Molecular basis for substrate recognition by the bacterial nucleoside transporter NupG

Chen Wang1,‡, Qingjie Xiao1,‡, Huaichuan Duan2,‡, Jinhong Li1, Jiying Zhang1, Qisheng Wang3, Li Guo3, Jianping Hu1,*, Bo Sun1,*, and Dong Deng1,*,‡

From the 1Key Laboratory of Birth Defects and Related Disease of Women and Children of MOE, State Key Laboratory of Biotherapy, Department of Obstetrics, West China Second Hospital, Sichuan University, Chengdu, China; 2School of Pharmacy, Sichuan Industrial Institute of Antibiotics, Chengdu University, Chengdu, China; 3Shanghai Synchrotron Radiation Facility, Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai, China

Edited by Mike Shipston

Nucleoside homeostasis, which is mediated by transporters and channels, is essential for all life on Earth. In *Escherichia coli*, NupG mediates the transport of nucleosides and was deemed to be the prototype of the nucleoside proton symporter (NHS) family and the major facilitator superfamily. To date, the substrate recognition and transport mechanisms of NHS transporters are still elusive. Here, we report two crystal structures of NupG (WT and D323A NupG) resolved at 3.0 Å. Both structures reveal an identical inward-open conformation. Together with molecular docking and molecular dynamics simulations and *in vitro* uridine-binding assays, we found that the uridine binding site, which locates in the central cavity between N and C domains of NupG, is constituted by R136, T140, F143, Q225, N228, Q261, E264, Y318, and F322. Moreover, we found that D323 is very important for substrate binding via *in vitro* uridine-binding assays using D323 mutations, although it does not have a direct contact with uridine. Our structural and biochemical data therefore provide an important framework for the mechanistic understanding of nucleoside transporters of the NHS family.

Nucleosides play important roles in all organisms on Earth. In addition to being the basic components of nucleic acids, they are also neurotransmitters for signaling translocation (1, 2). In addition, nucleoside analogs have been suggested to be antiviral and anticancer drugs and compounds with great antibiotic activity (3).

Nucleoside homeostasis, which is mediated by nucleoside uptake and export *via* transporters, is essential for cell division and growth (4). To date, two kinds of nucleoside transporters have been found in mammals: concentrative nucleoside transporters (CNTs) and equilibrative nucleoside transporters (5). Interestingly, the nucleoside-specific outer membrane transporter Tsx and the nucleoside proton symporter (NHS) NupG are restricted in bacteria (6–8). The NHS NupG belongs to the NHS family in the major facilitate superfamily (MFS) and has been recognized as one of the prototype MFS transporters (9).

The major facilitator superfamily (MFS) is the largest protein family of transporters, which facilitate the uptake and export of various molecules across the plasma membrane (10). According to their various coupling patterns, MFS transporters can be divided into three kinds of transporters: facilitators, symporters, and antiporters (11). In recent decades, breakthroughs in structural investigations have heavily promoted the elucidation of alternating access and substrate recognition of MFS transporters; these structural studies include the series of works of the classical lactose proton symporter LacY (12–14) and investigations of glucose transporters with different conformations (15–22).

In *Escherichia coli*, nucleosides are mainly delivered via NupC and NupG (23). NupC, a homolog of CNTs, was found to transport pyrimidine nucleosides, adenosine, and the antibiotic showdomycin but not guanosine or inosine (24). NupG has the capacity to transport both pyrimidine and purine nucleosides efficiently but not showdomycin (24, 25). Recently, structural investigations of CNTs have helped elucidate the transport and specific nucleoside recognition mechanisms of CNTs (26–29).

Here, we report two crystal structures of the NHS transporter NupG (both at a resolution of 3.0 Å). Together with the *in vitro* binding assay, molecular docking, and molecular dynamics (MD) simulations, we identified the substrate-binding site of NupG. Moreover, we found that a negatively charged residue (D323), which indirectly coordinates with uridine, is also very important for substrate binding. This study provides an important framework for the mechanistic understanding of nucleoside recognition of NHS nucleoside transporters.

