

## **SUPPORTING INFORMATION**

### **Mapping of Ion and Substrate Binding Sites in Human Sodium Iodide Symporter (hNIS)**

Hristina R. Zhekova<sup>1</sup>, Toshie Sakuma<sup>2,3</sup>, Ryan Johnson<sup>2</sup>, Susanna C. Concilio<sup>3,4</sup>, Patrycja J. Lech<sup>2</sup>, Igor Zdravkovic<sup>1</sup>, Mirna Damergi<sup>1</sup>, Lukkana Suksanpaisan<sup>2</sup>, Kah-Whye Peng<sup>2,3</sup>, Stephen J. Russell<sup>2,3</sup>, and Sergei Noskov<sup>1\*</sup>

1. Centre for Molecular Simulation, Department of Biological Sciences, University of Calgary, Calgary, AB, Canada
2. Imanis Life Sciences, Rochester, MN, USA
3. Department of Molecular Medicine, Mayo Clinic, Rochester MN, USA
4. Mayo Clinic Graduate School of Biomedical Sciences, Mayo Clinic, Rochester MN, USA

```

LeuT -----REHWATR-IGLILAMAGNAVCLGNFTKFPVQA-AEN-----GGGAFMIPYIIAFLLVGIPMWIEWAMGRYGA--QGHG 75
AdiC -----AHKVGL-IPVTLMVSGAIMCSGVFLPANLASTGG-----IAIYGVVTIIGAL-GLSMVYAKMSFLDPS-----PG 71
Mhp1 SLLNPSNAPTRYAERSVGP-FSLAAIWFAMAIQVAIFLAAGQMT-SSF-----QVWQVIQAIAGCT-IAIIDLFFFTQSAAIRWG-----I 88
BetP -----T-VSWISMMFAGMCIDLMFYGTTEPLTFYRNGVPGHDEHNVGVMSTTMFHWTLH--PWAQYAIVGLAIAYSTFRVGRKQ 215
dDAT -----DERETWSGK-VDFLLSVIGFAVDLANVWRFPYLC-YKN-----GGGAFIPYIGIMLAVGGIPFYMELALGQHNK-----K 93
vSGLT -----AGKSLPWWAVGASLIAMNSAEQFTGMSSGSGYSIG-----LAIASYEWMASAITLIDIGKY--FLPIFIEKGI--Y 112
hNIS -----GGRRLAAIFVGLSLSKSFMSAVQVLEVPSEAYRYG-----LKFVWMLGQLLNSVLTALL--FMPVFYRLGL--T 115

LeuT TTPAIFYLLW-----RNRFAKILGVFGLWIPLV-VAIVYVIEVSWTLGFAIKFLVGLVPEPPP??-----TDPDSIL-----RPFKEFLY-SYIG 153
AdiC GSYAYARRCF-----GPFLGYQTNVLYWLACWI-GNIAMVIVGVYLSYFFP-----IAIYGVVTIIGAL-GLSMVYAKMSFLDPS-----PG 71
Mhp1 NFTVAARMP-----GIRGSLIPITLKALISLF-WFGFQTWLGLALADEITRLLTG-----QVWQVIQAIAGCT-IAIIDLFFFTQSAAIRWG-----I 88
BetP LLSSAFVPLIGEKGAEWGLKLDILAIATVFGTACSIIGLALQIGAGLS-----AAN-----?IEDP 274
dDAT GAITCWGRIV-----PL-FKGIGYAVVLIAPY-VDFYVNVIIWSIRFFASFTNSLPWTSCNNIWNTPNCRPFE-??-GHVEGFQSAASEYFNRYILE 222
vSGLT TIPEFVEKRE-----NKKLKTILAVFWISLYIF-VNLTIVLYLGLALETILG-----LAIASYEWMASAITLIDIGKY--FLPIFIEKGI--Y 112
hNIS STYEYLEMRE-----SRAVRLCGTLQYIIVATM-LYTCVVIYAPALILNQVTG-----LAIASYEWMASAITLIDIGKY--FLPIFIEKGI--Y 112

LeuT VPKGDEP--ILKPSLFAYIVFLITMFINVSIIRGISKGIERFAKIAMPTLFIIDAVFIVIRVFLLETPNGTAADGLNF-----LW--TP 233
AdiC -----DPWVLTITCVVVLWVFLVLLNIVGPK-MITRQAVATVLALIPVIGIAVFGW-----FWFRG-----ET 177
Mhp1 -----FTNLPILWIVIFGALQVVTTFYGIT-FIRWMNVFASPVLLAMGYMVYMLDG-----ADVSLGEV-----MS--MG 201
