Expression, purification and preliminary characterisation of the choline transporter LicB from opportunistic bacterial pathogens

Angela TF Neves[1], Richard Stenner[1], Paul R Race[1] and Paul Curnow[1*]

[1]School of Biochemistry, University of Bristol, UK.
*Corresponding author. Email: p.curnow@bristol.ac.uk

Abstract

Many opportunistic bacteria that infect the upper respiratory tract decorate their cell surface with phosphorylcholine to support colonisation and outgrowth. These surface modifications require the active import of choline from the host environment, a process thought to be mediated by a family of dedicated integral membrane proteins that act as choline permeases. Here, we present the expression and purification of the archetype of these choline transporters, LicB from *Haemophilus influenzae*. We show that LicB can be recombinantly produced in *Escherichia coli* and purified to homogeneity in a stable, folded state using the detergent *n*-dodecyl-β-D-maltopyranoside. Equilibrium binding studies with the fluorescent ligand dansylcholine suggest that LicB is selective towards choline, with reduced affinity for acetylcholine and no apparent activity towards other small molecules including glycine, carnitine and betaine. We also identify a conserved sequence motif within the LicB family and show that mutations within this motif compromise protein structure and function. Our results are consistent with previous observations that LicB is a specific high-affinity choline transporter, and provide an experimental platform for further studies of this permease family.

Keywords

Introduction

Bacterial infections of the respiratory tract are a leading cause of illness and death worldwide [1, 2]. Many such infections are caused by bacteria that are generally benign, transient residents of the nose and throat but which can also proliferate and invade to cause a range of conditions including pneumonia, meningitis and sepsis [3]. Infections by these opportunistic pathogens require colonisation of the mucosal surface, subsequent bacterial growth within this environment, and evasion of the host immune system [4]. Components of the bacterial cell envelope can contribute to each of these aspects of infection, and so remain of active interest for both fundamental understanding and drug development.

One particular surface decoration found broadly among bacteria colonising the respiratory tract is the modification of the surface layer with phosphorylcholine (ChoP) [5]. ChoP is attached to lipopolysaccharide (LPS) at the outer cell wall of Gram-negative pathogens including *Haemophilus influenzae* [6], and introduced to teichoic or lipoteichoic acid within the peptidoglycan layer of Gram-positive strains such as *Streptococcus pneumoniae* [7]. Several lines of evidence point towards a role for ChoP in promoting bacterial persistence [8, 9], with ChoP contributing to epithelial adherence and invasion by binding to the endogenous receptor for platelet activating factor (rPAF) [10-14]. In terms of infection, the presence of ChoP actually stimulates the host immune system, eliciting a serum antibody response and enabling bacterial killing via C-reactive protein (CRP) [6, 8, 15, 16]. This means that the levels and accessibility of ChoP are likely to be modulated as bacteria move from colonisation to infection, and in *H. influenzae* this is an important part of LPS phase variation [17-19]. However, it has also emerged that ChoP can support evasion of the complement system by reducing antibody binding to the bacterial surface [20], and that differences in the exact location of ChoP within the LPS can suppress killing by CRP [21]. ChoP may provide additional benefits to the pathogen by conferring increased resistance to host antimicrobial peptides [22] and contributing to biofilm formation [23]. The biosynthesis and adaptation of ChoP are thus important aspects of bacterial pathogenesis.

The model system for understanding ChoP synthesis and regulation is the *lic1* operon of *H. influenzae*. This comprises four genes named *licA-D* that encode a metabolic pathway reliant upon the enzymatic activities of the four gene products LicA-D [6, 11, 17]. The first step in ChoP production is the active import of choline from the host environment by LicB, an integral membrane protein that is thought to act as a high-affinity choline permease. Once inside the

cell, choline is then phosphorylated by LicA, an ATP-dependent choline kinase. The resulting phosphocholine is transformed into cytidyldiphosphocholine (CDP-choline) by the cytidylyltransferase LicC. Finally, CDP-Choline is the substrate for LicD, which attaches phosphorylcholine to nascent glycans prior to their export to the cell surface. The mechanism and regulation of LicA and LicC are increasingly well-understood from studies of their cell biology, enzymology and structural biology (e.g. [24, 25]). In contrast, the structures and functions of LicD and LicB are yet to be fully elucidated. This current study aims to address this knowledge gap by developing new methods for the study of the permease LicB.

LicB is proposed as the type member of a small family of Choline Uptake Transporters that sit within the Drug/Metabolite Transporter (DMT) superfamily and are assigned Transporter Classification Database number 2.A.7.18 [26]. (This protein bears no structural or functional similarity to the lichenase of the same name). Bioinformatic sequence analysis predicts LicB to be a multipass integral membrane protein with 10 transmembrane helices [27]. LicB proteins are probably H^+/Choline symporters, harnessing the PMF of the plasma membrane to drive the simultaneous influx of both substrates, although this remains to be confirmed. Studies of *H. influenzae* mutant strains show that when choline is abundant LicB is not required for choline uptake [27]. However, under physiologically-relevant conditions where choline is scarce, *licB* is upregulated and LicB is essential for choline transport and subsequent ChoP synthesis [27]. LicB thus does not appear to be involved in osmoregulation like the Betaine/Carnitine/Choline Transporter (BCCT) family of transporters but instead is a high-affinity choline permease (K_M = 3.6 μM) that is the major route of choline entry into the cell under physiological conditions [27]. Because of the role of ChoP in supporting the bacterial lifecycle, *ΔlicB* strains of *H. influenzae* are less virulent than wild-type [28]. The homologous *licB* gene of *S. pneumoniae* also encodes an integral membrane protein [29], and has now been identified as essential for the proper growth and development of this pathogen [30].

Here, we present a study of purified recombinant *H. influenzae* LicB in a bid to further understand the structure and function of this enzyme. Such *in vitro* biophysical studies are a powerful means of investigating substrate affinity and specificity, and can be combined with site-directed mutagenesis to dissect protein activity in molecular detail. However, there are inherent challenges in using these techniques with an integral membrane protein such as LicB, and a major barrier concerns the reliable expression and purification of recombinant constructs.

We show that reasonable yields of *H. influenzae* LicB can be recovered from recombinant *Escherichia coli* once optimal conditions are determined, and that the purified protein is sufficiently robust for biophysical studies. Substrate binding assays utilising a fluorescent substrate suggest that LicB is selective for choline. Site-directed mutagenesis of a conserved sequence motif confirms the importance of these residues to protein structure and function. These results establish a basis for further *in vitro* studies of LicB.

Materials and Methods

Bioinformatics

The consensus transmembrane topology of LicB was predicted by TOPCONS (https://topcons.cbr.su.se) and a schematic topology diagram was generated by TOPO2 (Johns SJ, TOPO2 Transmembrane Protein Display Software; http://www.sacs.ucsf.edu/TOPO2/). Protein structures were generated by the Robetta server (https://robetta.bakerlab.org) using either the *ab initio* RoseTTAFold method [31] or by comparative modelling [32] with template PDB ID 5l20. Homologues of *H. influenzae* LicB (WP_005672105.1) were identified with BLASTP, with the search terms excluding the large number of depositions from *H. influenzae* strains. The top scorer, with 100% sequence identity, was designated as being from *S. pneumoniae* but this was considered misassigned and excluded. The top 100 results returned from BLASTP were then pruned to remove partial and redundant sequences. The remaining 34 sequences were aligned with ClustalOmega (https://www.ebi.ac.uk/Tools/msa/clustalo/). Ligand docking was performed with the HADDOCK webserver (https://wenmr.science.uu.nl/haddock2.4/) [33], nominating W14, Y84, W162, Y193, Y228, Y231, Y232, Y250, and W253 as being potentially active in substrate binding. Three of the best four structures from the top cluster had very similar ligand orientations and the best scorer was taken as being representative. Protein models were displayed in PyMOL.

