivity preferences exhibited by hRFC. We found that I48F substitution substantially shifted the selectivity profile compared with WT or Y126A hRFC in cold competition assays, as evident from the reduced block by unlabelled folinate (racemic) and PMX (Fig. 3g). E123A is non-functional for MTX uptake and so was not assessed in these experiments. These data confirm the structural observation that, while Y126 interacts with chemical features largely shared across different folates (PABA), I48 contacts a position on the substrate that is

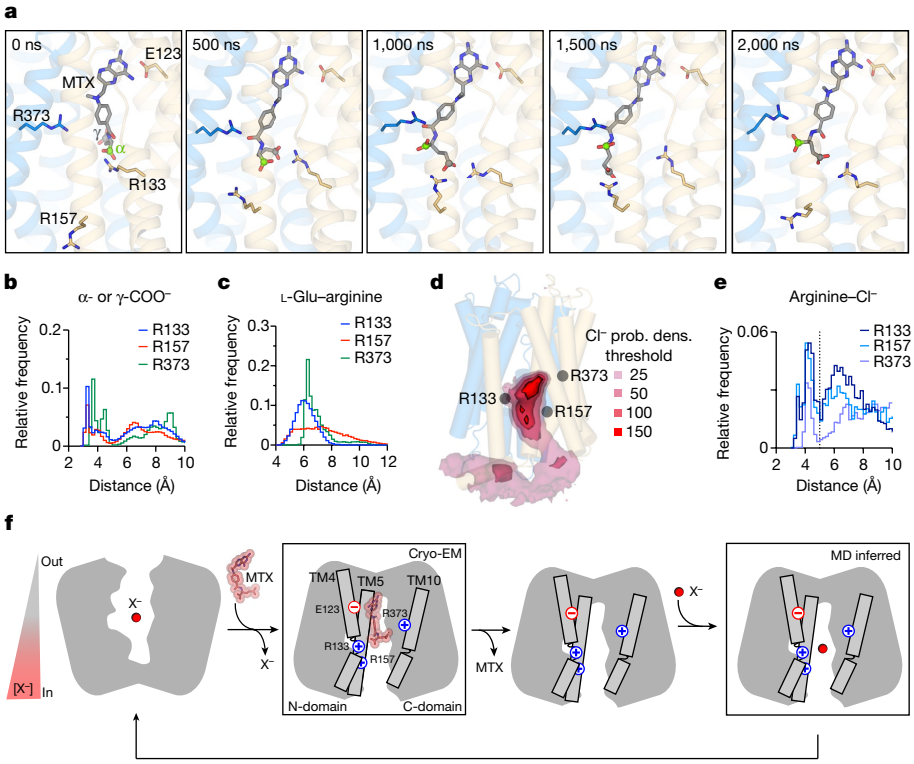


Fig. 4 | Dynamics and transport of MTX and anions by hRFC. **a**, Representative snapshots for MD simulations of MTX binding to hRFC sampled from a representative replicate at regular intervals of 500 ns. The arginine triad, E123 and MTX are depicted as sticks. The α- and γ-carboxylates of MTX are indicated by green and grey spheres, respectively. **b**, Histograms of distance from the α- or γ-carboxylate of MTX to the arginine triad for MTX-bound hRFC MD simulations (over all $n = 5$ replicates). **c**, Histograms of distance from the centre of mass of the MTX glutamate moiety to the arginine triad. **d**, WT hRFC simulations showing chloride probability densities carved at four different thresholds, where the value of 150 is equivalent to a sigma level of 69 (frames summed over $n = 3$ replicate simulations of 1 µs). **e**, Histograms of distance from chloride to the arginine triad from hRFC MD simulations. The vertical dashed line at 5 Å indicates a threshold for close interaction. **f**, Hypothetical mechanism for anion exchange and transport by hRFC, highlighting the roles of the arginine triad and E123. The ramp denotes the transmembrane thiamine phosphate concentration gradient that drives folate and antifolate uptake.

more variable across folates and antifolate drugs (pteridine or pterin) (Fig. 3e). The specificities exhibited by hRFC among different folates and antifolate drugs therefore appear to be largely dictated by the properties of the electronegative pocket. Accordingly, previous studies have shown that RFC substitutions at E45 exhibit a side chain size dependence in substrate preference and turnover of RFC[12], indicating some steric requirement in substrate selectivity. This may be through direct contact with substrate or ensuring correct placement of I48 over the pteridine ring by maintaining the partially unwound TM1 helix—a site of numerous variants associated with drug resistance, including G44E, G44R, E45K, S46I, S46N and I48F/W107G[11,35–42] (Extended Data Fig. 5 and Extended Data Table 2).

Dynamics of MTX and anion binding

To gain insights into the dynamics of MTX binding to hRFC, we pursued unrestrained all-atom molecular dynamics (MD) simulations of hRFC in a 1-palmitoyl-2-oleoyl-*sn*-phosphatidylcholine (POPC) membrane and solvated with 150 mM KCl (Extended Data Fig. 6a). As a control, 1-µs simulations were performed in $n = 5$ replicates using the $hRFC_{EM}$–MTX structure as a starting point. The pose of covalently linked MTX remained stable during the simulations (Extended Data Fig. 6b and Supplementary Video 1). Next, we removed the covalent bond between the MTX γ-carboxylate and K411 in the $hRFC_{EM}$–MTX structure and simulated over $n = 5$ replicates of 2 µs each (Extended Data Fig. 6c). While pteridine remains statically bound in the electronegative cavity (Extended Data Fig. 6c,d), the L-Glu moiety is highly dynamic within the electropositive cavity of hRFC (Fig. 4a and Extended Data Fig. 6c). Specifically, the anionic L-Glu group interacts with the three highly conserved and functionally important arginine residues (R133, R157 and R373), which we term the arginine triad (Fig. 4a, Extended Data Fig. 6e and Supplementary Video 2).

We analysed the distance distributions from the centre of mass for the α- or γ-carboxylate of MTX to the arginine guanidiniums and found that R133 and R157 make the closest contacts, with R373 making numerous but more distant interactions (Fig. 4b). Although R157 is located towards the cytoplasmic entrance of the cavity, it is highly dynamic and can interact with the γ-carboxylate (Fig. 4a). The shorter chain of lysine at this position would weaken the interaction with the MTX carboxylates, in line with R157K being non-functional (Fig. 2d). This highlights a previously unknown critical function of R157.

We also analysed the distances between the unrestrained MTX L-Glu centres of mass (calculated as the centre of mass of the six atoms comprising the two carboxylates) and the centre of mass for each arginine guanidinium (Fig. 4c). Distances of L-Glu to R133 form a Gaussian distribution, which is not the case for R157 or R373, indicating a central role for R133 in MTX L-Glu stabilization, with this residue acting like a vertex or central pivot in the arginine triad.

We also docked the more recently discovered antifolate PT523 (refs. [43,44]) according to our MTX structure and ran $n = 5$ MD simulations to ascertain binding stability (Extended Data Fig. 6f–h and Supplementary Video 3). PT523 is an MTX derivative that is heavily modified around the γ-carboxylate. Similar to the simulations of unlinked MTX, PT523 remains stably bound in the hRFC central cavity, liganded by

E123 and the arginine triad, but exhibits conformational heterogeneity in the modified L-Glu moiety. This leads to an often simultaneous interaction with all three arginines, providing a plausible explanation for its >10-fold-greater affinity over MTX, as well as explaining the length sensitivity of the modification (Extended Data Fig. 6h and Supplementary Video 3)[43,44]. More notably, it is known that hRFC functionally tolerates modifications of the γ-carboxylate of MTX but not of the α-carboxylate[32]. The binding pose of PT523 demonstrates that modifications to the γ-carboxylate can extend down the spacious electropositive ring of hRFC, whereas the α-carboxylate is not free to do so without compromising stable pteridine binding interactions. We caution that this is a highly putative binding mode for PT523—the initially placed PT523 used for MD simulations is one of many possibilities.

To probe general anion binding to hRFC, MD simulations were conducted of $hRFC_{EM}$ in the presence of KCl but without MTX. The chloride probability density was highest within the electropositive ring between R133 and R373 (Fig. 4d), with R133 and R157 making the closest contacts (Fig. 4e), similarly to MTX carboxylate binding. The more diffuse density around R157 reflects greater anion mobility in this region, indicating a role for R157 in drawing anionic species into the cavity (Fig. 4d). These data therefore suggest an important role for the arginine triad in substrate binding.

RFC is unique in the SLC19 family as it is the only member to transport anions, whereas SLC19A2 and SLC19A3 are cation carriers responsible for thiamine uptake[19]. The main differences in the highly conserved cavity (~70% similarity) are the charge conversions of R133E and K411Q (Extended Data Fig. 7a). On the basis of APBS surface electrostatics, R133E substitution is sufficient to substantially shift the electrostatic environment of the central cavity from electropositive to electronegative, with K411Q having only a slight impact (Extended Data Fig. 7b). MD simulations of R133E hRFC show that chloride occupancy in the central cavity near the electropositive ring is eliminated (Extended Data Fig. 7c). When further mutating towards the cation exchangers (K411Q, R373K, K393M), potassium ions can occupy the electronegative pocket (Extended Data Fig. 7c). This identifies a potential mechanism for substrate selectivity within the SLC19 family, where the residue at position 133 (arginine or glutamic acid) is largely responsible for tuning cavity electrostatics.

Implications for rational drug design

RFC is critical for sensitivity to MTX, a therapeutic that has found use not only in the treatment of cancers[6] but also as an anti-inflammatory agent for rheumatoid arthritis, psoriasis and inflammatory bowel disease[26,45,46]. PCFT is particularly active in the solid tumour microenvironment owing to its acidic pH optimum and is also a target for antifolate chemotherapeutics[23,24]. Selective targeting of RFC or PCFT through designer drugs would facilitate more effective treatments while mitigating toxic off-target effects[47]. We therefore compared MTX-labelled hRFC with the PMX-bound PCFT structure[25], in which PMX is bound in an alternate orientation with respect to the transporter (Extended Data Fig. 8a). First, the pyrrolopyrimidine of PMX is extensively stabilized by hydrogen bonds, whereas the pteridine of MTX is primarily stabilized by E123 (through the C4 amine) with an apparent hydrogen-bonding contribution from T49 and van der Waals interactions from I48 and Y126. This lack of many specific interactions probably explains why RFC can bind compounds with a variety of amine-rich heterocyclic head groups. Interestingly, I48 of hRFC and L196 of PCFT (I188 in the human protein) interact similarly with their respective substrates, with both appearing critical for substrate specificity and affinity[48]. Second, PCFT binds the linker benzene ring of PMX through a π-stacking interaction with a phenylalanine (Extended Data Fig. 8b), whereas MTX is simply closely packed into RFC by aromatic residues such as Y126 and Y286 (Fig. 3b). Indeed, modifying this particular moiety in PMX and related compounds results in their selective uptake by folate receptors and PCFT over RFC, indicating a size requirement for this drug moiety in RFC substrates[47,49]. Third, the L-Glu moiety of PMX is stabilized through the γ-carboxylate within the 'selectivity pocket' of PCFT[25]. On the basis of our cryo-EM structure and MD simulations, RFC appears to interact more closely with the α-carboxylate of MTX instead, which is in line with previous studies[32] (Extended Data Fig. 8c). The dynamic nature of MTX binding inferred from MD simulations suggests a role for the arginine triad in flexibly interacting with the L-Glu moiety, with the γ-carboxylate flexibly accommodated by either R133 or R157 and with modifications to the γ-carboxylate, as in PT523, accommodated within the spacious cavity. Selectively targeting one folate carrier over the other could therefore rely on modification of either glutamate carboxylate, the length of the PABA linker region and/or the nature of the polycyclic head group. Comparing drug binding between structures in two different conformations, however, has limitations and probably provides only part of the picture with regard to drug interactions. Regardless, the current structural, computational and functional data for hRFC and PCFT provide an initial framework for the rational design of improved antifolate therapeutics in the treatment of a wide variety of cancers and autoimmune disorders.

