Supplementary Information

Structure and mechanism of a bacterial sodium-dependent dicarboxylate transporter

Romina Mancusso, G. Glenn Gregorio, Qun Liu and Da-Neng Wang

## **Supplementary Discussion**

# **Transport mechanism**

Combination of the crystal structure and transport and binding assays of the vcINDY presented here, along with previous biochemical characterization of its bacterial, fly and, in particular, mammalian homologs <sup>1-4</sup>, suggest a transport mechanism for the Na<sup>+</sup>-dependent dicarboxylate transporter from Vibrio cholerae (Supplementary Fig. 17). In the "substrate-waiting," outwardfacing  $C_0$  conformation, the transporter adopts a structure in which its N-terminal half is in the "U-shape" while the C-terminal half is in the "V-shape", N(U)-C(V), with its Na<sup>+</sup>- and substratebinding sites facing the extracellular space. Following the induction of the proper substrate pocket by the binding of Na<sup>+</sup> ions, probably two, a substrate molecule binds to the transporter, vielding the  $C_0$ -S-Na<sup>+</sup> state. Using the binding energy and aided by Brownian motion, the transporter converts to the C<sub>i</sub>-S-Na<sup>+</sup> state. In this state, the palm of the N-terminal half moves toward the cytosol, whereas its thumb stays at its original position in the membrane, yielding an inverted V-shape. At the same time, the C-terminal half of the transporter changes from the Vshape to a U-shape. In this N(V)-C(U) structure, the substrate is exposed to the cytosolic space. Substrate release to the cytosol is then triggered by the escape of the Na2 ion, and is followed by the release of the Na1 ion. Finally, the transporter returns from this  $C_i$  state back to the  $C_0$  state.

Several pieces of published experimental data on various INDY proteins support the above transport mechanism. Transport kinetics studies on both the rabbit and the human NaDC1 support a binding order of three Na<sup>+</sup> ions from the extracellular space, followed by dicarboxylate <sup>5,6</sup>. Similar observations have been made for the sequence of binding of the two Na<sup>+</sup> ions and one substrate for bacterial INDY homologs <sup>7-9</sup>. Accessibility studies of cysteine substitution in

2

mammalian NaDC1 using a membrane-impermeable sulfonate reagent have shown that the tip of  $HP_{out}$ , including its SNT motif, is accessible from the extracellular space <sup>10-12</sup>. Such accessibility of the hairpin tip is increased in the presence of Na<sup>+</sup> ions, but access to the cysteine labeling is protected by the addition of substrate <sup>11,12</sup>. The equivalent of another segment near the substrate binding site, Ser200 – Pro202 in vcINDY, has also been found to be accessible from the extracellular side in the human sulfate transporter NaS1 <sup>13</sup>. All of the above evidence supports an induced-fit, alternating access model of Na<sup>+</sup>-dicarboxylate co-transport <sup>6,14</sup>.

For the vcINDY dimer, this transport mechanism implies that the two clusters of helices at the protomer-protomer interface (TMs 2, 3, 7 and 8 from both protomers) form an anchoring scaffold for the transporter in the membrane (Fig. 1d and Supplementary Fig. 17). Connected by TM4a and TM9a as their respective hinges, the two helical bundles in the N- and C-terminal hands move up and down relative to the anchoring scaffold in opposite directions along the membrane normal (Supplementary Fig. 9), realizing an alternating access model of substrate translocation across the membrane.

As the bound citrate in our vcINDY structure is exposed to the cytosolic space, the structure represents the transporter's inward-facing, substrate-bound conformation,  $C_i$ -S-Na<sup>+ 14</sup>. This interpretation is also in agreement with our observations that citrate interacts with the transporter in solution but inhibits its transport activity only slightly from outside of the cell (Fig. 1b and Supplementary Fig. 5). It follows that citrate is a  $C_i$ -conformation specific inhibitor. As an importer, the low energy state of vcINDY in the absence of substrate should be the outward-facing conformation,  $C_o$ . Following Na<sup>+</sup> and substrate binding, the transporter converts from the

3

 $C_o$ -S-Na<sup>+</sup> state to the  $C_i$ -S-Na<sup>+</sup> state for cytosolic release. In the present work, however, because a high concentration of citrate was present in the crystallization buffer, a pre-release  $C_i$  state of the transporter was captured in the crystal structure.

The vcINDY structure reported here represents its inward-facing,  $C_i$ -S-Na<sup>+</sup> conformation, a prerelease state (Supplementary Fig. 17). As the protein functions as an importer, its substrate selectivity depends on the atomic structure of the substrate-binding site in its outward-facing  $C_o$ or  $C_o$ -Na<sup>+</sup> conformation, a pre-binding state. Between the outward- and inward-facing states, the binding sites for Na<sup>+</sup> ions are likely to be conserved. The interactions between the two SNT motifs and the two carboxyl groups at the ends of the substrate molecule are also likely to be conserved. Additional residues at the binding site that interact with the substitutions at the middle carbons probably provide the fine tuning of the substrate specificity of the transporter <sup>9</sup> that, for example, allows it to distinguish between different kinds of dicarboxylate molecules.