Molecular biology

H. influenzae genomic DNA was a kind gift from Simon Newstead, University of Oxford. The *licB* gene was amplified by polymerase chain reaction using the forward primer 5'-GTTGTTCCATGGGTCGTGGCTATCTCTTTGGCATACTATCC-3' and the reverse primer 5'- ATGATGCTCGAGTTTTTGTTCCCCTTTGTAATAAAGTGTAACG-3'. These primers incorporated restriction sites for NcoI (forward) and XhoI (reverse) to enable subcloning of the amplified fragment into the expression vector pET28a by restriction-ligation. This cloning protocol introduces an extra glycine at the N-terminus and a His6 tag at the C-terminus for purification. Amplification used 25 thermal cycles of 30s at 95°C, 30 s at 55°C and 3 min at 72°C with the Phusion DNA polymerase (New England Biolabs). Site-directed mutagenesis used the plasmid amplification ('Quickchange') method with the same enzyme and annealing temperatures between 48-60°C. All constructs were verified by sequencing, and plasmids were transformed into *Escherichia coli* strains by standard heat-shock protocols.

Protein expression

All cell cultures were supplemented with 50 μg/ml kanamycin to maintain selective pressure for the expression plasmid. A single colony of the transformed strain from LB agar was inoculated into 10 ml LB media in a 30 ml glass vial and grown overnight at 37°C with orbital shaking at 250 rpm. This primary culture was used to inoculate 1 L of SB autoinduction media (Formedium AIMSB0210) in a baffled plastic flask (Tunair) and grown for 24 h at 37°C with shaking at 250 rpm. Cultures were harvested by centrifugation at 5000 × g, resuspended in 200 ml PBS and stored at -20°C overnight.

Cell lysis and membrane solubilisation

Harvested cells were thawed prior to lysis in a continuous-flow cell disruptor (Constant Systems) at 25 kpsi. Unbroken cells and large debris were pelleted at 10,000 × g for 12 min and the supernatant retained. Cell membranes were isolated from this clarified lysate by centrifugation at 170,000 × g for 1 hour at 4°C. The resulting membrane pellet was resuspended in 100 ml Membrane Buffer (50 mM sodium phosphate pH 7.4, 150 mM NaCl, 5% glycerol) and subjected to at least 30 passes in a hand-held glass homogenizer. This membrane suspension was then solubilized with detergents (see text) for 1h at room temperature. Insoluble membranes were pelleted at 170,000 × g for 1 hour at 4°C and discarded. The retained supernatant contained protein-detergent complexes that were subjected to further purification. Total protein concentration in cell fractions was determined using a commercial detergent-compatible Lowry assay (Bio-Rad DC Protein Assay), according to the manufacturer's instructions.

Protein purification

A 1 ml HisTrap column (GE Healthcare) was equilibrated in Column Buffer (Membrane Buffer plus detergent) with 20 mM imidazole. Detergent concentrations for purification were 0.05% DDM, 0.2% DM and 0.24% Cymal-5, reflecting the different critical micelle concentrations of these surfactants. In each case, 20 mM imidazole was added to the solubilized protein before being loaded onto the column at a flow rate of 1 ml/min. The column was washed in 40 ml Column Buffer with 75 mM imidazole at 1 ml/min and purified LicB was eluted in 2.5 ml Column Buffer with 0.5 M imidazole at 0.2 ml/min. Imidazole was immediately removed with a desalting column. The purified protein was concentrated to <500 μl in a centrifugal concentrator with 50 kDa molecular weight cut-off and applied to a Superdex 200 10/300 size

exclusion column equilibrated in 1.5 volumes of Column Buffer. Size exclusion was performed at 0.5 ml/min.

Protein characterisation

Concentration of purified LicB was determined by the absorption at 280 nm using a calculated extinction coefficient of 64,290 $M^{-1}cm^{-1}$ and a theoretical molecular weight of 34,120 Da (https://web.expasy.org/protparam/). SDS-PAGE was performed with unboiled samples on Tris-Glycine 4-20% acrylamide precast gels (NuSep). Western blotting with 40 µg total protein from each cell fraction used the Novex system (ThermoFisher) according to manufacturer's instructions, with anti-His-(C-term)-HRP (ThermoFisher R931-25, lot 1905395) at 1:5000 dilution and ECL detection using the SuperSignal West PicoPLUS reagents (ThermoFisher). Identification of recombinant LicB protein was confirmed using tryptic digest MALDI mass spectrometry of the major band. This only returned a single peptide (-SLSGLQVYFLR-) that unequivocally corresponded to LicB from *Haemophilus*. However, we have empirically found that many recombinant membrane proteins produce low numbers of tryptic peptides which prevents full coverage of the protein sequence by this method. Circular dichroism (CD) was performed on a Jasco J-1500 instrument in a 1 mm pathlength quartz cuvette at protein concentrations between 0.1–0.7 mg/ml. Data collection at 25°C involved 8 repetitions between 190 to 300 nm with a data interval of 1 nm. Buffer backgrounds were collected and subtracted from each protein sample, and wavelengths at which the high-tension (HT) voltage exceeded 700 V were excluded. Thermal melts were performed between 25°C and 95°C at a rate of 1°C/minute. Protein stability at 25°C was determined over 16 h with CD scans taken every 0.3 h. Helicity was estimated from the ellipticity at 222 nm. SEC-MALS data were collected with inline Wyatt Dawn Heleos II and Optilab rEX detectors, and was analysed with the Astra software and according to the method of Slotboom [34]. The calculated dn/dc of LicB was 0.194 and dn/dc of DDM was taken as 0.144.

Ligand binding assay

ITC experiments were performed at 10 µM LicB (titrand) in independent experiments using different choline (titrant) stock concentrations between 0.05 to 1 mM. Control experiments were performed to determine background noise from buffer mixing or titrant dilution. For fluorescence studies, dansylcholine was prepared according to the method of Weber [35], using 1 ml acetone at the step in the procedure calling for "additional acetone". Fluorimetry was performed on an Agilent Technologies Cary Eclipse instrument with excitation at 295 nm and

emission between 305 and 650 nm. All assays were performed in 50 mM sodium phosphate pH 7.4, 150 mM NaCl, 5% glycerol, 0.05% DDM at 25°C. For ligand binding experiments, 1.5 µM purified LicB or 9 µM N-acetyl-L-tryptophanamide (NATA) were pre-equilibrated for 10 min in assay buffer before being titrated with small volumes of dansylcholine. For competition assays, 1.5 µM LicB was similarly pre-incubated with dansylcholine at 100 µM prior to titration with other ligands.

Direct ligand binding data were fit to a hyperbolic binding equation:

$$Y = B_{max} * [L] / K_d + [L] \qquad\qquad \text{[Equation 1]}$$

Where [L] is ligand concentration. Displacement measurements were transformed to $\log_{10}[L]$ before fitting to the following sigmoidal function:

$$Y = Ymin + (Ymax-Ymin) / 1+10^{((\log IC50-[L]) * m))} \qquad \text{[Equation 2]}$$