Discussion

Using cryo-EM, we captured the hRFC structure in an MTX-bound, inward-facing state. Because of competition with anionic buffer components, transient exchanger–substrate interactions or the dynamic nature of MTX binding to hRFC, we relied on covalent linkage of MTX to hRFC, on the basis of decades of biochemical and cellular studies on the NHS–MTX reagent[30–32]. We acknowledge that the covalently linked MTX binding pose presented in this study cannot represent the entire ensemble of MTX binding conformations in hRFC. To overcome this limitation, we used MD simulations of unlinked MTX to obtain insights into MTX dynamics in hRFC. Together with our structure- and simulation-guided mutagenesis studies, as well as previous functional studies, we highlight several key features of folate, antifolate drug and anion recognition by hRFC. Adopting an MFS fold, RFC contains a spacious, highly polarized cavity in which a plethora of folates and antifolate drugs can bind. Like other MFS transporters, RFC is expected to transport substrate according to the 'rocker switch' mechanism of alternating access to the intracellular and extracellular sides of the membrane (Fig. 4f)[50], for which the unusual non-helical portion of TM1 probably has a role. The electropositive extracellular surface of hRFC and the membrane in the resting state may facilitate initial anion binding in the outward-facing state (Figs. 2b and 4f). During the transition from the outward to inward open state, the arginine triad stabilizes substrate binding, with the highly flexible R157 probably assisting with anion exchange in the inward open conformation. While our structural, functional and computational data have yielded rich insights into MTX and anion recognition by hRFC, interrogation of alternate conformations is needed to fully understand the mechanism of folate and antifolate drug transport by this important protein.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-022-05168-0.

1. Zheng, Y. & Cantley, L. C. Toward a better understanding of folate metabolism in health and disease. *J. Exp. Med.* **216**, 253–266 (2019).
2. Clare, C. E., Brassington, A. H., Kwong, W. Y. & Sinclair, K. D. One-carbon metabolism: linking nutritional biochemistry to epigenetic programming of long-term development. *Annu. Rev. Anim. Biosci.* **7**, 263–287 (2019).

3. Hou, Z. & Matherly, L. H. Biology of the major facilitative folate transporters SLC19A1 and SLC46A1. *Curr. Top. Membr.* **73**, 175–204 (2014).
4. Matherly, L. H., Hou, Z. & Deng, Y. Human reduced folate carrier: translation of basic biology to cancer etiology and therapy. *Cancer Metastasis Rev.* **26**, 111–128 (2007).
5. O'Connor, C. et al. Folate transporter dynamics and therapy with classic and tumor-targeted antifolates. *Sci. Rep.* **11**, 6389 (2021).
6. Kanarek, N. et al. Histidine catabolism is a major determinant of methotrexate sensitivity. *Nature* **559**, 632–636 (2018).
7. Kobayashi, H., Takemura, Y. & Ohnuma, T. Variable expression of RFC1 in human leukemia cell lines resistant to antifolates. *Cancer Lett.* **124**, 135–142 (1998).
8. Girardi, E. et al. A widespread role for SLC transmembrane transporters in resistance to cytotoxic drugs. *Nat. Chem. Biol.* **16**, 469–478 (2020).
9. Zhao, R. et al. Rescue of embryonic lethality in reduced folate carrier-deficient mice by maternal folic acid supplementation reveals early neonatal failure of hematopoietic organs. *J. Biol. Chem.* **276**, 10224–10228 (2001).
10. Svaton, M. et al. A homozygous deletion in the SLC19A1 gene as a cause of folate-dependent recurrent megaloblastic anemia. *Blood* **135**, 2427–2431 (2020).
11. Yang, R. et al. Sequence alterations in the reduced folate carrier are observed in osteosarcoma tumor samples. *Clin. Cancer Res.* **9**, 837–844 (2003).
12. Matherly, L. H. & Hou, Z. Structure and function of the reduced folate carrier a paradigm of a major facilitator superfamily mammalian nutrient transporter. *Vitam. Horm.* **79**, 145–184 (2008).
13. Yee, S. W. et al. SLC19A1 pharmacogenomics summary. *Pharmacogenet. Genomics* **20**, 708–715 (2010).
14. Guo, W. et al. Mechanisms of methotrexate resistance in osteosarcoma. *Clin. Cancer Res.* **5**, 621–627 (1999).
15. Luteijn, R. D. et al. SLC19A1 transports immunoreactive cyclic dinucleotides. *Nature* **573**, 434–438 (2019).
16. Ritchie, C., Cordova, A. F., Hess, G. T., Bassik, M. C. & Li, L. SLC19A1 is an importer of the immunotransmitter cGAMP. *Mol. Cell* **75**, 372–381.e5 (2019).
17. Alam, C., Hoque, M. T., Finnell, R. H., Goldman, I. D. & Bendayan, R. Regulation of reduced folate carrier (RFC) by vitamin D receptor at the blood-brain barrier. *Mol. Pharm.* **14**, 3848–3858 (2017).
18. Hou, Z., Ye, J., Haska, C. L. & Matherly, L. H. Transmembrane domains 4, 5, 7, 8, and 10 of the human reduced folate carrier are important structural or functional components of the transmembrane channel for folate substrates. *J. Biol. Chem.* **281**, 33588–33596 (2006).
19. Ganapathy, V., Smith, S. B. & Prasad, P. D. SLC19: the folate/thiamine transporter family. *Pflugers Arch.* **447**, 641–646 (2004).
20. Henderson, G. B. & Zevely, E. M. Anion exchange mechanism for transport of methotrexate in L1210 cells. *Biochem. Biophys. Res. Commun.* **99**, 163–169 (1981).
21. Zhao, R., Gao, F. & Goldman, I. D. Reduced folate carrier transports thiamine monophosphate: an alternative route for thiamine delivery into mammalian cells. *Am. J. Physiol. Cell Physiol.* **282**, C1512–C1517 (2002).
22. Goldman, I. D. The characteristics of the membrane transport of amethopterin and the naturally occurring folates. *Ann. NY Acad. Sci.* **186**, 400–422 (1971).
23. Zhao, R., Gao, F., Hanscom, M. & Goldman, I. D. A prominent low-pH methotrexate transport activity in human solid tumors: contribution to the preservation of methotrexate pharmacologic activity in HeLa cells lacking the reduced folate carrier. *Clin. Cancer Res.* **10**, 718–727 (2004).
24. Desmoulin, S. K., Hou, Z., Gangjee, A. & Matherly, L. H. The human proton-coupled folate transporter: biology and therapeutic applications to cancer. *Cancer Biol. Ther.* **13**, 1355–1373 (2012).
25. Parker, J. L. et al. Structural basis of antifolate recognition and transport by PCFT. *Nature* **595**, 130–134 (2021).
26. Jansen, G. et al. Sulfasalazine is a potent inhibitor of the reduced folate carrier: implications for combination therapies with methotrexate in rheumatoid arthritis. *Arthritis Rheum.* **50**, 2130–2139 (2004).
27. Straub, M. S., Alvadia, C., Sawicka, M. & Dutzler, R. Cryo-EM structures of the caspase-activated protein XKR9 involved in apoptotic lipid scrambling. *Elife* **10**, e69800 (2021).
28. Chun, E. et al. Fusion partner toolchest for the stabilization and crystallization of G protein-coupled receptors. *Structure* **20**, 967–976 (2012).
29. Gao, X. et al. Mechanism of substrate recognition and transport by an amino acid antiporter. *Nature* **463**, 828–832 (2010).
30. Henderson, G. B. & Zevely, E. M. Affinity labeling of the 5-methyltetrahydrofolate/methotrexate transport protein of L1210 cells by treatment with an N-hydroxysuccinimide ester of [^{3}H]methotrexate. *J. Biol. Chem.* **259**, 4558–4562 (1984).
31. Hou, Z., Stapels, S. E., Haska, C. L. & Matherly, L. H. Localization of a substrate binding domain of the human reduced folate carrier to transmembrane domain 11 by radioaffinity labeling and cysteine-substituted accessibility methods. *J. Biol. Chem.* **280**, 36206–36213 (2005).
32. Deng, Y. et al. Role of lysine 411 in substrate carboxyl group binding to the human reduced folate carrier, as determined by site-directed mutagenesis and affinity inhibition. *Mol. Pharmacol.* **73**, 1274–1281 (2008).
33. Liu, X. Y. & Matherly, L. H. Functional interactions between arginine-133 and aspartate-88 in the human reduced folate carrier: evidence for a charge-pair association. *Biochem. J.* **358**, 511–516 (2001).
34. Brigle, K. E., Spinella, M. J., Sierra, E. E. & Goldman, I. D. Characterization of a mutation in the reduced folate carrier in a transport defective L1210 murine leukemia cell line. *J. Biol. Chem.* **270**, 22974–22979 (1995).
35. Zhao, R., Sharina, I. G. & Goldman, I. D. Pattern of mutations that results in loss of reduced folate carrier function under antifolate selective pressure augmented by chemical mutagenesis. *Mol. Pharmacol.* **56**, 68–76 (1999).
36. Rothem, L. et al. Resistance to multiple novel antifolates is mediated via defective drug transport resulting from clustered mutations in the reduced folate carrier gene in human leukaemia cell lines. *Biochem. J.* **367**, 741–750 (2002).
37. Tse, A., Brigle, K., Taylor, S. M. & Moran, R. G. Mutations in the reduced folate carrier gene which confer dominant resistance to 5,10-dideazatetrahydrofolate. *J. Biol. Chem.* **273**, 25953–25960 (1998).
38. Gifford, A. J. et al. Role of the E45K-reduced folate carrier gene mutation in methotrexate resistance in human leukemia cells. *Leukemia* **16**, 2379–2387 (2002).
39. Zhao, R., Assaraf, Y. G. & Goldman, I. D. A mutated murine reduced folate carrier (RFC1) with increased affinity for folic acid, decreased affinity for methotrexate, and an obligatory anion requirement for transport function. *J. Biol. Chem.* **273**, 19065–19071 (1998).
40. Jansen, G. et al. A structurally altered human reduced folate carrier with increased folic acid transport mediates a novel mechanism of antifolate resistance. *J. Biol. Chem.* **273**, 30189–30198 (1998).
41. Drori, S., Jansen, G., Mauritz, R., Peters, G. J. & Assaraf, Y. G. Clustering of mutations in the first transmembrane domain of the human reduced folate carrier in GW1843U89-resistant leukemia cells with impaired antifolate transport and augmented folate uptake. *J. Biol. Chem.* **275**, 30855–30863 (2000).
42. Zhao, R., Assaraf, Y. G. & Goldman, I. D. A reduced folate carrier mutation produces substrate-dependent alterations in carrier mobility in murine leukemia cells and methotrexate resistance with conservation of growth in 5-formyltetrahydrofolate. *J. Biol. Chem.* **273**, 7873–7879 (1998).
43. Rosowsky, A., Wright, J. E., Vaidya, C. M. & Forsch, R. A. The effect of side-chain, *para*-aminobenzoyl region, and B-ring modifications on dihydrofolate reductase binding, influx via the reduced folate carrier, and cytotoxicity of the potent nonpolyglutamatable antifolate N^{α}-(4-amino-4-deoxypteroyl)-N^{δ}-hemiphthaloyl-L-ornithine. *Pharmacol. Ther.* **85**, 191–205 (2000).
44. Rhee, M. S., Galivan, J., Wright, J. E. & Rosowsky, A. Biochemical studies on PT523, a potent nonpolyglutamatable antifolate, in cultured cells. *Mol. Pharmacol.* **45**, 783–791 (1994).
45. Furst, D. E. The rational use of methotrexate in rheumatoid arthritis and other rheumatic diseases. *Br. J. Rheumatol.* **36**, 1196–1204 (1997).
46. Schroder, O. & Stein, J. Low dose methotrexate in inflammatory bowel disease: current status and future directions. *Am. J. Gastroenterol.* **98**, 530–537 (2003).
47. Hou, Z. et al. Dual targeting of epithelial ovarian cancer via folate receptor alpha and the proton-coupled folate transporter with 6-substituted pyrrolo[2,3-*d*]pyrimidine antifolates. *Mol. Cancer Ther.* **16**, 819–830 (2017).
48. Aluri, S. et al. Substitutions that lock and unlock the proton-coupled folate transporter (PCFT-SLC46A1) in an inward-open conformation. *J. Biol. Chem.* **294**, 7245–7258 (2019).
49. Desmoulin, S. K. et al. Targeting the proton-coupled folate transporter for selective delivery of 6-substituted pyrrolo[2,3-*d*]pyrimidine antifolate inhibitors of de novo purine biosynthesis in the chemotherapy of solid tumors. *Mol. Pharmacol.* **78**, 577–587 (2010).
50. Drew, D., North, R. A., Nagarathinam, K. & Tanabe, M. Structures and general transport mechanisms by the major facilitator superfamily (MFS). *Chem. Rev.* **121**, 5289–5335 (2021).
51. Jurrus, E. et al. Improvements to the APBS biomolecular solvation software suite. *Protein Sci.* **27**, 112–128 (2018).
52. Ashkenazy, H. et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44**, W344–W350 (2016).