BetP S-----DWT-IVGIVSVLTLAIFISFSAISGVGKGNANMVLALALAFVFFVVGPT-----VSILNLLPGSIGNYLSNFFQM 347
dDAT LNRSEGIHDLGAIKWDMAICLLIVYLCYFSLWKGIS-TSGKVWFTALFPYAVLILLIRGLTLP-----GSLGIQY-----YL--TP 299
vSGLT -----MYSILGLALFALVYSI?????V-----VWTDVIQVFFLVLGGEFMTTMAVSVFIGGTDGWFAGVSK-----MVDAAPG 230
hNIS -----WASLSTGIICTFYTA-----VGGMKAVVWTDVQVVMVMSGFWVVLARGVMLVG-----GPRQ-----VLTLAQN 225

LeuT DFE-----KLKDP-----GVWIAAVGQIFITL LGFGAITYASYVRKDKQ-----DIVLSGTAATNEKAEVILGGSISIPAAIAFF 306
AdiC YMA?????????-----AIQSTLNVTLWSEFISV-ESASVAAGVVKNPK-----RNVPIATGGVLAACVYVSTTAIMGMIP-----251
Mhp1 GEN-----P-GM-----PFSTAIMIFVGGWIVVVVSIHDIVKECKVDP-----NASREGQTKADARYATAQWLGMPASIFGFIGASMV-----277
BetP -----AGRTAMSADGTAGEWLGSWTIFYWAWWISWSPFGMFLARISRG-----RSIREFLGVLLVPAGVSTWFSIFGTALFEQ 425
dDAT NFS-----AIYKA-----EVWVDAATQVFSLCPGFGVLLAYASNYKYN-----NVYKDALTSFINSATSEI-AGFVIFSVLGYMA 371
vSGLT HFE-----MILD-----QSNPQYMNLPGI-AVLIGGLWVANIY WGFNQYIIRTLAAKS-----VSEAQGIQFAAFKLVPEFVVLPGIAAYIITS 313
hNIS HSR-----INLM-----DFNPPDRSRYTFWTFVVGGLVWLSMY-GVNCQAVQRYVACTR-----EKQAKLALINQVGLFLIVSS-AACCGIVMVFYFT 308

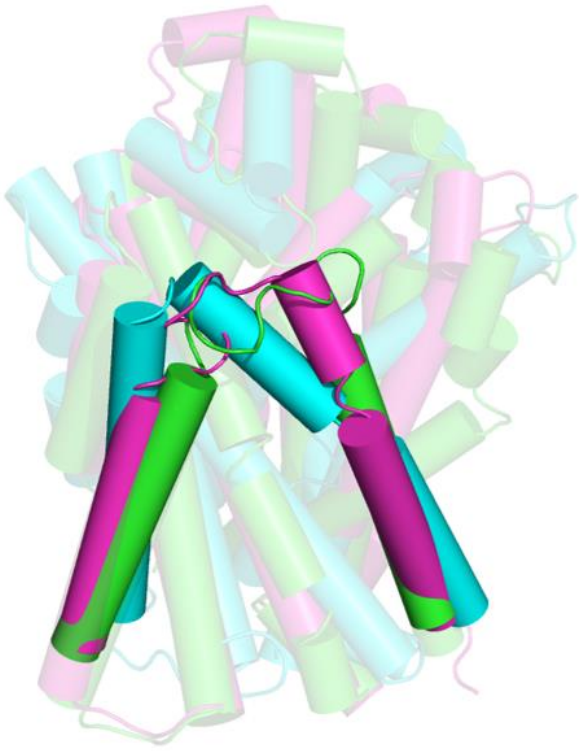
LeuT G--VA-----NAVAIAGKAGAFNLGFITLPAIFSQTAGGTFLGFVWFFLFEAGLTSIAIMQPMIAFL--EDE--LKL--S--RKHAV-379
AdiC ---NA-----ALRVS-----ASFPFGDAARMALGDTAGAVSFCAAAGCLGSLGGWTLLAGQTAKAADDGLFPPIFARVNKAGTPVAGL-327
Mhp1 -----VQ-----EWNPVIAITEVVGVS-IPMAILPQVVFVLLMTWSTNPAANLLSEAYTLCST-----FPRVF-T--FKTGV-340
BetP NGES-----IWC DGAA-----EEQLFGLLHALPGGQIMGITAMILLGTFFITTSADS-ASTVMGTMSQHGQ-----LEANKWVT-492
dDAT HTLGV-----RIEDVAT-EGPGLVVFVYPAAIATMPASTFWALIFFMMLATLGLDSSFG-GSEAIITALS--DE--FPKIKRN--REL FV-448
vSGLT DPQLMASLGDIAATN--LPSAANADKAYPWLTFQ-----LPVGKGVVFAALAAIVVSLASMLNSTATIFT--MDIYKEYISPDSDGDK--LVN V G-399
hNIS DCDPLL-LGRISAPD--QY--MPLLVDIFED-----LP-GVPGFLFLACAYSSTLSAST-SINAMAVTV--EDLIKPRLR-SLAPRK--LVIIS-387

LeuT LWTAIVFVFS AHLVMF--LNKSLDE--MDFWAGTIGVVFFGLTELIFFWIFGAD--- 430
AdiC IIVGIIMTIFQLSSISPNATKEFGL--VSSVSV-IFTLVPYLYTCAALLLGHGHF-- 380
Mhp1 IVS AVVGLLMM-----PWQFAG--VLNTFLNLLASALGPLAGIMISDYFLVRRR 388
BetP AAWGVATAAIGLTLTLLSGGDNALSN--LQNVT-IVAATPFLFVVIGLMFATVKDLSND 547
dDAT AGLFSLYFVVG LASCTQGGFYFFHL--LDRYAAGYSILVAVFFEAIVASWVYGTN-- 501
vSGLT RTAAVVVALIIACLIA---PMLGGI--GQAQYIIEYTGVLVSPGIIAVFLVGLFWK- 450
hNIS KGLSLIYG-SACLTVAAALSSLLGGVQLGSF---TVMGVISGPILLGAFILGMFLPA- 439

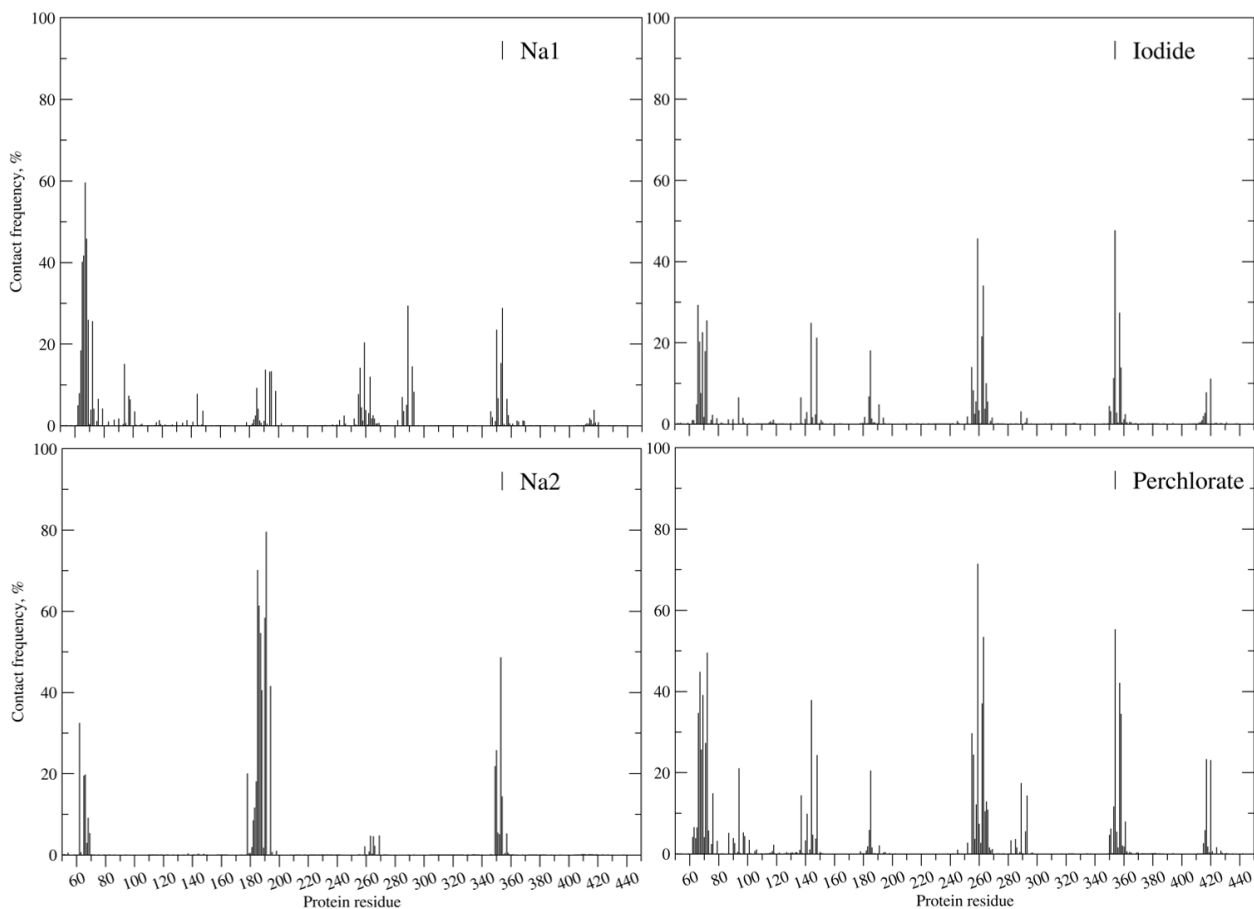
```