**Results**

**Characterization of recombinant NupG**

We overexpressed and purified recombinant NupG with monodisperse peaks in different detergent micelles (Fig. S1). To validate the function of NupG, isothermal titration calorimetry (ITC) was carried out to measure the binding between NupG and nucleosides (Fig. 1). As shown in the results, NupG
bonds to adenosine with a $K_d$ of 99.67 ± 14.57 μM. NupG also binds to guanosine, thymidine, cytidine, and uridine with $K_d$ values of 46.67 ± 6.66 μM, 162.5 ± 19.16 μM, 143.67 ± 19.66 μM, and 199.67 ± 15.01 μM, respectively. Nevertheless, NupG could not bind to xanthosine (Fig. 1, Table S1), and this was also observed in a previous biochemical study (30).

We further detected the binding affinity between NupG and uridine in different buffers at various pH values. NupG binds uridine with a binding affinity of approximately 200 μM in sodium citrate at pH 5.0 and 2-(N-Morpholino)ethanesulfonic acid (MES) at pH 6.0, whereas the binding affinity dramatically reduced to approximately 3 mM in a buffer containing 25-mM Tris at pH 8.0 (Fig. S2, Table S1).

A previous study reported that NupG has the capacity to transport various nucleosides (25). Our results indicate that NupG has broad substrate selectivity and are consistent with results from a reported transport assay.

**Overall structure of NupG**

We used X-ray crystallography to explore the substrate recognition mechanism of NupG. Extensive crystallization trials of NupG were carried out. Finally, we crystallized WT NupG and solved the structure of NupG in the apo state at a resolution of 3.0 Å (Fig. 2, Table S2). We used the molecular replacement method using the predicted structure as the search model to determine the structure of NupG (Fig. S3C).
Briefly, we crystallized NupG in two different space groups and collected diffraction data sets at 3.8 Å for P2₁ and 3.0 Å for P1. We succeeded in solving the structure of NupG with the P2₁ space group using molecular replacement as there was only one molecule in one asymmetric unit. In the P1 crystal, however, there are two molecules of NupG with reversed orientations in one asymmetric unit (Fig. S3A), and both molecules have an almost identical conformation with an RMSD of 0.21 Å over 387 Ca atoms (Fig. S3B). Therefore, we selected molecule A for further analyses.

Consistent with the topology prediction (25, 31), NupG has the typical MFS fold with 12 transmembrane helices (TM1-12). The N domain (TM1-6) and C domain (TM7-12) are linked by a flexible loop (Fig. 2A). The cavity between the N and C domains faces the cytoplasm (Fig. 2B). By comparing the NupG structure with well-characterized MFS transporters that have an inward-open conformation, XylE (17) and LacY (12), we found that NupG also has an inward-open conformation in the crystal structure (Fig. 2B).

Identification of potential residues involved in substrate binding

After extensive trials, we still failed to obtain a crystal of NupG in complex with a nucleoside. Fortunately, polar substrate-binding pockets were identified in previous studies of the MFS transporters XylE and LacY (12, 13, 17). The ribose moieties of nucleosides are polar carbohydrates that are structurally similar to the natural substrates of XylE (xylose) and LacY (lactose). We speculated that all three transporters partly share a similar substrate recognition pattern. Several polar or charged residues in the cavity should directly contact the polar carbohydrate. Therefore, the polar and charged residues in the cavity may also be important for nucleoside recognition by NupG.

Figure 2. Overall structure of NupG. A, the ribbon representation (left) and surface representation (right) of NupG. B, superposition of LacY (blue), XylE (gray), and NupG (green).

Crystal structures of nucleoside transporter NupG

To identify the residues involved in substrate binding, we first investigated several polar residues (Q225, N228, E264, and D323) in the cavity of NupG (Fig. 3A). Mutants were generated, and the uridine binding of these constructs was detected via ITC. As shown in the results, Q225A had a similar binding affinity for uridine ($K_d = 227.67 \pm 88.34$ μM) as WT NupG, whereas N228A and E264A dramatically reduced the uridine-binding ability, suggesting that N228 and E264 are essential for nucleoside binding. Surprisingly, NupGD323A bound to uridine with a $K_d$ of 9.67 ± 2.87 μM, which is a 20-fold increase in the binding affinity compared with that of NupGWT (Fig. 3B, Table S1). Considering the increased uridine-binding capacity of NupGD323A, we crystallized NupGD323A under nucleoside-containing conditions and solved the structure of NupGD323A at a resolution of 3.0 Å (Table S2). The superposition of the NupGD323A and NupGWT structures showed that they had identical conformations, with an RMSD of 0.3 Å over 384 Ca atoms (Fig. S3D). Unfortunately, no visible electron density of uridine was observed in the central cavity (Fig. S3E).