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

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2022

Methods

Oocyte radiotracer uptake assays

[^{3}H]MTX was purchased from American Radiolabeled Chemicals or Moravek. Defolliculated oocytes from *X. laevis* were purchased from Ecocyte Bioscience. cDNAs corresponding to full-length WT hRFC or hRFC$_{EM}$ were transferred into the pGEM-HE vector. Mutants of WT hRFC were generated by site-directed mutagenesis (PfuTurbo, Agilent). DNA templates for in vitro transcription reactions (mMESSAGE mMACHINE T7 Transcription kit, Invitrogen) were generated by SbfI linearization of the pGEM-HE clones. Oocytes were injected with 30 ng cRNA or an equal volume of water as a background control. Expression was carried out at 17 °C for 3–4 d in ND-96 solution (20 mM HEPES, 96 mM NaCl, 2 mM KCl, 1.8 mM $CaCl_2$, 1 mM $MgCl_2$, 100 U ml^{-1} penicillin-streptomycin, pH 7.4). All uptake assays were carried out at room temperature. Oocytes were combined in batches of five per reaction vessel and briefly prewashed with RFC uptake buffer (20 mM HEPES, 225 mM sucrose, pH 6.8 with MgO) to remove residual anion left over from the ND-96 solution. Uptake was initiated by replacement of RFC uptake buffer with 200 µl assay buffer containing [^{3}H]MTX. Oocyte batches were collected after 30 min unless otherwise noted and rapidly washed four times with 2.0 ml ND-96 buffer. The oocytes were then transferred to scintillation vials containing 500 µl of 10% SDS and incubated overnight for complete lysis and solubilization. Then, 5.0 ml of liquid scintillation fluid (EcoLume, MP Biomedicals) was added, and samples were subjected to liquid scintillation counting. [^{3}H]MTX was used at specific radioactivities of 1–29.7 Ci mmol^{-1}. For assessment of mutant activity relative to WT (Figs. 2d and 3d), [^{3}H]MTX concentrations of 0.5–1.0 µM were used. For cold competition by select mutants, uptake of 0.5 µM [^{3}H]MTX was measured in the presence or absence of 5 µM unlabelled competitor (Fig. 3h). Values presented in Figs. 2c and 3d were from individual biological replicates background corrected and normalized using the average signal from water-injected and WT hRFC-injected controls from each assay date. All graphical representations of data from uptake assays in this study were prepared in Prism 8.

Surface expression analysis of hRFC, hRFC$_{EM}$ and non-functional mutants

cRNA and oocyte injection were performed as described in the preceding section. Surface biotinylation was conducted as previously described with the following modifications[53]. Batches of 15–35 oocytes were biotinylated for 15 min at 25 °C in 0.5 ml of 10 mM triethanolamine (pH 9.0), 150 mM NaCl, 2 mM $CaCl_2$ and 2 mg ml^{-1} EZ-link Sulfo-NHS-SS-biotin (Thermo Scientific), and reactions were then quenched with 1 M Tris-HCl (pH 8.0) and washed in ND-96 containing 192 mM glycine and 25 mM Tris-HCl (pH 8.0). Oocytes were solubilized in 50 µl per oocyte of lysis buffer (40 mM *n*-dodecyl-β-D-maltoside (DDM), 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 5 mM EDTA, 10 µg ml^{-1} each of aprotinin, leupeptin and pepstatin, 2 mg ml^{-1} iodoacetamide and 0.2 mM PMSF) for 1 h at 4 °C. Clarified lysates were incubated overnight with 50 µl Neutravidin high-capacity resin (Pierce) at 4 °C, washed three times with 500 mM NaCl lysis buffer and then eluted with 35 µl of 4× SDS–PAGE sample buffer (Bio-Rad) containing 100 mM dithiothreitol (DTT). Following SDS–PAGE (Genscript), protein was transferred onto 0.45-µm PVDF membrane, blocked with 5% BSA in TBS and probed with monoclonal mouse anti-FLAG M2 antibody diluted 1:1,000 (Sigma-Aldrich) and then polyclonal goat anti-mouse IgG HRP-conjugated antibody diluted 1:10,000 (Rockland), and detected using the SuperSignal West Pico PLUS reagent kit (Thermo Fisher). Uncropped blots from the main figures can be found in Supplementary Fig. 2.

Synthesis of NHS–MTX

1-Ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDCI) was purchased from Chem-Impex. NHS was purchased from Sigma-Aldrich. MTX was purchased from Santa Cruz Biotechnology. Deuterated dimethyl sulfoxide (DMSO-d_6) was purchased from Cambridge Isotope Laboratories. All ^{1}H nuclear magnetic resonance (NMR) spectra were recorded with a Bruker 500 (500-MHz) spectrometer at 25 °C and calibrated to the residual isotopomer solvent signal (DMSO-d_6: δ = 2.50 ppm). NHS–MTX was synthesized as follows[54]. MTX (16.0 mg, 0.035 mmol) was treated with EDCI (10.1 mg, 0.053 mmol) in DMSO-d_6 (1.5 ml). The resulting reaction mixture was stirred for 5 min at 25 °C. NHS (6.1 mg, 0.053 mmol) was added, and the resulting reaction mixture was stirred for 1 h at 25 °C. The reaction was monitored by ^{1}H NMR. After completion, the reaction mixture was used in biological assays without further purification.

Construct design, expression and purification for hRFC$_{EM}$

All hRFC constructs were cloned into the pEG-BacMam vector[55] as a fusion with C-terminal mEGFP and a Flag-6×His tag, with a protease-cleavable linker (PreScission) between the transporter and EGFP. Constructs were prescreened with fluorescence size-exclusion chromatography (FSEC)[55] to assess detergent solution behaviour. A loop truncation of TM6–TM7 was identified that retained good FSEC behaviour (Δ215–241); we replaced these residues with BRIL[28], yielding hRFC$_{EM}$ (~75 kDa) (Extended Data Fig. 1a). Baculovirus was generated as previously described[55] and amplified to P_3. Cultures of HEK293S GnTI$^{-/-}$ (300–600 ml) were grown in Freestyle 293 medium to cell densities of 2–3 million cells per ml and infected with 6% P_3 baculovirus. Following growth at 37 °C, 8% CO_2 and 80% humidity for 16 h, 10 mM sodium butyrate was added to enhance protein expression. The temperature was then decreased to 30 °C, and the culture was incubated for an additional 48 h. Cells were collected at 2,000 r.p.m. for 10 min at 4 °C and resuspended in 50 mM Tris-HCl (pH 8.0) and 150 mM NaCl to 4–5 ml g^{-1} (cell wet weight) and then supplemented with 10 µg ml^{-1} aprotinin, 10 µg ml^{-1} pepstatin, 0.2 mM PMSF, 2 mg ml^{-1} iodoacetamide and bovine DNase I. The cell suspension was probe sonicated on ice three times using 30 pulses with 1-min breaks. Extraction was initiated by adding 40 mM solid DDM (Anatrace) directly to the lysate, and samples were incubated at 4 °C with gentle agitation for 1 h. Insoluble debris was then pelleted at 16,000 r.p.m. for 30 min at 4 °C, and the clarified lysate was incubated with 3 ml of anti-FLAG M2 resin for 1 h at 4 °C with gentle agitation. The resin was gently pelleted at 2,000 *g* for 10 min at 4 °C, transferred to a gravity-flow column casing and washed with 5 column volumes (CV) of 20 mM Tris-HCl (pH 8.0), 550 mM NaCl and 1 mM DDM followed by 5 CV of 20 mM Tris-HCl (pH 8.0), 150 mM NaCl and 1 mM DDM. Protein was eluted with 5 CV of wash buffer 2 supplemented with 0.2 mg ml^{-1} Flag peptide. Following concentration of the eluent ~5-fold with a 50-kDa molecular weight cut-off (MWCO) spin concentrator, the concentrate was supplemented with 1 mM DTT and 20:1 (wt/wt) PreScission protease and incubated at 4 °C for 2 h with gentle rotation to cleave GFP–Flag-10×His from the C terminus of hRFC$_{EM}$. The eluent was then further concentrated to ~500 µl and injected onto a Superdex 200 Increase (Cytiva) gel filtration column equilibrated in 20 mM Tris-HCl (pH 8.0), 150 mM NaCl and 0.08% digitonin (Sigma-Aldrich). To ensure complete detergent exchange, peak fractions were then collected, reconcentrated to ~500 µl and re-injected onto the same column (Extended Data Fig. 1b). The two peak fractions were concentrated to ~4–5 mg ml^{-1} (A_{280}) and then immediately used for grid preparation. SDS–PAGE analysis of the resulting protein confirmed the purity of the protein. Note that many transporters, including hRFC, are prone to aggregation in SDS–PAGE running buffer irrespective of their detergent solution stability, hence the laddering effect seen on the SDS–PAGE gel in Extended Data Fig. 1b.

Preparation and characterization of hRFC–MTX

To effectively label hRFC$_{EM}$ with NHS–MTX (Extended Data Fig. 1c), elimination of chloride was necessary. To ensure protein stability, labelling was therefore conducted on washed membranes. Following cell collection and lysis by probe sonication, membranes were washed in

labelling buffer (20 mM HEPES, 225 mM sucrose, pH 6.8 with MgO). Membranes were then extensively homogenized in 0.4 g (starting cell wet weight) per ml of labelling buffer and then labelled in the presence of 0.3 mM NHS–MTX at 37 °C for 30 min with shaking at 100 r.p.m. Membranes were then pelleted and resuspended in 50 mM Tris-HCl (pH 8.0) and 150 mM NaCl, and protein extraction and purification proceeded as described above for the non-labelled protein, except that PreScission protease treatment proceeded overnight at 4 °C. The final purified protein was characterized by SDS–PAGE, and labelling stoichiometry was determined by UV–visible spectrophotometry (NanoDrop 2000c) deconvolution using equation (1) (Extended Data Fig. 1c)[56].

$$\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{S}^T\boldsymbol{S})^{-1}\boldsymbol{S}^T \quad (1)$$

where $\boldsymbol{C}$ is the column matrix of the concentrations of hRFC and MTX for hRFC–MTX, the ratio of which yields the labelling stoichiometry. $\boldsymbol{A}$ is the column matrix of the observed hRFC–MTX spectrum and $\boldsymbol{S}$ is the column matrix for the reference spectra of MTX and hRFC. All spectra were baseline corrected at 340 nm. Typically, we observed labelling ratios of MTX:hRFC$_{EM}$ between 1 and 2, with the superstoichiometric labelling attributed to non-specific labelling of surface-exposed lysine residues or the protein N terminus.

Cryo-EM grid preparation and data collection

Following gel filtration and concentration, the hRFC$_{EM}$ sample was incubated at 4 °C for 2 h with 2 mM MTX (2% DMSO final concentration), although no MTX density was observed in the final reconstruction. The apo hRFC$_{EM}$ and hRFC$_{EM}$–MTX samples were directly used for grid preparation following centrifugation at 16,900 *g* for ~30 min at 4 °C, with 2% DMSO added immediately before grid freezing. In both cases, 3 µl of sample was applied and incubated for 60 s on freshly glow-discharged UltrAufoil R1.2/1.3 300-mesh grids (Quantifoil) and then blotted with a Leica EM GP2 plunge freezing station for 1.5 or 2 s at 4 °C and 85% humidity before plunge freezing into liquid ethane. One dataset was collected for hRFC$_{EM}$ on Quantifoil R1.2/1.3 300 Au mesh grids (Quantifoil), for which blotting lasted 3 s.

Datasets for hRFC$_{EM}$, apo hRFC$_{EM}$ and hRFC$_{EM}$–MTX were collected on a Titan Krios transmission electron microscope (Thermo Fisher) operating at 300 kV equipped with a K3 detector (Gatan) in counting mode with a BioQuantum GIF energy filter (slit width of 20 eV), using the Latitude-S (Gatan) single-particle data acquisition program. Data were collected at a magnification of ×81,000 with a pixel size of 1.08 Å at the specimen level. For the hRFC$_{EM}$ and hRFC$_{EM}$–MTX datasets, movies contained 60 frames with a 4.6-s exposure time and a dose rate of ~15 e^- per pixel per second, for a total accumulated dose of ~60 e^- Å^{-2}. For the apo hRFC$_{EM}$ dataset, movies contained 40 frames with a 2.3-s exposure time and a dose rate of ~30 e^- per pixel per second, for a total accumulated dose of ~60 e^- Å^{-2}. The nominal defocus values were set from −0.8 to −1.8 µm. One dataset was collected for apo hRFC$_{EM}$ while five and four datasets from multiple preparation sessions were collected for hRFC$_{EM}$–MTX and hRFC$_{EM}$, respectively.