**Figure S1** Multiple sequence alignment of the LeuT-folds of NIS, vSGLT (PDB code 3DH4), Mhp1 (PDB code 4D1B), LeuT (PDB code 2A65), BetP (PDB code 4DOJ), dDAT (PDB code 4XPA), AdiC (PDB code 3L1L).

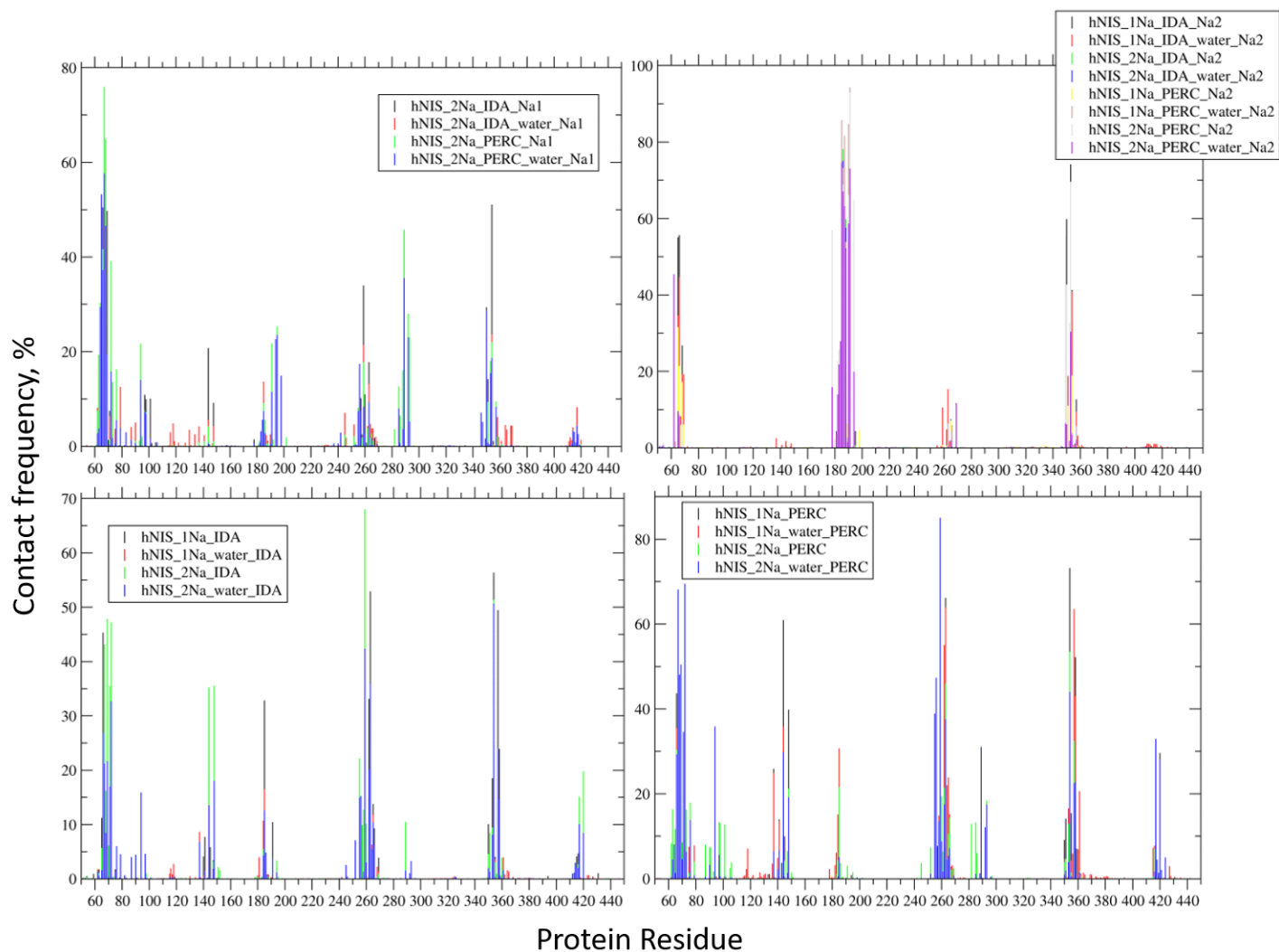
The color coding is as follows: magenta (Na<sup>2</sup> binding site), orange (Na<sup>1</sup> binding site), green (substrate binding site), green highlight (intracellular gate), red highlight (extracellular gate), grey and black highlights (conserved residues), yellow highlights with question marks (missing residues in crystal structure, affects numbering).