## Substrate Specificity of sulfate transporters in the SLC13 family

The substrate-binding pocket in our crystal structure also explains the substrate preference between carboxylate transporters and sulfate transporters among SLC13 proteins <sup>2-4</sup>. In the sulfate transporters the substrate probably binds at the same location. At the third residue of the C-terminal SNT motif, in both NaS1 and NaS2, there is a proline residue there instead of a threonine or a valine as found in the carboxylate transporters (Fig. 4c, Supplementary Fig. 3). Interestingly, when we mutated Thr379 in vcINDY into a proline, the transport of succinate became sensitive to sulfate competition (Fig. 4d, Supplementary Fig. 15). In addition, there is a conserved serine residue at the vcINDY-P201 position among sulfate transporters. Previous

4

studies have shown that a substitution of the serine by alanine in the human NaS1 resulted in an increased  $K_m$  for sulfate transport <sup>13</sup>.

# **Comparison with CNT**

Although the amino acid sequence varies, the structure of the hairpin tip - capping loop motif is conserved between the two proteins, suggesting that this structural motif may be a common Na<sup>+</sup>binding motif in membrane transporter proteins. While vcINDY and CNT share a similar core structure, the ways by which each core is linked to its respective scaffold are completely different. CNT forms a trimer, with three pairs of long helices forming a scaffold around the threefold axis, which links to the core from each protomer (Supplementary Fig. 16). In contrast, in the dimeric vcINDY, the two palms in the core are each connected by a hinge to the anchoring thumbs, which cluster around the center of the dimer. The helical bundle of the palm forms a relatively well-defined structure for Na<sup>+</sup> binding, with the substrate sitting between the two oppositely-inverted palms. While the flexible linker between the palm and the thumb allows the two hands to change shape in a symmetrical manner<sup>15</sup>, the anchoring scaffold formed by the thumbs at the center of the dimer allows the N- and C-terminal palms to move up and down in opposite directions in the membrane to propel substrate translocation. Such domain movement relative to an anchor scaffold during substrate translocation has been observed previously in the trimeric Na<sup>+</sup>-dependent glutamate transporter <sup>16</sup>.

## References

1	Wright, E.M., Transport of carboxylic-acids by renal membrane-vesicles. Ann Rev
	<i>Physiol</i> <b>47</b> , 127-141 (1985).

Pajor, A.M., Sodium-coupled transporters for Krebs cycle intermediates. *Ann Rev Physiol* 61, 663-682 (1999).

- <sup>3</sup> Markovich, D. & Murer, H., The SLC13 gene family of sodium sulphate/carboxylate cotransporters. *Pflug Arch Eur J Phy* **447**, 594-602 (2004).
- <sup>4</sup> Pajor, A.M., Molecular properties of the SLC13 family of dicarboxylate and sulfate transporters. *Pflug Arch Eur J Phy* **451**, 597-605 (2006).
- <sup>5</sup> Wright, S.H., Hirayama, B., Kaunitz, J.D., Kippen, I., & Wright, E.M., Kinetics of sodium succinate cotransport across renal brush-border membranes. *J Biol Chem* **258**, 5456-5462 (1983).
- <sup>6</sup> Yao, X. & Pajor, A.M., The transport properties of the human renal Na<sup>+</sup>-dicarboxylate cotransporter under voltage-clamp conditions. *Am J Physiol Renal Physiol* **279**, F54-64 (2000).
- <sup>7</sup> Hall, J.A. & Pajor, A.M., Functional characterization of a Na<sup>+</sup>-coupled dicarboxylate carrier protein from *Staphylococcus aureus*. *J Bacteriol* **187**, 5189-5194 (2005).
- <sup>8</sup> Hall, J.A. & Pajor, A.M., Functional reconstitution of SdcS, a Na<sup>+</sup>-coupled dicarboxylate carrier protein from *Staphylococcus aureus*. *J Bacteriol* **189**, 880-885 (2007).
- <sup>9</sup> Strickler, M.A., Hall, J.A., Gaiko, O., & Pajor, A.M., Functional characterization of a Na<sup>+</sup>-coupled dicarboxylate transporter from *Bacillus licheniformis*. *Bba-Biomembranes* **1788**, 2489-2496 (2009).
- <sup>10</sup> Pajor, A.M., Krajewski, S.J., Sun, N., & Gangula, R., Cysteine residues in the Na<sup>+</sup>/dicarboxylate co-transporter, NaDC-1. *Biochem J* 344, 205-209 (1999).
- <sup>11</sup> Pajor, A.M., Conformationally sensitive residues in transmembrane domain 9 of the Na<sup>+</sup>/dicarboxylate co-transporter. *J Biol Chem* **276**, 29961-29968 (2001).
- Pajor, A.M. & Randolph, K.M., Conformationally sensitive residues in extracellular loop 5 of the Na<sup>+</sup>/dicarboxylate co-transporter. *J Biol Chem* 280, 18728-18735 (2005).
- <sup>13</sup> Li, H. & Pajor, A.M., Serines 260 and 288 are involved in sulfate transport by hNaSi-1. *J Biol Chem* **278**, 37204-37212 (2003).
- <sup>14</sup> Law, C.J., Maloney, P.C., & Wang, D.N., Ins and outs of major facilitator superfamily antiporters. *Annu Rev Microbiol* **62**, 289-305 (2008).
- <sup>15</sup> Forrest, L.R. & Rudnick, G., The rocking bundle: a mechanism for ion-coupled solute flux by symmetrical transporters. *Physiology (Bethesda, Md* **24**, 377-386 (2009).
- <sup>16</sup> Reyes, N., Ginter, C., & Boudker, O., Transport mechanism of a bacterial homologue of glutamate transporters. *Nature* **462**, 880-885 (2009).