Cryo-EM data processing

hRFC$_{EM}$. Beam-induced motion correction and dose weighting were performed with MotionCor2 (ref. [57]). Corrected micrographs from datasets A, B, C and D (corresponding to 3,731, 3,521, 9,159 and 4,591 initial movies, respectively) were then imported into cryoSPARC[58] for contrast transfer function (CTF) estimation with CTFFIND4 (ref. [59]), and micrographs exhibiting an estimated CTF resolution of less than 5 Å were discarded. Particle picking was performed with Template Picker in cryoSPARC. Particles were then extracted with a 64-pixel 4× binned box size at 4.32 Å per pixel and subjected to two-dimensional (2D) classification to remove obvious junk. A total of 4,538,955 particles were then subjected to classification with cryoSPARC ab initio reconstruction. This classification consisted of two rounds of two-class ab initio reconstruction jobs (default settings except for the following: initial minibatch size, 150; final minibatch size, 600; class similarity, 0), in which particles corresponding to the better 3D volume from the first round were used as input for the second round. Owing to the fact that small membrane proteins exhibit low signal-to-noise levels, we found 3D classifications to perform poorly. To address this and retain more good particles throughout classification, three parallel classifications were performed and the good classes from the second ab initio jobs were combined with duplicate particles removed before the next round of ab initio reconstruction classification. The aforementioned procedure was then iteratively performed, with initial and maximum alignment resolution settings gradually increased with each classification run (see Extended Data Fig. 3a for details). Good particles corresponding to classes showing clear protein features were re-extracted with a 2× binned 100-pixel box size at 2.16 Å per pixel after the first classification run. In the fourth run of ab initio reconstruction classification, four rounds of ab initio classification were performed per replicate. The good resulting 3D volumes from each of the three replicates were combined for a total of 298,876 particles (after removing duplicates) and transferred to RELION[60,61] for $k = 1$ 3D classification, gold-standard refinement in 3D auto-refine with a tight transmembrane domain (TMD) mask (0.5° local angular search) and Bayesian polishing, followed by non-uniform refinement and local refinement in cryoSPARC, which yielded a reconstruction to 3.8-Å resolution. To improve map quality, local resolution estimation and local filtering were applied to the final reconstruction, with a map-sharpening *B*-factor of −120.

Apo hRFC$_{EM}$. Beam-induced motion correction and dose weighting were performed with MotionCor2 (ref. [57]). Motion-corrected micrographs originating from a single imaging session (12,201 movies initially) were then imported into cryoSPARC[58] for CTF estimation with CTFFIND4 (ref. [59]), and micrographs exhibiting an estimated CTF resolution of less than 5 Å were discarded. Particle picking was performed with Template Picker in cryoSPARC. Particles were then extracted with a 64-pixel 4× binned box size at 4.32 Å per pixel and subjected to 2D classification to remove obvious junk. A total of 2,536,392 particles were then subjected to the 3D classification protocol (described above for hRFC$_{EM}$), which resulted in 138,522 good particles. The clean particle stack was re-extracted with a 200-pixel box size at 1.08 Å per pixel and subjected to non-uniform refinement in cryoSPARC (default settings except for the following: initial reference low-pass filtered to 12 Å, minimization over per-particle scale enabled, five additional final passes). The particles were transferred to RELION[60,61] and gold-standard refinement in 3D auto-refine with a tight TMD mask (0.9° local angular search) was performed, followed by Bayesian polishing. The shiny particles were re-imported into cryoSPARC and subjected to non-uniform refinement, as well as rounds of CTF refinement and local refinement (Extended Data Fig. 4a). The final reconstruction was of good quality, with a cryoSPARC-reported resolution of 3.6 Å (Extended Data Fig. 4b–d). To improve map quality, local resolution estimation and local filtering were applied to the final reconstruction, with a map-sharpening *B*-factor of −75.

hRFC$_{EM}$–MTX. Beam-induced motion correction and dose weighting were performed with MotionCor2 (ref. [57]). Corrected micrographs were then imported into cryoSPARC[58] for CTF estimation with CTFFIND4 (ref. [59]). Micrographs exhibiting an estimated CTF resolution of less than 5 Å were discarded. Curated micrographs originating from datasets A, B and C (initially consisting of 7,007, 6,349 and 2,608 movies, respectively) were processed first. Particles from template picking were extracted with 4× binning, resulting in a 64-pixel box size at 4.32 Å per pixel, and subjected to 2D classification to remove obvious junk. After 2D classification clean-up, 2,882,398 particles were subjected to the 3D classification protocol (described above for hRFC$_{EM}$), which resulted in 249,804 good particles. Template picking, 2D classification and ab

initio 3D classification were then repeated on datasets A, B, C and D in combination (dataset D initially consisting of 3,317 movies), resulting in 291,421 good particles. These two particle stacks were combined, duplicates were removed and particles were re-extracted with a 200-pixel box size at 1.08 Å per pixel, resulting in 501,313 good particles. The particles were then subjected to three rounds of non-uniform refinement in cryoSPARC, using a good 3D volume from an earlier ab initio classification as a 3D reference (default settings except for the following: initial reference low-pass filtered to 12 Å, minimization over per-particle scale enabled, five additional final passes). Local refinement was then performed using a tight TMD mask, followed by particle stack transfer to RELION[60,61] and gold-standard refinement in 3D auto-refine with a tight TMD mask (0.5° local angular search) and then Bayesian polishing. The resulting shiny particles were then subjected to a masked 3D classification with no image alignment ($k = 5$, $t = 12$), which resulted in a single class of 331,064 particles exhibiting higher-resolution features. Dataset E (from 6,226 initial movies) was processed separately. Good micrographs were subjected to template picking, 2D classification and ab initio 3D classification, resulting in 161,057 good particles, which were combined with the 331,064-particle stack from earlier processing, resulting in 492,117 total particles. This particle set was subjected to rounds of non-uniform refinement, Bayesian polishing, CTF refinement and local refinement, yielding a final reconstruction of 3.3-Å resolution (Extended Data Fig. 4a–d). To improve map quality, local resolution estimation and local filtering were applied to the final reconstruction, with a map-sharpening *B*-factor of −53.6.

Model building and refinement

All model building was performed in Coot with ideal geometry restraints[62]. An initial model was built de novo into an early 3D reconstruction of $hRFC_{EM}$. This model was then manually placed and rigid body fit into the $hRFC_{EM}$–MTX map, and further manual model building and adjustments were performed. The restraints for ligand were generated as follows: an isomeric SMILES string for MTX-modified lysine was used as input in eLBOW (in the Phenix software suite)[63]. The atom identifiers corresponding to lysine in the restraint files were then manually edited for appropriate recognition as part of the protein chain in Coot. Further manual adjustments of the ligand restraint file were performed to ensure correct stereochemistry and good geometries. Ligand placement and incorporation into the protein chain at $hRFC_{EM}$ position 411 was performed with real-space refinements in Coot. Real_space_refinement jobs were then carried out in Phenix[63] after model building, with global minimization, local grid search and secondary structure restraints. Molprobity[64] was used to help identify problematic regions (http://molprobity.biochem.duke.edu). Model building and refinement were performed in a similar manner for apo $hRFC_{EM}$. Owing to the higher resolution and quality of the apo $hRFC_{EM}$ reconstruction relative to the initial $hRFC_{EM}$ reconstruction, coordinates after real-space refinement against the apo $hRFC_{EM}$ data were placed into the final $hRFC_{EM}$ maps, and minor adjustments were made before running a final real_space_refinement job in Phenix[65].

Conservation analysis with ConSurf

Conservation within the hRFC cavity was analysed using the ConSurf server[52] and visualized with PyMOL. $hRFC_{EM}$ was used as an input structure, and a manually curated MAFFT[66] sequence alignment of 251 SLC19A1 sequences retrieved from PSI-BLAST[67] was used as the input multiple sequence alignment (non-redundant sequence database with less than 90% sequence identity used for PSI-BLAST).

MD simulations

Simulation system preparation. Apo hRFC, non-covalently bound hRFC–MTX and PT523 simulation systems were prepared using CHARMM-GUI[68]. The following force fields were used for all-atom MD simulations: ff19SB for protein[69], generalized Amber force field (GAFF) 2.2 for ligand[70] and Lipid17 for lipid. In each system, hRFC protein was embedded in a POPC bilayer solvated by 0.15 M KCl with a TIP3P water model (Extended Data Fig. 6a)[71]. For the comparison, a complex system with MTX covalently bound to K411 was prepared using LEaP and the non-covalently bound hRFC–MTX system. GAFF2.2 was used to parametrize covalently bound MTX–lysine. Extended Data Table 3 summarizes the simulation system information.

MD simulation protocol. MD simulations were performed with the pmemd.cuda module of AMBER20 using the simulation system and input files generated by CHARMM-GUI[65,72–75]. All systems were minimized for 5,000 steps, of which the first 2,500 steps used the steepest descent method and the following 2,500 steps used the conjugated gradient method. Equilibrations with weak restraints were conducted before running production MD, following standard CHARMM-GUI membrane equilibration steps[68]. Pressure was regulated by semi-isotropic Monte Carlo barostat with a pressure relaxation time of 1.0 ps for equilibration steps with the NPT (constant particle number, pressure and temperature) ensemble. All production MD simulations were performed in the NPT ensemble at 310 K and 1 atm. The weak restraints (0.1 kcal mol^{-1} $Å^{-2}$) were applied to protein for the first 10 ns of production MD simulations; see the simulation production time of each system in Extended Data Table 3. The hydrogen mass repartitioning scheme was applied for all systems, which permits a time step of 4 fs (refs. [76,77]). All bond lengths involving hydrogens were constrained using the SHAKE algorithm. Long-range electrostatics in solution were treated with the p method, and van der Waals interactions were calculated with a cut-off distance of 9.0 Å (refs. [78,79]).

To measure the distance between MTX and three crucial arginine residues (R133, R157 and R373), we used the centre of mass of the NH1, NH2 and CZ atoms in each arginine and the centre of mass of the α- and γ-carboxylates (COO^-) in MTX. To explore the relative importance of α- and γ-carboxylates in MTX binding, we also measured the distance of the α- or γ-carboxylate to each arginine residue separately. CPPTRAJ was used for trajectory analysis, and an ion density map was generated with Chimera[80,81]. VMD was used for visualization[82].

Drug resistance-associated mutation mapping. Previously reported drug resistance-associated mutations were mapped onto the $hRFC_{EM}$–MTX structure (Extended Data Fig. 5). In addition to the drug resistance-associated mutations discussed in the main text because of their close proximity to the MTX-binding site, the following mutants were also included in this analysis: D56H, L143P, A147V, R148G, S301N and D522N (Extended Data Table 2)[83–85].

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

Coordinates have been deposited in the Protein Data Bank with IDs 7TX6 ($hRFC_{EM}$–MTX), 7XT7 ($hRFC_{EM}$) and 8DEP (apo $hRFC_{EM}$). The cryo-EM maps have been deposited in the Electron Microscopy Data Bank with IDs EMD-26155 ($hRFC_{EM}$–MTX), EMD-26156 ($hRFC_{EM}$) and EMD-27394 (apo $hRFC_{EM}$). Source data are provided with this paper.

53. Harris, M., Firsov, D., Vuagniaux, G., Stutts, M. J. & Rossier, B. C. A novel neutrophil elastase inhibitor prevents elastase activation and surface cleavage of the epithelial sodium channel expressed in *Xenopus laevis* oocytes. *J. Biol. Chem.* **282**, 58–64 (2007).
54. Di Francesco, V., Di Francesco, M., Decuzzi, P., Palomba, R. & Ferreira, M. Synthesis of two methotrexate prodrugs for optimizing drug loading into liposomes. *Pharmaceutics* **13**, 332 (2021).
55. Goehring, A. et al. Screening and large-scale expression of membrane proteins in mammalian cells for structural studies. *Nat. Protoc.* **9**, 2574–2585 (2014).