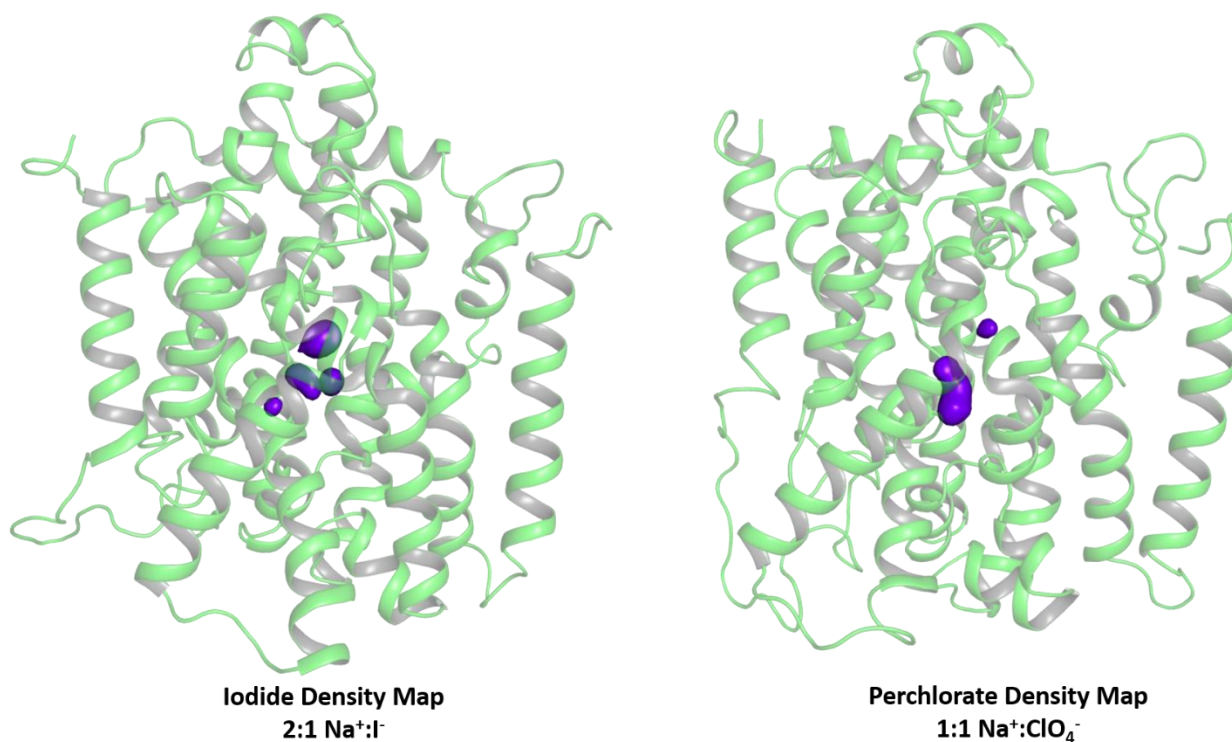
**Figure S2** Overlap of three different LeuT-fold models of NIS: green (based on vSGLT and Mhp1 crystal structures), cyan (based on LeuT, AdiC, dDAT, and BetP structures), and magenta (mixed structure based on all 6 templates). Transmembrane domains 10 and 11 which include the extracellular gate are highlighted.



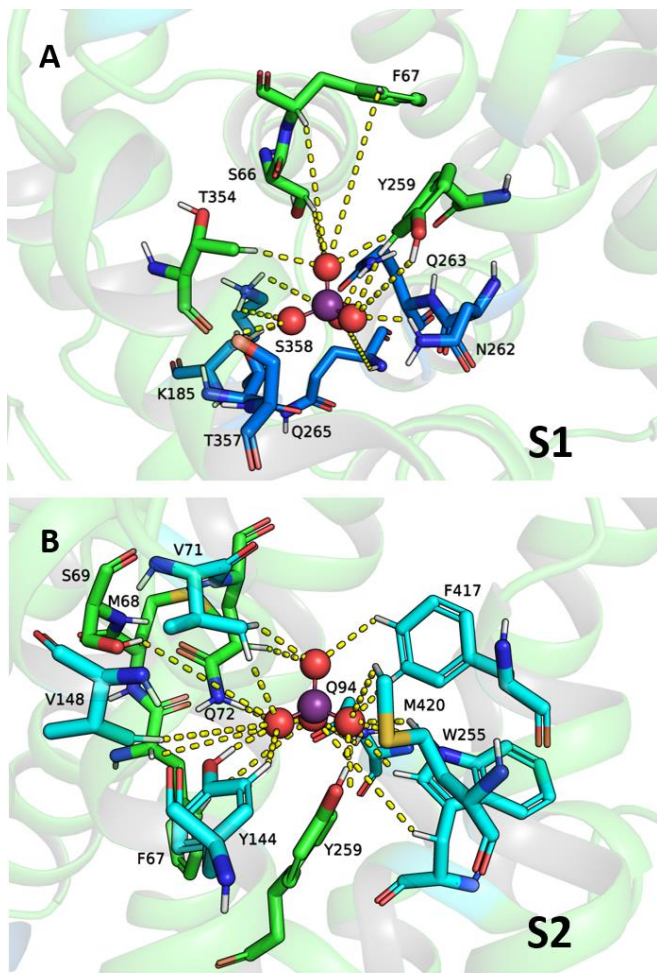
**Figure S3** Cumulative contact frequencies for Na1, Na2, I<sup>-</sup> and ClO<sub>4</sub><sup>-</sup> ions with NIS residues evaluated from MD trajectories within a distance cutoff of 5Å. As example, the Na2 contact frequency map contains information from all MD trajectories since all systems subjected to MD simulations feature a Na2 ion. The Iodide map has contributions from the MD simulations which contained iodide (i.e. systems with and without water present in the binding cavity in the beginning of the MD simulations and with 1:1 Na<sup>+</sup>:I<sup>-</sup> and 2:1 Na<sup>+</sup>:I<sup>-</sup> ion binding stoichiometries). Individual contact frequency maps can be found in Fig.S4.



