

Int. j. adv. multidisc. res. stud. 2023; 3(2):888-910

Received: 24-02-2023 **Accepted:** 04-04-2023

International Journal of Advanced Multidisciplinary Research and Studies

ISSN: 2583-049X

Elevator mechanism of alternating access in the *Escherichia coli* concentrative nucleoside transporter NupC

¹Lijie Sun, ²Simon G Patching

^{1, 2} School of Biomedical Sciences and Astbury Centre, University of Leeds, Leeds, United Kingdom

Corresponding Author: Simon G Patching

Abstract

Members of the concentrative nucleoside transporter (CNT) family of proteins mediate uptake of nucleosides into cells driven by a cation gradient, which then enter salvage pathways for nucleic acid synthesis. In humans, they also transport hydrophilic anticancer and antiviral nucleoside analogue drugs into cells and tissues where they exert their pharmacological effects. Escherichia coli CNT NupC (400 residues) is pyrimidine-specific and driven by a proton gradient. We have used computational, biochemical, and biophysical characterize methods to evolutionary relationships, conservation of residues, structural domains, transmembrane helices, and residues involved in nucleoside binding and/or transport activity in NupC compared with those of sodium-driven Vibrio cholerae CNT (vcCNT) and human CNTs (hCNT1-3). As in the crystal structure of vcCNT, NupC appears to contain eight transmembranespanning *a*-helices. Wild-type NupC and single-cysteinecontaining mutants were tested for transport activity in energized E. coli whole cells and for binding of nucleosides in non-energized native inner membranes using novel crosspolarization magic-angle spinning solid-state nuclear magnetic resonance methods. Wild-type NupC had an apparent affinity of initial rate transport (K_m^{app}) for $[^{14}C]$ uridine of 22.2 \pm 3.7 μ M and an apparent binding affinity (K_d^{app}) for [1'- ¹³C]uridine of 1.8–2.6 mM. Mutant S142C retained transport and binding affinities similar to those of the wild type. Mutants G146C and E149C had no transport activity but retained varying degrees of partial binding activity with affinities decreasing in the following order: wild type > S142C > G146C > E149C. Results were explained with respect to a homology model of NupC based on the structure of vcCNT and a hypothetical elevator-type mechanism of alternating access membrane transport in NupC.

Keywords: Alternating Access Mechanism, Binding Affinity, Homology Model, Membrane Topology, Nucleoside Transport, Site-Directed Mutagenesis, Solid-State NMR

1. Introduction

The concentrative nucleoside transporter (CNT) family of proteins, found in both prokaryotic and eukaryotic organisms, catalyzes inward movement of nucleosides across cell membranes against their concentration gradient, driven by a cation gradient moving in the same direction. In lower organisms, such as bacteria, CNTs scavenge nucleosides from the environment, and these enter salvage pathways for nucleic acid synthesis. ^[1] In higher organisms, including humans, CNTs control cellular concentrations of physiological nucleosides; they also provide routes of entry for hydrophilic antiviral (e.g., zidovudine and ribavirin) and anticancer (e.g., Ara-C and gemcitabine) nucleoside analogue drugs. ^[2-7] Nucleoside transporters are therefore important pharmacological targets, including sites in the brain.^[8] While this paper presents new methods and results with Escherichia coli proton-driven CNT NupC (400 residues), [9-13] designs of experiments and discussions of results were also influenced by human CNTs and by other bacterial CNTs with available high-resolution crystal structures. Human CNTs are classified as solute carrier family SLC28; ^[14-18] its three members (hCNT1-3) are sodium-driven and have selective substrate specificity and tissue distribution. ^[19-21] hCNT1 and hCNT2 transport pyrimidine nucleosides (e.g., uridine, thymidine, and cytidine) and purine nucleosides (e.g., adenosine, guanosine, and inosine), respectively, coupling sodium:nucleoside transport with 1:1 stoichiometry. hCNT3 has broad selectivity for purine and pyrimidine nucleosides, coupling sodium:nucleoside transport with 2:1 stoichiometry. hCNT3 can also couple transport of uridine to a proton gradient with 1:1 stoichiometry. ^[22, 23] The first structurally characterized CNT was from the bacterium Vibrio cholerae. ^[24] Sodium-driven vcCNT (418 residues) shares a high degree of sequence homology with hCNTs, the highest being 38.8% identity with hCNT3. The 2.4 Å crystal structure of vcCNT revealed a homotrimeric structural organization with each protomer comprising eight transmembranespanning α -helices (TM1-TM8), three interfacial helices (IH1-IH3) that run parallel to the membrane, and two hairpin reInternational Journal of Advanced Multidisciplinary Research and Studies

www.multiresearchjournal.com

entrant loops (HP1 and HP2) (Fig 1). Similar hairpin reentrant loops were also seen in structures of the Pyrococcus horikoshii sodium-coupled aspartate transporter, GltPh, which also has eight transmembrane helices and exists as a homotrimer. [25-27] The overall shape of the trimer is like an inverted triangular basin with its opening facing the intracellular side and a knoblike structure facing the extracellular side. [24] The N- and C-terminal ends are both periplasmic, and each protomer has a novel overall fold divided into two subdomains based on locations at outer and inner regions relative to its center. The "scaffold domain" (TM1, TM2, IH1, EH, TM3, and TM6), in the outer region, maintains the overall architecture, including trimerization contacts. TM6 is exceptionally long (42 residues, ~60 Å) and tilted almost 60° with respect to the membrane normal. The "transport domain" comprises two structural groups (IH2, HP1, TM4, and TM5 and IH3, HP2, TM7, and TM8) on either side of TM6 that are related by internal two-fold pseudosymmetry. The tips of HP1 and HP2 and unwound regions of TM4 and TM7 lie at the center of the transport

domain, located slightly below the middle of the membrane plane. These regions contain residues involved in direct binding interactions with nucleosides. Based on crystal structures of vcCNT in complex with uridine, [24,28] HP1 residues Gly153, Gln154, Thr155, and Glu156 and TM4 residue Val188 interact with the base moiety, while HP2 residue Glu332 and TM7 residues Phe366, Asn368, and Ser371 interact with ribose. The single sodium-binding site (Asn149, Val152, Ser183, and Ile184) is located between the tip of HP1 and the unwound region of TM4, which is near the nucleoside-binding site. [24] Crystal structures of a sodium-coupled CNT from Neisseria wadsworthii (nwCNT) (425 residues) have been determined in outward-facing, multiple intermediate, and inward-facing conformations.^[29] nwCNT shares 66.5% sequence identity with vcCNT, and its structures have confirmed that CNTs appear to adopt an elevator-type mechanism of alternating access membrane transport, ^[30-36] first hypothesized from crystal structures of GltPh. ^[25, 26, 37]



Fig 1: Structural organization of the V. cholerae concentrative nucleoside transporter vcCNT. Crystal structure of homotrimeric Na⁺-coupled nucleoside transporter vcCNT from V. cholerae in complex with uridine viewed (A) from the periplasm and (B) from the membrane plane. In panel B, the grey horizontal lines indicate the boundaries of the membrane. The structure was drawn with the N-terminus coloured purple and the C-terminus coloured red using Protein Data Bank entry 4PD6 and PDB Protein Workshop 3.9. ^[53] (C) Membrane topology of vcCNT. Schematic illustration of transmembrane-spanning helices (TM1–TM8), interfacial helices (IH1–IH3), extracellular helices (EH), and hairpin re-entrant loops (HP1 and HP2) in vcCNT based on the crystal structure. The "transport domain" is highlighted in pink

NupC transports the pyrimidines uridine, cytidine, and thymidine and also the purine adenosine. [9-13] While NupC does not transport guanosine and inosine, there is evidence to suggest transport of xanthosine with a very low affinity. ^[38] Heterologous expression of NupC in Xenopus laevis oocytes demonstrated a nucleoside selectivity resembling that of human and rat CNT1, including an ability to transport adenosine and the therapeutic drugs 3'-azido-3'deoxythymidine (AZT, zidovudine), 2',3'- dideoxycytidine (ddC, zalcitabine), and 2'-deoxy-2',2'-difluorocytidine (gemcitabine). ^[39] Apparent affinities for transport (K_m^{app} values) were 1.5-6.3 μ M for adenosine, uridine, and gemcitabine and 112 and 130 μ M for AZT and ddC, respectively. Competitive inhibition of [14C]-uridine transport by recombinant NupC in E. coli whole cells with 46 unlabelled physiological nucleosides and nucleobases, nucleoside analogues, and other compounds facilitated rigorous investigation of ligand recognition. [40] This assay identified specific structural motifs necessary for binding of ligands by NupC compared with *E. coli* major facilitator superfamily nucleoside transporter NupG. ^[41–44] The overall pattern of inhibition demonstrated the close functional relationship of NupC with hCNTs and confirmed NupC as a useful experimental model for the human proteins. ^[40] Using nucleosides synthesized with nuclear magnetic resonance (NMR)-active isotope labels (¹³C and ¹⁵N), ^[40, 45] solid-state NMR methods were developed for observing and characterizing direct binding of nucleosides to NupC with amplified expression in native *E. coli* inner membranes. Under these non-energized conditions, the apparent K_d value for binding of [1'-¹³C]uridine to NupC was 2.6 mM. ^[46]

This work explores evolutionary relationships, sequence homologies, and structural and functional characteristics of NupC compared with those of vcCNT, nwCNT, and hCNTs. This includes analyses of NupC residues putatively involved in direct binding interactions with nucleosides and the predicted membrane topology and structural organization of NupC based on the structure of vcCNT. Site-specific mutants of NupC were constructed and characterized, including measurement of transport activities and nucleoside binding affinities using novel solid-state NMR methods. Results are explained with respect to a homology model of NupC and a hypothetical elevator-type mechanism of alternating access concentrative nucleoside transport.

2. Materials and Methods

2.1 Computational methods

Protein sequences were taken from the UniProt KnowledgeBase (http://www.uniprot.org/) and aligned using Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/). ^[47] Phylogenetic analysis was performed by exporting the resultant neighbour-joining phylogenetic tree in Newick format and redrawing using iTOL: Interactive Tree of Life (http://itol.embl.de/). [48] Membrane topology predictions performed using were TOPCONS [49] (http://topcons.cbr.su.se/pred/), TMHMM [50] (http://www.cbs.dtu.dk/services/TMHMM/), and CCTOP (http://cctop.enzim.ttk.mta.hu/). [51] Amino acid compositions were determined using the ExPASy tool ProtParam (http://web.expasy.org/protparam/). [52] Protein structures were drawn using PDB Protein Workshop 3.9 and NGL viewer. [53, 54] The homology model of NupC was ExPASy tool using SWISS-MODEL produced (https://swissmodel.expasy.org/).^[55]

2.2 Site-directed mutagenesis

Site-directed mutagenesis used the QuikChange polymerase chain reaction (PCR) method (Stratagene). "Touchdown" PCR was typically performed over 18-20 cycles using the following conditions: denaturation at 94 °C for 15 s, annealing at 60-55 °C for 30 s, and extension at 72 °C for 30 s per kilobase of template DNA followed by final extension at 72 °C for 10 min. An aliquot (10 µL) of the analyzed by agarose mixture was reaction gel electrophoresis, and a further aliquot (20 μ L) was treated with 10 units of DpnI at 37 °C for 1-2 h to degrade parental methylated (5'-Gm⁶ ATC-3') and hemimethylated DNA. An aliquot (1-2 μ L) of the DpnI-treated reaction mixture was used to transform competent E. coli Top10 cells (50 μ L). These were transferred to LB-agar plates containing carbenicillin (100 μ g/mL) and incubated overnight at 37 °C. Colonies containing putative mutant plasmids were grown overnight in LB medium (5 mL) and DNA minipreps performed. The presence of the required mutation was confirmed by DNA sequencing. Cysteineless NupC (C96A) was constructed first and used as the parental protein for construction of all further mutants.

2.3 Cell growth

E. coli strain BL21(DE3) harboring isopropyl β -D-1thiogalactopyranoside (IPTG)-inducible plasmid pGJL16 for amplifying expression of wild-type (WT) NupC or constructs with the necessary mutations for producing NupC C96A/S142C, C96A/G146C, mutants (C96A, or C96A/E149C) were grown in M9 minimal medium supplemented with glucose (20 mM) in batch culture (500 mL volumes in 2 L baffled flasks) at 37 °C while being shaken at 200 rpm. At the optimal stage of growth established in pilot studies, IPTG (0.5 mM) was added to induce amplified expression. Typically, cells were grown to an A_{600} of ~0.6 over 4 h, followed by a further period of 3 h for induction to a final A_{600} of ~1.8–2.0.

2.4 Transport studies on whole cells

Measurements of the rate of uptake of [¹⁴C]uridine (Amersham Biosciences) into energized *E. coli* whole cells expressing WT NupC or mutant forms were performed using the method described previously.^[40] Data were obtained during initial rates of [¹⁴C]uridine transport after an uptake period of 15 s, and the resultant curves were used to estimate values of K_m and V_{max} by fitting to the Michaelis–Menten equation using GraphPad Prism. For expression tests, aliquots of the same cells used for transport measurements were subjected to mixed membrane preparation by the water lysis method ^[56, 57] and analysis by Western blotting.