56. Shinkarev, V. P., Crofts, A. R. & Wraight, C. A. Spectral analysis of the bc_1 complex components in situ: beyond the traditional difference approach. *Biochim. Biophys. Acta* **1757**, 67–77 (2006).
57. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
58. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
59. Rohou, A. & Grigorieff, N. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
60. Asarnow, D., Palovcak, E., Cheng, Y. UCSF pyem v0.5. *Zenodo* https://doi.org/10.5281/zenodo.3576630 (2019).
61. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *Elife* **7**, e42166 (2018).
62. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
63. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
64. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).
65. Lee, J. et al. CHARMM-GUI Input Generator for NAMD, Gromacs, Amber, Openmm, and CHARMM/OpenMM simulations using the CHARMM36 Additive Force Field. *Biophys. J.* **110**, 641a–641a (2016).
66. Katoh, K., Rozewicki, J. & Yamada, K. D. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief Bioinform.* **20**, 1160–1166 (2019).
67. Gabler, F. et al. Protein sequence analysis using the MPI Bioinformatics Toolkit. *Curr. Protoc. Bioinformatics* **72**, e108 (2020).
68. Jo, S., Kim, T. & Im, W. Automated builder and database of protein/membrane complexes for molecular dynamics simulations. *PLoS ONE* **2**, e880 (2007).
69. Tian, C. et al. ff19SB: amino-acid-specific protein backbone parameters trained against quantum mechanics energy surfaces in solution. *J. Chem. Theory Comput.* **16**, 528–552 (2020).
70. He, X., Man, V. H., Yang, W., Lee, T. S. & Wang, J. A fast and high-quality charge model for the next generation general AMBER force field. *J. Chem. Phys.* **153**, 114502 (2020).
71. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
72. Case, D. A. et al. *AMBER 2021* (Univ. California, 2021).
73. Gotz, A. W. et al. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 1. Generalized Born. *J. Chem. Theory Comput.* **8**, 1542–1555 (2012).
74. Salomon-Ferrer, R., Gotz, A. W., Poole, D., Le Grand, S. & Walker, R. C. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. Explicit solvent particle mesh Ewald. *J. Chem. Theory Comput.* **9**, 3878–3888 (2013).
75. Lee, J. et al. CHARMM-GUI supports the Amber force fields. *J. Chem. Phys.* **153**, 035103 (2020).
76. Hopkins, C. W., Le Grand, S., Walker, R. C. & Roitberg, A. E. Long-time-step molecular dynamics through hydrogen mass repartitioning. *J. Chem. Theory Comput.* **11**, 1864–1874 (2015).
77. Gao, Y. et al. CHARMM-GUI supports hydrogen mass repartitioning and different protonation states of phosphates in lipopolysaccharides. *J. Chem. Inf. Model.* **61**, 831–839 (2021).
78. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald - an *N*.Log(*N*) method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092 (1993).
79. Essmann, U. et al. A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577–8593 (1995).
80. Roe, D. R. & Cheatham, T. E. 3rd PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* **9**, 3084–3095 (2013).
81. Pettersen, E. F. et al. UCSF Chimera-a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
82. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
83. Kaufman, Y., Ifergan, I., Rothem, L., Jansen, G. & Assaraf, Y. G. Coexistence of multiple mechanisms of PT523 resistance in human leukemia cells harboring 3 reduced folate carrier alleles: transcriptional silencing, inactivating mutations, and allele loss. *Blood* **107**, 3288–3294 (2006).
84. Kaufman, Y. et al. Reduced folate carrier mutations are not the mechanism underlying methotrexate resistance in childhood acute lymphoblastic leukemia. *Cancer* **100**, 773–782 (2004).
85. Roy, K., Tolner, B., Chiao, J. H. & Sirotnak, F. M. A single amino acid difference within the folate transporter encoded by the murine RFC-1 gene selectively alters its interaction with folate analogues. Implications for intrinsic antifolate resistance and directional orientation of the transporter within the plasma membrane of tumor cells. *J. Biol. Chem.* **273**, 2526–2531 (1998).

Acknowledgements Cryo-EM data were screened and collected at the Duke University Shared Materials Instrumentation Facility (SMIF) and at the Pacific Northwest Center for Cryo-EM (PNCC) at Oregon Health & Science University (OHSU). We thank N. Bhattacharya at SMIF and J. Myers at the PNCC for assistance with microscope operation. This research was supported by National Institutes of Health grant R01GM137421 (S.-Y.L. and J.H.), American Heart Association fellowship 20PRE35210058 (N.J.W.) and National Science Foundation grant MCB-2111728 (W.I.). A portion of this research was supported by National Institutes of Health grant U24GM129547, performed at the PNCC at OHSU and accessed through EMSL (grid.436923.9), a Department of Energy Office of Science User Facility sponsored by the Office of Biological and Environmental Research. The Duke University SMIF is affiliated with the North Carolina Research Triangle Nanotechnology Network, which is in part supported by the National Science Foundation (ECCS-2025064).

Author contributions J.G.F. conducted biochemical preparation, sample freezing, grid screening and surface accessibility analysis; N.J.W. performed single-particle 3D reconstruction as well as radiotracer uptake assays; and Y.S. and J.G.F. collected data and J.Y. performed initial biochemical characterization, all under the guidance of S.-Y.L. N.J.W. and S.-Y.L. performed model building and refinement. H.Z. carried out all MD simulations under the guidance of W.I. P.J. synthesized NHS–MTX under the guidance of J.H. N.J.W., J.G.F. and S.-Y.L. wrote the manuscript.

Competing interests The authors declare no competing interests.

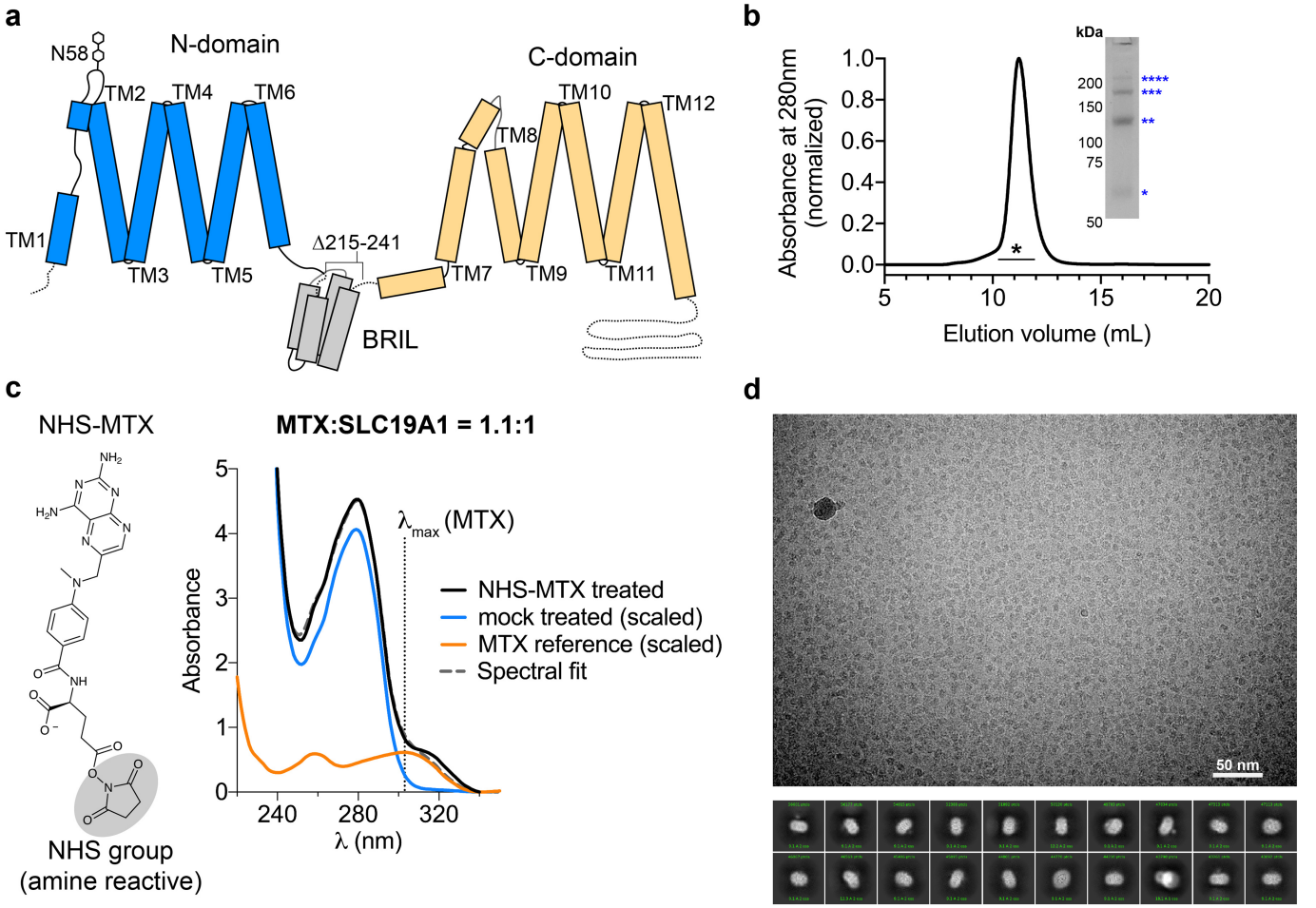
Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-022-05168-0.

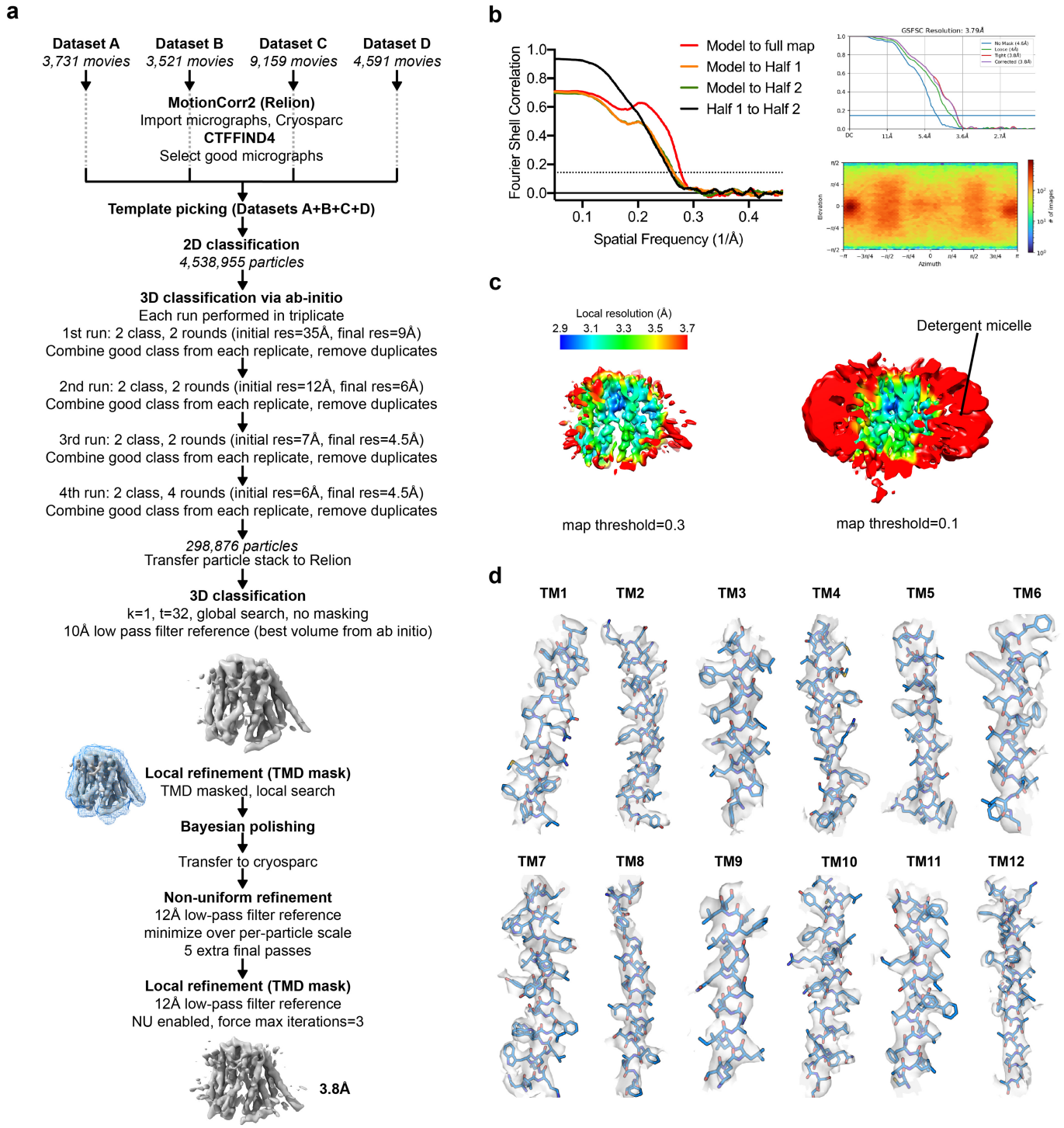
Correspondence and requests for materials should be addressed to Seok-Yong Lee.

