Structural insights into the mechanism of the sodium/iodide symporter (NIS)

Silvia Ravera¹, Juan Pablo Nicola², Glicella Salazar de Simone³, Fred J. Sigworth⁴, Erkan Karakas¹, L. Mario Amzel^{5#}, Mario A. Bianchet⁶, and Nancy Carrasco^{1&*}

- ¹ Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN 37232, USA,
- ² Department of Clinical Biochemistry, National University of Córdoba, Córdoba, Argentina,
- ³ Department of Obstetrics and Gynecology, Columbia University Medical Center, New York, NY 10032, USA,
- ⁴ Department of Cellular and Molecular Physiology, Yale School of Medicine, New Haven, CT 06510, USA,
- ⁵ Department of Biophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA, and
- ⁶ Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA.

[#] L.M.A. passed away on August 28, 2021.

^{*}Correspondence and requests for materials should be addressed to N.C. (email: nancy.carrasco@vanderbilt.edu)

This file contains Supplementary Figures 1-11 and Supplementary Table 1.



Supplementary Fig. 1 | Electrophoretic pattern of purified NIS. Lanes 1 and 3: molecular weight markers; lane 2: NIS after affinity purification; lane 4: NIS after size exclusion chromatography. The gel was stained with Coomassie Blue. Dotted rectangle is shown in Extended Data Fig.1.



Supplementary Fig. 2 Alignment of NIS sequences from different species. α -helices are represented by cylinders. Glycosylation sites are marked with a *; critical residues with inverted triangles; non-WT residues found in IDD patients with diamonds; and the deletion Δ 439-443 found in IDD patients with a crescent moon. The alignment was done using Clustaw Omega and displayed using Jalview.



Supplementary Fig. 3 |Alignment of the inverted structural repeats. NIS has a LeuT fold with two 5-helix bundle domains (TMSs 2-6 and 7-11) related by a pseudo-two-fold symmetry. The alignment between the two repeats shows a root mean square deviation (RMSD) of 3.0 Å.



Supplementary Fig. 4 | Structural alignment between NIS in grey, vSGLT (pdb accession code 3DH4) in green (a), hSGLT1 (pdb accession code 7SLA) in wheat (b), and hSGLT2 (pdb accession code 7VSI) in olive (c). *the stars in g, h, and i indicate the extra TMSs present in vSGLT, hSGLT1, and hSGLT2.









Supplementary Fig. 5 NIS mutant proteins are properly targeted to the plasma membrane. HEK cells transfected with an empty plasmid (i.e., not containing the cDNA of interest) and HEK cells transfected with WT NIS or the indicated NIS mutants were incubated under nonpermeabilized conditions with an anti-HA antibody that recognizes the extracellular N-terminus HA epitope and analyzed by flow cytometry. The x axes show the intensity of the fluorescence of each single cell; the y axes the values of the side scatter parameters. For each experiment, cells transfected with an empty plasmid were used as a reference to identify the negative cells and determine the percentage of cells expressing WT NIS or the respective mutants.



Supplementary Fig. 6 Effect of single amino acid substitutions at position 67 on I⁻ and ReO₄⁻ transport. a. NISmediated I⁻ uptake at steady state. cDNA constructs coding for NIS mutants in which F67 is replaced with the residues indicated were transfected into COS7 or HEK cells. I⁻ uptake by these NIS mutants was measured at 20 µM (white bars) and 200 μM (gray bars) I⁻ at 140 mM Na⁺ for 30 min with or without the NIS-specific inhibitor ClO₄⁻ (values obtained in the presence of ClO₄-, which are < 10% of the values obtained in its absence, have already been subtracted). Results are given as pmols of I- accumulated/µg DNA ± s.e.m. Values represent averages of the results from two or three different experiments, each of which was carried out in triplicate (n = 6-9). b and d. Kinetic analysis of initial rates of I⁻ uptake (2-min time points) determined at 140 mM Na⁺ and varying concentrations of I⁻ (b), and at varying concentrations of Na⁺ and 500 μ M I⁻ (d). c and e. I⁻ K_M and Na⁺ K_M values determined from b and d, respectively. The error bars represent the standard deviation of the Michaelis Menten (b), and Hill equation (d) analysis. f. NIS-mediated ReO₄-uptake at steady state. ReO₄-uptake by these NIS mutants was measured at 3 µM (white bars) and 30 μ M (gray bars) ReO₄- at 140 mM Na⁺ for 30 min with or without the NIS-specific inhibitor ClO₄-(values obtained in the presence of CIO_4^- , which are < 10% of the values obtained in its absence, have already been subtracted). Results are given as pmols of I accumulated/µg DNA ± s.e.m. Values represent averages of the results from two or three different experiments, each of which was carried out in triplicate. g and i. Kinetic analysis of initial rates of ReO₄- uptake (2-min time points) determined at 140 mM Na⁺ and varying concentrations of ReO₄- (g), and at varying concentrations of Na⁺ and 100 μ M ReO₄⁻ (i). **h** and **j**. ReO₄⁻ and Na⁺ K_Ms values determined from g and i, respectively. The error bars represent the standard deviation of the Michaelis Menten (g), and Hill equation (i) analysis.