2.5 Western blotting

Samples of mixed membranes containing 10 μ g of total protein were resolved by sodium dodecyl sulfatepolyacrylamide gel electrophoresis (SDS-PAGE), and proteins were detected by Western blotting using a primary antibody raised against a 15-residue sequence of NupC (SEENIQMSNLHEGQS). ^[58] Proteins were transferred from the gel to a nitrocellulose membrane by blotting at a constant voltage (100 V) for 1 h using a Bio-Rad Mini Protean II apparatus and a buffer containing glycine (190 mM), Tris (25 mM), and methanol (20%). The blot was washed twice for 10 min in Tris-buffered saline {TBS [20 mM Tris-HCl and 500 mM NaCl (pH 7.5)]} and then twice for 20 min in TTBS (TBS containing 0.1% Tween 20). The blot was blocked by being washed with non-fat dried milk (5%) in TTBS for 1 h and then washed twice for 10 min in TTBS. The blot was incubated overnight with a 1:1000 dilution of the primary antibody in antibody buffer (1% nonfat dried milk in TTBS) and then washed three-times for 15 min in TTBS. The blot was incubated for 1 h with a 1:5000 dilution of the secondary antibody (anti-rabbit IgG horseradish peroxidise-conjugated antibody, Sigma) in antibody buffer and then washed three-times for 15 min in TTBS. The blot was removed from the wash, drained, and then transferred to a sheet of Saran wrap on a flat surface. A solution of chemiluminescent substrate (Supersignal West Pico chemiluminescent substrate, Pierce) (~5 mL) was transferred to the blot to cover it evenly and then left at room temperature for 5 min. The blot was drained of solution, enveloped in Saran wrap, and exposed to X-ray film for 1 min.

Removal of any contaminating antibodies from the primary antibody preparation was achieved by treatment with an acetone powder of *E. coli* total protein before use. An aliquot (5 μ L) of a primary antibody solution was diluted to 1 mL with TBS, and 10 mg [1% (w/v)] of a total protein acetone powder from *E. coli* BL21(DE3) cells was added. The mixture was incubated at 4 °C on a blood-wheel for 30 min and then spun in a microcentrifuge (10000 rpm, 10 min) to remove the protein. The supernatant containing the primary antibody was diluted to the required titer and then used to probe the Western blot as described above.

2.6 Membrane preparation

Inside-out vesicles were prepared from harvested cells by explosive decompression in a French press, and the inner membrane fraction was isolated by sucrose density gradient ultracentrifugation. ^[57] Membrane vesicles were washed three-times by suspension in 20 mM Tris-HCl buffer (pH 7.5) followed by ultracentrifugation before final suspension International Journal of Advanced Multidisciplinary Research and Studies

in the same buffer for storage at -80 °C after rapid freezing. The total protein concentration in the vesicle suspension was determined by the method of Schaffner and Weissmann, and the percent NupC content was estimated by densitometry measurements on the proteins resolved by SDS–PAGE and stained with Coomassie Brilliant Blue. Control membranes without amplification of NupC expression were prepared in the same way from cells containing plasmid pTTQ18 without the transporter gene insert.

2.7 NMR conditions

¹³C-observed cross-polarization magic-angle spinning (CP-MAS) NMR measurements were preliminarily performed at the University of Leeds using a Varian InfinityPlus 300 instrument operating at a magnetic field of 7 T with a Varian double-bearing MAS probe tuned to 75.32 MHz for ¹³C and 299.5 MHz for ¹H. Measurements were repeated and continued at the University of Manchester Institute of Science and Technology (UMIST) (which was subsumed by the University of Manchester) and at the University of Liverpool using a Bruker Avance 400 spectrometer operating at a magnetic field of 9.3 T, corresponding to Larmor frequencies of 400.13 MHz for ¹H and 100.13 MHz for ¹³C. All experiments used a sample temperature of 4 °C. At 400 MHz, standard ¹H-¹³C CP-MAS experiments were performed using a double-tuned probe head at a sample spinning frequency of 4 kHz. CP was initiated by a ¹H 90° pulse length of 3.5-4.0 µs, followed by Hartmann-Hahn cross-polarization from ¹H to ¹³C at a ¹H field of 65 kHz. Continuous wave proton decoupling at a field of 85 kHz was applied during signal acquisition. Variable contact time experiments used contact times of 0.1, 0.5, 1, 2, 3, 5, 8, and 10 ms. Spectra were acquired over the given number of scans; chemical shifts were referenced to adamantane at 37.8 ppm, and spectra were processed with 40 Hz exponential line broadening.

Spectra representing [1'-¹³C]uridine interacting specifically with NupC were obtained by difference after displacement of the bound substrate with either 40- or 80-mM unlabelled uridine. Spectra were first obtained at the different contact times from membranes containing 5 mM [1'-13C]uridine. Unlabelled uridine was then added to the membranes, and another series of spectra was obtained under identical NMR conditions. In the second series of spectra, the [1'-¹³C]uridine was fully displaced from NupC but a residual peak from [1'-13C]uridine was present, which was assigned as labelled substrate interacting non-specifically with the membranes. Each of the second series of spectra was subtracted from the corresponding spectrum obtained at the same contact time before unlabelled uridine was added. This produced a series of spectra exhibiting a peak from [1'-¹³C]uridine representing only the fraction of the labelled substrate interacting with NupC. These difference spectra are termed "displacement spectra".

2.8 Analysis of NMR data

Values of the dissociation constant (K_d) representing the affinity of $[1'-{}^{13}C]$ uridine for NupC were estimated from two independent NMR experiments. In one experiment, peak intensities for $[1'-{}^{13}C]$ uridine were measured from ${}^{13}C$ CP-MAS NMR spectra of membranes containing 0–80 mM unlabelled uridine at a fixed contact time. Intensities were plotted as a function of unlabelled uridine concentration and apparent K_d values calculated using a standard

Marquardt-Levenberg curve fitting procedure. The second experiment used variable contact time CP-MAS measurements, whereby peak intensities for specifically bound [1'-13C]uridine were measured from displacement spectra obtained with a range of contact times. Values of the apparent K_d and the rate constant for substrate dissociation (k_{off}) were estimated by comparing experimental intensities with simulations obtained using the Monte Carlo method.^[46] Simulated curves were a function of k_{off} and K_d , contact time (t_c), substrate and protein concentrations, ¹H T_1 relaxation times in the rotating frame $(T_{1\rho})$ for the free and bound substrate, and the CP rate constant for the bound substrate $(R_{\rm HC})$. A series of curves was calculated for contact times of 0-10 ms at different combinations of k_{off} and K_d values, keeping all the other input parameters constant. Values of $T_{1\rho}$ and $R_{\rm HC}$ were measured experimentally or estimated from natural abundance signals of other membrane components as described previously. [46] Curves in closest agreement with experimental data were determined by minimizing the χ^2 function.

3. Results and Discussion

3.1 Evolutionary relationships and conservation of residues in NupC and CNT family proteins

The close evolutionary relationship between NupC and other characterized CNT family proteins (vcCNT, nwCNT, and hCNTs) is demonstrated by a nearest-neighbor phylogenetic tree annotated with levels of sequence homology (Fig S1). NupC is most closely related to vcCNT and nwCNT (30.3 and 32.3% identical and 28.8 and 27.3% highly similar, respectively; combined 59.1 and 59.6%, respectively) and then with the human proteins in the following order: hCNT3 > hCNT1 > hCNT2. Based on separate sequence alignments, human CNTs share 35.9% identical residues and 25.0% highly similar residues (combined 60.9%) (Fig S2), with the closest relationship between hCNT1 and hCNT2 (63.3% identical, 17.7% highly similar, combined 81.0%) (Table S1). Like NupC, vcCNT is most closely related to hCNT3 (38.8% identical, 25.4% highly similar, combined 64.2%) (Table S1). A BLAST search for NupC against all Eukaryota in the UniProt KnowledgeBase identified two very close homologues from the soil fungus Beauveria bassiana (Fig S3). Proteins A0A0A2W106 (395 residues) and A0A0A2V7L6 (380 residues) have sequences that are 92.7 and 73.4% identical to that of NupC, respectively. No other reliable close homologues of NupC were identified, with all other eukaryotic proteins having sequences less than 30% identical with that of NupC.

A sequence alignment between NupC and vcCNT highlights conserved residues, locations of vcCNT structural domains (TM1-TM8, IH1-IH3, HP1, and HP2) and vcCNT residues involved in direct binding interactions with nucleosides (Fig S4). The region of vcCNT structural domains involved in nucleoside binding (HP1, TM4, HP2, and TM7) is highly conserved with NupC and hCNTs (Fig S5). Of nine vcCNT residues involved in direct binding interactions with nucleosides, seven are identically conserved in NupC and two are highly similar (Table 1). Also note that all nine vcCNT residues are identically conserved in nwCNT. The corresponding residues in NupC are Gly146, Gln147, Ser148, Glu149, Ile181, Glu321, Phe350, Asn352, and Ser355. Except for hCNT1 residues corresponding to Gly153 and Val188 and the hCNT2 residue corresponding to Gln154, all vcCNT residues involved in nucleoside

binding are identically conserved in hCNTs (Table 1). Of four vcCNT sodium-binding residues, three (Asn149, Val152, and Ile184) are identically conserved in hCNTs but not in NupC. The position corresponding to vcCNT sodiumbinding Ser183 is occupied by threonine in NupC and hCNTs. These differences are consistent with NupC being proton-driven and vcCNT, nwCNT, and hCNTs being sodium-driven.

Table 1: Conservation of Nucleoside-Binding Residues in Concentrative Nucleoside Transporters. Residues in *V. cholerae* vcCNT structural domains (HP1, TM4, HP2, and TM7) that interact directly with nucleosides based on crystal structures of vcCNT^{24,28} are compared with residues at the corresponding positions in NupC from *E. coli* (POAFF2) and human concentrative nucleoside transporters hCNT1 (O00337), hCNT2 (O43868), and hCNT3 (Q9HAS3) based on a sequence alignment (Fig S3). Residues that are identically conserved are shown in bold and those that are highly similar in italic. Binding site residues in vcCNT are all identical with those in nwCNT

vcCNT domain		NupC	hCNT1	hCNT2	hCNT3
	Gly153	Gly146	Ser319	Gly313	Gly340
UD1	Gln154	Gln147	Gln320	Met314	Gln341
прі	Thr155	Ser148	Thr321	Thr315	Thr342
	Glu156	Glu149	Glu322	Glu316	Glu343
TM4	Val188	Ile181	Leu354	Val348	Val375
HP2	Glu332	Glu321	Glu498	Glu492	Glu519
	Phe366	Phe350	Phe542	Phe536	Phe563
TM7	Asn368	Asn352	Asn544	Asn538	Asn565
	Ser371	Ser355	Ser547	Ser541	Ser568

Residues responsible for nucleoside specificities in hCNT1 and hCNT2 are those corresponding to Ser319/Gln320 and Ser353/Leu354 in hCNT1. Conversion to the corresponding residues in hCNT2 (Gly313/Met314 and Thr347/Val348) changes hCNT1 from a pyrimidine-specific to a purine-specific transporter. ^[19] Furthermore, hCNT1 mutation S353T produces a profound decrease in cytidine transport efficiency and, in combination with L354V (S353T/L354V), produces a novel uridine-preferring transport phenotype. ^[59] On this basis, similar mutations at corresponding positions in NupC are predicted to change NupC from pyrimidinespecific to purine-specific. The position in NupC corresponding to hCNT1 residue Ser319 is already glycine (Gly146), and also glycine in hCNT3 (Gly340), which may explain the ability of NupC to transport adenosine. Other NupC mutations predicted to alter nucleoside specificity would be Q147M, S180T, and I181V. Roles for hCNT1 residues Glu308, Glu322, and Glu498 in sodium-coupled nucleoside transport have also been identified, with Glu322 having the strongest influence on hCNT1 transport function. ^[60] Corresponding residues at all three of these positions are identically conserved as glutamate in NupC, vcCNT, hCNT2, and hCNT3 (Fig S6).

3.2 Predictions of transmembrane helices and distributions of amino acids in NupC and vcCNT

To demonstrate challenges associated with defining the structural organization of NupC, its sequence was subjected to analysis by three of the most popular and/or best performing tools that predict transmembrane helices: TOPCONS, ^[49] TMHMM, ^[50] and CCTOP ^[51] (Fig S7). TOPCONS has been updated to efficiently separate any N-terminal signal peptides from transmembrane regions and can predict re-entrant loops. ^[61] Also, the ability of

TOPCONS to predict 1551 transmembrane helices in 235 membrane proteins with high-resolution structures has recently been assessed, giving a 94.8% success rate. [62] For NupC, the TOPCONS consensus result predicted nine transmembrane helices with a periplasmic N-terminus and a cytoplasmic C-terminus (Fig S7A). TMHMM predicted eight full transmembrane helices with periplasmic N- and Ctermini and also a ninth partial helix (Fig S7B). The CCTOP consensus result predicted nine transmembrane helices with a cytoplasmic N-terminus and a periplasmic C-terminus (Fig S7C). The positions of the nine transmembrane helices predicted by these three methods are roughly in agreement with each other or at least show significant overlap (Fig S8). Arrangements predicted by TOPCONS and TMHMM are most plausible based on the "positive-inside rule". [63] Fig S8 also shows putative locations of transmembrane helices in NupC based on sequence homology and the structure of vcCNT. While transmembrane helices based on vcCNT all show some overlap with those coming from the prediction tools, there are significant differences in their lengths and in their start and end positions. This is partly because the prediction tools generally give the same number of residues in all transmembrane helices, for example, 22 and 21 residues from TOPCONS and TMHMM, respectively. [62] The prediction tools were therefore never going to identify the full length of a 42-residue helix for TM6 as in vcCNT. A further limitation of the prediction tools is that they tend to miss transmembrane helices containing internal positively charged residues. [64] The presence of positively charged residues may therefore have affected their ability to recognize locations for many of the putative transmembrane helices in NupC based on those provided by the structure of vcCNT (Fig S8). The extra "ninth" helix in NupC given by the prediction tools appears in the region of residues 278-307 (Fig S8), corresponding to IH3 in vcCNT (Fig S5). All three prediction tools gave a full transmembrane helix for NupC at this location. A high-resolution crystal structure of NupC is clearly desirable to confirm whether it shares the same structural organization as vcCNT.