Peer review information *Nature* thanks Larry Matherly and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

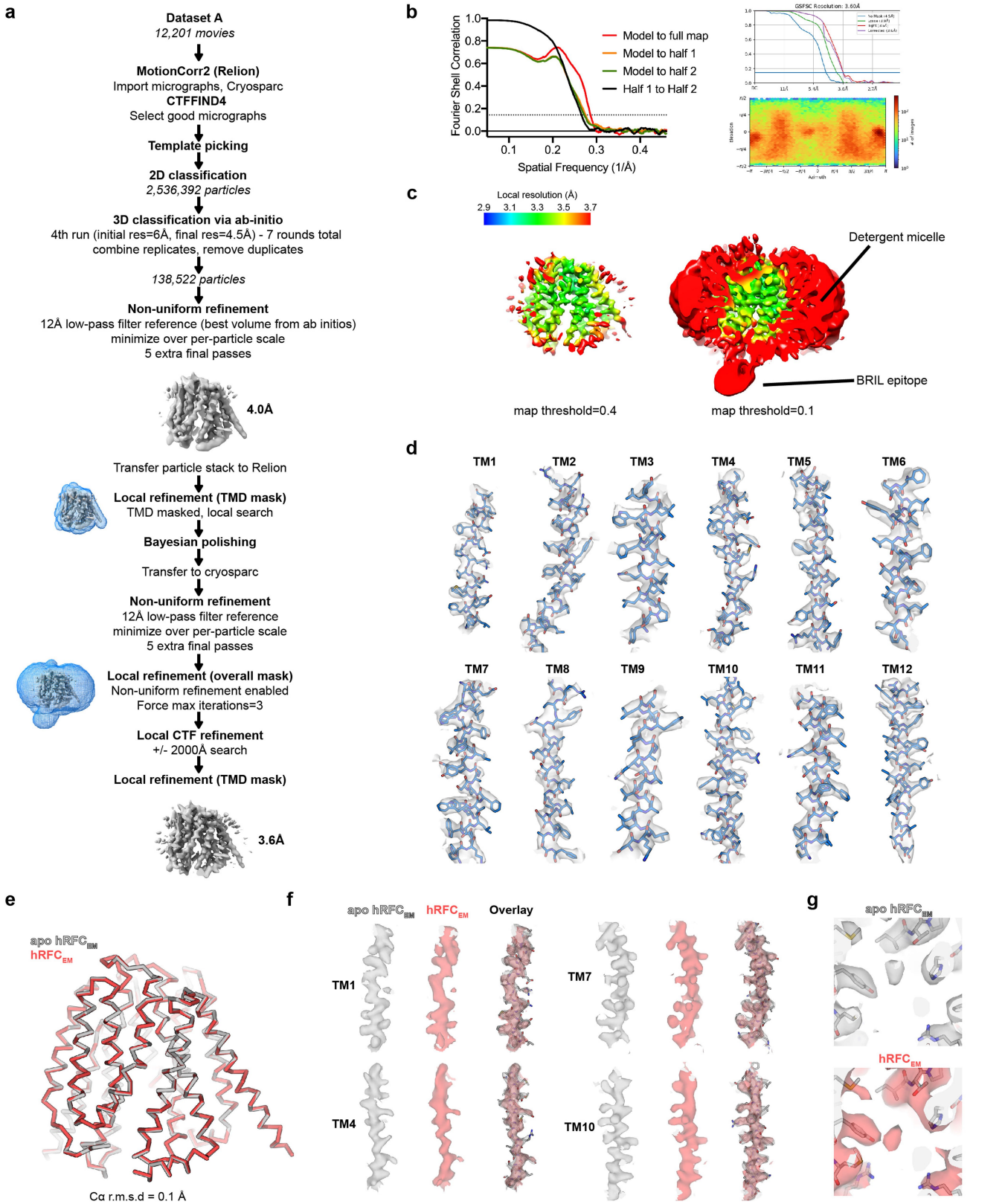
Reprints and permissions information is available at http://www.nature.com/reprints.



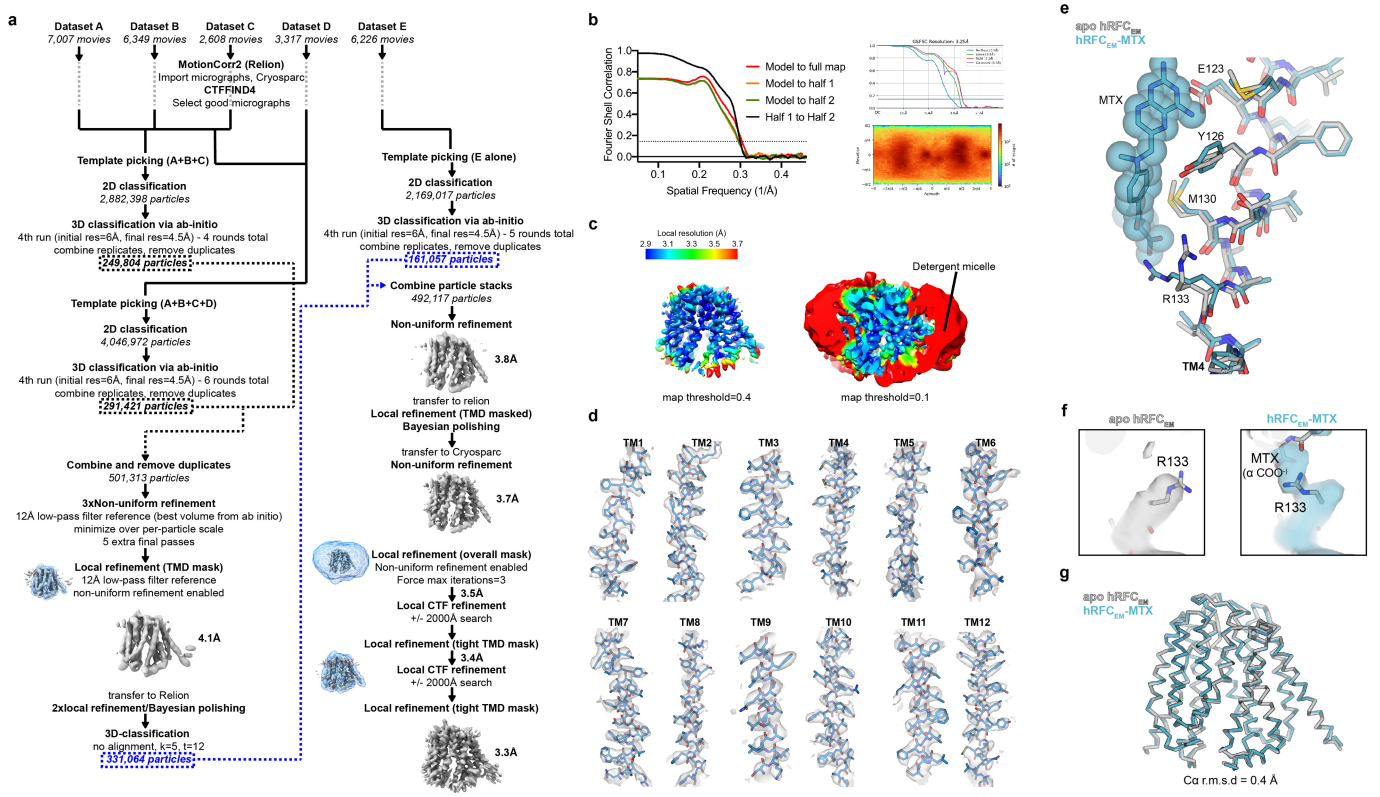
Extended Data Fig. 1 | Protein biochemistry, NHS-MTX protein modification and cryo-EM analysis of hRFC$_{EM}$. a, Topology diagram of hRFC$_{EM}$ used for structural elucidation. **b**, Representative gel-filtration profile for the final purification step and representative SDS-PAGE analysis (Coomassie stained) of purified protein used for cryo-EM grid preparation. Protein laddering during SDS-PAGE is common for small membrane proteins, with degree of non-specific oligomerization denoted by the number of asterisks (* monomer, ** dimer, *** trimer, **** tetramer). This purification is performed routinely with very similar results, reliably yielding pure and biochemically stable protein sample. **c**, Characterization of MTX modification of RFC by NHS-MTX. A representative spectral deconvolution of the MTX-RFC UV-vis spectrum into MTX and pure unlabeled RFC yields a labelling ratio of 1.1:1 MTX:RFC (graph prepared in Prism 8). **d**, Cryo-EM micrograph for hRFC$_{EM}$ sample (of representative quality for all collected cryo-EM data reported in this study) and 2D-classes of hRFC$_{EM}$.



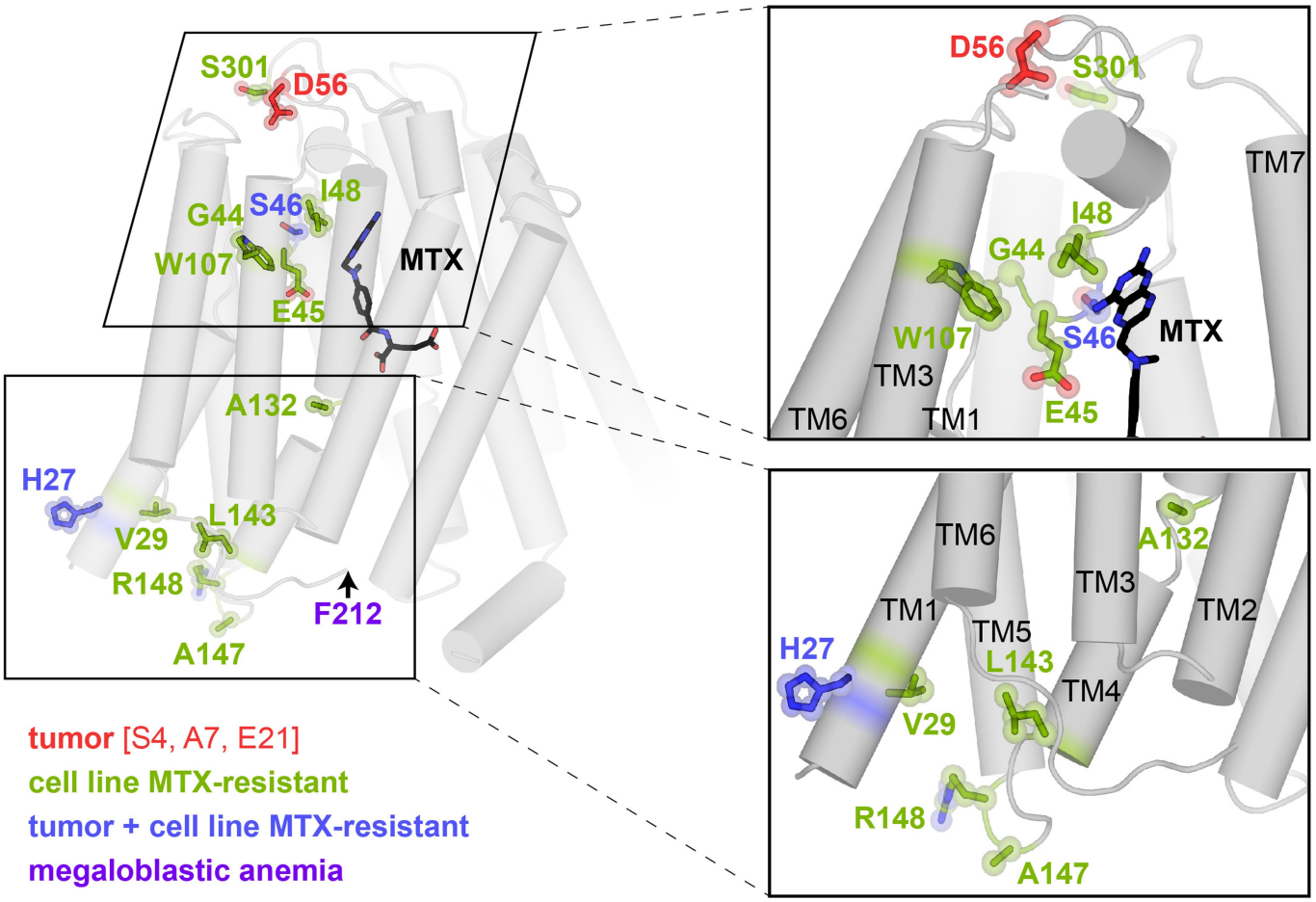
Extended Data Fig. 2 | Cryo-EM data processing of hRFC$_{EM}$. a, Processing workflow for hRFC$_{EM}$. **b**, Phenix and cryoSPARC reported Fourier shell correlations, and particle angular distribution for the final focused map. **c**, Local resolution analysis. **d**, Cryo-EM density corresponding to hRFC$_{EM}$ TM1-12 (map threshold = 0.20).