Putative substrate-binding site

We continued to explore the nucleoside-binding site of NupG via molecular docking (Fig. 4). Considering the higher binding affinity between NupGD323A and uridine, we first generated a uridine docking model based on the NupGD323A structure using Discovery Studio 3.5 software (Fig. 4A). Then, we confirmed the stability of the system by analyzing a series of conventional structural convergence parameters based on the MD simulation trajectory (Fig. 4B). As shown in Figure 4B, the potential energy reaches equilibrium soon after removing the constraints, and its mean value and SD are $-8.93 \times 10^5$ and $0.07 \times 10^5$ kcal mol$^{-1}$, respectively; the uridine-bound system tends to be stable after 10 ns with an RMSD of 0.19 ± 0.08 nm; according to the change in the gyration radius over the simulation time, the system tends to equilibrate after 30 ns, with a small fluctuation in the amplitude (approximately 0.01 nm), which indicates that the system experienced a slight expansion and contraction during this period; regarding the distribution of the RMSF on different amino acid residue Cα atoms, the overall RMSF value is low and the atomic motion range is not large. All the above simulation results suggest that uridine-bound NupG is relatively stable.

In the uridine-bound model, the uridine molecule was docked into the central cavity, which has direct contact with the side chains of residues from the N and C domains (Fig. 4A). Nucleosides are comprised of two basic components, namely, the ribose and base. In the docking model, R136 and T140 from the N domain and E264 from the C domain form hydrogen bonds with the ribose moiety of uridine. Interestingly, these three residues are identical in three members of the NHS family (NupG, XapB, and YegT). These three transporters have different nucleoside selectivity, but this is understandable for the recognition of invariable ribose by identical residues in the cavity of NHS transporters. The further ConSurf evolutionary analysis using 150 homologs of NupG also reveals that R136, T140, and E264 are highly...
conserved (Fig. S6). Furthermore, Q225, N228, Q261, and Y318 from the C domain form hydrogen bonds with the base of uridine. Unlike the contact with ribose, the residues involving base recognition are relatively variable. With the exception of Q261, the Q225, N228, and Y318 residues of NupG are A226, Y310, and H315 in YegT (Fig. S5). The conservation scores of Q225 and N228 are 6 and 7, respectively, as shown by the ConSurf evolutionary analysis (Fig. S6). This observation indicates that these residues determine the specificities of nucleoside recognition. In addition, there are two conserved aromatic residues around uridine (Fig. 4A, Fig. S6).

To further verify the docking results, in vitro binding assays were carried out. We constructed a series of single point mutations of NupG: R136A, T140A, F143A, Q225A, N228A, Q261A, E264A, Y318A, and F322A. As the ITC results showed, all mutant versions of NupG, except NupGQ225A, abolished the binding affinity between NupG and uridine (Fig. 5, Table S1). This result indicated that the hydrogen-bound interactions between these residues and uridine are essential for substrate recognition. Interestingly, F143A and F322A also eliminated the binding affinity between NupG and uridine, indicating that F143 and F322 are also essential for substrate binding. The surrounding aromatic residues involved in substrate recognition of nucleoside transporters, such as F366 of vcCNT (27) and F564 of hCNT3 (29), are a common feature. F143 and F322 might form π–π interactions with uridine and restrict the specific orientation of the nucleosides.

**Discussion**

In a previous study, N114 was predicted to be important in the recognition of nucleoside transport by NupG (31, 32). However, in our study, N114 is far away from uridine. In addition, N114A also did not influence the binding affinity between NupG and uridine (Fig. S4B). In the structure of NupG, N114 is in TM4 and corresponds to E126 in LacY, which is indispensable for substrate binding and might form a charge pair with R144 of LacY (12, 33).

Interestingly, there are three NHS transporters, NupG (8), XapB (34), and YegT (9), that are found in E. coli. XapB shares 58% sequence identity and 76% sequence similarity with NupG (Fig. S5) and was identified as a xanthosine-specific transporter (30); however, the natural substrate of YegT should be further investigated, although YegT shares 27% sequence identity and 50% sequence similarity with NupG (Fig. S5). The residues
involved in the uridine binding pocket of NupG are identical to those of XapB. A previous biochemical study suggested that the preferred substrate of XapB is xanthosine, whereas NupG could not bind or transport xanthosine (Fig. 1). Therefore, other residues surrounding the pocket also contribute to substrate recognition.