For vcCNT, both TOPCONS and TMHMM predicted ten transmembrane helices with periplasmic N- and C-termini (Fig S9A,B). CCTOP predicted eight transmembrane helices with periplasmic N- and C-termini (Fig S9C). Of ten transmembrane helices predicted by TOPCONS and TMHMM, eight show good agreement or significant overlap, while the others show significant discrepancies. These eight helices in the TOPCONS and TMHMM results all show significant overlap with the eight transmembrane helices predicted by CCTOP (Fig S8). While the eight common helices coming from the prediction tools all show some overlap with transmembrane helices in the structure of vcCNT, they all have different lengths and different start and end positions. The differences may partly be due to the presence of internal positively charged residues in some of the transmembrane helices in vcCNT. The structure of vcCNT does conform to the "positive-inside rule", [63] with the majority of positively charged residues confined to the cytoplasmic side (Fig 2). Negatively charged residues have an approximately equal distribution at the periplasmic and cytoplasmic sides of vcCNT (Fig 2). In terms of other types of amino acids in vcCNT, hydrophobic residues predominate in transmembrane regions, as expected. Hydroxyl-containing residues and amido-containing

residues are scattered throughout the transmembrane helices and the periplasmic and cytoplasmic regions (Fig 2).



Fig 2: Distribution of amino acids in the structure of vcCNT from *V. cholerae*. The crystal structure of vcCNT from *V. cholerae* in complex with uridine was drawn using Protein Data Bank entry 4PD6 and PDB Protein Workshop 3.9. ^[53] Residues are coloured to show those that are positively charged at physiological pH (arginine and lysine) (*red*), negatively charged at physiological pH (aspartic acid and glutamic acid) (*blue*), hydrophobic (alanine, isoleucine, leucine, phenylalanine, valine, and glycine) (*green*), hydroxyl-containing (serine, threonine, and tyrosine) (*pink*), and amido-containing (*orange*)

Amino acid compositions in NupC, vcCNT, nwCNT, and hCNTs were also calculated (Table S2). Percentage contents for individual amino acids and for groupings of amino acids with similar physicochemical properties (hydrophobic, positively charged, negatively charged, hydroxyl-containing, and amido-containing) in NupC, vcCNT, and nwCNT were similar to overall average values in 235 secondary transport proteins from *E. coli*. ^[65] Percentage contents in hCNTs were similar to overall average values in 336 human secondary transport proteins. ^[64] Hence, none of the CNTs contain significantly high contents of any individual type of amino acid or group of amino acids with similar physicochemical properties.

3.3 Characterization of NupC cysteine mutants at positions involved in nucleoside binding and transport activity

3.3.1 Details and construction of NupC cysteine mutants

Cysteine scanning mutagenesis investigated roles of specific NupC residues in nucleoside binding and transport activity. ^[66] This work deals with three residues (Ser142, Gly146, and Glu149) in a region of NupC corresponding to HP1 in the structure of vcCNT (Fig 3). Ser142 corresponds to Asn149 in vcCNT (and nwCNT), which is involved in coordinating the sodium ion (Fig 3), and this position is conserved as asparagine in hCNTs (Fig S5). Gly146 is conserved in vcCNT, nwCNT, hCNT2, and hCNT3 and corresponds to Ser319 in hCNT1 (Fig S5) that when

mutated to glycine allows purine nucleoside transport. ^[19] The corresponding vcCNT residue (Gly153) is close to both the base moiety of uridine ^[28] and the sodium-binding site (Fig 3). Glu149 corresponds to Glu156 in vcCNT, which coordinates with the base moiety of uridine via water molecules, ^[28] and this position is conserved as glutamate in hCNTs (Fig S5). Interestingly, mutant E149D retained partial transport activity, suggesting an important role for a carboxyl group at this position. ^[66] NupC contains one native and non-essential cysteine residue (Cys96) for which the corresponding position in vcCNT, nwCNT, and hCNTs is occupied by proline (Fig S6). This residue was mutated to alanine to create cysteineless NupC (C96A), which retained the same transport activity as WT, [66] and this was used as the parental protein for construction of the cysteine mutants. Three single-cysteine-containing mutants (S142C, G146C, and E149C) were constructed, expressed in E. coli (Fig S10), and characterized for nucleoside transport activity, nucleoside binding, and the effect of thiol-specific reagents.



Fig 3: Locations of vcCNT and nwCNT residues corresponding to NupC mutated residues Ser142, Gly146, and Glu149. (A) Model of a vcCNT protomer with bound uridine indicating the positions of Asn149, Gly153, and Glu156 in HP1 that correspond to mutated NupC residues. The position of crucial TM7b residue Phe366 is also indicated. (B) Model of the vcCNT uridine-binding site showing positions of residues Asn149, Gly153, and Glu156 in relation to the base moiety of uridine. The position of important Phe366 is also indicated. (C) Model of the vcCNT sodium-binding site showing the positions of residues Asn149, Gly153, and Glu156.

(D) Model of the nwCNT sodium-binding site showing the positions of residues Asn149 and Gly153. Models were produced using NGL viewer ^[54] and the relevant Protein Data Bank entries

3.3.2 Transport activities of NupC cysteine mutants



Fig 4: Transport activities of NupC cysteine mutants. Concentration dependence of uptake of [¹⁴C]uridine after 15 s by transport energized *E. coli* BL21(DE3) cells harboring plasmid pTTQ18 with no transporter gene insert (*blue*) or plasmid pGJL16 for expressing WT NupC (*black*) or constructs for expressing mutants S142C (*red*), G146C (*green*), and E149C (*pink*) following growth in minimal medium and induction with IPTG. For WT and S142C, data points represent the average of duplicate measurements. Kinetic parameters for WT and S142C were obtained by fitting data to the Michaelis– Menten equation using GraphPad Prism. The inset is a Western blot analysis of membrane preparations from the same cells used in the transport assay and also membranes from cells expressing cysteineless NupC (C96A)

Using an assay for NupC-mediated uptake of [¹⁴C]uridine into energized E. coli whole cells, [40] transport activities of cysteine mutants were compared to that for WT (Fig 4). Transport by WT was concentration-dependent and saturable, and the resultant curve was fitted to obtain values for the affinity of initial rate transport (K_m^{app}) and maximal velocity (V_{max}) of 22.2 ± 3.7 μ M and 13.5 ± 0.8 nmol (mg of cells)⁻¹ (15 s)⁻¹, respectively. According to data obtained over a range of time points with the lowest (5 μ M) and highest (100 μ M) concentrations of [¹⁴C]uridine used with WT NupC and mutant S142C, 15 s is in the linear range and not approaching equilibrium (Fig S11). A time point of 15 s is one that has routinely been used for transport assays in energized whole cells with a range of similar secondary active bacterial transporters, and this has proven to be in the initial linear range. Also, $K_{\rm m}^{\rm app}$ values in the low micromolar range are expected in energized whole cells, which is consistent with values obtained for other bacterial secondary active transporter substrates, for example, [³H]galactose with GalP (42–59 μ M), ^[67] [¹⁴C]uridine and [¹⁴C]adenosine with NupG (23.6 \pm 3.1 and 20.6 \pm 3.4 μ M, respectively), ^[42] $[^{14}C]$ phenyl-thio- β -D-glucuronide and $[^{14}C]-p$ acetamidophenyl- β -D-glucuronide with GusB (198 ± 69 and 36 \pm 4 μ M, respectively), ^[68] [³H]myo-inositol with IolT $(111 \pm 22 \ \mu M)$, ^[69] [¹⁴C]allantoin with PucI (24 ± 3 \ \mu M), ^[70] and [³H]-L-glutamate with GltP (22.6 \pm 5.5 μ M).^[71] Mutant S142C had a $K_{\rm m}^{\rm app}$ value (18.6 \pm 2.5 μ M) that was not significantly different from that of WT, while its V_{max} [10.7 ± 0.4 nmol (mg of cells)⁻¹ (15 s)⁻¹] was slightly lower (Fig 4). Treatment with the membrane-permeable thiol-specific reagent MTSEA reduced the rate of S142C-mediated uptake of [14C]uridine by around 50%. [66] Mutants G146C and E149C both had uptake levels similar to those obtained for cells with no amplified expression of any transport protein,

as demonstrated using cells transformed with plasmid pTTQ18 containing no transporter gene insert (Fig 4). This background uptake into negative control cells, which does fit to the Michaelis-Menten curve, is presumed to originate from naturally occurring levels of potentially multiple endogenous proteins capable of transporting or binding uridine. Measurements of uptake of [14C]adenosine and ¹⁴C]uridine after 15 s and 2 min by cells transformed with empty pTTO18 compared with cells transformed with NupC-expressing plasmid pGJL16 (Fig S12) confirm the background levels of uptake seen in Fig 4. Hence, mutants G146C and E149C had both lost the ability to transport $[^{14}C]$ uridine, even up to a concentration of 400 μ M. Loss of transport activity was not due to a decrease in the level of protein expression, because Western blot analysis of membrane preparations from the same cells used in the transport assay demonstrated comparable levels of expression for WT NupC and the three mutants (Fig 4). Analysis of transport activity by the same NupC mutants expressed in X. laevis oocytes produced results broadly consistent with observations made in E. coli whole cells. S142C was functional and inhibited by PCMBS; G146C was non-functional, and E149C was borderline functional with a flux too low to characterize. [72] The next consideration was whether mutations G146C and E149C had affected just transport activity or the ability of NupC to recognize and bind a substrate. The transport defective mutants may have lost the ability to undergo a conformational change required to complete the transport cycle but retained nucleoside binding activity.

3.3.3 Characterization of nucleoside binding by NupC cysteine mutants using CP-MAS solid-state NMR

We next investigated whether the different transport activities of NupC mutants were a consequence of impaired transport function, a decrease in the affinity for the nucleoside-binding site, or both. Purified NupC was not available, and nucleosides absorb wavelengths of radiation in regions that can hamper binding measurements using techniques such as fluorescence spectroscopy and nearultraviolet circular dichroism spectroscopy. ^[73, 74] Analysis of nucleoside binding was therefore conducted using ¹³C CP-MAS NMR spectroscopy measurements on the proteins with amplified expression in un-energized E. coli inner membranes (Fig S10). These experiments used and expanded on methods already developed to observe and quantify binding of [1'-13C]uridine to WT NupC. [46] 13C CP-MAS NMR spectra of NupC-enriched membranes containing [1'-13C]uridine exhibit a peak from the labelled substrate at a chemical shift position of ~91 ppm. At high substrate concentrations (≥ 5 mM), the peak intensity at a defined Hartmann-Hahn contact time reflects both specific interactions of uridine with NupC and interactions at other, undefined sites within the membranes. Fractions of total peak intensity that reflect specific and undefined interactions can be determined by displacing all of the specifically bound [1'-13C]uridine with a non-labelled competitor. ^[46] In the work presented here, displacement NMR spectra representing only [1'-¹³C]uridine interacting specifically with NupC were obtained by the subtraction method described in Materials and Methods. Transport proteins can display relatively weak binding for their substrates to accommodate transient interactions that take place during compound recognition and the transport cycle. Apparent dissociation constants (K_d^{app} values) in the low millimolar range are therefore expected for substrate binding in nonenergized native membranes. This is consistent with similar solid-state NMR measurements on other bacterial secondary active transporters, including GalP, GusB, NupG, ZitB, and GltP. ^[42, 46, 71, 75–79]



Fig 5: Relative affinities of uridine binding to wild-type NupC and mutants in native inner membranes. (A) Displacement ¹³C CP-MAS NMR spectra obtained from un-energized *E. coli* inner membranes containing 5 mM [1'-¹³C]uridine and 80 nmol of WT NupC or mutants S142C, G146C, and E149C. Displacement spectra were obtained by the procedure described in Materials and Methods, in which spectra of each membrane sample were obtained before and after the addition of 40- or 80-mM unlabelled uridine to displace the labelled substrate. Each spectrum was obtained using a Hartmann–Hahn contact time of 10 ms and was the result of the accumulation of 6144 scans. Spectra from membranes with amplified expression of WT, S142C, G146C, or E149C all exhibit a peak at 91 ppm from the substrate interacting specifically with NupC. Also shown is a control displacement spectrum from membranes containing natural abundance levels of NupC, which indicates that the peak intensity was not displaced. Difference spectra were normalized to the intensity of the lipid peak at 32 ppm in the original spectra. (B) Curves of ¹³C CP-MAS peak intensities for [1'-¹³C]uridine in membranes containing WT NupC or mutants S142C, G146C, and E149C at increasing concentrations of unlabelled *D*-uridine. Peak intensities (•) were obtained from spectra recorded using a Hartmann–Hahn contact time ($_{\circ}$). Solid lines are curves of best fit calculated according to Equation 2 corresponding to the values of the apparent *K*_d shown in each plot

Normalized ¹³C CP-MAS displacement spectra were recorded for membrane preparations containing 5 mM [1'-¹³C]uridine and 80 nmol of WT NupC or mutants S142C, G146C, and E149C (Fig 5A). The background signal from natural abundance ¹³C in the membranes was removed by subtraction, and each spectrum shows only a single peak at 91 ppm representing the labelled substrate displaced from the membranes by 40 mM unlabelled uridine. Peaks representing interactions of [1'-13C]uridine with WT and S142C are considerably more intense than those for interactions with G146C and E149C. Peak intensities in displacement spectra are related to the specific association of [1'-13C]uridine and NupC only if the labelled substrate is not displaced from other, undefined sites by unlabelled uridine. When control membranes that contain NupC at the natural abundance level of <1% of total membrane protein were used, specific binding of [1'-¹³C]uridine to NupC was negligible compared with the undefined interactions. No peaks were observed in the displacement spectrum from control membranes, confirming that $[1'-{}^{13}C]$ uridine was not displaced from sites other than with NupC. Initial visual inspection of peak intensities in displacement spectra suggested that the affinity of uridine for G146C and E149C was somewhat different from that for WT. It is not possible to obtain quantitative information from single spectra, so further experiments were necessary to establish how the mutations had affected binding affinity.