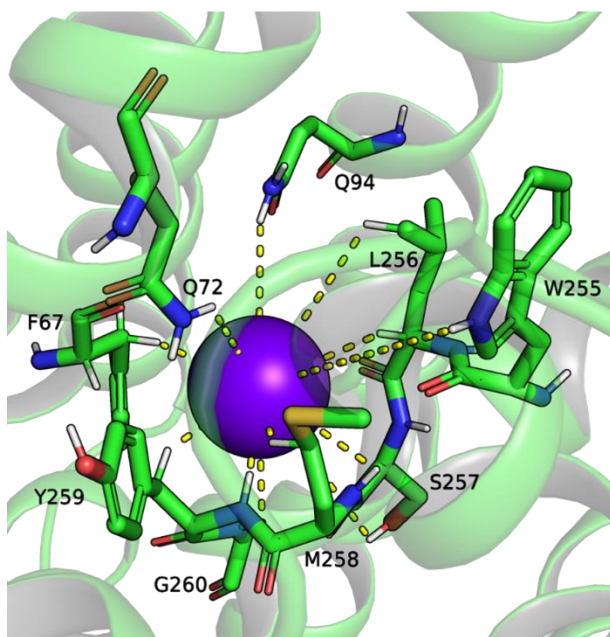
**Figure S4** Contact frequencies for Na1, Na2, I<sup>-</sup> (IDA), and ClO<sub>4</sub><sup>-</sup> (PERC) evaluated from different groups of MD trajectories (from NIS models with and without water in the binding cavity at MD onset, with 1:1 and 2:1 Na<sup>+</sup>:anion binding stoichiometry).



**Figure S5 (Left)** Anion density map evaluated from 13 MD simulations of the I<sup>-</sup> bound NIS, at 2:1 Na<sup>+</sup>:I<sup>-</sup> binding stoichiometry and without water at the MD simulations onset. **(Right)** Anion density map evaluated from 13 MD simulations of the ClO<sub>4</sub><sup>-</sup> bound NIS, at 1:1 Na<sup>+</sup>:ClO<sub>4</sub><sup>-</sup> binding stoichiometry and without water at the MD simulations onset.

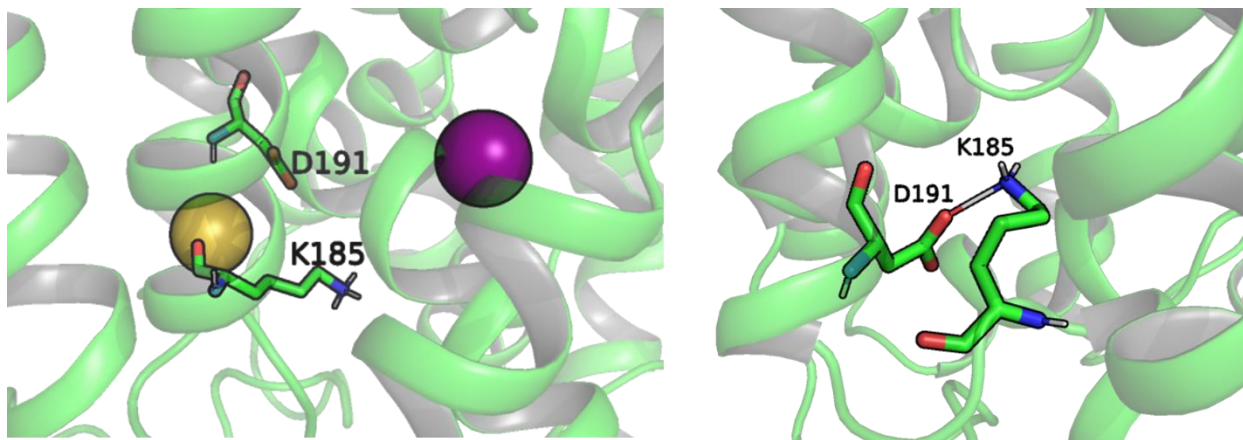


**Figure S6** Representative figures of the S1 and S2 binding patterns of  $\text{ClO}_4^-$  (purple and red spheres) in NIS, consistent with the contact frequency maps in Figures S3 and S4, the  $\text{ClO}_4^-$  density map in Figure S5 and the residues listed in Table S2. The residues are color coded: green for residues involved in both S1 and S2, dark blue for residues unique to S1, and cyan for residues unique to S2.



**Figure S7** Putative allosteric binding site for I<sup>-</sup> (purple sphere) extracted from the I<sup>-</sup> density map in Fig.S5.





**Figure S8** Formation of D191-K185 salt bridge upon Na<sup>+</sup> and I<sup>-</sup> exit from NIS: (left) Na<sup>+</sup> and I<sup>-</sup> are still present; (right) all ions have exited NIS.

**Table S1** In-house developed CHARMM force field parameters for perchlorate used in the reported MD simulations.

Nonbonded parameters		
	Cl	O
q, electron	1.32	-0.58
E <sub>min</sub> , kcal/mol	-0.272	-0.150
R <sub>min</sub> /2, Å	2.347	1.785
Bonded parameters		
	Cl-O	O-Cl-O
Equilibrium value	1.483 Å	109.4 <sup>o</sup>
Force constant	600.0 kcal/(mol x Å <sup>2</sup> )	105.0 kcal/(mol x rad <sup>2</sup> )

**Table S2** Comparison of the residues of the putative binding sites in NIS determined from MD simulations in the present work and the residues of the putative binding sites in vSGLT, hSGLT1, PutP, Mhp1, and LeuT, determined from X-ray crystallography (marked with \*) or from functional mutagenesis (bold text). The analogy between the residues from the different proteins has been made on the basis of the multiple sequence alignment in Fig.S1 and the NIS, vSGLT, hSGLT1, PutP alignment published in Ref.<sup>1</sup>