Results

A bioinformatic sequence analysis confirmed earlier predictions [27] that LicB would be a multipass integral membrane protein comprising 10 transmembrane (TM) α-helices connected by short extramembrane loops (Fig. 1a). Computational modelling of LicB by both *ab initio* and homology methods converged on the same overall helix-bundle fold as other members of the DMT superfamily including PDB IDs 5I20, 5OGE and 6OH3 (Fig. 1b). This differs from the 12TM architecture common to the BCCT transporters [36]. Multiple sequence alignment with a non-redundant set of 34 full-length bacterial homologues revealed 30 conserved amino acid residues (Supplementary Figure S1), shown in red in Fig. 1a and 1b. Fig. 1c shows an illustrative section of this sequence alignment comparing *H. influenzae* LicB to five other homologues. This highlights the particular high sequence conservation of residues within TMs 1 and 3, and at the cytoplasmic ends of TMs 8 and 9. In the computational models many of these conserved residues had their sidechains oriented into the centre of the protein bundle, consistent with an important role for these amino acids in protein folding, ligand binding or ligand transport.
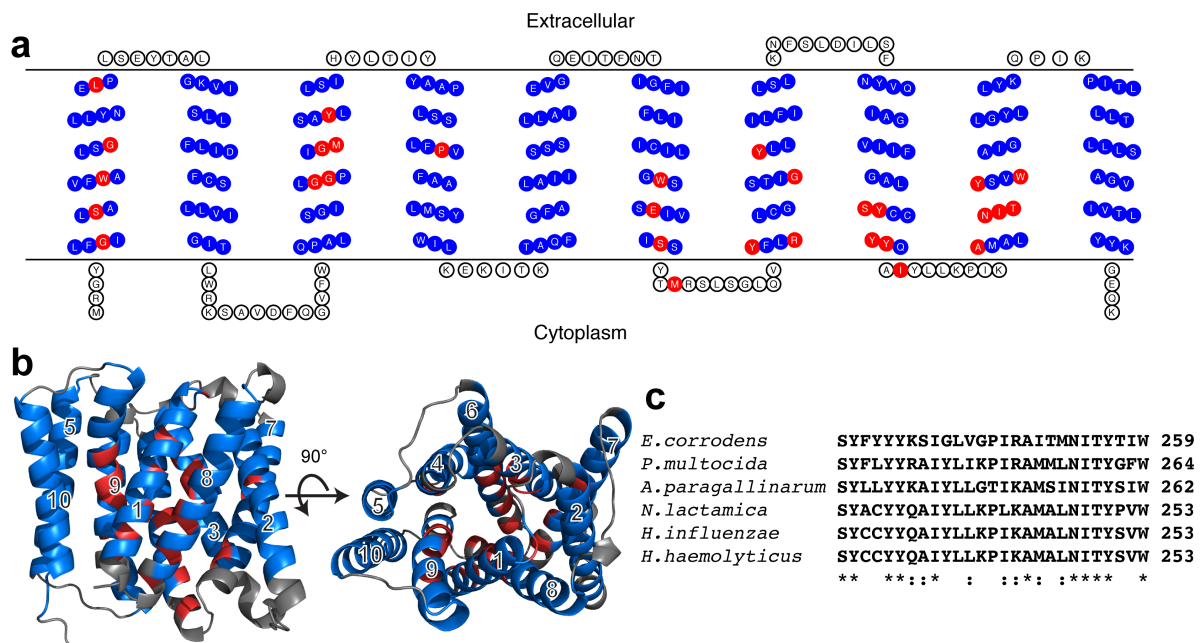

Figure 1. Transmembrane topology, structure prediction and sequence conservation of *H. influenzae* LicB. (**a**) Schematic drawing showing the predicted number and topology of membrane-spanning helices in LicB. Residues predicted to be in transmembrane segments are shown in blue and conserved residues are coloured red. (**b**) *ab initio* computational model of LicB, using the same colour scheme as panel (**a**). (**c**) Illustrative section of a multiple sequence alignment of *H. influenzae* LicB against some homologues arbitrarily selected for display, highlighting the sequence conservation in transmembrane helices 8 and 9. Annotations below the alignment relate to the full MSA of 34 homologues, which is provided as Supplementary Figure S1.

Protein expression and purification

Recombinant *H. influenzae* LicB was cloned into the T7 expression vector pET28a with a polyhistidine affinity tag at the C-terminus. This cloning strategy deliberately avoided any N-terminal tags since the first helix of polytopic integral membrane proteins generally acts as a membrane localisation signal [37]. The resulting construct was successfully expressed in the *E. coli* strain T7 Express cultured in SB autoinduction media (Fig. 2a). Very little or no expression was observed in the strain BL21Star or the Walker strain C43(DE3). Using autoinduction broth resulted in protein yields approximately 3-fold higher than standard IPTG induction protocols in LB media. Reducing the temperature to 25°C during induction surprisingly appeared to increase protein aggregation during purification. Thus our expression conditions were routinely T7 Express, 37°C, 24h in autoinduction SB media.

LicB was purified via immobilised metal affinity chromatography from recombinant cell membranes solubilised in either 1% *n*-dodecyl-β-D-maltopyranoside (DDM), 1% *n*-decyl-β-D maltopyranoside (DM) or 2.4% 5-cyclohexyl-1-pentyl-β-D-maltopyranoside (Cymal-5). The maltosides were chosen since they are mild, non-denaturing detergents that are known to solubilise membrane proteins which are properly folded after overexpression in *E. coli* [38, 39]. Overall protein yields from these experiments were similar with 1.8, 2.4 and 1.7 mg purified LicB obtained per litre of expression culture for DDM, DM and Cymal-5, respectively. Coomassie-stained SDS-PAGE gels showed that in all cases a major band corresponding to LicB was recovered from the affinity columns, running slightly faster than the expected molecular weight of 34.1 kDa (Fig. 2a). Many membrane proteins migrate atypically on SDS-PAGE gels, and applying the Rath-Deber correction factor of 0.82 for fast-migrating proteins gave an apparent weight close to the theoretical value [40]. This same band was seen in Western blots of fractionated recombinant cell membranes, although intense high-molecular weight bands were also observed suggesting non-specific binding, protein aggregation, or both. The LicB band disappeared when the purified sample was boiled before loading, implying that LicB irreversibly aggregated at high temperatures and could no longer enter the gel bed. It thus appeared that all three of the different detergents tested could support purification of LicB. However, preparative size-exclusion chromatography showed that LicB purified in Cymal-5 was susceptible to extensive non-specific aggregation (Fig. 2b), with this heterogeneity making this detergent unsuitable for further work. Some aggregation was also apparent in chromatograms of DDM and DM, but the strong single peak eluting at later volumes suggested that these surfactants did sustain a stable population of folded and disperse LicB. Fractions

corresponding only to this well-resolved single peak were collected and used for all further analyses.
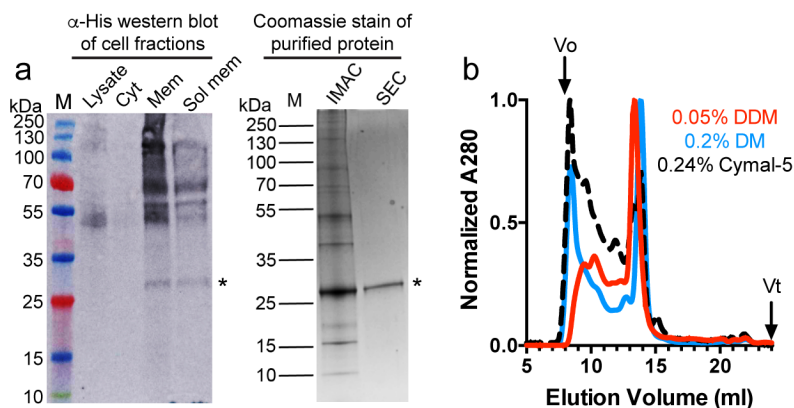

Figure 2: Purification of recombinant *H. influenzae* LicB. (**a**) Anti-His-tag western blot of *E. coli* cell fractions and Coomassie-stained SDS-PAGE of affinity-purifed LicB. *Lysate*, total cell lysate; *Cyt*, cytosol; *Mem*, membranes; *Sol mem*, membranes after solubilisation in 1% DDM. *IMAC*, eluent from Ni-NTA affinity column; *SEC*, pooled fractions corresponding to the well-resolved peak in size exclusion chromatogram. (**b**) Size exclusion profiles of LicB in each of the three different detergents DDM, DM and Cymal-5. Absorption data are normalized for direct comparison. *A280*, in-line detection of protein absorption at 280 nm.

Protein characterisation

We next assessed the secondary structure of purified LicB in 0.05% DDM and 0.2% DM. Circular dichroism (CD) confirmed that purified LicB was an α-helical protein (Fig. 3a), in agreement with predictions from sequence analysis. DDM supported a greater degree of secondary structure formation than DM. The overall helicity of ~62% in DDM was also consistent with the bioinformatic predictions that approximately 206 of the 305 amino acids sit within transmembrane helices. Thermal denaturation curves revealed a major cooperative loss of secondary structure with T_m of ~66°C, above which point LicB irreversibly aggregates (Fig. 3b). In either detergent, LicB appeared to be stable at room temperature and retained secondary structure over 16h continuous incubation at 25°C (Fig. 3c).