Dissociation constants for substrates of bacterial secondary transporters have been estimated previously using ¹³C CP-

MAS NMR measurements of displacement by a competitive substrate. ^[42, 46, 71, 76, 78] When the rate of dissociation of the substrate is slow, the NMR peak intensity for a labelled substrate is related to the concentration of the competitive substrate by an inverse rectangular hyperbolic function. The total peak intensity (I_T) from a labelled ligand (L*) is a function of contact time (t_c) and the concentration of unlabelled competitor ligand (L). In the case of a labelled ligand such as $[1'_{-13}C]$ uridine, I_T has components I_S from the specifically bound substrate and I_{NS} from the remaining nonspecific or undefined interactions of the substrate. For example.

$$I_{\rm T}(t_c, [{\rm L}]) = I_{\rm S}(t_c, [{\rm L}]) + I_{\rm NS}(t_c)$$
(1)

If total peak intensity is normalized to unity when the concentration of unlabelled ligand [L] is zero, i.e., $I_T(t_c,[L]_0) = 1$, it follows that

$$I_{\rm T}(t_{\rm c}, [{\rm L}]) = F \frac{[{\rm L}^*]_{\rm S}/K_{\rm d}}{1 + [{\rm L}^*]_{\rm S}/K_{\rm d} + [{\rm L}]/K_{\rm d}} + I_{\rm NS}(t_{\rm c})$$
(2)

Where

$$F = 1 + \frac{K_d [1 - I_{NS}(t_c)]}{[L^*]_S} - I_{NS}(t_c)$$
(3)

 $I_{\rm NS}(t_{\rm c})$ is measured as the peak intensity remaining after the specifically bound ligand has been fully displaced by the unlabelled ligand. Values of K_d and $[L^*]_S$ can be estimated by nonlinear least-squares fitting of Equation 2 to NMR peak intensity curves measured in a titration experiment. Equations 1-3 assume that dilution effects because of titrating unlabelled substrate solution into the membranes are negligible. Equations 1-3 hold only when a linear relationship exists between $I_{\rm S}$ and the concentration of the specifically bound ligand, which occurs only when the rate of ligand dissociation is slow. [46] When the off rate of the ligand is not known, the measured K_d represents a lower limit value and further experiments are necessary. It is also assumed that component $I_{NS}(t_c)$ from the non-specifically bound ligand remains constant and is not diminished by addition of the unlabelled substrate. The validity of this assumption was supported by the lack of signal observed in the displacement spectrum from the control membranes (Fig 5A). Dissociation constants for interaction between [1'-¹³C]uridine and NupC were estimated from curves of ¹³C NMR peak intensities at 91 ppm measured after titrating unlabelled uridine into the membrane sample to a final concentration of 40 mM (Fig 5B). A decrease in peak intensity with increasing concentrations of unlabelled substrate was observed for WT NupC, for which curve fitting (Fig 5B) yielded an apparent K_d of 1.8 mM and a concentration for specifically bound [1'-13C]uridine of 3.5 mM (i.e., 70% of the total substrate concentration). The K_d value is slightly lower than the value of 2.6 mM measured previously.^[46] Similar displacement curves for the mutants produced apparent K_d values of 2.8, 6.5, and 15.3 mM for S142C, G146C, and E149C, respectively (Fig 5B). In all three cases, best fitting curves corresponded to a concentration of specifically bound uridine of 3.5 mM. In a control experiment with WT NupC containing [1'-¹³C]uridine and the unlabelled non-substrate L-uridine (Fig 5B), the peak intensity did not decrease upon titration with the stereoisomer. This experiment confirmed that the decrease in peak intensity produced by addition of the Duridine substrate was a direct result of displacing [1'-¹³C]uridine from NupC.

 $K_{\rm d}$ values obtained in the titration experiments serve only as estimates because it is not known whether the dissociation rates of uridine are slow enough for Equation 2 to hold. Binding affinities and dissociation rate constants (k_{off}) for [1'-13C]uridine in NupC membranes were therefore estimated using variable contact time ¹³C CP-MAS NMR as an alternative method. Peak intensity curves were measured for the labelled substrate in the membranes over a range of contact times of ≤ 10 ms. These were compared with curves calculated for different K_d and k_{off} values to determine the combination of values that minimized the χ^2 function. Experimental peak intensity curves for membranes containing 5 mM [1'-13C]uridine and 80 nmol of WT NupC are shown in Fig 6. The closest agreement between experimental and simulated curves was for a K_d of 2.6 mM and a k_{off} of 333 s⁻¹. It should be noted, however, that curves for K_d values between 0.5 and 5.5 mM fell within the experimental error. Similar experiments with the mutants yielded χ^2 minima for K_d values of 3.1, 11.9, and 13.4 mM for S142C, G146C, and E149C, respectively (Fig 6). In each case, k_{off} decreased at values of 300-400 s⁻¹. There is reasonable agreement between the K_d values obtained by the variable contact time and displacement methods, although

values for G146C (6.5 and 11.9 mM) are somewhat different. The quality of NMR data that can be obtained in a practicable experimental time (12 h) undoubtedly limits the precision and accuracy of measurements, so caution must be exercised in drawing conclusions from the absolute values. Both NMR methods of analysis identify the same trend in binding affinity of [1'-13C]uridine for WT NupC and the mutants (i.e., WT > S142C > G146C > E149C), however, and conclusions about relative substrate affinities can be made with greater confidence. In summary, mutation S142C appears to have only a weak effect on the affinity for uridine, while G146C and E149C decrease the affinity for uridine. The latter two mutants are still able to bind uridine, however, while the transport of [14C]uridine in whole *E. coli* cells was abolished (Fig 4). Gly146 and Glu149 therefore appear to be important for uridine binding but are essential for uridine transport.



Fig 6: Variable contact time analysis of uridine binding to wildtype NupC and mutants in native inner membranes. Experimental curves and simulations of ¹³C peak intensities at 91 ppm from variable contact time displacement spectra of 5 mM [1'-¹³C]uridine in *E. coli* inner membranes containing 80 nmol of WT NupC or mutants S142C, G146C, and E149C. Displacement spectra were obtained as described in Materials and Methods using 40 mM unlabelled uridine as the displacing substrate. The points of the curve (left, circles) represent specific binding and were measured as the decrease in peak intensity at each contact time after the labelled substrate was displaced by unlabelled uridine. Error bars represent the level of noise in the spectra. Simulated curves were computed for different combinations of K_d and dissociation rate

constant k_{off} using the Monte Carlo method described previously. ^[46] The goodness of fit of each curve to the experimental data was assessed from the function $\chi^2 = \sum_i [(S_i - E)/\sigma_{ii}]^2$. The contour plot

(right) is shaded according to the χ^2 values (shown in the key) obtained for different combinations of K_d and k_{off} . For example, in the case of WT, the χ^2 minimum (denoted by the darkest region on the contour plot) was obtained for a K_d of 2.6 mM and a k_{off} of 3.1 ms. The curve corresponding to the χ^2 minimum is shown on the intensity plot (right) as a solid line, alongside curves for K_d values of 0.5 mM (dotted line) and 5.5 mM (dashed line) for comparison. Fixed simulation parameters were as follows: ¹H $T_{1\rho} = 100$ ms for free substrate, ¹H $T_{1\rho} = 8$ ms for bound substrate, $R_{HC} = 5 \times 10^{-3} \text{ s}^{-1}$, protein content = 80 nmol, and substrate concentration = 5 mM

Experiments were also conducted to compare the ability of WT NupC and the mutants to bind the uridine-competitive substrate thymidine. Interaction of thymidine with NupC was observed indirectly through displacement of 5 mM [1'-¹³C]- uridine with 50 mM unlabelled thymidine (Fig 7). The peak intensity for WT was reduced by \sim 50%, similar to that observed when specifically bound [1'-¹³C]uridine was fully displaced by 40 mM unlabelled uridine (Fig 5B). From these spectra, it is reasonable to conclude that all, or most, of the labelled uridine is displaced from WT NupC by 50 mM thymidine, as would be expected for competitive substrates having a similar affinity for NupC. Spectra for S142C and G146C also show an ~50% drop in peak intensity for [1'-¹³C]uridine upon addition of thymidine (Fig 7), again suggesting that uridine and thymidine have similar affinities for these mutants. By contrast, the spectrum for E149C shows little or no loss of peak intensity following addition of thymidine (Fig 7). Of the three mutants, E149C has the lowest affinity for uridine ($K_d > 15$ mM) and an inability of 50 mM thymidine to displace specifically bound uridine, indicating that this mutant has an even lower affinity for thymidine.



Fig 7: Competitive binding of thymidine to wild-type NupC and mutants in native inner membranes. Regions of 13 C CP-MAS NMR spectra of $[1'-^{13}$ C]uridine in *E. coli* inner membranes containing WT NupC or mutants S142C, G146C, and E149C. Spectra show the peak at 91 ppm from the labelled substrate in the absence (*left*) or presence (*right*) of 50 mM unlabeled thymidine. Spectra were recorded using a Hartmann–Hahn contact time (*t*_c) of 2 ms after 6144 scans

A further experiment investigated the effect of cysteine reactive iodoacetate on [1'-¹³C]uridine binding by mutant E149C and thus the proximity of Glu149 to the uridinebinding site (Fig 8). Control ¹³C CP-MAS measurements were first performed on membranes with amplified expression of cysteineless NupC (C96A), whereby [1'-¹³C]uridine binding was displaced by 50 mM unlabelled uridine but was not affected by pre-treatment with 10 mM iodoacetate. In the case of membranes with amplified expression of E149C, [1'-¹³C]- uridine binding was displaced by 50 mM unlabelled uridine and also blocked by pre-treatment with 10 mM iodoacetate. This result confirms the direct involvement of Glu149 in uridine binding by NupC.



Fig 8: Elimination of binding of uridine to NupC mutant E149C in native inner membranes by treatment with iodoacetate. Expansions of the signals from ¹³C CP-MAS NMR spectra are shown for binding of 5 mM [1'-¹³C]uridine in inner membranes with amplified expression of NupC mutants C96A (cysteineless NupC) and E149C for samples in the absence of any other additions and after pre-treatment with 10 mM iodoacetate (IA) and/or competition with 50 mM unlabelled uridine. For pre-treatment with iodoacetate, membranes containing NupC protein at a concentration of ~0.1 mM were incubated with the following successive additions at 4 °C: 2 mM DTT for 30 min, 10 mM iodoacetate for 1 h in the dark, 20 mM DTT for 15 min (to mop up excess iodoacetate), and then 5 mM [1'-¹³C]uridine for 30 min. Spectra were recorded using a Hartmann–Hahn contact time (*t*_c) of 2 ms after 7200 and 6000 scans for C96A and E149C, respectively

3.4 Explanation of results with respect to an elevatortype mechanism of alternating access membrane transport in NupC

A homology model of a NupC protomer based on the crystal structure of vcCNT in complex with uridine predicts that mutated residues Ser142, Gly146, and Glu149 are located at the tip of HP1 and close to the putative nucleoside-binding site (Fig 9A). While the position corresponding to Ser142 is conserved as asparagine in vcCNT, nwCNT, and hCNTs and involved in coordinating the sodium ion in these proteins, the hydroxyl on Ser142 is not essential for substrate binding or transport activity in NupC. The NMR results suggest that mutations G146C and E149C have disrupted interactions between the substrate and its binding site. The abolished transport activity by G146C and E149C also suggests that these mutations have blocked substrate binding interactions and/or conformational changes required for completing the transport cycle.



Fig 9: Homology model of a NupC protomer and cartoon representation of a hypothetical elevator-type mechanism of concentrative nucleoside transport by NupC. (A) Homology model of a NupC protomer based on the crystal structure of vcCNT in complex with uridine (Protein Data Bank entry 4PD6) produced using SWISS-MODEL (https://swissmodel.expasy.org/). [55] The model includes a blue N-terminus and a red C-terminus, and the positions of mutated HP1 residues Ser142, Gly146, and Glu149 are indicated. The position of Phe350 that corresponds to the crucial vcCNT TM7b residue Phe366 is also indicated. (B) Cartoon representation of a hypothetical elevator-type mechanism of alternating access concentrative nucleoside transport by a NupC protomer. The external nucleoside (red circle) binds to an outwardopen state that undergoes conformational change to an outwardoccluded state in which the nucleoside packs between the tips of HP1 (mid green) and HP2 (light orange), TM4b (dark green), and TM7b (dark orange) behind the center of TM6 (light green). Elevator-like movement of the nucleoside-binding site (including HP1, HP2, TM4b, and TM7b) across TM6 produces an inwardoccluded state that undergoes conformational change to an inwardopen state from which the nucleoside is released to inside the cell. The outward-open state is then recovered. TM4b and HP1b act as gates for the outward-facing and inward-facing end states, respectively. The horizontal grey lines represent the approximate outer and inner limits of the lipid bilayer

On the basis of sequence homology, predictions of transmembrane helices, homology modeling, and the results of substrate binding and transport experiments, we have no reason to believe that NupC does not also adopt an elevatortype mechanism of alternating access membrane transport similar to that described for vcCNT and nwCNT. [24, 28, 29] In a hypothetical transport cycle for NupC (Fig 9B), the external nucleoside binds to an outward-open state that undergoes a conformational change to an outward-occluded state in which the nucleoside packs between the tips of HP1 and HP2, TM4b, and TM7b behind the center of TM6. Elevator-like movement of the nucleoside-binding site (including HP1, HP2, TM4b, and TM7b) across TM6 produces an inward-occluded state that undergoes a conformational change to an inward-open state from which the nucleoside is released. The outward-open state is then recovered, possibly via multiple intermediate conformations.

TM4b and HP1b act as gates for the outward-facing and inward-facing end states, respectively. It is therefore easy to envisage how mutations G146C and E149C at the tip of HP1 could block conformational changes in the transport cycle.