Site	NIS	vSGLT* <sup>1</sup>	hSGLT1	PutP	Mhp1* <sup>2</sup>	LeuT* <sup>3</sup>
<b>Na2</b>	S62	S59	S73	<b>S50</b> <sup>4</sup>	I35/S31	L16
	A65	A62* <sup>1</sup>	A76	<b>A53</b> <sup>4</sup>	A38* <sup>2</sup>	G20/A19* <sup>3</sup>
	<b>S66</b> <sup>5</sup>	A63	S77	<b>S54</b> <sup>4,6</sup>	M39	N21
	<b>M68</b>	I65* <sup>1</sup>	I79	<b>M56</b> <sup>4</sup>	I41/M39* <sup>2</sup>	V23* <sup>3</sup>
	Y178	Y176	Y191	Y174	Q153/F149	N179
	M184	L182	L197	F180	M167	F194
	<b>K185</b>	S183	A198	L181	N168	A195/K196/K189
	A186	A184	A199	A182	V169/A171	K196/A195/A198
	V187	V185	V200	V183	V169/F170	I197
	V188	V186	I201	S184	A171/V174	A198
	T190	T188	T203	T186	S172	T201
	<b>D191</b> <sup>5</sup>	D189	<b>D204</b> <sup>7-8</sup>	<b>D187</b> <sup>9</sup>	V174/S172/D229	T201/E112/E192
	<b>Q194</b> <sup>5</sup>	Q192	Q207	<b>Q190</b> <sup>9</sup>	A177	I204
	<b>Q263</b> <sup>5</sup>	Q268	Q295	<b>Q251/P252</b> <sup>10-11</sup>	S226/H228	Y265/Y268/S267
	<b>S349</b> <sup>12</sup>	A360	A388	<b>A336</b> <sup>13</sup>	L308	F350
	G350	A361* <sup>1</sup>	S389	<b>A337</b> <sup>13</sup>	A309* <sup>2</sup>	A351* <sup>3</sup>
	<b>S353</b> <sup>12</sup>	S364* <sup>1</sup>	<b>S392</b> <sup>8,14</sup>	<b>S340</b> <sup>13</sup>	S312* <sup>2</sup>	T354* <sup>3</sup>
	<b>T354</b> <sup>12</sup>	<b>S365</b> * <sup>1</sup>	<b>S393</b> <sup>14</sup>	<b>T341</b> <sup>13</sup>	T313* <sup>2</sup>	S355* <sup>3</sup>
<b>Na1</b>	<b>S64</b> <sup>15</sup>	I61	F75	<b>G52</b> <sup>4</sup>	F37	A19
	A65	A62* <sup>1</sup>	A76	<b>A53</b> <sup>4</sup>	A38* <sup>2</sup>	G20/A19* <sup>3</sup>
	<b>S66</b> <sup>5</sup>	A63	S77	<b>S54</b> <sup>4,6</sup>	M39	N21
	F67	<b>N64</b> /F70 <sup>16</sup>	<b>N78</b> /F75 <sup>14</sup>	<b>D55</b> /W59 <sup>4,6,10</sup>	A40/Q42/F46	A22/F28* <sup>3</sup>
	<b>M68</b>	I65* <sup>1</sup>	I79	<b>M56</b> <sup>4</sup>	I41/M39* <sup>2</sup>	V23* <sup>3</sup>
	<b>S69</b>	S66	S81/G80	<b>S57</b> <sup>4</sup>	Q42* <sup>2</sup>	G24
	<b>Q72</b>	<b>Q69</b> * <sup>1</sup>	<b>H83</b> <sup>8</sup>	<b>L60</b> <sup>4</sup>	Q42/I45* <sup>2</sup>	N27* <sup>3</sup>
	<b>Q94</b>	S91/S86/ <b>E88</b> * <sup>1,10</sup>	A105/N104/G100/ <b>E102</b> <sup>14</sup>	L82	C69	F51
	Y144	Y138/N142	Y153/K157	<b>Y140</b> <sup>6</sup>	W117* <sup>2</sup>	Y108/Y107
	<b>K185</b>	S183	A198	L181	N168	A195/K196
	A186	A184	A199	A182	V169/A171	K196/A195/A198
	<b>D191</b> <sup>5</sup>	D189	<b>D204</b> <sup>7-8</sup>	<b>D187</b> <sup>9</sup>	V174/S172/D229	T201/E112/E192
	<b>Q194</b> <sup>5</sup>	Q192	Q207	<b>Q190</b> <sup>9</sup>	A177	I204
	V195	V193	T208	A191	L176	L205
	<b>M198</b>	L196	M211	M194	Y181/M182	F208
	<b>W255</b> <sup>17</sup>	W257/ <b>N260</b> /W264* <sup>1</sup>	W289/ <b>T287</b> <sup>8</sup>	<b>W244</b> <sup>6,10-11</sup>	W220* <sup>2</sup>	F253
	L256	L261	L288	<b>G243</b> <sup>11</sup>	V223	L257
	<b>Y259</b> <sup>17</sup>	<b>Y263</b> /W264* <sup>1,16</sup>	<b>Y290</b> /W291 <sup>8,14</sup>	<b>Y248</b> <sup>6,10-11</sup>	S226	F259/Y265/Y268* <sup>3</sup>
	<b>Q263</b> <sup>5</sup>	Q268	Q295	<b>Q251/P252</b> <sup>10-11</sup>	S226/H228	Y265/Y268/S267
	L289	L293	L320	M278	V261	E287