Supplementary Fig. 7 MD simulations. a. RMSD plot of the accumulated trajectories from 13 different MD simulations of various lengths, showing the RMSDs between C- α atoms in the experimental and simulated structures. In each of the thirteen runs, the structure was equilibrated using a different distribution of initial velocities at a temperature of 310.15 K. The red dashed line indicates the approximate starting value of the equilibrate structure used in each leg of the production stage. The time-series plot of the RMSDs clearly shows that they plateaued, indicating that there is no systematic change in the dynamic structure (left panel). Histogram of the RMSD distribution (right panel). **b**. Cluster of tunnels identified in 1059 snapshots from the MD simulations carried out with NIS-I⁻. A pathway analysis performed on 1059 snapshots from the MD simulations with the NIS-I⁻ structure sampled every 5 ns revealed a cluster of tunnels similar to the one in the apo-NIS structure in 95 ± 0.03% of the snapshots. The analysis was carried out as follows: the snapshots were divided into 10 groups, the tunnels in each group were counted, and the average of the counts computed with its standard deviation. **c-e.** section of the tunnel cluster taken respectively at 3.5 Å (a), 5 Å (b), and 10 Å (c) from the position of the I⁻ binding in the structure. We used CAVER Analyst 2.0 (BETA 2) with the default parameters, with the exception of a probe radius of 1.2 (Å), a clustering threshold of 5.3 (Å) , and a shell depth of 5 (Å); the ions were excluded.



Supplementary Fig. 8 F67 conformations open up the ion pocket toward the exit pathway. Ramachandran plot of the chi1 and chi2 side chain dihedral angles of F67 visited during the MD simulations with NIS-I⁻. The dihedral angles selected are the principal determinants of the position of the side chain. The excursions of these dihedral angles (during the MD simulations) away from the conformational basins corresponding to the cryoEM structure (green dot in basin 2) and toward conformational basins (blue dot in basins 1) open up the exit path. In these histograms, the frequency of a given conformational state is indicated by a rainbow gradient from deep purple (0 frequency) to red (highest frequency).