Since DDM appeared to be the detergent best-suited to the proper folding of LicB, the oligomeric state of DDM-solubilised protein was determined by reapplying this sample to a size-exclusion column connected to an in-line light scattering instrument (SEC-MALS; Fig. 3d). Using the three-detector method described by Slotboom [34], the calculated mass of LicB within the micelle was 29 kDa, close to the theoretical monomer mass of 34.1 kDa. The mass ratio of detergent to protein (δ) was found to be 3.4. Collectively, these data suggest that LicB

is a well-folded monomer within the DDM micelle and that this monomer state persists in the long term at ambient temperatures.
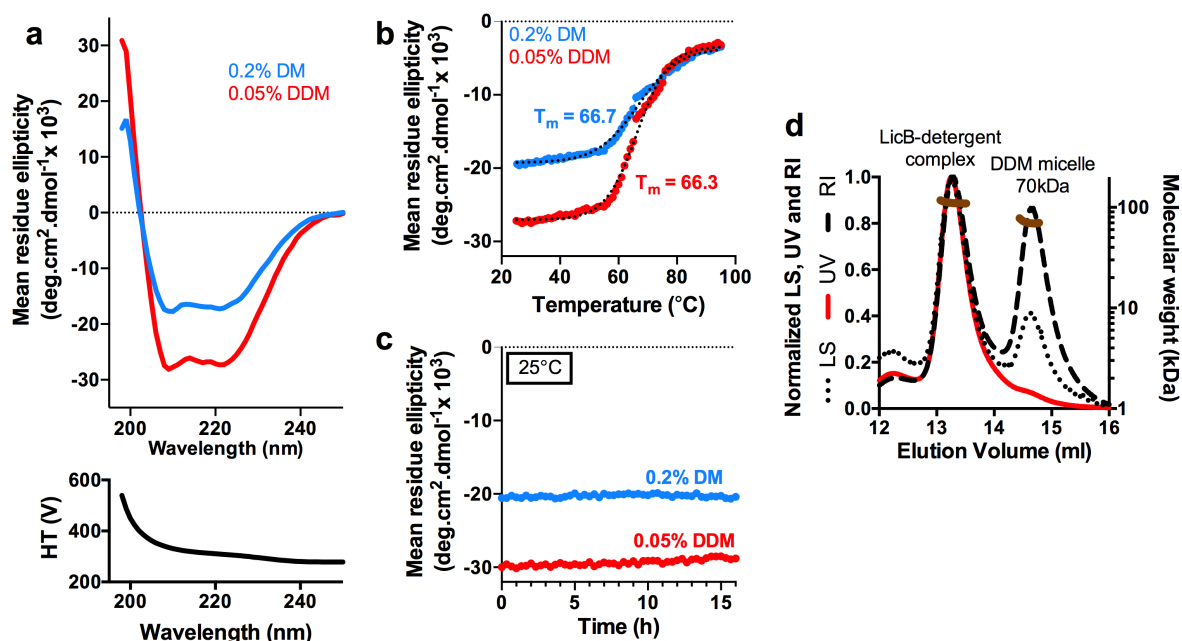
Figure 3: Characterisation of LicB purified in the two maltoside surfactants DDM and DM. (**a**) Circular dichroism spectroscopy determines that LicB has greater helicity when solubilised in DDM versus DM. (**b**) In either detergent, LicB shows cooperative thermal unfolding with a midpoint temperature of 66°C and (**c**) is stable at room temperature. (**d**) SEC-MALS in DDM confirms that purified LicB is a monodisperse monomer. See text for discussion.
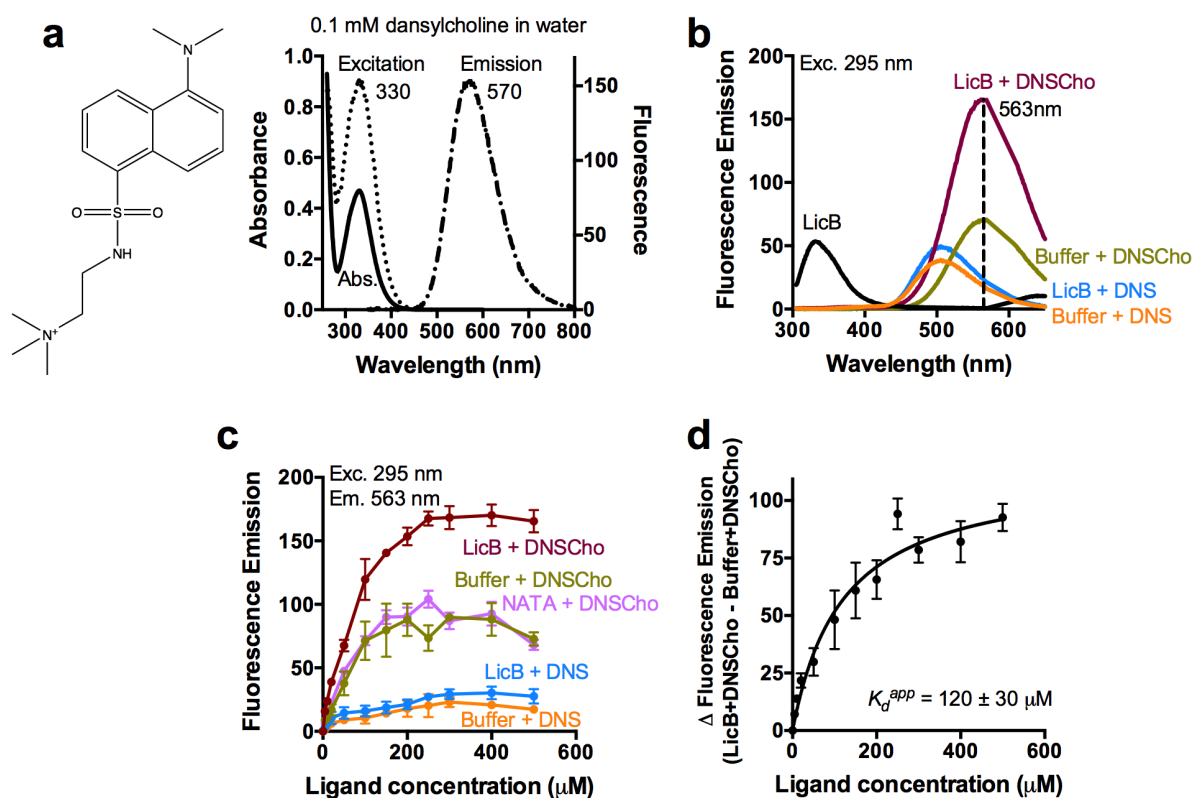
Binding of a fluorescent ligand

Since LicB is proposed to recognise choline with relatively high affinity (K_m = 3.6 μM; [27]) it ought to be possible to detect substrate binding at equilibrium. We initially attempted to study choline binding to LicB by isothermal titration calorimetry (ITC) but were unable to detect any heat changes associated with ligand binding (Supplementary Figure S2). The amount of protein required for each ITC run precluded more extensive testing. We also pursued choline transport assays, but these were inconclusive (not shown). Instead, we turned to simple and sensitive fluorescence methods to assess protein-ligand interactions. Previously both the melibiose and lactose permeases were studied using resonant energy transfer (FRET) from intrinsic tryptophan residues to dansylated sugars [41-44]. An analogous fluorescent probe is available for LicB: dansylcholine (1-dimethylaminonapthalene-5-sulfonamidoethyltrimethylammonium; hereafter, DNSCho) which was originally developed to study cholinergic receptors [35, 45]. This probe retains the positively-charged trimethylammonium "head" of choline, which is known to be critical for substrate binding to

other choline transporters [36], and replaces the hydroxyl group at the end of the choline alkyl 'tail' with the dansyl fluorophore (Fig. 4a). The Förster distance of the Trp-Dansyl pair is 21 Å [46], and Trp→Dansyl FRET is plausible given the relatively high number of Trps in LicB, including those facing the bundle interior that might be involved in substrate binding. However, if specific interactions or shape compatibility with the choline 'tail' are an important part of the binding mode then DNSCho would be expected to bind to LicB with weaker overall affinity, if at all. The successful synthesis of DNSCho [35] was confirmed by mass spectrometry (Supplementary Figure S3).

As expected, the DNSCho probe had different spectral properties to underivatized dansyl chloride with fluorescence excitation at 330 nm producing an emission band that was shifted to 570 nm (Fig 4a). Introducing DNSCho into LicB resulted in an increase in probe fluorescence upon excitation at 295 nm, which specifically excites the protein Trp donor(s) (Fig. 4b). Under these conditions the emission intensity of the DNSCho acceptor was enhanced nearly 2-fold over control experiments that either titrated DNSCho into the reaction buffer or which replaced LicB with the soluble tryptophan analogue N-acetyl-L-tryptophanamide (Fig. 4b-c; Supplementary Fig. S4). The observed increase in DNSCho emission was not accompanied by any obvious shift in the emission wavelength. A substantial background level of fluorescence emission could be observed even in the absence of LicB, reflecting the very broad excitation band of the dansyl group (Fig. 4b). No corresponding increase in emission was observed in control experiments with underivatized dansyl (Fig. 4b). These results are consistent with a specific protein-ligand interaction allowing energy transfer from Trp donor(s) to the DNSCho acceptor, resulting in increased acceptor emission. This ought to coincide with a decrease in donor emission, and this was observed; however, the donor signal was influenced significantly by inner filter effects, making interpretation challenging (Supplementary Figure S4). We thus instead used the changes in DNSCho emission intensity to assess ligand binding to LicB.