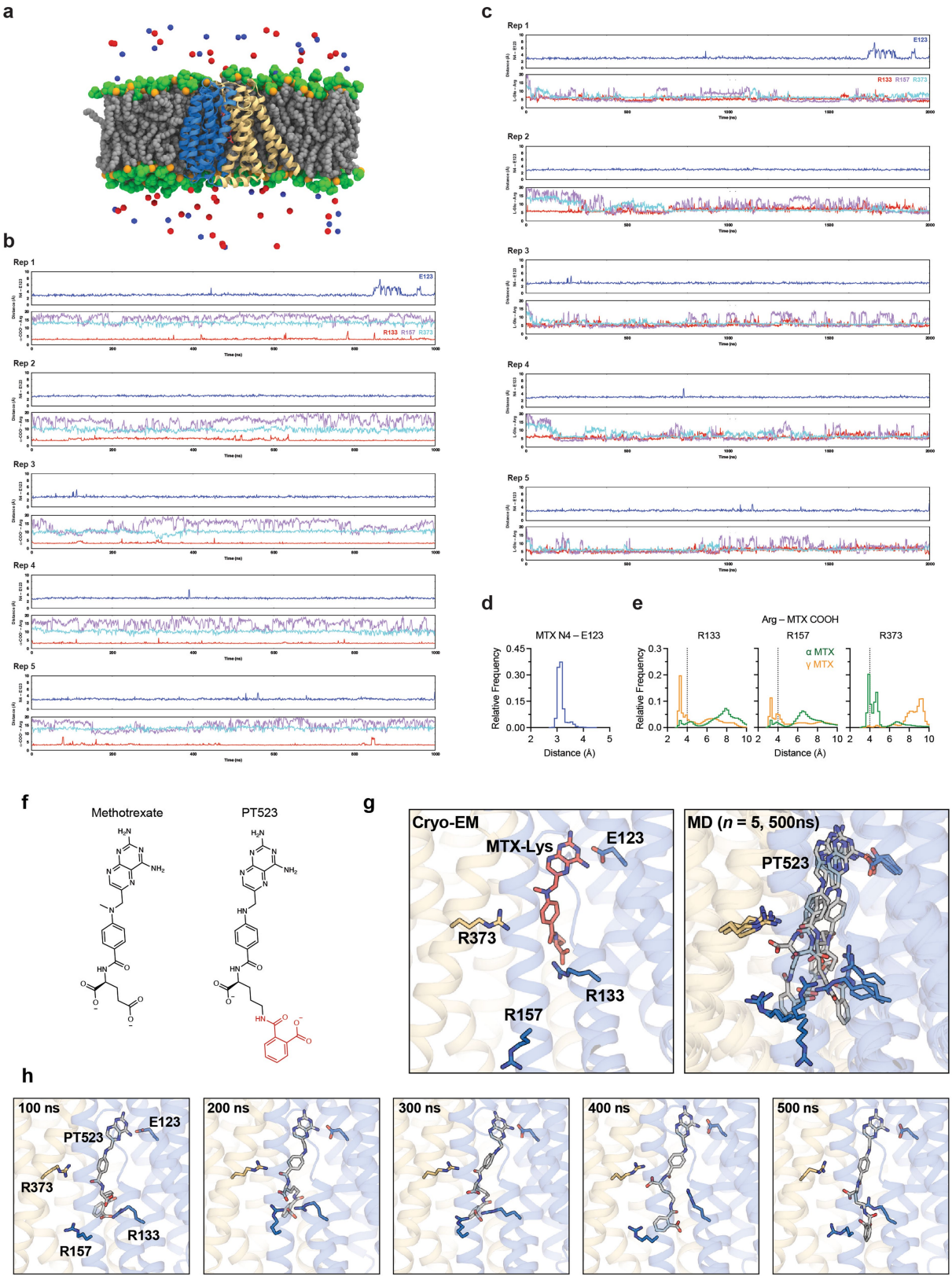
Extended Data Fig. 3 | Cryo-EM data processing of Apo hRFC$_{EM}$. a, Processing workflow for Apo hRFC$_{EM}$. **b**, Phenix and cryoSPARC reported Fourier shell correlations, and particle angular distribution for the final map. **c**, Local resolution analysis. **d**, Cryo-EM density corresponding to Apo hRFC$_{EM}$ TM1–12 (map threshold = 0.25). **e**, Structural superposition of the final refined coordinates for apo hRFC$_{EM}$ and hRFC$_{EM}$ **f**, Overlay of the final cryo-EM reconstructions (sharpened maps) for apo hRFC$_{EM}$ and hRFC$_{EM}$ (final coordinates for apo hRFC$_{EM}$ shown for reference, hRFC$_{EM}$ map resampled relative to the Apo hRFC$_{EM}$ map, with map threshold shown at 0.3). **g**, Weak, spurious cryo-EM density in the transporter central cavity present in both apo hRFC$_{EM}$ and hRFC$_{EM}$ sharpened maps (Map threshold shown at 0.15 for apo hRFC$_{EM}$, 0.10 for hRFC$_{EM}$).



Extended Data Fig. 4 | Cryo-EM data processing of hRFC$_{EM}$-MTX. a, Processing workflow for MTX modified hRFC$_{EM}$ **b**, Fourier shell correlation and particle angular distribution for the final reconstruction **c**, Local resolution analysis of the final reconstruction at two different map thresholds **d**, Cryo-EM density corresponding to hRFC$_{EM}$-MTX TM1–12 (map threshold = 0.2). **e**, Structural superposition of the final refined coordinates for apo hRFC$_{EM}$ (grey) and hRFC$_{EM}$-MTX (blue), highlighting apparent ligand induced changes in the TM4 conformational state **f**, Cryo-EM density corresponding to R133 in the apo hRFC$_{EM}$ and hRFC$_{EM}$-MTX reconstructions (hRFC$_{EM}$ map resampled relative to the hRFC$_{EM}$-MTX map, with map threshold shown at 0.1) **g**, Ribbon depiction of superposed structures.

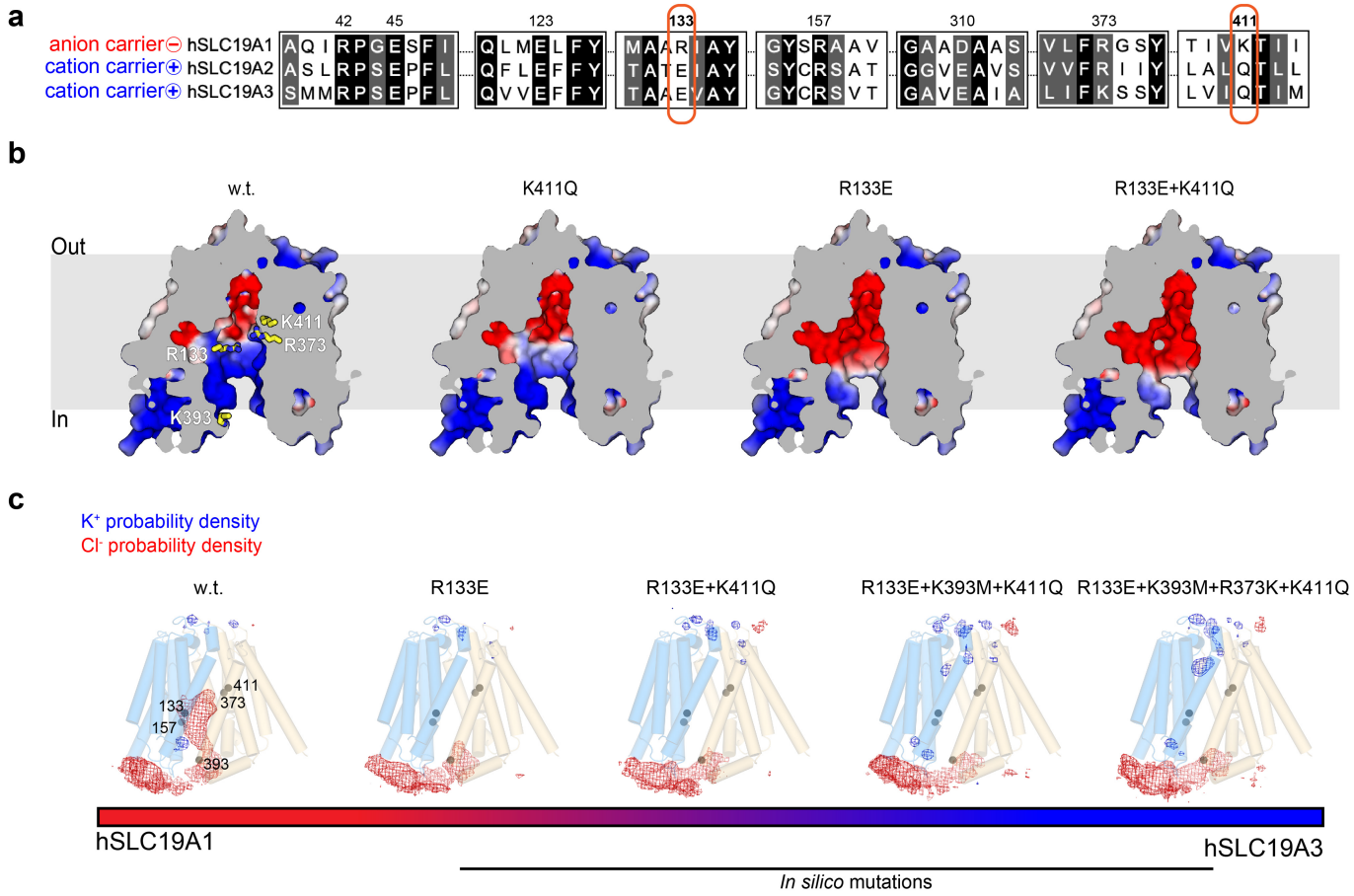


Extended Data Fig. 5 | Human disease and drug resistance associated mutations in hRFC. Mapping of clinically relevant mutations of full-length RFC[11,35–42,83–85] onto the $hRFC_{EM}$-MTX structure fall into two general regions. MTX-resistance associated mutations observed in tumor samples (red), in cell lines (green), or observed in both tumors and in cell lines (blue). The putative position of the megaloblastic anemia-associated mutation of ΔPhe212 is shown in purple as the model only extends to residue 211. Other residues not resolved in the structure are listed in parenthesis in the legend.