01.0545		0.0
310345		83
SLC5A1	1 · · · MDSSTWSPKTTAVTRPVETHELIRN · · · · · AADISIIVIYFVVVMAVGLWAMFST · NRGTVGG · · · · FLAGRSMVWPIGASLFASNIGSGHFVGLAGTGAASGIAIGGFE · · WN · · · · ·	104
SLC5A2		101
SI C543		85
81.054.4		10.4
010540		105
510346	1. MSVBVSTSAPLSPTSBTSVDMSTFSIMDTVVFVLLVLSLATBLTHACHOWDRHTVGELLMADRKMGCLPVALSLLATFUSAVALLBVPSETTRFBTUTVFLL	100
SLC5A7		90
SLC5A8	1 · · · · · · · · · · · MDTPRGIGTFV · · · · · · · · · · · · · · · · · · ·	91
SLC5A9		118
\$1.05440		0.0
3103A70	I	90
SLC5A11		104
SLC5A12		87
	Y144	
SI C545		196
010544		200
31C347	105ALVLVVVLOWLFVFITIKAGVVIMPETERKRFBGURIUVTLSLESTLESTIFTKISAUFSGATFINLALGENETLATALTHIBGLAAVITIDILUTV	209
SLC5A2		206
SLC5A3	86 · · · · · · · · · · · · · · · ALLLULLULGWVFIPIYIRSGVYTMPEYLSK FGGHRIQVYFAALSLILYIFTKLSVDLYSGALFIQESLGWNLYVSVILLIGMTALLTVTGGLVAVIYTDTLQAL	190
SLC5A4		209
SLC5A6		208
\$10547	91	103
01.0540		100
51C348	82	194
SLC5A9	119 S GGDRG I HPRSHGRT GVRSQATWLLLALGWVFVPVY I AAGVVTMPQYLKK FGGQR I QVYMSVLSLI LY I FTK I STDIFSGALFI QMALGWNLYLSTGI LLVVT AVYTI AGGLMAVI YTDALQTV	243
SLC5A10	97 · · · · · · · · · · · · · · · · ATYVLLALAWVFVPIYISSEIVTLPEYIQKRYGGQRIRMYLSVLSLLLSVFTKISLDLYAGALFVHICLGWNFYLSTILTLGITALYTIAGGLAAVIYTDALQTL	201
SLC5A11		209
\$1.05412		100
0200002		
SLC5A5		290
81.0544		220
3103A7	210 INCOSETE TO AFREVO	320
SLC5A2	207 VILGGACTEMBYAFHEVGGYSGLFDKYLGAATSETVSEDPAVGNTSSFCYR-PHPDSYHLERHVUGDLFWPALLEGETTVSGWWGSDQVTVQHCLAGKSETHTMAGGTECGYEKLTPMF	326
SLC5A3		311
SLC5A4	210 IMLIGSFILMGFAFNEVG····GYESFTEKYVNATPSVVEGDNLT···ISASCYT·PRADSFHIFRDAVTGDIPWPGIIFGMPITALWWWCTNQVIVORCLCGKDMSHVKAACIMCAYLKLLPMF	326
SI C5A6		301
\$10547		202
310540		202
SLC5A8	185 INVAGE AS VITUAVVNDG	288
SLC5A9	244 IMVGGALVLMFLGFQDVG····WYPGLEQRYRQAIPNVTV······PNTTCHL·PRPDAFHILRDPVSGDIPWPGLIFGLTVLATWCWCTDQVIVQRSLSAKSLSHAKGGSVLGGYLKILPMF	355
SLC5A10	202 IMVVGAVILTIKAFDOIG GYGOLEAAYAOA IPSRTI ANTTCHL. PRTDAMHMFRDPHTGDLPWTGMTFGLTIMATWWCTDOVIVORS LSARDLNHAKAGS I LASYLKMLPMG	313
SLC5A11		321
\$1.05412		294
3203472		204
51C3A3	291 LIVSSAAC. CGI WWA	380
SLC5A1	327 IMMP-GM-ISRILYTEKIACVVPSECEKYCGTKVGCTNIAYPTLVVELMPN-GLRGLMLSVMLASLMSSLTSIFNSASTLFTMDIYA-KVRKRASEKELMIAGRLFI	430
SLC5A2	327 LMVMP - GM- ISRILYP DEVACVVP EVCRVCCTEVCCSNIAYPRLVVKLMPN - GLRGLMLAVMLAALMSSLASIFNSSSTLFTMDIYT - RLRPRAGDRELLLVCRUW	430
SLC5A3		415
\$1.054.4		420
BLOSAG		400
310346	302 dvalcvodeli bevine en seara milande en sos es instantis antis en seara milande en seara milande e	402
SLC5A7	283 VMATPATL-IGATGASTOWNOTAYGLPDPKTTEEADMILPIVLQYLCPVYTSFFGLGAVSAAVMSSADSSILSASSMFARNTYQLSFBQNASDKETVWVMBI.T	384
SLC5A8	289 A I LTCSVF · CGLALY · · · · · · · · · · · · · · · · · · ·	388
SLC5A9	356 F IVMP-GM-ISRALFPDEVGCVDPDVCQRICGARVGCSNIAYPKLVMALMPVGLRGLMIAVIMAALMSSLTSIFNSSSTLFTIDVWQ-RFRKSTEGELMV/GRVFV	459
SLC5A10	314 LIIMP-GM-ISRALFPGAHVYEERHQVSVSRTDDVGCVVPSECLRACGAEVGCSNIAYPKLVMELMPIGLRGLMIAVMLAALMSSLTSIFNSSSTLFTMDIWR-RLRPRSGERELLLVGRLVI	433
\$1.05411	322 IMA/EP, GM, VERTLEP,	425
SLOSA //		204
3103H12		304
	10 Free	
01.0545		507
210343	391 SLITGSACLIVAALSSELGGGVEGGSFTVMGVISGFEEGAFTEGMFEFACNTPGVEAGEGAGEAESEMVAEGATETPFSEETMKVEFSSAARVAESVASGEEGPAEE	5U/
SLC5A1	431 LVLTGTSTAWVPTVQSAQSGQLFDYTQSTTSYLGPPTAAVFLLATFWKRVNEPGAFWGLTLGLLTGTSRMTTEFAYGTGS	513
SLC5A2	431 VF I VVVSVAWLPVVQAAQGGQLFDY I QAVSSYLAPPVSAVFVLALFVPRVNEQGAFWGL I GGLLMGLARLI PEFSFGSGS	513
SLC5A3		498
SI C544		513
SI C546	403 AFGYRLLCLEMAY ISSOM, GPVLOAALS IF BW/GGPLLELECLEMEFPCANPPCAN/GLLAGLMAAFWLELESLIVTSMGSSMPPCPSNGSSSIPTNIT/ATVTTLA	500
SLCSAT		400
3LO34/		400
SLC5A8	389 SVYTGALCISMAALAS LM- GALLQAALSVF GMVGGP LMGLFALGILVPF ANS IGALVGLMAGFAISLWVGIGAU	505
SLC5A9	460 VF LVVISILWIPIIQSSNSGQLF DYIQAVTSYLAPPITALFLLAIFCKRVTEPGAFWGLVFGLGVGLLRMILEFSYPAPA	542
SLC5A10		516
SI C5A11		508
\$1.05412		485
3103472		400
SLC5A5	508 PSS BMDAS PALADSFYAISYLYYGALGTLTTVLCGALISCLTGPTKRSTLAPGLLWWDLARQTASVAPKEVAIL····DDNLVKGPEEL······PTG······NKKPP·······	602
SLC5A1		599
SLC5A2	514 PS ···· ACPAFLCGVH ··· YLYFAIVLFFCSGLUTLTVSLCTAPIPRKHLHRLVFSLEHSKEEREDUDADE00GSSLPV · ONGCPESAMEM ····· NEP ····· OAPAP	604
SLC5A3		611
010544		500
320344		280
SLC5A6		601
SLC5A7		548
SLC5A8		590
SLC5A9		641
SI C5410		595
SLOSA IV		000
3103A11	OUG TO ET VERASITA TET PSMILSTVELITERSKEMUSHEINHIGHUPVUKEUAPPAAPLSEILS. UNGMPEASSSSSVUFEMVUENISKTHSCDMTPKQ	014
SLC5A12	4% TMP. VLSS MPGTAD I WMTSIS MET AV GCLGCIVAGVIISLITGRORGED I OPLLIRPVCNLFCFWSKKYKTLCWCGVOHDSGTEOENUENGENGENG	580
SLC5A5	603 ····································	643
SLC5A1	600 ···································	664
SLC5A2	605 · · · · · · · · · · · · · · · · · · ·	672
\$1.0543		710