Figure 4. Putative substrate-binding site of NupG. A, the docking model of the NupG-D323A–uridine complex. Left, the overall structure is represented as ribbons, and uridine is represented as spheres. Right, detailed interaction between putative substrate binding sites (yellow sticks) and uridine (white stick). B, MD simulations of the docking model. MD, molecular dynamics.
During this study, we found that the D323A mutant enhances the binding affinity between NupG and uridine (Fig. 2B). This result indicated that D323 is an important residue for nucleoside transport, although D323 does not directly contact the substrate-binding site. Further investigations showed that the D323A mutant maintained a similar uridine-binding affinity in different buffers with various pH values, while NupGWT decreased its binding affinity at a basic pH (Fig. S2) (Fig. S4). This property is similar to that of the mutated lactose proton symporter LacY E325A (35, 36). In the case of LacY, E325 and K319 were identified as protonation sites, and the uncharged side chains of E325 and K319 retain a high affinity with lactose, even at an alkaline pH (35, 36). Therefore, we speculated that D323 may be the proton-escaping site of NupG during the proton-coupling procedure. Interestingly, NupG_{D323N}, which mimics the protonation state at position 323, bound to uridine with a reducing K_d of 1109 ± 192.21 μM (Fig. S4B, Table S1). Therefore, more structural and biochemical investigations should be carried out to elucidate how D323 couples substrate binding and protonation.

In this study, we solved the first structure of an NHS transporter NupG. Together with the MD simulations and biochemical assays, we identified the nucleoside binding site of NupG. Our study provides a framework for understanding the transport and substrate recognition mechanisms of NHS transporters.

**Experimental procedures**

**Protein purification**

Full-length nupG was cloned from *E. coli* K12 and subcloned into pET21b. Overexpression of NupG was induced in *E. coli* C43(DE3) cells by 0.2-mM IPTG when the cell density reached an absorbance at 600 nm of 1.0. After incubation for 15 h at 18°C, the cells were harvested and homogenized in the buffer containing 25-mM MES (pH 6.0) and 150-mM NaCl. The membranes were collected and incubated for 2 h with 2% (w/v) dodecyl-β-D-maltopyranoside (DDM, Anatrace) at 4°C. After centrifugation at 17,000 rpm for 30 min, the supernatant was collected and loaded onto Ni-NTA affinity resin (Qiagen) and rinsed with the buffer containing 25-mM MES (pH 6.0), 150-mM NaCl, 20-mM imidazole, and 0.02% DDM. The protein was eluted with a buffer containing 25-mM MES (pH 6.0), 150-mM NaCl, 250-mM imidazole and 0.4% NG. The protein was concentrated to approximately 20 mg/ml before further purification by gel filtration (Superdex200 10/300 increase column; GE Healthcare) in the buffer containing...
Crystal structures of nucleoside transporter NupG

The atomic coordinates have been deposited in the Protein Data Bank (PDB accession code: 7DL9 for NupGWT and 7DLA for NupGD323A). All other data are contained within the article.

Supporting information—This article contains supporting information.

Acknowledgments—We thank the staff at Shanghai Synchrotron Radiation Facility (SSRF) beamline BL18U1. This work was supported by fund from the National Natural Science Foundation of China (31971132).

Author contributions—D. D. conceived the project. J. H., B. S., and D. D. designed all of the experiments. C. W., J. L., and J. Z. expressed, purified, and crystallized NupG. Q. X., Q.W., and B. S. expressed, purified, and crystallized NupG.
Crystal structures of nucleoside transporter NupG

collected and processed the data and solved the structure. H. D. and J. H. performed molecular docking and MD simulation. L. G., J. H., B. S., and D. D. wrote the manuscript. All authors discussed the results and commented on the manuscript.

Conflict of interest—The authors declare that no competing interests exist.

Abbreviations—The abbreviations used are: CNT, concentrative nucleoside transporter; DDM, dodecyl-β-D-maltopyrano-sulfonic acid; ITC, isothermal titration calorimetry; MES, 2-(N-Morpholino) ethane-sulfonic acid.

References