Equilibrium binding experiments were performed by titrating DNSCho into LicB or a buffer-only control (Fig. 4c). The further analysis of these data was complicated by the relatively high and non-linear background signal, but subtracting the buffer-only experiment did result in a hyperbolic binding curve (Fig. 4d) with apparent K_d of 120 ± 30 μM. This result is consistent with a relatively high-affinity permease-ligand interaction; however it is much weaker than the K_m for the native ligand, choline, of 3.6 μM [27]. This likely arises from differences in

equilibrium binding in detergent micelles versus membrane transport, the detergent-solubilised protein adopting a conformation that is suboptimal for substrate binding, and the presence of the dansyl moiety. Nonetheless this assay ought to be useful for the qualitative comparison of the binding of different ligands.


Figure 4. Equilibrium binding of the fluorescent ligand dansylcholine (DNSCho) to purified recombinant LicB. (**a**) Absorption and emission spectrum of 0.1 mM DNSCho in water. (**b**) Representative scans of LicB and equivalent buffer controls in the absence and presence of either DNSCho or underivatized dansyl chloride (DNS). Protein 1.5 μM, ligand 500 μM. The increased emission of DNSCho is consistent with FRET between intrinsic Trps of LicB (donor) and DNSCho (acceptor). (**c**) Titrations of DNSCho show a non-linear increase in signal both with and without LicB. Protein at 1.5μM where used. *NATA*, additional control replacing LicB with N-acetyl-L-tryptophanamide. (**d**) Subtracting the buffer control gives an apparent hyperbola, fit to Equation 1. Panels (**c**) and (**d**) show mean ± s.d. from at least two experiments.

Competition, or displacement, assays were then used to screen a panel of small molecules for interactions with LicB. The protein was pre-incubated with sub-saturating concentrations of DNSCho (100 μM; still at large excess over protein) so that a small protein donor peak remained at 330 nm that could be used as an internal control to correct for non-specific signal

changes and dilution. Data are thus presented as the ratio of the acceptor emission at 563 nm to the donor emission at 330 nm, normalised to the same ratio obtained at zero ligand ($F^{563/330}/F_0^{563/330}$). When choline was used as the titrant, a reproducible decrease was observed in the acceptor emission consistent with the displacement of specifically-bound DNSCho and loss of FRET (Figs. 5a and 5c). Almost no change was observed in donor emission at 340 nm. This was distinct from controls in which equivalent volumes of buffer were added, which showed a slight decrease in the overall background fluorescence at both 560 and 340 nm but gave no overall corrected signal change (Figs. 5b and 5c).

Figs 5c and 5d show that both choline and acetylcholine partially displaced DNSCho from LicB, with the midpoint of the titration curve (IC_{50}) being at 0.5 mM and 1.2 mM ligand, respectively. Results for other small molecules are shown in Fig. 5e-i. Titrations with benzoylcholine were inconclusive, and neither L-carnitine nor L-betaine could displace DNSCho despite the reasonably close chemical similarity of these compounds to choline. Experiments with L-glycine and L-cysteine also showed no displacement. Thus, LicB is apparently selective for binding choline and has markedly reduced affinity for other possible substrates, with substitutions to the alkyl chain being disfavoured. Such selectivity would be congruent with the proposed function of LicB.
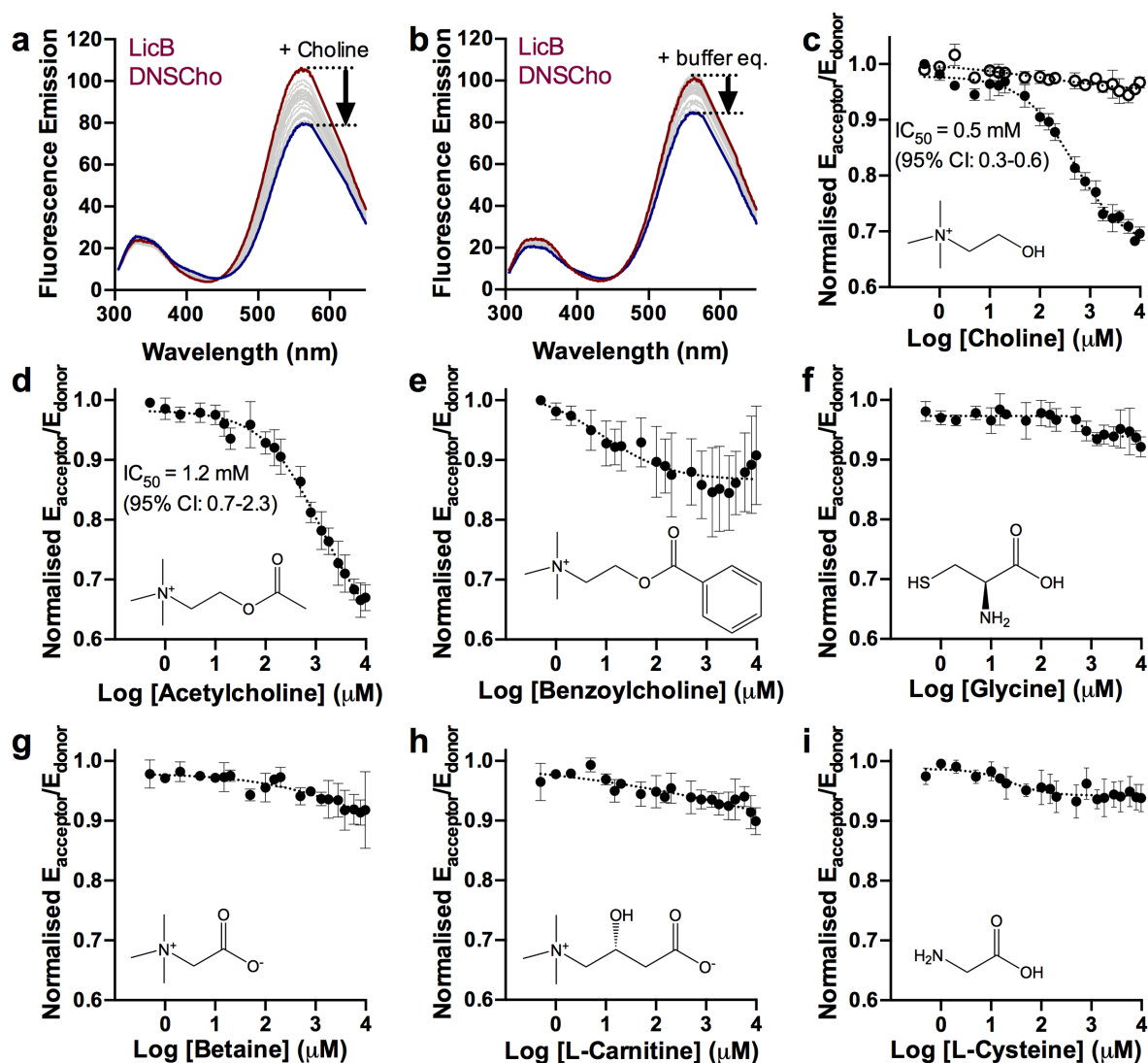
Figure 5. DNSCho displacement assays show that LicB is selective for choline and acetylcholine. (**a**) DNSCho acceptor emission is diminished by titrating choline into pre-equilibrated protein-DNSCho complexes. (**b**) This displacement effect is less pronounced when equivalent volumes of buffer are used rather than ligand. (**c**) Data from panels (**a**) and (**b**) plotted as a titration curve. *Filled circles*, protein plus choline; *open circles*, protein plus equivalent buffer volumes. The ratio of the acceptor and donor peaks is used after both sets of data are normalised to the initial emission intensity. (**d-i**) Ligand displacement measurements using other small molecules as shown. All data are mean ± s.d of three independent repeats. *95% CI*, range of the 95% confidence interval. Dotted lines show the fit to Equation 2.