4. Conclusion

E. coli CNT family protein NupC shares the highest degree of sequence homology with crystallographically defined vcCNT and nwCNT (30.3 and 32.3% identical and 28.8 and 27.3% highly similar, respectively; combined 59.1 and 59.6%, respectively) and then with the human proteins in the following order: hCNT3 > hCNT1 > hCNT2. Regions of NupC corresponding to vcCNT structural domains involved in nucleoside binding (HP1, TM4, HP2, and TM7) contain highly conserved motifs. Of nine vcCNT residues involved in direct binding interactions with nucleosides, seven are identically conserved in NupC and the other two are highly similar. The corresponding residues in NupC are Gly146, Gln147, Ser148, Glu149, Ile181, Glu321, Phe350, Asn352, and Ser355. A large majority of these residues are also conserved in hCNTs. The four sodium-binding site residues in vcCNT are not conserved in NupC, reflecting their differences in cation selectivity. Based on mutations that change hCNT1 from a pyrimidine-specific to a purinespecific transporter, three mutations (Q147M, S180T, and I181V) are predicted to have the same effect on NupC.

Analysis of the NupC sequence by membrane topology prediction tools suggested nine putative transmembrane helices (or eight with a partial helix). The N-terminus was predicted to be periplasmic or cytoplasmic, but the distribution of positively charged residues was more consistent with a periplasmic N-terminus as it is in vcCNT. The putative locations of transmembrane helices in NupC based on sequence homology and the structure of vcCNT all showed some overlap with those coming from prediction tools, but there were significant differences in their lengths and in their start and end positions. The extra "ninth" helix in NupC given by prediction tools appears in a region corresponding to IH3 in vcCNT. In comparison, analysis of the vcCNT sequence by membrane topology prediction tools suggested ten or eight transmembrane helices with periplasmic N- and C-termini. While eight of these transmembrane helices showed some overlap with those in the structure of vcCNT, significant differences in their lengths and in their start and end positions emphasized the challenges in predicting transmembrane helices in CNTs.

Cysteine scanning mutagenesis of NupC investigated roles of specific residues in nucleoside binding and/or transport activity. WT NupC had an apparent transport affinity (K_m^{app}) for $[^{14}C]$ uridine of 22.2 ± 3.7 µM in energized *E. coli* whole cells and an apparent binding affinity (K_d^{app}) for [1'-¹³C]uridine of 1.8-2.6 mM in non-energized inner membranes. The position corresponding to Ser142 is occupied by asparagine in vcCNT, nwCNT, and hCNTs; in the structure of vcCNT, this is involved in coordinating the sodium ion. Consistent with this, NupC mutant S142C retained transport and binding affinities similar to those of WT. Gly146 is conserved in vcCNT, nwCNT, hCNT2, and hCNT3 and corresponds with Ser319 in hCNT1 that when mutated to glycine allows purine nucleoside transport. Glu149 corresponds to Glu156 in vcCNT, which is involved in nucleoside binding, and this position is conserved as glutamate in nwCNT and hCNTs. Consistent with these observations, mutants G146C and E149C both had abolished transport activity. These mutants did retain varying degrees of partial uridine binding activity, however, with affinities decreasing in the following order: WT > S142C > G146C > E149C. Gly146 and Glu149 therefore appear to be important for uridine binding but are essential for uridine transport. WT NupC and mutants S142C and G146C had similar respective affinities for binding uridine and thymidine. Mutant E149C had an even lower affinity for thymidine than for uridine, and treatment of E149C with iodoacetate blocked uridine binding, thus confirming the direct involvement of Glu149 in binding nucleosides.

In addition to vcCNT and nwCNT, NupC is a viable model for concentrative nucleoside transport, including the human proteins. Hence, there is a compelling requirement for highresolution crystal structures of NupC determined in multiple conformations and in complex with different ligands. These would confirm the putative structural organization of NupC and determine if it adopts the same elevator-type mechanism of alternating access membrane transport proposed for vcCNT and nwCNT. They would also demonstrate structural differences in substrate binding and the transport cycle for proton-driven versus sodium-driven nucleoside transport. Such structures could help explain how hCNT3 can use a proton gradient as well as a sodium gradient to drive uridine transport. NupC mutants with abolished transport activity, such as G146C and E149C from this study, may be locked in a single conformation of the transport cycle and therefore potentially more tractable to crystallization than the WT protein. Crystal structures of NupC could be combined with sub-angstrom resolution structural information about nucleoside conformation and interactions in the substrate-binding site using isotope labels and solid-state NMR methods that measure internuclear distances and torsion angles. [45, 58, 77, 80-82] Also, construction of further site-specific mutants and their characterization using a range of computational, chemical, biochemical, and biophysical methods should be pursued.

5. Funding

This work was funded by the Biotechnology and Biological Sciences Research Council (BBSRC), the Engineering and Physical Sciences Research Council (EPSRC), the Medical Research Council (MRC), the Wellcome Trust, the European Drug Initiative on Channels and Transporters (EDICT) consortium, and the University of Leeds. Isotopelabelled chemicals were funded by the Royal Society and GlaxoSmithKline. LS thanks the MRC and the University of Leeds for a PhD studentship.

6. Conflict of interest

The authors declare that they have no conflict of interest.

7. References

- 1. Munch-Peterson A, Mygind B. in *Metabolism of* nucleotides, nucleosides and nucleobases in microorganisms (Munch-Peterson A, Ed.), Academic Press, London, 1983.
- Mackey JR, Baldwin SA, Young JD, Cass CE. Nucleoside transport and its significance for anticancer drug resistance. Drug Resistance Updates. 1998; 1(5):310-324.
- 3. Baldwin SA, Mackey JR, Cass CE, Young JD. Nucleoside transporters: Molecular biology and

implications for therapeutic development. Molecular Medicine Today. 1999; 5(5):216-224.

- Damaraju VL, Damaraju S, Young JD, Baldwin SA, Mackey J, Sawyer MB, Cass CE. Nucleoside anticancer drugs: The role of nucleoside transporters in resistance to cancer chemotherapy. Oncogene. 2003; 22(47):7524-7536.
- King AE, Ackley MA, Cass CE, Young JD, Baldwin SA. Nucleoside transporters: From scavengers to novel therapeutic targets. Trends in Pharmacological Sciences. 2006; 27(8):416-425.
- 6. Choi JS, Berdis AJ. Nucleoside transporters: Biological insights and therapeutic applications. Future Medicinal Chemistry. 2012; 4(11):1461-1478.
- Saidijam M, Afshar S, Ahmad I, Patching SG. Nucleoside transporters in PET imaging of proliferating cancer cells using 3'-deoxy-3'-[¹⁸F]fluoro-L-thymidine. Journal of Diagnostic Imaging in Therapy. 2018; 5(1):1-13.
- 8. Parkinson FE, Damaraju VL, Graham K. Yao SY, Baldwin SA, Cass CE, *et al.* Molecular biology of nucleoside transporters and their distributions and functions in the brain. Current Topics in Medicinal Chemistry. 2011; 11(8):948-972.
- 9. Mygind B, Munch-Petersen A. Transport of pyrimidine nucleosides in cells of *Escherichia coli* K12. European Journal of Biochemistry. 1975; 59(2):365-372.
- Munch-Petersen A, Mygind B. Nucleoside transport systems in *Escherichia coli* K12: Specificity and regulation. Journal of Cellular Physiology. 1976; 89(4):551-559.
- 11. Munch-Petersen A, Mygind B, Nicolaisen A, Pihl NJ. Nucleoside transport in cells and membrane vesicles from *Escherichia coli* K12. Journal of Biological Chemistry. 1979; 254(10):3730-3737.
- Zhang Y, Craig JE, Gallagher MP. Location of the *nupC* gene on the physical map of *Escherichia coli* K-12. Journal of Bacteriology. 1992; 174(17):5758-5759.
- Craig JE, Zhang Y, Gallagher MP. Cloning of the *nupC* gene of *Escherichia coli* encoding a nucleoside transport system, and identification of an adjacent insertion element, IS 186. Molecular Microbiology. 1994; 11(6):1159-1168.
- 14. Gray JH, Owen RP, Giacomini KM. The concentrative nucleoside transporter family, SLC28. Pfluegers Archives. 2004; 447(5):728-734.
- Aymerich I, Duflot S, Fernàndez-Veledo S, Guillèn-Gòmez E, Huber-Ruano I, Casado FJ, Pastor-Anglada M. The concentrative nucleoside transporter family (SLC28): New roles beyond salvage? Biochemical Society Transactions. 2005; 33(Pt 1):216-219.
- Pastor-Anglada M, Cano-Soldado P, Errasti-Murugarren E, Casado FJ. SLC28 genes and concentrative nucleoside transporter (CNT) proteins. Xenobiotica. 2008; 38(7-8):972-994.
- Young JD, Yao SY, Baldwin JM, Cass CE, Baldwin, SA. The human concentrative and equilibrative nucleoside transporter families, SLC28 and SLC29. Molecular Aspects of Medicine. 2013; 34(2-3), 529-547.
- Young JD. The SLC28 (CNT) and SLC29 (ENT) nucleoside transporter families: A 30-year collaborative odyssey. Biochemical Society Transactions. 2016; 44(3):869-876.
- 19. Loewen SK, Ng AM, Yao SY, Cass CE, Baldwin SA,

Young JD. Identification of amino acid residues responsible for the pyrimidine and purine nucleoside specificities of human concentrative Na⁺ nucleoside cotransporters hCNT1 and hCNT2. Journal of Biological Chemistry. 1999; 274(35):24475-24484.

- Ritzel MW, Ng AM, Yao SY, Graham K, Loewen SK, Smith KM, *et al.* Molecular identification and characterization of novel human and mouse concentrative Na⁺-nucleoside co-transporter proteins (hCNT3 and mCNT3) broadly selective for purine and pyrimidine nucleosides (system cib). Journal of Biological Chemistry. 2001; 276(4):2914-2927.
- 21. Zhang J, Smith KM, Tackaberry T, Visser F, Robins MJ, Nielsen LP, *et al.* Uridine binding and transportability determinants of human concentrative nucleoside transporters. Molecular Pharmacology. 2005; 68(3):830-839.
- 22. Smith KM, Slugoski MD, Loewen SK, Ng AM, Yao SY, Chen XZ, Karpinski E, Cass CE, Baldwin SA, and Young JD. The broadly selective human Na⁺/nucleoside co-transporter (hCNT3) exhibits novel cation-coupled nucleoside transport characteristics. Journal of Biological Chemistry. 2005; 280(27):25436-25449.
- Smith KM, Slugoski MD, Cass CE, Baldwin SA, Karpinski E, Young JD. Cation coupling properties of human concentrative nucleoside transporters hCNT1, hCNT2 and hCNT3. Molecular Membrane Biology. 2007; 24(1):53-64.
- 24. Johnson ZL, Cheong CG, Lee SY. Crystal structure of a concentrative nucleoside transporter from *Vibrio cholerae* at 2.4 Å. Nature. 2012; 483(7390):489-493.
- 25. Yernool D, Boudker O, Jin Y, Gouaux E. Structure of a glutamate transporter homologue from *Pyrococcus horikoshii*. Nature. 2004; 431(7010):811-888.
- 26. Boudker O, Ryan RM, Yernool D, Shimamoto K, Gouaux E. Coupling substrate and ion binding to extracellular gate of a sodium-dependent aspartate transporter. Nature. 2007; 445(7126):387-393.
- 27. Rahman M, Ismat F, Jiao L, Baldwin JM, Sharples DJ, Baldwin SA, Patching SG. Characterisation of the DAACS family *Escherichia coli* glutamate/aspartateproton symporter GltP using computational, chemical, biochemical and biophysical methods. Journal of Membrane Biology. 2017; 250(2):145-162.
- 28. Johnson ZL, Lee JH, Lee K, Lee M, Kwon DY, Hong J, Lee SY. Structural basis of nucleoside and nucleoside drug selectivity by concentrative nucleoside transporters. eLife. 2014; 3:e03604.
- 29. Hirschi M, Johnson ZL, Lee SY. Visualizing multistep elevator-like transitions of a nucleoside transporter. Nature. 2017; 545(7652):66-70.
- Vergara-Jaque A, Fenollar-Ferrer C, Kaufmann D, Forrest LR. Repeat-swap homology modeling of secondary active transporters: Updated protocol and prediction of elevator-type mechanisms. Frontiers in Pharmacology. 2015; 6:183.
- Wöhlert D, Grötzinger MJ, Kühlbrandt W, Yildiz Ö. Mechanism of Na⁺-dependent citrate transport from the structure of an asymmetrical CitS dimer. eLife. 2015; 4:e09375.
- 32. Coincon M, Uzdavinys P, Nji E, Dotson DL, Winkelmann I, Abdul-Hussein S, Cameron AD, Beckstein O, Drew D. Crystal structures reveal the molecular basis of ion translocation in sodium/proton

antiporters. Nature Structural and Molecular Biology. 2016; 23(3):248-255.