	I292	I296	M323	C281	I265	E290* <sup>3</sup>
	G350	A361* <sup>1</sup>	S389	<b>A337/A336</b> <sup>13</sup>	A309* <sup>2</sup>	A351* <sup>3</sup>
	<b>T354</b> <sup>12</sup>	<b>S365</b> * <sup>1</sup>	<b>S393</b> <sup>14</sup>	<b>T341</b> <sup>13</sup>	T313* <sup>2</sup>	S355* <sup>3</sup>
	<b>F417</b>	<b>F424/Q428</b> * <sup>1, 10</sup>	<b>F453/Q457</b> <sup>8</sup>	W405	F362	F405/W406/D404* <sup>3</sup>
<b>S1</b>	A65	A62* <sup>1</sup>	A76	<b>A53</b> <sup>4</sup>	A38* <sup>2</sup>	G20/A19
	<b>S66</b> <sup>5</sup>	A63	S77	<b>S54</b> <sup>4, 6</sup>	M39	N21
	F67	<b>N64</b> <sup>16</sup>	<b>N78/F75</b> <sup>14</sup>	<b>D55</b> <sup>4, 6, 10</sup>	A40/Q42/F46	A22/F28* <sup>3</sup>
	<b>M68</b>	I65* <sup>1</sup>	I79	<b>M56</b> <sup>4</sup>	I41/M39* <sup>2</sup>	V23* <sup>3</sup>
	<b>S69</b>	S66	S81/G80	<b>S57</b> <sup>4, 10</sup>	Q42* <sup>2</sup>	G24
	<b>Q72</b>	<b>Q69</b> * <sup>1</sup>	<b>H83</b> <sup>8</sup>	<b>L60</b> <sup>4</sup>	Q42/I45* <sup>2</sup>	N27* <sup>3</sup>
	Y144	Y138/N142	Y153/K157	<b>Y140</b> <sup>6</sup>	W117* <sup>2</sup>	Y108/Y107* <sup>3</sup>
	<b>K185</b>	S183	A198	L181	N168	A195/K196
	<b>Y259</b> <sup>17</sup>	<b>Y263/W264</b> * <sup>1, 16</sup>	<b>Y290/W291</b> <sup>8, 14</sup>	<b>Y248</b> <sup>6, 10-11</sup>	S226	F259/Y265/Y268* <sup>3</sup>
	N262	N267	D294	<b>Q251</b> <sup>11</sup>	D229	T264
	<b>Q263</b> <sup>5</sup>	Q268	Q295	<b>Q251/P252</b> <sup>10-11</sup>	S226/H228	Y265/Y268/S267
	<b>Q265</b>	Y269/I270/Q272	I297/Q299	<b>I254</b> <sup>11</sup>	E233	S267
	<b>S353</b> <sup>12</sup>	S364* <sup>1</sup>	<b>S392</b> <sup>8</sup>	<b>S340</b> <sup>13</sup>	S312* <sup>2</sup>	T354* <sup>3</sup>
	<b>T354</b> <sup>12</sup>	<b>S365</b> * <sup>1</sup>	<b>S393</b> <sup>14</sup>	<b>T341</b> <sup>13</sup>	T313* <sup>2</sup>	S355* <sup>3</sup>
	<b>T357</b> <sup>12</sup>	<b>S368</b> <sup>10</sup>	S396	<b>C344</b> <sup>13</sup>	N318* <sup>2</sup>	S356
	<b>S358</b> <sup>12</sup>	M369	I397	<b>Q345</b> <sup>13</sup>	N318* <sup>2</sup>	S356
<b>S2</b>	<b>S66</b> <sup>5</sup>	A63	S77	<b>S54</b> <sup>4, 6</sup>	M39	N21
	F67	<b>N64</b> <sup>16</sup>	<b>N78/F75</b> <sup>14</sup>	<b>D55</b> <sup>4, 6, 10</sup>	A40/Q42/F46	A22/F28* <sup>3</sup>
	<b>S69</b>	S66	S81/G80	<b>S57</b> <sup>4, 10</sup>	Q42* <sup>2</sup>	G24
	V71	E68	G82	<b>W59</b> <sup>4, 6</sup>	V43/A44	G26* <sup>3</sup>
	<b>Q72</b>	<b>Q69</b> * <sup>1</sup>	<b>H83</b> <sup>8</sup>	<b>L60</b> <sup>4</sup>	Q42/I45* <sup>2</sup>	N27* <sup>3</sup>
	<b>Q94</b>	S91/S86/ <b>E88</b> * <sup>1, 10</sup>	A105/N104/G100/ <b>E102</b> <sup>14</sup>	L82	C69	F51
	Y144	Y138/N142	Y153/K157	<b>Y140</b> <sup>6</sup>	W117* <sup>2</sup>	Y108/Y107* <sup>3</sup>
	V148	V146	D161/A160	G144	Q121* <sup>2</sup>	V109
	<b>W255</b> <sup>17</sup>	W257/ <b>N260/W264</b> * <sup>1</sup>	W289/ <b>T287</b> <sup>8</sup>	<b>W244</b> <sup>6, 10-11</sup>	W220* <sup>2</sup>	F253* <sup>3</sup>
	<b>Y259</b> <sup>17</sup>	<b>Y263/W264</b> * <sup>1, 16</sup>	<b>Y290/W291</b> <sup>8, 14</sup>	<b>Y248</b> <sup>6, 10-11</sup>	S226	F259/Y265/Y268* <sup>3</sup>
	<b>S353</b> <sup>12</sup>	S364* <sup>1</sup>	<b>S392</b> <sup>8</sup>	<b>S340</b> <sup>13</sup>	S312* <sup>2</sup>	T354* <sup>3</sup>
	<b>T354</b> <sup>12</sup>	<b>S365</b> * <sup>1</sup>	<b>S393</b> <sup>14</sup>	<b>T341</b> <sup>13</sup>	T313* <sup>2</sup>	S355* <sup>3</sup>
	<b>F417</b>	<b>F424/Q428</b> * <sup>1, 10</sup>	<b>F453/Q457</b> <sup>8</sup>	W405	F362	F405/W406/D404* <sup>3</sup>
	M420	T431	T459	F408/W405	S368	V413/F414

**Table S3** Duration (in ns) of the 50 ns long MD trajectories during which all ions remain bound to NIS in the 104 simulated systems, constructed as described in the Methodology section. A duration of less than 50 ns signifies that at least one ion has exited the protein during the MD simulation. The first ion to exit is indicated in brackets. In some systems with 2:1 Na<sup>+</sup>:substrate binding stoichiometry, when a Na2 ion exits NIS, it is replaced by the Na1 ion (also indicated in brackets). 13 anion binding sites, determined from GCMC+PB calculations were studied. The color coding of the anion binding sites corresponds to the colors in Figure 5: green for areas of protein involved both in S1 and S2 binding sites (labeled as S1/S2 here), dark blue for site S1, cyan for site S2.

Site	No water in binding cavity at MD onset		Water in binding cavity at MD onset	
	1Na	2Na	1Na	2Na
Iodide				
S2(a)	50	50	50	37.5 (Na1)
S2(b)	50	50	50	50
S2(c)	7.5 (IOD)	50	25.5 (IOD)	29.5 (Na2)
S1/S2(a)	50	50	50	50
S1/S2(b)	50	18.5 (Na2)	16.5 (IOD)	50
S1/S2(c)	50	50	4.5 (IOD-Na pair)	50
S1/S2(d)	50	50	35.5 (IOD)	50
S1(a)	50	15 (Na2, replaced by Na1)	50	50
S1(b)	50	25 (Na1)	12 (IOD)	50
S1(c)	50	44 (Na1)	50	50
S1(d)	50	34 (Na1)	22 (Na2)	33.5 (Na2, replaced by Na1)
S1(e)	10.5 (IOD)	50	20.5 (IOD)	3 (Na1)
S1(f)	23.5 (IOD)	50	18 (IOD)	31.5 (Na1)
Perchlorate				
S2(a)	50	50	50	50
S2(b)	50	50	50	12.5 (Na2)
S2(c)	50	50	50	27 (Na1)
S1/S2(a)	50	50	50	45 (Na1)
S1/S2(b)	50	50	50	21.5 (Na2)
S1/S2(c)	11.5 (Na2)	50	50	50
S1/S2(d)	50	50	22 (Na2)	50
S1(a)	50	42 (Na1)	14.5 (PERC)	50
S1(b)	50	50	50	22 (Na2, replaced by Na1)
S1(c)	50	50	50	50
S1(d)	45 (Na2)	50	50	50
S1(e)	50	50	50	50
S1(f)	50	50	50	11.5 (Na1)