Site-directed mutagenesis of conserved residues

Figure 1 highlights a highly conserved region found at the cytoplasmic end of TM9 that features the motif NITYxxW. Site-directed mutagenesis was used to probe the role of each of these conserved residues on the folding, assembly and function of LicB. This resulted in a set of mutant proteins N247A, I248A, T249A, Y250A and W253A, numbered according to their sequence position in *H. influenzae* LicB.

Each of these mutations was recombinantly expressed according to the protocols used for WT LicB. In every case the LicB mutants were recovered at a lower yield than WT (Fig. 6a). W253 was produced at an almost undetectable level and was not pursued further. T249A showed substantial aggregation during size-exclusion chromatography (Fig. 6b). All mutants showed reduced helicity in CD (Fig. 6c) and this was particularly acute for T249A and Y250A; however, the precision of this measurement was likely to have been affected by uncertainty around exact protein concentrations given the low levels of mutant protein available. T249A also showed a loss of cooperative unfolding when heated, providing further evidence that this protein was poorly folded (Fig. 6d). Other mutants behaved similarly to wildtype when heated, with a midpoint transition of approximately 65°C.
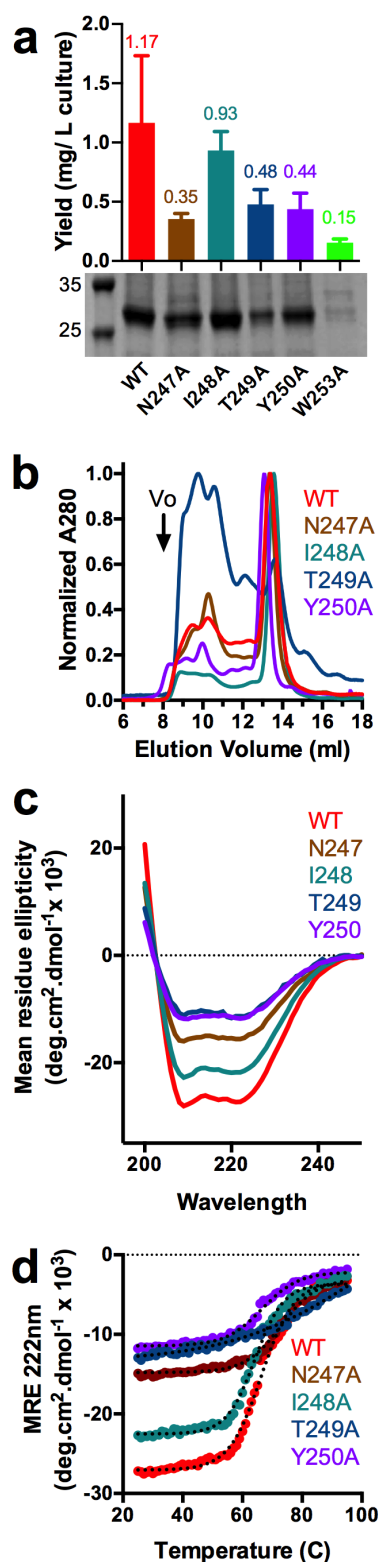
Figure 6. Recombinant expression and characteristion of LicB mutants. (**a**) All mutants express at lower levels than WT LicB. Data are mean ± s.d of at least three protein preparations. (**b**) Mutant size exclusion profiles are similar to WT except for T249A, which is aggregation-prone. (**c**) Mutants do not adopt the same degree of secondary structure as WT, and this appears to be severely compromised in N247, T249A and Y250A although the exact determination of MRE is uncertain; see text. (**d**) Thermal unfolding of mutant proteins. Note the loss of cooperativity in T249A.

Competition assays with choline (Fig. 7) showed that all of these mutants had a lower affinity for choline, with any such displacement being abolished completely in T249A. This confirms the apparent loss of structure in this mutant. In the context of these findings, we were surprised to find that the direct DNSCho binding signal from T249A (*i.e.* equivalent to Fig. 4b) was only slightly lower than WT. This implies that non-specific background effects may be more prevalent in this mutant, perhaps because the apparent unfolding of T249A exposes hydrophobic patches on the protein that interact with the label. It may also be that such non-specific effects may contribute more significantly than was originally suspected to the direct binding data presented in Fig. 4.

Overall, mutation I248A had a minimal effect on both folding and ligand binding, being rather similar to WT. Other mutations compromised the structural integrity of LicB, and these were associated with changes in ligand binding capability. The most surprising finding was for Y250A, which appeared to lose a substantial degree of secondary structure (Fig. 6c) yet still retain some vestiges of ligand binding (Fig. 7d). Further characterisation will be required to understand the true impact of this mutation on LicB.
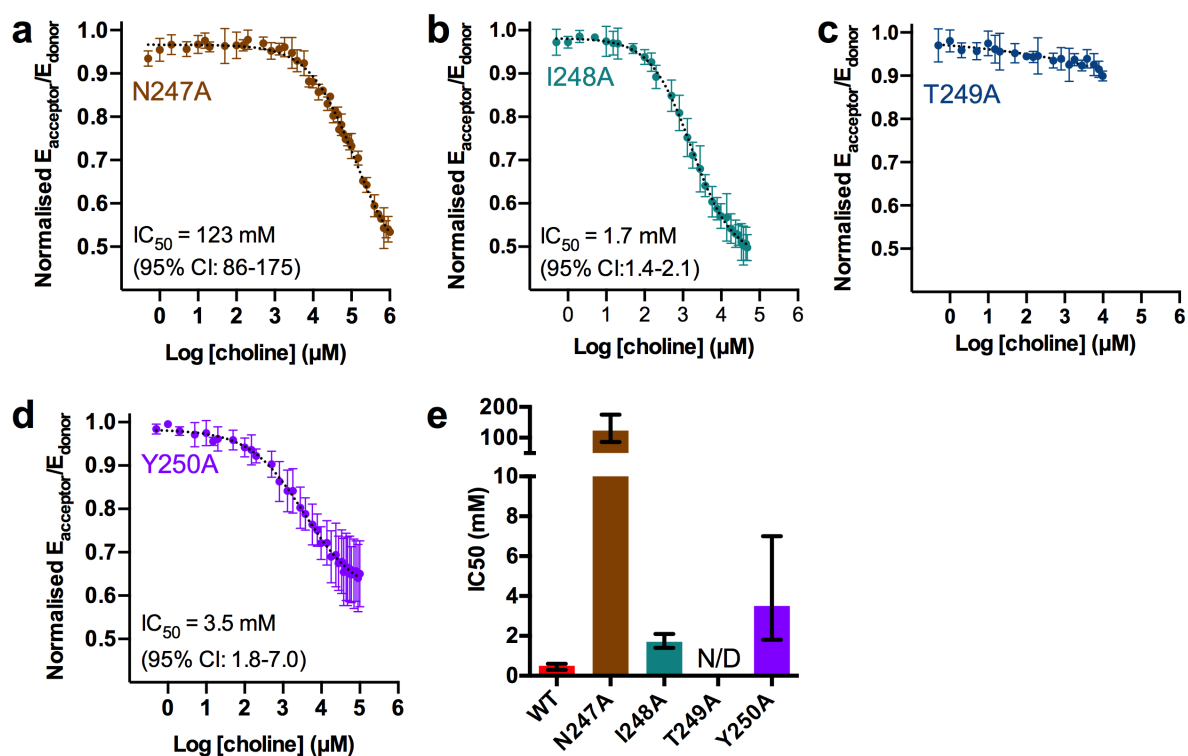
Figure 7. DNSCho displacement assays of LicB mutants. (**a-d**) Data collected on different mutant as shown. Data are mean ± s.d., n=3. (**e**) Comparison of mutant and WT data, shown as $IC_{50} \pm 95\%$ CI.

Discussion

This report concerns the first *in vitro* studies of a member of the LicB family of choline permeases. This protein is of interest because its biological function supports the colonisation of the human airways by a range of bacteria, including pathogenic strains that are responsible for critical illness. The role of LicB in synthesising the virulence factor ChoP makes this protein a potential target for anti-infective drugs.