- Drew D, Boudker O. Shared molecular mechanisms of membrane transporters. Annual Review of Biochemistry. 2016; 85:543-572.
- 34. McCoy JG, Ren Z, Stanevich V, Lee J, Mitra S, Levin EJ, Poget S, *et al.* The structure of a sugar transporter of the glucose EIIC superfamily provides insight into the elevator mechanism of membrane transport. Structure. 2016; 24(6):956-964.
- 35. Mulligan C, Fenollar-Ferrer C, Fitzgerald GA, Vergara-Jaque A, Kaufmann D, Li Y, *et al.* The bacterial dicarboxylate transporter VcINDY uses a two-domain elevator-type mechanism. Nature Structural and Molecular Biology. 2016; 23(3):256-263.
- 36. Garaeva AA, Slotboom DJ. Elevator-type mechanisms of membrane transport. Biochemical Society Transactions. 2020; 48(3):1227-1241.
- Reyes N, Ginter C, Boudker O. Transport mechanism of a bacterial homologue of glutamate transporters. Nature. 2009; 462(7275):880-885.
- 38. Nørholm MH, Dandanell G. Specificity and topology of the *Escherichia coli* xanthosine permease, a representative of the NHS subfamily of the major facilitator superfamily. Journal of Bacteriology. 2001; 183(16), 4900-4904.
- Loewen SK, Yao SY, Slugoski MD, Mohabir NN, Turner RJ, Mackey JR, *et al.* Transport of physiological nucleosides and anti-viral and anti-neoplastic nucleoside drugs by recombinant *Escherichia coli* nucleoside-H⁺ co-transporter (NupC) produced in *Xenopus laevis* oocytes. Molecular Membrane Biology. 2004; 21(1):1-10.
- Patching SG, Baldwin SA, Baldwin AD, Young JD, Gallagher MP, Henderson PJ, *et al.* The nucleoside transport proteins, NupC and NupG, from *Escherichia coli*: Specific structural motifs necessary for the binding of ligands. Organic and Biomolecular Chemistry. 2005; 3(3):462-470.
- 41. Westh Hansen SE, Jensen N, Munch-Petersen A. Studies on the sequence and structure of the *Escherichia coli* K-12 *nupG* gene, encoding a nucleoside-transport system. European Journal of Biochemistry. 1987; 168(2):385-391.
- 42. Xie H, Patching SG, Gallagher MP, Litherland GJ, Brough AR, Venter H, *et al.* Purification and properties of the *Escherichia coli* nucleoside transporter NupG, a paradigm for a major facilitator transporter sub-family. Molecular Membrane Biology. 2004; 21(5):323-336.
- 43. Holyoake J, Caulfeild V, Baldwin SA, Sansom MS. Modeling, docking, and simulation of the major facilitator superfamily. Biophysical Journal. 2006; 91(10):L84-L86.
- Vaziri H, Baldwin SA, Baldwin JM, Adams DG, Young JD, Postis VL. Use of molecular modelling to probe the mechanism of the nucleoside transporter NupG. Molecular Membrane Biology. 2013; 30(2):114-128.
- 45. Patching SG, Middleton DA, Henderson PJ, Herbert RB. The economical synthesis of [2'- ¹³C, 1,3-¹⁵N₂]uridine; Preliminary conformational studies by solid state NMR. Organic and Biomolecular Chemistry. 2003; 1(12):2057-2062.
- 46. Patching SG, Brough AR, Herbert RB, Rajakarier JA, Henderson PJ, Middleton DA. Substrate affinities for

membrane transport proteins determined by ¹³C crosspolarization magic-angle spinning nuclear magnetic resonance spectroscopy. Journal of the American Chemical Society. 2004; 126(10):3072-3080.

- 47. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular Systems Biology. 2011; 7:539.
- Letunic I, Bork P. Interactive tree of life (iTOL) v3: An online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Research. 2016; 44(W1):W242-W245.
- Bernsel A, Viklund H, Hennerdal A, Elofsson A. TOPCONS: Consensus prediction of membrane protein topology. Nucleic Acids Research. 2009; 37(Web Server issue), W465-W468.
- 50. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. Journal of Molecular Biology 2001; 305(3):567-580.
- Dobson L, Remènyi I, Tusnàdy GE. CCTOP: A Consensus Constrained TOPology prediction web server. Nucleic Acids Research. 2015; 43(W1):W408-W412.
- 52. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins M, Appel RD, *et al.* Protein identification and analysis tools on the ExPASy Server. In *The proteomics protocols handbook* (Walker JM, Ed.). Humana Press, Totowa, NJ, 2005, 571-607.
- 53. Moreland JL, Gramada A, Buzko OV, Zhang Q, Bourne PE. The Molecular Biology Toolkit (MBT): A modular platform for developing molecular visualization applications. BMC Bioinformatics. 2005; 6:21.
- 54. Rose AS, Hildebrand PW. NGL Viewer: A web application for molecular visualization. Nucleic Acids Research. 2015; 43(W1):W576-W579.
- 55. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, *et al.* SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Research. 2014; 42(Web Server issue):W252-W258.
- 56. Witholt B, Boekhout M, Brock M, Kingma J, van Heerikhuizen H, De Leij L. An efficient and reproducible procedure for the formation of spheroplasts from variously grown *Escherichia coli*. Analytical Biochemistry. 1976; 74(1):160-170.
- 57. Ward A, Sanderson NM, O'Reilly J, Rutherford NG, Poolman B, Henderson PJF. The amplified expression, identification purification, assay and properties of hexahistidine-tagged bacterial membrane transport proteins. In *Membrane transport: A practical approach* (Baldwin SA, Ed.), Oxford University Press, Oxford, U.K, 2000, 141-166.
- 58. Patching SG. Chemical and solid-state NMR approaches to determine structure-activity relationships for substrates bound to the nucleoside transport proteins, NupC and NupG, of *Escherichia coli*. Ph.D. Thesis, University of Leeds, Leeds, U.K, 2002.
- Slugoski MD, Loewen SK, Ng AM, Smith KM, Yao SY, Karpinski E, Cass CE, *et al.* Specific mutations in transmembrane helix 8 of human concentrative Na⁺/nucleoside co-transporter hCNT1 affect permeant selectivity and cation coupling. Biochemistry. 2007;

46(6):1684-1693.

- 60. Yao SY, Ng AM, Slugoski MD, Smith KM, Mulinta R, Karpinski E, Cass CE, Baldwin SA, Young JD. Conserved glutamate residues are critically involved in Na⁺/nucleoside co-transport by human concentrative nucleoside transporter 1 (hCNT1). Journal of Biological Chemistry. 2007; 282(42):30607-30617.
- Tsirigos KD, Peters C, Shu N, Käll L, Elofsson A. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. Nucleic Acids Research. 2015; 43(W1):W401-W407.
- 62. Saidijam M, Azizpour S, Patching SG. Comprehensive analysis of the numbers, lengths and amino acid compositions of transmembrane helices in prokaryotic, eukaryotic and viral integral membrane proteins of high-resolution structure. Journal of Biomolecular Structure and Dynamics. 2018; 36(2):443-464.
- 63. von Heijne G. Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. Journal of Molecular Biology. 1992; 225(2):487-494.
- 64. Saidijam M, Azizpour S, Patching SG. Amino acid composition analysis of human secondary transport proteins and implications for reliable membrane topology prediction. Journal of Biomolecular Structure and Dynamics. 2017; 35(5):929-949.
- 65. Saidijam M, Patching SG. Amino acid composition analysis of secondary transport proteins from *Escherichia coli* with relation to functional classification, ligand specificity and structure. Journal of Biomolecular Structure and Dynamics. 2015; 33(10):2205-2220.
- 66. Sun L. Molecular mechanisms of nucleoside transport. Ph.D. Thesis, University of Leeds, Leeds, U.K, 2007.
- McDonald TP, Walmsley AR, Martin GE, Henderson PJ. The role of tryptophans 371 and 395 in the binding of antibiotics and the transport of sugars by the Dgalactose-H⁺ symport protein (GalP) from *Escherichia coli*. Journal of Biological Chemistry. 1995; 270(51):30359-30370.
- Liang WJ, Wilson KJ, Xie H, Knol J, Suzuki S, Rutherford NG, *et al.* The *gusBC* genes of *Escherichia coli* encode a glucuronide transport system. Journal of Bacteriology. 2005; 187(7):2377-2385.
- 69. Bettaney KE, Sukumar P, Hussain R, Siligardi G, Henderson PJ, Patching SG. A systematic approach to the amplified expression, functional characterization and purification of inositol transporters from *Bacillus subtilis*. Molecular Membrane Biology. 2013; 30(1):3-14.
- 70. Ma P, Patching SG, Ivanova E, Baldwin JM, Sharples D, Baldwin SA, *et al.* Allantoin transport protein, PucI, from *Bacillus subtilis*: Evolutionary relationships, amplified expression, activity and specificity. Microbiology. 2016; 162(5):823-836.
- 71. Rahman M, Ismat F, Jiao L, Baldwin JM, Sharples DJ, Baldwin SA, *et al.* Characterisation of the DAACS family *Escherichia coli* glutamate/aspartate-proton symporter GltP using computational, chemical, biochemical and biophysical methods. Journal of Membrane Biology. 2017; 250(2):145-162.
- 72. Mulinta R, Yao SY, Ng AM, Cass CE, Young JD. Substituted cysteine accessibility method (SCAM) analysis of the transport domain of human concentrative nucleoside transporter 3 (hCNT3) and other family

members reveals features of structural and functional importance. Journal of Biological Chemistry. 2017; 292(23):9505-9522.

- 73. Patching SG, Edara S, Ma P, Nakayama J, Hussain R, Siligardi G, *et al.* Interactions of the intact FsrC membrane histidine kinase with its pheromone ligand GBAP revealed through synchrotron radiation circular dichroism. Biochimica et Biophysica Acta, Biomembranes. 2012; 1818(7):1595-1602.
- 74. Siligardi G, Hussain R, Patching SG, Phillips-Jones MK. Ligand- and drug-binding studies of membrane proteins revealed through circular dichroism spectroscopy. Biochimica et Biophysica Acta, Biomembranes. 2014; 1838(1 Pt A), 34-42.
- 75. Patching SG, Herbert RB, O'Reilly J, Brough AR, Henderson PJ. Low ¹³C-background for NMR-based studies of ligand binding using ¹³C-depleted glucose as carbon source for microbial growth: ¹³C-labelled glucose and ¹³C-forskolin binding to the galactose-H⁺ symport protein GalP in *Escherichia coli*. Journal of the American Chemical Society. 2004; 126(1):86-87.
- 76. Patching SG, Psakis G, Baldwin SA, Baldwin J, Henderson PJ, Middleton DA. Relative substrate affinities of wild-type and mutant forms of the *Escherichia coli* sugar transporter GalP determined by solid-state NMR. Molecular Membrane Biology. 2008; 25(6-7):474-484.
- 77. Patching SG, Henderson PJ, Herbert RB, Middleton DA. Solid-state NMR spectroscopy detects interactions

between tryptophan residues of the *E. coli* sugar transporter GalP and the alpha-anomer of the D-glucose substrate. Journal of the American Chemical Society. 2008; 130(4):1236-1244.

- 78. Rahman M, Patching SG, Ismat F, Henderson PJ, Herbert RB, Baldwin SA, *et al.* Probing metal ion substrate-binding to the *E. coli* ZitB exporter in native membranes by solid state NMR. Molecular Membrane Biology. 2008; 25(8):683-690.
- 79. Patching SG, Henderson PJ, Sharples DJ, Middleton DA. Probing the contacts of a low-affinity substrate with a membrane-embedded transport protein using ¹H-¹³C cross-polarisation magic-angle spinning solid-state NMR. Molecular Membrane Biology. 2013; 30(2):129-137.
- Patching SG, Edwards R, Middleton DA. Structural analysis of uniformly ¹³C-labelled solids from selective angle measurements at rotational resonance. Journal of Magnetic Resonance. 2009; 199(2):242-246.
- Edwards R, Madine J, Fielding L, Middleton DA. Measurement of multiple torsional angles from onedimensional solid-state NMR spectra: Application to the conformational analysis of a ligand in its biological receptor site. Physical Chemistry Chemical Physics. 2010; 12(42):13999-14008.
- Patching SG. Solid-state NMR structures of integral membrane proteins. Molecular Membrane Biology. 2015 32(5-8): 156-178.

Supporting Information

(Tables S1 and S2, Figures S1-S12)

Table S1: Sequence homology in concentrative nucleoside transporters. Amino acid sequences of NupC from *Escherichia coli* (P0AFF2), vcCNT from *Vibrio cholera* (Q9KPL5), nwCNT from *Neisseria wadsworthii* (G4CRQ5) and human concentrative nucleoside transporters hCNT1 (O00337), hCNT2 (O43868) and hCNT3 (Q9HAS3) were taken from the UniProt KnowledgeBase (http://www.uniprot.org/) and separately aligned with each other using the online multiple sequence alignment tool Clustal Omega

(http://www.ebi.ac.uk/Tools/msa/clustalo/).¹ The percentage sequence homology is given in terms of identical residues (*red*), highly similar residues (*blue*) and the combined total of these (*black*).

	NupC		NupC vcCNT			nwCNT		hCNT1		hCNT2		hCNT3						
NupC				30.3	28.8	58.1	32.3	27.3	59.6	27.0	26.3	53.3	22.3	31.3	53.5	25.3	29.3	54.5
vcCNT	30.3	28.8	59.1				66.5	18.2	84.7	36.4	26.8	63.2	36.1	26.6	62.7	38.8	25.4	64.2
nwCNT	32.3	27.3	59.6	66.5	18.2	84.7				35.5	25.6	61.1	34.6	25.9	60.5	34.6	27.3	61.9
hCNT1	27.0	26.3	53.3	36.4	26.8	63.2	35.5	25.6	61.1				63.3	17.7	81	43.9	23.4	67.3
hCNT2	22.3	31.3	53.5	36.1	26.6	62.7	34.6	25.9	60.5	63.3	17.7	81				42.1	24.3	66.4
hCNT3	25.3	29.3	54.5	38.8	25.4	64.2	34.6	27.3	61.9	43.9	23.4	67.3	42.1	24.3	66.4			

Table S2: Amino acid compositions of concentrative nucleoside transporters. Amino acid sequences of NupC from *Escherichia coli* (POAFF2), vcCNT from *Vibrio cholera* (Q9KPL5), nwCNT from *Neisseria wadsworthii* (G4CRQ5) and human concentrative nucleoside transporters hCNT1 (O00337), hCNT2 (O43868) and hCNT3 (Q9HAS3) were taken from the UniProt KnowledgeBase (http://www.uniprot.org/) and analysed by the online ExPASy tool ProtParam (http://web.expasy.org/protparam/).² The contents of individual amino acids and of groupings of amino acids with similar physicochemical properties in each nucleoside transporter were compared with overall average values in 235 secondary transport proteins from *Escherichia coli* and in 336 human secondary transport proteins.^{3,4} The groupings of amino acids with similar physicochemical properties are: hydrophobic (alanine, isoleucine, leucine, phenylalanine and valine); positively charged at physiological pH (arginine and lysine); negatively charged at physiological pH (aspartic acid and glutamic acid); hydroxyl-containing (serine and threonine); amido-containing (asparagine and glutamine).