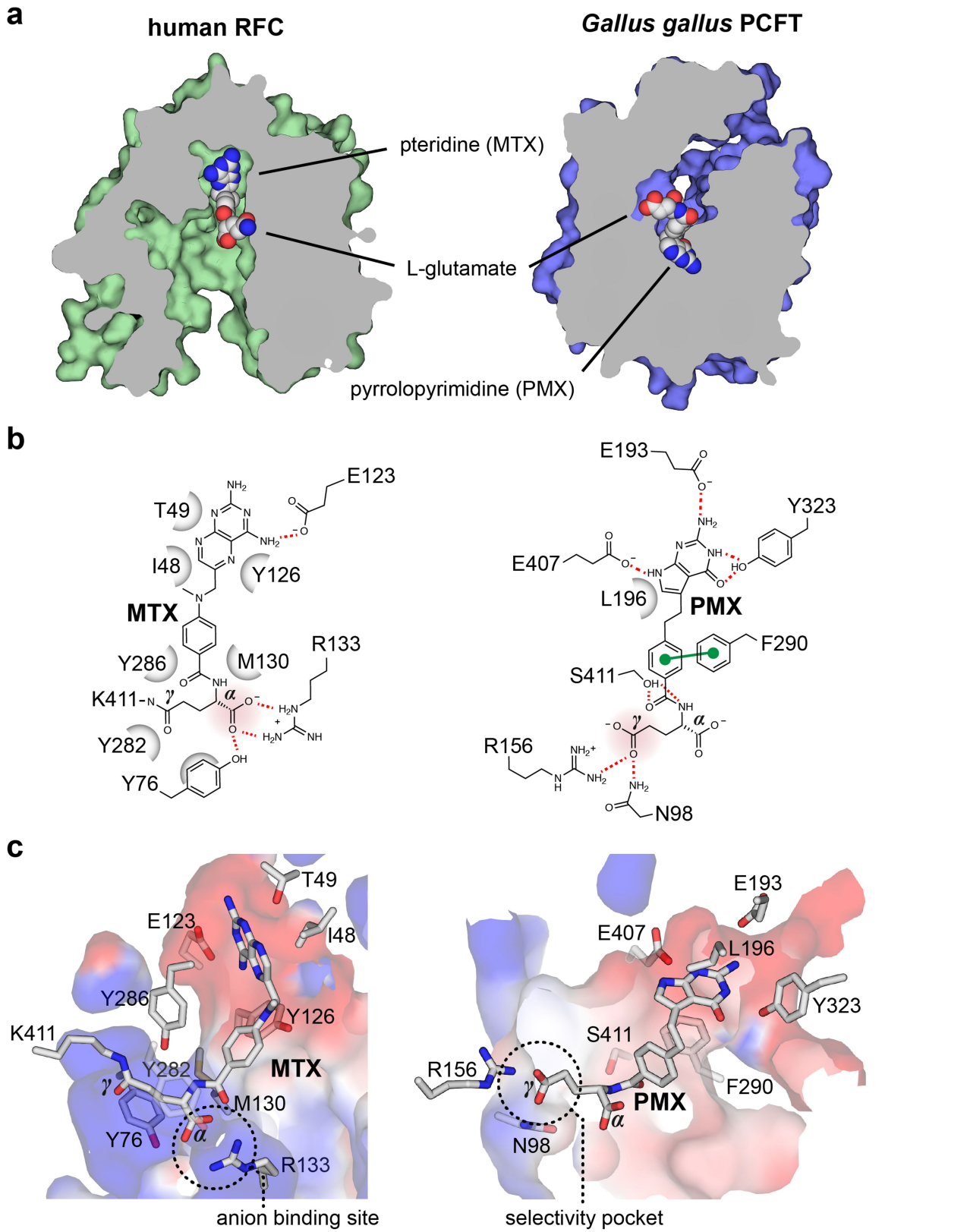


Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | MD simulations of hRFC$_{EM}$ with MTX and PT523. **a**, The all-atom molecular dynamics system setup for MTX-bound hRFC embedded in a POPC membrane and solvated with 150 mM KCl (Extended Data Table 3). hRFC is shown in cartoon, the N- and C- terminal domains colored in blue and yellow, respectively. MTX is shown as sticks, in pink. Lipids are depicted as spheres with glycerol-palmitoyl and -oleoyl groups colored gray, phosphates in orange, and choline in green. Red and blue spheres represent Cl^- and K^+, respectively. **b**, Timecourse traces for $n = 5$ replicates for MD simulations of MTX-K411 hRFC (MD system "MTX-LYS"). **c**, Timecourse traces for $n = 5$ replicates for MD simulations of hRFC with unlinked MTX (MD system "MTX"). For b and c, distances from the MTX N4 to E123 carboxylate center-of-mass (blue), and the L-Glu center-of-mass of MTX to Arg guanidiniums (R133, red; R157, purple; R373, cyan), are plotted as a function of time. **d**, Histogram plot of MTX N4 to E123 distances over $n = 5$ replicates. **e**. Histogram plots of MTX α- and γ- carboxylates to Arg guanidiniums over $n = 5$ replicates. **f**–**h**, MD simulations of PT523 docked into hRFC. **f**, Chemical structure of PT523 compared to MTX with the structural difference highlighted in red. **g**, the cryo-EM structure of MTX-labelled hRFC versus the MD simulation of docked PT523 (MD system "PT523"), with snapshots taken at 500 ns for $n = 5$ replicates. **h**, Snapshots from an MD simulation of docked PT523 sampled at various timepoints.



Extended Data Fig. 7 | Evolutionary determinants of ion selectivity within the SLC19 family. a, Select regions of a multiple sequence alignment of hSLC19A1, hSLC19A2, hSLC19A3. Numbering consistent with sequence position for hSLC19A1 **b**,The mutational effects of R133E and K411Q on the surface electrostatics of $hRFC_{EM}$ by APBS[51] calculations in PyMOL. **c**, Ion probability densities from MD simulations of $hRFC_{EM}$ with and without *in silico* introduced mutants, becoming gradually more thiamine transporter-2 (hSLC19A3) like. Simulations performed in the presence of 150 mM KCl, with a threshold value of 25 shown for chloride or potassium for each simulation.



Extended Data Fig. 8 | Antifolate drug recognition by human RFC and chicken proton-coupled folate transporter (PCFT). a, Surface representation of the cavities for inward-facing $hRFC_{EM}$-MTX and outward-facing pemetrexed (PMX)-bound PCFT. The orientation flip of the drugs in the cavities is highlighted by labelling the chemical groups. 0**b**, Chemical structure and immediate coordinating environment for MTX and PMX, with hydrogen bonds shown as red dashed lines and charged interactions highlighted by denoting charges. α- and γ- carboxylates are highlighted based on their extent of interactions in either system. **c**, The electrostatic environment of MTX and PMX as calculated by APBS[61] of the $hRFC_{EM}$-MTX and PCFT, respectively. Residues from **b** are shown as sticks with the anion binding site of RFC and selectivity pocket of PCFT as denoted.

Extended Data Table 1 | Cryo-EM data collection, refinement and validation statistics

	hRFC$_{EM}$-MTX (EMDB-26155) (PDB 7TX6)	hRFC$_{EM}$ (EMDB-26156) (PDB 7TX7)	Apo hRFC$_{EM}$ (EMDB-27394) (PDB 8DEP)
Data collection and processing			
Magnification	81,000	81,000	81,000
Voltage (kV)	300	300	300
Electron exposure (e–/Å^2)	60	60	60
Defocus range (μm)	-0.8 to -1.8	-0.8 to -1.8	-0.8 to -1.8
Pixel size (Å)	1.08	1.08	1.08
Symmetry imposed	C1	C1	C1
Initial particle images (no.)	9,098,387	4,538,955	2,536,392
Final particle images (no.)	492,117	298,876	138,522
Map resolution (Å)	3.3	3.8	3.6
FSC threshold	0.143	0.143	0.143
Refinement			
Initial model used (PDB code)	7TX7 (hRFC$_{EM}$)	none	7TX7 (hRFC$_{EM}$)
Map sharpening *B* factor (Å^2)	-53.6	-120.0	-75.0
Model composition			
Non-hydrogen atoms	3,169	3,203	3,223
Protein residues	397	397	398
Ligands	AJP: 1, MTX: 1	AJP: 4	AJP: 4
B factors (Å^2)			
Protein	81.36	116.42	108.42
Ligand	82.37	113.43	117.98
R.m.s. deviations			
Bond lengths (Å)	0.005	0.003	0.003
Bond angles (°)	0.870	0.965	0.976
Validation			
MolProbity score	1.48	1.32	1.28
Clashscore	5.44	5.80	5.28
Poor rotamers (%)	0.63	0.00	0.65
Ramachandran plot			
Favored (%)	96.93	98.22	99.49
Allowed (%)	3.07	1.78	0.51
Disallowed (%)	0.00	0.00	0.00

Extended Data Table 2 | MTX and antifolaxe resistance mutations in RFC

Residue	Variants	Patient / Cell Line Derivation	Reference
S4	P	Osteosarcoma patients	ref[11]
A7	V	Osteosarcoma patients	ref[11]
E21	K	Osteosarcoma patients	ref[11]
R27	H	Common coincident variant in osteosarcoma patients and cell lines	ref[11]
V29	L	Human cell line	ref[41]
G44	E	Murine cell line	ref[35]
	R	Human cell lines	ref[36]
E45	K	Murine and human cell lines	ref[38-41]
S46	I	Human cell line	ref[41]
	N	Murine cell line, osteosarcoma patients	ref[11,42]
I48	F	Murine cell line	ref[37]
D56	H	Acute lymphoblastic leukemia patients	ref[84]
W107 (105 mouse)	G	Murine cell line	ref[37]
A132 (130 mouse)	P	Murine cell line	ref[34]
L143	P	Human cell line	ref[83]
A147	V	Human cell line	ref[83]
R148	G	Human cell line	ref[83]
F212	del	Megaloblastic anemia patients	ref[10]
S301 (297 mouse)	N	Murine cell line	ref[85]
D522	N	Acute lymphoblastic leukemia patients	ref[84]

Extended Data Table 3 | Simulation system information

System Name	Mutation	Number of atoms	Replicas	Simulation time (μs)
Apo	WT	142,640	3	1.0
	R133E	142,180	3	1.0
	R133E+K411Q	143,714	3	1.0
	R133E+R373K+K411Q	142,085	3	1.0
	R133E+K393M+K411Q	143,777	3	1.2
	R133E+R373K+K393M+K411Q	142,219	3	1.2
MTX	WT	143,572	5	2.0
MTX-LYS	WT	143,560	5	1.0
PT523	WT	143,245	5	0.7

nature portfolio

Corresponding author(s): Seok-Yong Lee

Last updated by author(s): Jul 17, 2022

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Latitude S 1.3
Data analysis	MotionCor2 1.2.6 - published and freely available, CTFFIND4 - published and freely available, Relion 3.1/4.0 - published and freely available, UCSF pyem tools - published and open source, CryoSPARC 3.0/3.1 - commercially available, Coot 0.8 - published and freely available, PHENIX 1.18 - published and freely available, UCSF Chimera - published and freely available, Prism 8 -commercially available, PyMOL 2.4.0 - open source, APBS Tools 2.1 - open source, CHARMM-GUI - published and freely available, AMBER20 - commercially available, VMD 1.9.4 - published and freely available

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Coordinates have been deposited in the Protein Data Bank with the PDB IDs 7TX7 (hRFCEM) and 7TX6 (hRFCEM-MTX), and 8DEP (hRFCEM-apo), respectively. The cryo-EM maps have been deposited in the Electron Microscopy Data Bank with the IDs EMD-26156 (hRFCEM), EMD-26155 (hRFCEM-MTX), and EMD-27394 (hRFCEM-apo), respectively.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Dataset size for cryo-EM structure determination were decided based on grid quality and resource availability, in order to achieve the highest quality and resolution 3D-reconstruction possible. For the oocyte experiments, batch sizes per replicate were determined based on signal-to-noise optimization, and replicates were accumulated to ensure accurate and reproducible determination of data values. Since our functional assays were performed in a recombinant system which exhibits low background, sample sizes of at least n=3 (biologically independent replicates) were sufficient in this regard. Predetermination of sample size was not required for these structural and functional studies.
Data exclusions	For functional/biochemical assays no data were excluded. Junk particles were removed during cryo-EM data processing by means of image classifications, in order to identify a subset of particle images that led to the highest resolution 3D-reconstruction.
Replication	All biochemical assays were repeated at least in triplicate (n ≥ 3) and were reproducible (sample size is denoted in all cases in the figure legends).
Randomization	Randomization was not utilized nor relevant for the structural and functional studies presented in this paper.
Blinding	Blinding was not utilized nor relevant for the structural and functional studies presented in this paper.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Monoclonal mouse anti-FLAG, clone M2 (Sigma-Aldrich, F3165) was used 1:1000 dilulted, and polyclonal goat anti-mouse IgG HRP-conjugated (Rockland, 610-1302) was used 1:10,000 diluted.
Validation	Primary antibody was validated against a blot of the purified target protein and, as shown in the manuscript, target-expressing samples exhibit a clear band, whereas the negative control (water injected ooctes) do not. From the manufacturer: Sensitivity Test Conforms: Detects 2 ng of FLAG-BAP fusion protein by dot blot using chemiluminescent detection. Specificity Confirmed: Detects a single band of protein on a western blot from an E. coli crude cell lysate.

Eukaryotic cell lines

Policy information about cell lines

Cell line source(s)	Human embryonic kidney (HEK) 293S GnTI- cells are purchased from ATCC, X. laevis oocytes are purchased from Ecocyte (Autsin, Texas, USA)
Authentication	Cell lines were not authenticated.

Mycoplasma contamination	HEK293S GnTI- cells were tested negative for mycoplasma contamination by Duke University Cell Culture Facility.
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell line was used in this study.