The recombinant technology employed here is potentially a powerful tool for understanding LicB structure and function away from the complex cell environment. However, a major barrier is the isolation of sufficient quantities of monodisperse, well-folded protein-detergent complexes which are a prerequisite for biophysical analysis [47]. We have determined that this challenge can be overcome successfully for LicB by identifying an appropriate expression system (T7 Express strain in autoinduction broth) and by employing DDM as the solubilising surfactant (Fig. 2). DDM-solubilised LicB is a stable monomer that largely resists aggregation and appears to be fully-folded and compact (Fig. 3). DDM is a widely-used detergent that is compatible with many other protein analyses, including reconstitution into synthetic lipid systems [48] and protein crystallisation [49]. Improved protein yields would add considerable impetus to further studies of LicB, and future work should screen a panel of LicB homologues to identify alternative sequences with higher levels of recombinant production.

Attempts to understand the function of LicB were frustrated by a lack of signal in ITC experiments using protein-detergent complexes (Fig. S2), and inconclusive evidence of choline transport in proteoliposomes (not shown). While calorimetry is known to be challenging for membrane permeases, LicB is expected to have an affinity for choline in the low micromolar range [27] which ought to be compatible with ITC [50]. Since it was thus unclear whether purified LicB was folded into an active state, the fluorescent probe DNSCho was used in equilibrium binding studies. While the precise details of the putative Trp→DNSCho FRET system remain to be determined, the results allow for cautious and provisional conclusions about the activity of LicB. Equilibrium binding and displacement assays do suggest that LicB has an affinity towards choline in the micromolar range, and that LicB discriminates for choline over related small molecules (Figs. 4 and 5). Mutations to conserved residues appear to severely disrupt protein folding and function, consistent with their presumed importance in LicB family proteins (Figs 6 and 7). The exception was the mutation I248A, which was better-tolerated and resembled WT. These studies pave the way for more detailed insights into ligand

binding and transport by LicB. This should include the LicB-mediated transport of radiolabelled choline in proteoliposomes, to assess whether LicB requires an energetic proton gradient for substrate transport. This reconstituted system could also be used to provide further validation of the ligand binding assays carried out in micelles here.

Insights from structural biology of the BCCT family have suggested the following general principles for their interaction with substrates such as choline [36]. As with other membrane transporters, the substrate binding region is a solvent-accessible pocket located in the membrane centre that is alternatively exposed from one side of the membrane to the other. The substrate 'headgroup' is coordinated through cation-pi and Van der Waals interactions with the indole sidechains of a set of highly conserved tryptophans. This has been dubbed an 'aromatic cage' [36]. Our computational model of LicB shows a series of aromatic residues that face the interior cavity of the protein and so may be candidates for forming a similar aromatic box. Computational ligand docking showed that choline may be accommodated between these residues in LicB, nestled between W14, Y84, W162 and Y250. In this position the substrate could make cation-pi interactions to W162 and in this particular simulation forms a hydrogen bond with Y250 (Fig. 8). The choline 'tail' points out of the pocket, which could provide the space required to allow DNSCho binding.
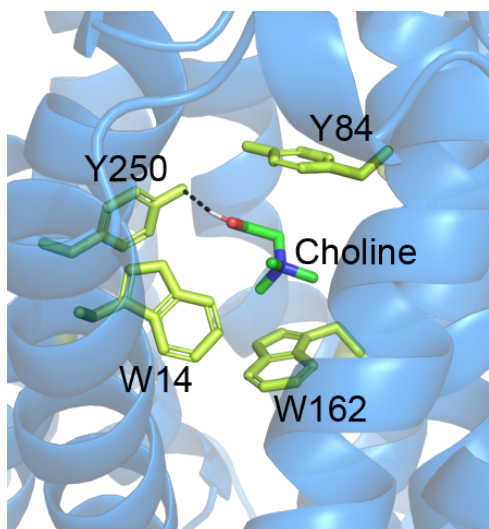
Fig. 8. View of the putative 'aromatic box' binding mode for choline docked into the LicB model. Docking was performed with HADDOCK online [33].

Conclusions

The current model for the biological role of the LicB permeases rests upon the ability of this protein family to bind choline at low concentrations. Our first attempts to study LicB *in vitro* broadly support this model, and report that LicB is relatively selective in recognising choline.

Further work is now required to elucidate whether LicB-type transporters coordinate substrate via the aromatic box mechanism found in osmoregulatory choline transporters, and whether LicB is indeed a proton-coupled symporter. A more complete view of the structure and function of LicB remains of interest for the understanding and treatment of communicable respiratory infections.

References

1.   Wahl, B., K.L. O'Brien, A. Greenbaum, A. Majumder, L. Liu, Y. Chu, et al., (2018) *Burden of Streptococcus pneumoniae and Haemophilus influenzae type b disease in children in the era of conjugate vaccines: global, regional and national estimates for 2000-2015.* Lancet Glob Health, **6**: e744-57.

2.   GBD 2016 Lower Respiratory Infections Collaborators, (2018) *Estimates of the global, regional, and national morbidity, mortality and aetiologies of lower respiratory infections in 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016.* Lancet Infect Dis, **18**: 1191-210.

3.   Clark, S.E., (2020) *Commensal bacteria in the upper respiratory tract regulate susceptibility to infection.* Curr Opin Immunol, **66**: 42-49.

4.   Siegel, S.J. and J.N. Weiser, (2015) *Mechanisms of bacterial colonization of the respiratory tract.* Annu Rev Microbiol, **69**: 425-44.

5.   Clark, S.E. and J.N. Weiser, (2013) *Microbial modulation of host immunity with the small molecule phosphorylcholine.* Infect Immun, **81**(2): 392-401.

6.   Weiser, J.N., M. Shchepetov, and S.T.H. Chong, (1997) *Decoration of lipopolysaccharide with phosphorylcholine: a phase-variable characteristic of Haemophilus influenzae.* Infect Immun, **65**(3): 943-50.

7.   Fischer, W., T. Behr, R. Hartmann, J. Peter-Katalinić, and H. Egge, (1993) *Teichoic and lipoteichoic acid of Streptococcus pneumoniae possess identical chain structures.* Eur J Biochem, **215**: 851-7.

8.   Weiser, J.N., N. Pan, K.L. McGowan, D. Musher, A. Martin, and J. Richards, (1998) *Phosphorylcholine on the lipopolysaccharide of Haemophilus influenzae contributes to persistence in the respiratory tract and sensitivity to serum killing mediated by C-reactive protein.* J Exp Med, **187**(4): 631-40.

9.   Pang, B., D. Winn, R.W. Johnson, W. Hong, S. West-Barnette, N. Kock, et al., (2008) *Polysaccharides containing phosphorylcholine delay pulmonary clearance of nontypeable Haemophilus influenzae.* Infect Immun, **76**(5): 2037-43.

10.   Swords, W.E., B.A. Buscher, K. Ver Steeg Li, A. Preston, W.A. Nichols, J.N. Weiser, et al., (2002) *Non-typeable Haemophilus influenzae adhere to and invade bronchial epithelial cells via an interaction of lipooligosaccharide with the PAF receptor.* Mol Microbiol, **37**(1): 13-27.

11.   Young, N.M., S.J. Foote, and W.W. Wakarchuk, (2013) *Review of phosphocholine substituents on bacterial pathogen glycans: synthesis, structures and interactions with host proteins.* Mol Immunol, **56**(4): 563-573.

12.   Cundell, D.R., N.P. Gerard, C. Gerard, I. Idanpaan-Heikkila, and E. Tuomanen, (1995) *Streptococcus pneumoniae anchor to activated human cells by the receptor for platelet-activating factor.* Nature, **377**: 435-8.

13.   Weiser, J.N., D.M. Ferreira, and J.C. Paton, (2018) *Streptococcus pneumoniae: transmission, colonization and invasion.* Nat Rev Microbiol, **16**: 255-67.

14.   Hergott, C.B., A.M. Roche, N.A. Naidu, C. Mesaros, I.A. Blair, and J.N. Weiser, (2015) *Bacterial exploitation of phosphorylcholine mimicry suppresses inflammation to promote airway infection.* J Clin Invest, **125**(10): 3878-90.

15.   Goldenberg, H.B., T.L. McCool, and J.N. Weiser, (2004) *Cross-reactivity of human immunoglobulin G2 recognizing phosphorylcholine and evidence for protection against major bacterial pathogens of the human respiratory tract* J Infect Disease, **190**(7): 1254-63.