Amino acid	E. coli	Human	NupC	vcCNT	nwCNT	hCNT1	hCNT2	hCNT3
Ala	10.8%	8.2%	9.8%	13.2%	13.6%	10.3%	9.4%	7.7%
Arg	3.5%	4.2%	2.8%	2.6%	2.6%	5.4%	3.6%	3.9%
Asn	2.8%	3.0%	4.2%	3.3%	2.8%	2.3%	3.0%	4.3%
Asp	2.3%	3.1%	1.8%	2.1%	2.1%	2.8%	2.1%	2.6%
Cys	1.1%	2.1%	0.2%	0.9%	0.2%	3.1%	3.0%	2.0%
Gln	2.6%	3.4%	2.5%	2.1%	2.6%	3.7%	3.2%	3.0%
Glu	2.6%	4.4%	4.2%	3.8%	4.0%	5.2%	5.0%	4.9%

International Journal of Advanced Multidisciplinary Research and Studies

Gly	8.9%	8.0%	7.8%	13.2%	11.8%	6.6%	7.9%	7.4%
His	1.4%	1.8%	0.8%	0.0%	0.5%	1.2%	1.1%	2.3%
Ile	8.1%	6.2%	9.8%	6.8%	7.5%	5.4%	6.1%	6.9%
Leu	13.6%	12.6%	12.8%	13.0%	12.5%	14.3%	12.0%	10.6%
Lys	2.9%	3.8%	3.8%	3.3%	3.5%	2.9%	4.3%	3.8%
Met	4.0%	2.8%	4.5%	4.2%	3.8%	2.5%	2.6%	3.3%
Phe	6.1%	5.8%	7.8%	7.3%	5.2%	6.5%	6.8%	6.8%
Pro	4.0%	5.1%	1.8%	4.0%	3.1%	3.5%	3.2%	3.6%
Ser	6.6%	7.5%	9.2%	6.4%	7.1%	8.0%	7.8%	8.8%
Thr	5.5%	5.5%	2.8%	3.3%	3.3%	4.2%	6.2%	5.5%
Trp	2.1%	1.6%	0.8%	1.2%	0.9%	2.2%	2.0%	2.3%
Tyr	2.7%	3.2%	2.8%	1.9%	2.1%	2.0%	2.4%	3.0%
Val	8.4%	7.8%	10.2%	7.3%	11.1%	7.9%	8.2%	7.1%
Hydrophobic	47.0%	40.5%	50.4%	47.6%	49.9%	44.4%	42.5%	39.1%
Positive	6.4%	8.0%	6.6%	5.9%	6.1%	8.3%	7.9%	7.7%
Negative	4.8%	7.4%	6.0%	5.9%	6.1%	8.0%	7.1%	7.5%
Hydroxyl	12.1%	13.0%	12.0%	9.7%	10.4%	12.2%	14.0%	14.3%
Amido	5.5%	6.4%	6.7%	5.4%	5.4%	6.0%	6.2%	7.3%



Fig S1: Evolutionary relationships of NupC and concentrative nucleoside transporters. Amino acid sequences of NupC from *Escherichia coli* (POAFF2), vcCNT from *Vibrio cholera* (Q9KPL5), nwCNT from *Neisseria wadsworthii* (G4CRQ5) and human concentrative nucleoside transporters hCNT1 (O00337), hCNT2 (O43868) and hCNT3 (Q9HAS3) were taken from the UniProt KnowledgeBase (http://www.uniprot.org/) and aligned using the online multiple sequence alignment tool Clustal Omega

(http://www.ebi.ac.uk/Tools/msa/clustalo/).¹ The resultant neighbour-joining phylogenetic tree was exported in Newick format and drawn using the online tool iTOL: Interactive Tree of Life (http://itol.embl.de/).⁵ Sequence homologies between NupC and the other proteins are given in terms of identical residues (*red*), highly similar residues (*blue*) and the combined total of these (*black*). These values were calculated from individual sequence alignments between NupC and the other proteins

hCNT1	MENDPSRRRESISLTPVAKGLENMGADFLESLEEG
hCNT2	MEKASGRQSIALSTVETGTVNPGLELMEKEVEP
hCNT3	MELRSTAAPRAEGYSNVGFQNEENFLENENTSGNNSIRSRAVQSREHTNTKQDEEQVTVE
hCNT1	Q-LPRSDLSPAEIRSSWSEAAPKPFSRWRNLQPALRARSFCREHMQLFRWIGTGLLCTGL
hCNT2	EGSKRTDAQGHSLGDGLGPSTY-QRRSRWPFSKARSFCKTHASLFKKILLGLLCLAY
hCNT3	QDSPRNREHMEDDDEEMQQKGCLERRYDTVCGFCRKHKTTLRHIIWGILLAGY
hCNT1	SAFLLVACLLDFQRALALFVLTCVVLTFLGHRLLKRLLGPKLRRFLKPQGHPRLLLWF
hCNT2	AAYLLAACILNFQRALALFVITCLVIFVLVHSFLKKLLGKKLTRCLKPFENSRLRLWT
hCNT3	LVM VISACVLNFHRAL PL FVIT VAAIFF V VWDHLMAKYEH RI DEMLS P GRRLLNSHW FW L
hCNT1	KRGLALAAFLGLVLWLSLDTSQR-PEQLVSFAGICVFVALLFACSKHHCAVSWRAVSWGL
hCNT2	KWVFAGVSLVGLILWLALDTAQR-PEQLIPFAGICMFILILFACSKHHSAVSWRTVFSGL
hCNT3	KWVIWSSLVLAVIFWLAFDTAKLGQQQLVSFGGLIMYIVLLFLFSKYPTRVYWRPVLWGI
hCNT1	GLOFVLGLLVIRTEPGFIAFEWLGEOIRIFLSYTKAGSSFVFGEALVKDVFAFOVLPIIV
hCNT2	GLOFVFGILVIRTDLGYTVFOWLGEOVOIFLNYTVAGSSFVFGDTLVKDVFAFOALPIII
hCNT3	GLOFLLGLLILRTDPGFIAFDWLGROVOTFLEYTDAGASFVFGEKYKDHFFAFKVLPIVV
hCNT1	FFSCVISVLYHVGLMOWVILKIAWLMOVTMGTTATETLSVAGNIFVSOTEAPLLIRPYLA



Fig S2: Protein sequence alignment between human concentrative nucleoside transporters. Amino acid sequences of the human concentrative nucleoside transporters hCNT1 (000337), hCNT2 (043868) and hCNT3 (Q9HAS3) were taken from the UniProt KnowledgeBase (http://www.uniprot.org/) and aligned using the online multiple sequence alignment tool Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/).¹ Residues are coloured to indicate those that are identical (*red*) and highly similar (*blue*)



Fig S3: Evolutionary relationships of NupC with Eukaryota. A. The amino acid sequence of NupC from *Escherichia coli* (P0AFF2) was subjected to a BLAST search against all Eukaryota in the UniProt KnowledgeBase (http://www.uniprot.org/). Sequences for the top 500 closest proteins were extracted and aligned using the online multiple sequence alignment tool Clustal Omega

(http://www.ebi.ac.uk/Tools/msa/clustalo/).¹ The resultant neighbour-joining phylogenetic tree was exported in Newick format and drawn using the online tool iTOL: Interactive Tree of Life (http://itol.embl.de/).⁵ B. Expansion of the branch containing NupC that also contains two close homologues of NupC, A0A0A2W106 and A0A0A2V7L6, from the soil fungus *Beauveria bassiana*

NupC	MDRVLHFVLALAVVAILALLVSSDRKKIRIRYVIQLLVIEVLLAWFFLNSDVGLG	55
VCCNT	<mark>MSLFMSCCGMAVLL</mark> GIAVLLSSNRKAINLRTVGGAFAIQFSLGAFILYVPWGQE	54
NupC	FVKGFSEMFEKLLGFANEGTNFVFGSMNDQGLAFFFLKVLCPIVFISALIGI	107
VCCNT	LLRGFSDAVSNVINYGNDGTSFLFGGLVSGKMFE <mark>VFGGGGFIFAFRVL</mark> PTLIFFSALIS <mark>V</mark>	114
NupC	LQHIRVLPVIIRAIGFLLSKVNGMGKLESFNAVSSLILGQSENFIAYKDILGKISRNRMY	167
VCCNT	<mark>LYYLGVMQWVIRILGGGLQ</mark> KALGTSRAESMSAAANIFV <mark>GQTE</mark> APLVVRPF <mark>VPKMTQSELF</mark>	174
NupC	TMAATAMSTVSMSIVGAYMTM-LEPKYVVAALVLNMFSTFIVLSLINPYRVD-ASEENIQ	225
VCCNT	AVMCGGLASIAGG <mark>yl</mark> ag <mark>yasmgvkieylvaa</mark> sfmaapggllfaklmmpetekpqdnedit	234
NupC	MSNL-HEGQSFFEMLGEYIL AGFKVAI IVA AMLI GFIALIAALNALFATVTGWFGYS-IS	283
VCCNT	LDGG <mark>DDKPANVIDAAAGGASAGLQLALNVGAMLIAFIGLIALINGM</mark> LGGIG <mark>GWFGMPELK</mark>	294
NupC	FQGILGYIFYPIAWVMGVPSSEALQVGSIMATKLVSNEFVAMMDLQKIASTLSPR	338
VCCNT	LEMLLGWLFAPLAFLIG <mark>VPWNEA</mark> TVAGEFIGLKTVANEFVAYSQFAPYLTEA <mark>APVVLSEK</mark>	354
NupC	AEGIISVFLVSFANFSSIGIIAGAVKGLNEEQGNVVSRFGLKLVYGSTLVSVLSASIAAL	398
VCCNT	TKAIISFALCG <mark>FANLSS</mark> IAILL <mark>G</mark> GLGS <mark>LAPKRRGDIARMGVKAVIAGTLSNLMAAT</mark> IAGF	414
NupC	VL 400	
VCCNT	FLSF 418	
Red	Identical residues	
Blue	Highly similar residues	
TM1-TM8	Transmembrane helices in vcCNT	
IH1-IH3	Interfacial helices in vcCNT	
HPI-HP2 Double	Hairpin re-entrant loops in vcCNT	
BINGING	NUCLEOSIAE DINAING IN VCUNT	

Fig S4: Conservation of residues in NupC from *Escherichia coli* and vcCNT from *Vibrio cholerae*. Amino acid sequences of NupC from *Escherichia coli* (P0AFF2) and vcCNT from *Vibrio cholera* (Q9KPL5) were taken from the UniProt KnowledgeBase

(http://www.uniprot.org/) and aligned using the online multiple sequence alignment tool Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/).¹ Residues are coloured to indicate those that are identical (*red*) or highly similar (*blue*). Residues in vcCNT are highlighted to indicate those that contribute to transmembrane spanning helices (TM1-TM8) (*green*), interfacial helices (IH1-IH3) (*cyan*), hairpin re-entrant loops (HP1 and HP2) (*yellow*) and in direct binding interactions with nucleosides (Gly153, Gln154, Thr155, Glu156, Val188, Glu332, Phe366, Asn368 and Ser371) (*pink*) based on crystal structures of vcCNT.^{6,7}

HP1		
NupC	KVN G MGKL ES FNAVSS LILGQSE NF	151
VCCNT	KALGTSRAESMSAAANIFV <mark>GQTE</mark> AP	158
hCNT1	VTMGTTATETLSVAGNIFVSQTEAP	324
hCNT2	ITMGTTATETLAVAG NIFVG MTEAP	318
hCNT3	VTTGSSPIESVVASGNIFVGQTESP	345
TM4		
NupC	LGKISRNRMYTMAATAMSTVSMSIV	182
vcCNT	VPKMTQSELFAVMCGGLASIAGG <mark>V</mark> L	189
hCNT1	LADMTLSEVHVVMTGGYATIAGSLL	355
hCNT2	LGDMTLSEIHAVMTGGFATISGTVL	349
hCNT3	LPYITKSELHAIMTAGFSTIAGSVL	376
HP2		
NupC	VPSSEALQVGSIMATKLVSNEFVAM	325
vcCNT	VPWNEATVAGEFIGLKTVAN <mark>E</mark> FVAY	336
hCNT1	VAWEDCPVVAELLGIKLFLNEFVAY	502
hCNT2	VEWTDCPMVAEMVGIKFFINEFVAY	496
hCNT3	VEWQDSFMVARLIGYKTFFNEFVAY	523
TM7		
NupC	STLSPRAEGIISVFLVSFANFSS	IGI IA 360
VCCNT	APVV LSEKTKAIISFALCG<mark>F</mark>ANLS<mark>S</mark>I	CAILL 376
hCNT1	RKQWISVRAEVLTTFALCGFANFSS	IGIML 552
hCNT2	EKQWISVRAEIITTFSLCGFANLSS	IGI TL 546
hCNT3	VQQY ISIRSEIIAT YA LCGFANI G S I	L <mark>GI</mark> VI 573
Red	Identical residues	
Blue	Highly similar residues	
HP1+HP	<mark>2</mark> Hairpin re-entrant loop	s in vcCN1
TM4+TM	7 Transmembrane helices i	n vcCNT
Bindin	g Nucleoside binding in v	CCNT

Fig S5: Conservation of residues in nucleoside binding regions of concentrative nucleoside transporters. Sections representing the structural domains HP1, TM4, HP2 and TM7 in vcCNT taken from a complete amino acid sequence alignment between NupC from *Escherichia coli* (P0AFF2), vcCNT from *Vibrio cholera* (Q9KPL5) and human concentrative nucleoside transporters hCNT1 (000337), hCNT2 (043868) and hCNT3 (Q9HAS3). Sequences were taken from the UniProt KnowledgeBase (http://www.uniprot.org/) and aligned using the online multiple sequence alignment tool Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/).¹ Residues are coloured to indicate those that are identical (*red*) or highly similar (*blue*). Residues in vcCNT are highlighted to indicate those that contribute to the given structural domain (HP1 and HP2) (*yellow*), (TM4 and TM7) (*green*) and in direct binding interactions with nucleosides (Gly153, Gln154, Thr155, Glu156, Val188, Glu332, Phe366, Asn368 and Ser371) (*pink*) based on crystal structures of vcCNT. ^[6,7]