## References

1. Faham, S.; Watanabe, A.; Besserer, G. M.; Cascio, D.; Specht, A.; Hirayama, B. A.; Wright, E. M.; Abramson, J., The crystal structure of a sodium galactose transporter reveals mechanistic insights into Na<sup>+</sup>/sugar symport. *Science* **2008**, *321* (5890), 810-4.
2. Simmons, K. J.; Jackson, S. M.; Brueckner, F.; Patching, S. G.; Beckstein, O.; Ivanova, E.; Geng, T.; Weyand, S.; Drew, D.; Lanigan, J.; Sharples, D. J.; Sansom, M. S.; Iwata, S.; Fishwick, C. W.; Johnson, A. P.; Cameron, A. D.; Henderson, P. J., Molecular mechanism of ligand recognition by membrane transport protein, Mhp1. *EMBO J* **2014**, *33* (16), 1831-44.
3. Yamashita, A.; Singh, S. K.; Kawate, T.; Jin, Y.; Gouaux, E., Crystal structure of a bacterial homologue of Na<sup>+</sup>/Cl<sup>-</sup>-dependent neurotransmitter transporters. *Nature* **2005**, *437* (7056), 215-23.
4. Pirch, T.; Landmeier, S.; Jung, H., Transmembrane domain II of the Na<sup>+</sup>/proline transporter PutP of *Escherichia coli* forms part of a conformationally flexible, cytoplasmic exposed aqueous cavity within the membrane. *J Biol Chem* **2003**, *278* (44), 42942-9.
5. Ferrandino, G.; Nicola, J. P.; Sanchez, Y. E.; Echeverria, I.; Liu, Y.; Amzel, L. M.; Carrasco, N., Na<sup>+</sup> coordination at the Na<sub>2</sub> site of the Na<sup>+</sup>/I<sup>-</sup> symporter. *Proc Natl Acad Sci U S A* **2016**, *113* (37), E5379-88.
6. Olkhova, E.; Raba, M.; Bracher, S.; Hilger, D.; Jung, H., Homology model of the Na<sup>+</sup>/proline transporter PutP of *Escherichia coli* and its functional implications. *J Mol Biol* **2011**, *406* (1), 59-74.
7. Quick, M.; Loo, D. D.; Wright, E. M., Neutralization of a conserved amino acid residue in the human Na<sup>+</sup>/glucose transporter (hSGLT1) generates a glucose-gated H<sup>+</sup> channel. *J Biol Chem* **2001**, *276* (3), 1728-34.
8. Loo, D. D.; Jiang, X.; Gorraitz, E.; Hirayama, B. A.; Wright, E. M., Functional identification and characterization of sodium binding sites in Na symporters. *Proc Natl Acad Sci U S A* **2013**, *110* (47), E4557-66.
9. Amin, A.; Ando, T.; Saijo, S.; Yamato, I., Role of Asp187 and Gln190 in the Na<sup>+</sup>/proline symporter (PutP) of *Escherichia coli*. *J Biochem* **2011**, *150* (4), 395-402.
10. Li, Z.; Lee, A. S.; Bracher, S.; Jung, H.; Paz, A.; Kumar, J. P.; Abramson, J.; Quick, M.; Shi, L., Identification of a second substrate-binding site in solute-sodium symporters. *J Biol Chem* **2015**, *290* (12), 7361.
11. Bracher, S.; Schmidt, C. C.; Dittmer, S. I.; Jung, H., Core Transmembrane Domain 6 Plays a Pivotal Role in the Transport Cycle of the Sodium/Proline Symporter PutP. *J Biol Chem* **2016**, *291* (50), 26208-26215.
12. De la Vieja, A.; Reed, M. D.; Ginter, C. S.; Carrasco, N., Amino acid residues in transmembrane segment IX of the Na<sup>+</sup>/I<sup>-</sup> symporter play a role in its Na<sup>+</sup> dependence and are critical for transport activity. *J Biol Chem* **2007**, *282* (35), 25290-8.
13. Raba, M.; Baumgartner, T.; Hilger, D.; Klempahn, K.; Hartel, T.; Jung, K.; Jung, H., Function of transmembrane domain IX in the Na<sup>+</sup>/proline transporter PutP. *J Mol Biol* **2008**, *382* (4), 884-93.
14. Sala-Rabanal, M.; Hirayama, B. A.; Loo, D. D.; Chaptal, V.; Abramson, J.; Wright, E. M., Bridging the gap between structure and kinetics of human SGLT1. *Am J Physiol Cell Physiol* **2012**, *302* (9), C1293-305.
15. Zhang, Z.; Liu, Y. Y.; Jhiang, S. M., Cell surface targeting accounts for the difference in iodide uptake activity between human Na<sup>+</sup>/I<sup>-</sup> symporter and rat Na<sup>+</sup>/I<sup>-</sup> symporter. *J Clin Endocrinol Metab* **2005**, *90* (11), 6131-40.
16. Watanabe, A.; Choe, S.; Chaptal, V.; Rosenberg, J. M.; Wright, E. M.; Grabe, M.; Abramson, J., The mechanism of sodium and substrate release from the binding pocket of vSGLT. *Nature* **2010**, *468* (7326), 988-91.
17. Paroder-Belenitsky, M.; Maestas, M. J.; Dohan, O.; Nicola, J. P.; Reyna-Neyra, A.; Follenzi, A.; Dadachova, E.; Eskandari, S.; Amzel, L. M.; Carrasco, N., Mechanism of anion selectivity and stoichiometry of the Na<sup>+</sup>/I<sup>-</sup> symporter (NIS). *Proc Natl Acad Sci U S A* **2011**, *108* (44), 17933-8.