S69 Q72

000 002 EBNPVASLEHSEAETPVDAYSNGQAALMGEKERKKETDDOGRYWKFIOWFCOFKSKSL-SKRSLRDLMEEEAVCLQMLEETRQVKVILNIGLFAVCSLGIFMFVYFSL 507 CLKKAVDLFCOLQKOP ... KLUKEEEEALSKKLTDTSER SWRTIVNINAILLLAVVVFIHGYYA-602 NOVLGDSRDKEAMALDDTAYOGSSSTCILQETSL 509 OFMILSSTFTNKEAFLDVDSPCGSGTEDNLQ 501 AFNHIELN 718 659 SLC5A3 SLC5A4 635 580 SLC5A6 SLC5A7
0-44
COMMILES IFINKEAF LUVOS VELSIS IE LUNLO

051
AFN HIELN

642
SWIGK LUWSWFCGLS GTP

646
SWIGK LUWSWFCGLS GTP

647
SWIGK LUWSWFCGLS GTP

648
SWIGK LUWSWFCGLS GTP

649
SWIGK LUWSWFCGLS GTP

640
SWIGK LUWSWFCGLS GTP

641
SWIGK LUWSWFCGLS GTP

645
SWIGK LUWSWFCGLS GTP

646
SWIGK LUWSWFCGLS GTP

647
SWIGK LUWSWFCGLS GTP

648
SWIGK LUWSWFCGLS GTP

649
SWIGK LUWSWFCGLS GTP

6415
SKVVK AL LUWGYFA

645
SKVVK AL LUWCGI GEKG

581
GAESVLQN 610 706 SLC5A8 SLC5A9 SLC5A10 612 SLC5A11 675 SLC5A12 618

Supplementary Fig. 9 Alignment of the sequences of the members of the SLC5 family. The red inverted triangles indicate the residues interact with Na1; the blue inverted triangles, the residues that interacts with Na2; and the green inverted triangles, the residues whose side chains contribute to forming the canonical Na2 site. The alignment was done using Clustaw Omega and displayed using Jalview.



Supplementary Fig. 10 Local densities in the Apo-NIS (top), NIS-I (middle), and NIS-ReO₄ (bottom) maps show that the Q72 side chain has a different conformation in each of the three structures.



Supplementary Fig. 11 Close-up of the NIS region containing the β -hydroxyl residues that form the canonical Na2 site. S353 and T354, highlighted in green—in the NIS-I⁻ (top) and NIS-ReO₄⁻ (bottom) structures.

Supplemental Information Table S1. Composition of, and production times for, the MD simulation system.

Complex	Atoms	Lipids	Number of ions	Ligands	Water	Simulation
		(DOPC)			molecules	time
I⁻/Na⁺	161318	405	87 Na ⁺ and 92 Cl ⁻	2Na⁺ and 1I⁻	32586	5390 ns