NupC	
vcCNT	
hCNT1	MENDPSRRRESISLTPVAKGLENMGADFLESLEEG 35
hCNT2	MEKASGRQSIALSTVETGTVNPGLELMEKEVEP 33
hCNT3	$MELRSTAAPRAEGYSNVGFQNEENFLENENTSGNNSIRSRAVQSREHTNTKQDEEQVTVE\ 60$
N G	
NupC	
vcCNT	
hCNT1	Q-LPRSDLSPAEIRSSWSEAAPKPFSRWRNLQPALRARSFCREHMQLFRWIGTGLLCTGL 94
hCNT2	EGSKRTDAQGHSLGDGLGPSTY-QRRSRWPFSKARSFCKTHASLFKKILLGLLCLAY 89
hCNT3	QDSPRNREHMEDDDEEMQQKGCLERRYDTVCGFCRKHKTTLRHIIWGILLAGY 113
NupC	
vcCNT	
hCNT1	SAFLLVACLLDFQRALALFVLTCVVLTFLGHRLLKRLLGPKLRRFLKPQGHPRLLLWF 153
hCNT2	AAYLLAACILNFQRALALFVITCLVIFVLVHSFLKKLLGKKLTRCLKPFENSRLRLWT 147
hCNT3	I VMVISACVI NEHRAI DI EVITVA AIEEVVWDHI MAKVEHRIDEMI SPCPRI I NSHWEWI 173
	L VIVI VISAC V LIVITIKALI LI VITI VAAITT V V VIDILLIVIAK TEHKIDEMLSI OKKLENSIIVITVIL 175
	L VM VISAC V LIVITIKALI LI VITI VAAILT V V W DILLWAK I LIIKIDEWLSI OKKLENSIIWI WL 175
NupC	MDRVLHFVLALAVVAILALLVSSDRKKIRIRYVIQL 36
NupC vcCNT	MDRVLHFVLALAVVAILALLVSSDRKKIRIRYVIQL 36 GPAVPRMSLFMS-CCGMAVLLGIAVLLSSNRKAINLRTVGGA 41
NupC vcCNT hCNT1	MDRVLHFVLALAVVAILALLVSSDRKKIRIRYVIQL 36 GPAVPRMSLFMS-CCGMAVLLGIAVLLSSNRKAINLRTVGGA 41 KRGLALAAFLGLVLWLSLDTSQR-PEQLVS-FAGICVFVALLFACSKHHCAVSWRAVSWG 211
NupC vcCNT hCNT1 hCNT2	MDRVLHFVLALAVVAILALLVSSDRKKIRIRYVIQL 36 GPAVPRMSLFMS-CCGMAVLLGIAVLLSSNRKAINLRTVGGA 41 KRGLALAAFLGLVLWLSLDTSQR-PEQLVS-FAGICVFVALLFACSKHHCAVSWRAVSWG 211 KWVFAGVSLVGLILWLALDTAQR-PEQLIP-FAGICMFILILFACSKHHSAVSWRTVFSG 205
NupC vcCNT hCNT1 hCNT2 hCNT3	MDRVLHFVLALAVVAILALLVSSDRKKIRIRYVIQL 36 GPAVPRMSLFMS-CCGMAVLLGIAVLLSSNRKAINLRTVGGA 41 KRGLALAAFLGLVLWLSLDTSQR-PEQLVS-FAGICVFVALLFACSKHHCAVSWRAVSWG 211 KWVFAGVSLVGLILWLALDTAQR-PEQLIP-FAGICMFILILFACSKHHSAVSWRTVFSG 205 KWVIWSSLVLAVIFWLAFDTAKLGQQQLVS-FGGLIMYIVLLFLFSKYPTRVYWRPVLWG 232

NupC LVIEVLLAWFFLNSDVGLGFVKGFSEMFEKLLGFANEGTNFVFGSMNDQ------GLA 88

vcCNT FAIQFSLGAFILYVPWGQELLRGFSDAVSNVINYGNDGTSFLFGGLVSGKMFEVFGGGGF 101	
hCNT1 LGLQFVLGLLVIRTEPGFIAFEWLGEQIRIFLSYTKAGSSFVFGEALVKD 261	
hCNT2 LGLOFVFGILVIRTDLGYTVFOWLGEOVOIFLNYTVAGSSFVFGDTLVKD 255	
hCNT3 IGLQFLLGLLILRTDPGFIAFDWLGRQVQTFLEYTDAGASFVFGEKYKDH 282	
NupC FFFLKVLCPIVFISALIGILQHIRVLPVIIRAIGFLLSKVNGMGKLESFNAVSSLILGQS 148	
vcCNT IFAFRVLPTLIFFSALISVLYYLGVMOWVIRILGGGLOKALGTSRAESMSAAANIFVGOT 161	
hCNT1 VFAFQVLPIIVFFSCVISVLYHVGLMQWVILKIAWLMQVTMGTTATETLSVAGNIFVSQT 321	
hCNT2 VFAFOALPIIIFFGCVVSILYYLGLVOWVVOKVAWFLOITMGTTATETLAVAGNIFVGMT 315	
hCNT3 FFAFKVLPIVVFFSTVMSMLYYLGLMQWIIRKVGWIMLVTTGSSPIESVVASGNIFVGQT 342	
NupC ENFIAYKDILGKISRNRMYTMAATAMSTVSMSIVGAYMTMLEP-KYVVAALVLNMFSTFI 207	
vcCNT EAPLVVRPFVPKMTOSELFAVMCGGLASIAGGVLAGYASMGVKIEYLVAASFMAAPGGLL 221	
hCNT1 EAPLLIRPYLADMTLSEVHVVMTGGYATIAGSLLGAYISFGIDATSLIAASVMAAPCALA 381	
hCNT2 EAPLLIRPYLGDMTLSEIHAVMTGGFATISGTVLGAFIAFGVDASSLISASVMAAPCALA 375	
hCNT3 ESPLLVRPYLPYITKSELHAIMTAGFSTIAGSVLGAYISFGVPSSHLLTASVMSAPASLA 402	
NupC VLSLINPYRVDASEENIOMSNLHEGOSFFEMLGEYILAGFKVAIIVAAMLIGFIALI 264	
vcCNT FAKLMMPETEKPODNEDITLDGG-DDKPANVIDAAAGGASAGLOLALNVGAMLIAFIGLI 280	
hCNT1 LSKLVYPEVEESKFRREEGVKLT-YGDAONLIEAASTGAAISVKVVANIAANLIAFLAVL 440	
hCNT2 SSKLAYPEVEESKFKSEEGVKLP-RGKERNVLEAASNGAVDAIGLATNVAANLIAFLAVL 434	
hCNT3 AAKLFWPETEKPKITLKNAMKME-SGDSGNLLEAATOGASSSISLVANIAVNLIAFLALL 461	
NupC AALNALFATVTGWFGYS-ISFQGILGYIFYPIAWVMGVPSSEALQVGSIMATKLVSNEFV 323	
vcCNT ALINGMLGGIGGWFGMPELKLEMLLGWLFAPLAFLIGVPWNEATVAGEFIGLKTVANEFV 340	
hCNT1 DFINAALSWLGDMVDIQGLSFOLICSYILRPVAFLMGVAWEDCPVVAELLGIKLFLNEFV 500	
hCNT2 AFINAALSWLGELVDIOGLTFOVICSYLLRPMVFMMGVEWTDCPMVAEMVGIKFFINEFV 494	
hCNT3 SFMNSALSWFGNMFDYPOLSFELICSYIFMPFSFMMGVEWODSFMVARLIGYKTFFNEFV 521	
NupC AMMDLOKIASTLSPRAEGIISVFLVSFANFSSIGIIAGAVKGLNE 368	
vcCNT AYSOFAPYLTEAAPVVLSEKTKAIISFALCGFANLSSIAILLGGLGSLAP 390	
hCNT1 AYQDLSKYKORRLAGAEEWVGDRKQWISVRAEVLTTFALCGFANFSSIGIMLGGLTSMVP 560	
hCNT2 AYOOLSOYKNKRLSGMEEWIEGEKOWISVRAEIITTFSLCGFANLSSIGITLGGLTSIVP 554	
hCNT3 AYEHLSKWIHLRKEGGPKFVNGVOOYISIRSEIIATYALCGFANIGSLGIVIGGLTSMAP 581	
NupC EOGNVVSRFGLKLVYGSTLVSVLSASIAALVL 400	
vcCNT_KRRGDIARMGVKAVIAGTLSNLMAATIAGFFLSF424	
hCNT1_ORKSDFSOIVLRALFTGACVSLVNACMAGILYMPRGAEVDCMSLLNTTLSSSSFEI 616	
hCNT2 HRKSDLSKVVVRALFTGACVSLISACMAGILYVPRGAEADCVSFPNTSFTNRTYET 610	
hCNT3_SRKRDIASGAVRALIAGTVACFMTACIAGILSST-PVDINCHHVLENAFNSTFPGNTTKV 640	
NupC	
vcCNT	
hCNT1 YOCCREAFOSVNPEFSPEALDNCCRFYNHTICAO 650	
hCNT2 YMCCRGLFOSTSLNGTNPPSFSGPWEDKEFSAMALTNCCGFYNNTVCA 658	
hCNT3 IACCOSLLSSTVAKGPGEVIPGGNHSLYSLKGCCTLLNPSTFNCNGISNTF 691	
Fig S6: Protein sequence alignment between concentrative nucleoside transporters. Amino acid sequences of NupC from Escheric	ch

Fig S6: Protein sequence alignment between concentrative nucleoside transporters. Amino acid sequences of NupC from *Escherichia coli* (P0AFF2), vcCNT from *Vibrio cholera* (Q9KPL5) and human concentrative nucleoside transporters hCNT1 (O00337), hCNT2 (O43868) and hCNT3 (Q9HAS3) were taken from the UniProt KnowledgeBase (http://www.uniprot.org/) and aligned using the online multiple sequence alignment tool Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/).¹ Residues are coloured to indicate those that are identical (*red*) and highly similar (*blue*). vcCNT residues are highlighted to show those that contribute to transmembrane spanning helices (TM1-TM8, *green*), interfacial helices (IH1-IH3, *cyan*) and hairpin re-entrant loops (HP1 and HP2, *yellow*) based on crystal structures of vcCNT.^{6,7}



Fig S7: Membrane topology predictions for NupC. The amino acid sequence of NupC from *Escherichia coli* (P0AFF2) was taken from the UniProt KnowledgeBase (http://www.uniprot.org/) and analysed by the membrane topology prediction tools TOPCONS (http://topcons.cbr.su.se/pred/)⁸ (A), TMHMM (http://www.cbs.dtu.dk/services/TMHMM/)⁹ (B) and CCTOP (http://cctop.enzim.ttk.mta.hu/)¹⁰ (C)



Fig S8: Predictions of transmembrane helices in NupC from *Escherichia coli* and in vcCNT from *Vibrio cholerae*. Amino acid sequences of NupC from *Escherichia coli* (P0AFF2) (*left*) and vcCNT from *Vibrio cholera* (Q9KPL5) (*right*) were taken from the UniProt KnowledgeBase (http://www.uniprot.org/) and analysed by the membrane topology prediction tools TOPCONS (http://topcons.cbr.su.se/pred/),⁸ TMHMM (http://www.cbs.dtu.dk/services/TMHMM/)⁹ and CCTOP (http://cctop.enzim.ttk.mta.hu/).¹⁰ Coloured highlighting is used to show the positions of predicted full transmembrane spanning helices (*grey*), the positions of transmembrane helices based on the crystal structure of vcCNT (Figure 3) (*green*) and the positions of positively charged residues at physiological pH (arginine and lysine) (*red*)



Fig S9: Membrane topology predictions for vcCNT. The amino acid sequence of vcCNT from Vibrio cholera (Q9KPL5) was taken from the UniProt KnowledgeBase (http://www.uniprot.org/) and analysed by the membrane topology prediction tools TOPCONS (http://topcons.cbr.su.se/pred/)⁸ (A), TMHMM (http://www.cbs.dtu.dk/services/TMHMM/)⁹ (B) and CCTOP (http://cctop.enzim.ttk.mta.hu/)¹⁰ (C)



Fig S10: Amplified expression of wild-type NupC and mutants in inner membrane preparations. Coomassie blue-stained SDS-PAGE separation of proteins (30 μg) in inner membrane preparations from *E. coli* BL21(DE3) cells harbouring plasmid pGJL16 for expressing WT NupC or constructs for expressing mutants S142C, G146C and E149C following growth in minimal medium and induction with IPTG. The mobilities of molecular weight (MW) marker proteins are shown and the arrow indicates the position of NupC



Fig S11: Initial rates of NupC-mediated [¹⁴C]uridine uptake into energised *E. coli* cells. Uptake of [¹⁴C]uridine (5 μ M and 100 μ M) after time points of 5, 10, 15 and 20 seconds into transport energised *E. coli* BL21(DE3) cells harbouring plasmid pGJL16 for expressing WT NupC from a culture grown in minimal medium and induced with IPTG. Data points represent the average of duplicate measurements



Fig S12: NupC-mediated uptake of [¹⁴C] adenosine and [¹⁴C] uridine into energised *E. coli* cells. Uptake of [¹⁴C]adenosine (50 μM) and [¹⁴C]uridine (50 μM) after 15 seconds and 2 minutes into transport energised *E. coli* BL21(DE3) cells harbouring empty plasmid pTTQ18 (no transporter gene insert) induced with IPTG (0.05 mM) (A) or plasmid pGJL16 for expressing WT NupC without induction (B) or induced with IPTG (C) from cultures grown in minimal medium. Data points represent the average of duplicate measurements