16.   Casey, R., J. Newcombe, J. McFadden, and K.B. Bodman-Smith, (2008) *The acute-phase reactant C-reactive protein binds to phosphorylcholine-expressing Neisseria*

meningitidis and increases uptake by human phagocytes Infect Immun, **76**(3): 1298-304.

17. Weiser, J.N., J.M. Love, and E.R. Moxon, (1989) *The molecular mechanism of phase variation of H. influenzae lipopolysaccharide.* Cell, **59**(4): 657-65.

18. Humphries, H.E. and N.J. High, (2002) *The role of licA phase variation in the pathogenesis of invasive disease by Haemophilus influenzae type b.* FEMS Immunol Med Microbiol, **34**(3): 221-30.

19. Weiser, J.N. and N. Pan, (2002) *Adaptation of Haemophilus influenzae to acquired and innate humoral immunity based on phase variation of lipoolysaccharide.* Mol Microbiol, **30**(4): 767-75.

20. Clark, S.E., J. Snow, J. Li, T.A. Zola, and J.N. Weiser, (2012) *Phosphorylcholine allows for evasion of bactericidal antibody by Haemophilus influenzae.* PLOS Pathogens, **8**: e1002521.

21. Lysenko, E., J.C. Richards, A.D. Cox, A. Stewart, A. Martin, M. Kapoor, et al., (2000) *The position of phosphorylcholine on the lipopolysaccharide of Haemophilus influenzae affects binding and sensitivity to C-reactive protein-mediated killing.* Mol Microbiol, **35**(1): 234-45.

22. Lysenko, E., J. Gould, R. Bals, J.M. Wilson, and J.N. Weiser, (2000) *Bacterial phosphorylcholine decreases susceptibility to the antimicrobial peptide LL-37/hCAP18 expressed in the upper respiratory tract.* Infect Immun, **68**: 1664-71.

23. Hong, W., B. Pang, S. West-Barnette, and W.E. Swords, (2007) *Phosphorylcholine expression by nontypeable Haemophilus influenzae correlates with maturation of biofilm communities in vitro and in vivo.* J Bacteriol, **189**(22): 8300-7.

24. Wang, L., Y.L. Jiang, J.R. Zhang, C.Z. Zhouc, and Y. Chen, (2015) *Structural and enzymatic characterisation of the choline kinase LicA from Streptococcus pneumoniae.* PLoS One, **10**(3): e0120467.

25. Kwak, B.Y., Y.M. ZHang, M. Yun, R.J. Heath, C.O. Rock, S. Jackowski, et al., (2002) *Structure and mechaninism of CTP:phosphocholine cytidyltransferase (LicC) from Streptococcus pneumoniae.* J Biol Chem, **277**(6): 4343-50.

26. Saier, M., V. Reddy, G. Moreno-Hagelsieb, K. Hendargo, Y.M. ZHang, V. Iddamsetty, et al., (2021) *The transporter classification database (TCDB): 2021 update.* Nucleic Acids Research, **49**: D461-7.

27. Fan, X., C.D. Pericone, E. Lysenko, H. Goldfine, and J.N. Weiser, (2003) *Multiple mechanisms for choline transport and utilization in Haemophilus influenzae.* Mol Microbiol, **50**(2): 537-548.

28. Bondy, J., S. Osharovich, J. Storm, G. Durning, T. McAuliffe, and X. Fan, (2016) *Haemophilus influenzae LicB contributes to lung damage in an aged mouse co-infection model.* Microb Pathogen, **90**: 1-6.

29. Eberhardt, A., L.J. Wu, J. Errington, W. Vollmer, and J.-W. Veening, (2009) *Cellular localization of choline-utilization proteins in Streptococcus pneumoniae using novel fluorescent reporter systems.* Mol Microbiol, **74**(2): 395-408.

30. Liu, X., C. Gallay, M. Kjos, A. Domenech, J. Slager, S.P. van Kessel, et al., (2017) *High-throughput CRISPRi phenotyping identifies new essential genes in Streptococcus pneumoniae.* Mol Syst Biol, **13**(5): 931.

31. Baek, M., F. DiMaio, I. Anishchenko, J. Dauparas, S. Ovchinnikov, G.R. Lee, et al., (2021) *Accurate prediction of protein structures and interactions using a three-track neural network.* Science, **373**(6557): 871-876.

32. Song, Y., F. DiMaio, R.Y.-R. Wang, D. Kim, C. Miles, T.J. Brunette, et al., (2013) *High resolution comparative modeling with RosettaCM.* Structure, **21**(8): 1735-42.

33. van Zundert, G.C.P., J.P.G.L.M. Rodrigues, M. Trellet, C. Schmitz, P.L. Kastritis, E. Karaca, et al., (2016) *The HADDOCK2.2 webserver: user-friendly integrative modeling of biomolecular complexes.* J Mol Biol, **428**(4): 720-725.

34. Slotboom, D., R. Duurkens, K. Olieman, and G. Erkens, (2008) *Static light scattering to characterize membrane proteins in detergent solution.* Methods, **46**: 73-82.

35. Weber, G., D.P. Borris, E. De Robertis, F.J. Barrantes, J.L. La Torre, and M.C. Llorente de Carlin, (1971) *The use of a cholinergic fluorescent probe for the study of the receptor phospholipid.* Mol Pharmacol, **7**: 530-537.

36. Ziegler, C., E. Bremer, and R. Krämer, (2010) *The BCCT family of carriers: from physiology to crystal structure.* Mol Microbiol, **78**(1): 13-34.

37. Goder, V. and M. Spiess, (2001) *Topogenesis of membrane proteins: determinants and dynamics.* FEBS Letters, **504**(3): 87-93.

38. Mathieu, K., W. Javed, S. Vallet, C. Lesterlin, M.-P. Candusso, F. Ding, et al., (2019) *Functionality of membrane proteins overexpressed and purified from E. coli is highly dependent upon the strain.* Sci Rep, **9**: 2654.

39. Stetsenko, A. and A. Guskov, (2017) *An overview of the top ten detergents used for membrane protein crystallization.* Crystals, **7**(7): 197.

40. Rath, A. and C.M. Deber, (2013) *Correction factors for membrane protein molecular weight readouts on sodium dodecyl sulfate-polyacrylamide gel electrophoresis.* Anal Biochem, **434**: 67-72.

41. Maehrel, C., E. Cordat, I. Mus-Vuteau, and G. Leblanc, (1998) *Structural studies of the melibiose permease of Escherichia coli by fluorescence resonance energy transfer.* J Biol Chem, **273**(50): 33192-7.

42. Guan, L., S. Nurva, and S.P. Ankeshwarapu, (2011) *Mechanism of melibiose/cation symport of the melibiose permease of Salmonella typhimurium.* J Biol Chem, **286**(8): 6367-74.

43. Smirnova, I.N., V.N. Kasho, and H.R. Kaback, (2006) *Direct sugar binding to LacY measured by resonance energy transfer.* Biochemistry **45**: 15279-87.

44. Guan, L. and P. Hariharan, (2021) *X-ray crystallography reveals molecular recognition mechanism for sugar binding in a melibiose transporter MelB.* Communications Biology, **4**.

45. Ressl, S., A.C. Terwisscha van Scheltinga, C. Vonrhein, V. Ott, and C. Ziegler, (2009) *Molecular basis of transport and regulation in the Na$^+$/betaine symporter BetP.* Nature, **458**(7234): 47-52.

46. Wu, P. and L. Brand, (1994) *Resonance energy transfer: methods and applications.* Anal Biochem, **218**: 1-13.

47. Seddon, A.M., P. Curnow, and P.J. Booth, (2004) *Membrane proteins, lipids and detergents: not just a soap opera.* Biochim. Biophys. Acta, **1666**: 105-117.

48. Skrzypek, R., S. Iqbal, and R. Callaghan, (2018) *Methods of reconstitution to investigate membrane protein function.* Methods, **147**: 126-141.

49. Parker, J. and S. Newstead, (2016) *Membrane protein crystallisation: current trends and future perspectives.* Adv Exp Med Biol, **922**: 61-72.

50. Boudker, O. and S. Oh, (2015) *Isothermal titration calorimetry of ion-coupled membrane transporters.* Methods, **76**: 171-82.