Data set	1	2	3	4	All (1-4)
Unit-cell parameters a, b, c (Å) β (°)	101.42, 101.41, 165.12 101.96	101.56, 101.64, 165.73 102.12	101.73, 101.04, 165.50 101.58	102.83, 100.85, 165.74 100.78	101.88, 101.24, 165.53 101.61
Number of frames	360	360	360	360	1,440
Bragg spacings (Å)	40-3.5 (3.59-3.5)	40-3.7 (3.8-3.7)	40-4.0 (4.1-4.0)	40-4.0 (4.1-4.0)	40-3.5 (3.59-3.5)
Measurements	313,099	258,201	213,160	199,285	1,002,530
Unique reflections	41,571	34,565	28,103	28,068	41,621
Multiplicity	7.5 (7.3)	7.5 (6.2)	7.6 (7.4)	7.1 (7.4)	24.1 (7.4)
Completeness (%)	99.3 (93.2)	98.2 (74.9)	99.8 (100.0)	99.6 (100.0)	99.4 (94.1)
$R_{merge}^{1}$	0.090 (0.609)	0.146 (0.906)	0.145 (0.650)	0.181 (0.727)	0.246 (0.632)
$I/\sigma(I)^2$	19.1 (3.8)	12.6 (2.6)	17.6 (4.5)	11.9 (3.8)	19.3 (3.8)
$\Delta F/\sigma (\Delta F)^3$	1.97 (0.91)	1.32 (0.72)	1.43 (0.68)	1.18 (0.74)	1.91 (0.92)
Anomalous CC <sup>4</sup> (%)	85.6 (17.6)	73.2 (4.4)	63.5 (4.8)	60.1 (0.4)	83.0 (14.2)

Supplementary Table 1. Crystallographic data collection and reduction statistics for phasing

Notes:

 ${}^{1}R_{merge} = \sum_{hkl} \sum_{i} |I_{i}^{hkl} - \langle I^{hkl} \rangle| / \sum_{hkl} \sum_{i} |I_{i}^{hkl}|$ , where I is the intensity of a reflection hkl and  $\langle I \rangle$  is the average over measurements of hkl.

 $^{2}$ I/ $\sigma$ (I) = <<  $I^{hkl}$  > /  $\sigma$ (< $I^{hkl}$  >)> where <  $I^{hkl}$  > is the weighted mean of all measurements for a reflection hkl and  $\sigma$ (< $I^{hkl}$  >) is the standard deviation of the weighted mean. The values are as reported from *SCALA* as Mn(I/sd).

 ${}^{3}\Delta F/\sigma(\Delta F)$  is average anomalous signal from data truncated to  $d_{min} = 4$  Å. The values are derived by using CCP4 programs and are computed as  $< |\Delta F| / \sigma(\Delta F) >$  where  $\Delta F = |F(h)| - |F(-h)|$ .

<sup>4</sup>Anomalous correlation coefficient evaluated from data truncated to  $d_{min} = 4$  Å.

Values in parentheses are from the highest resolution shell.

	Crystal #1
Data collection	
Space group	$P2_1$
Unit-cell dimensions	
<i>a</i> , <i>b</i> , <i>c</i> (Å)	101.42, 101.41, 165,12
$\alpha, \beta, \gamma$ (°)	90, 101.96, 90
Resolution (Å)	50.00-3.21 (3.27-3.21)*
R <sub>merge</sub>	9.8 (92.3)
$I / \sigma(I)$	11.71 (1.51)
Completeness (%)	97.9 (83.3)
Redundancy	3.9 (3.5)
Refinement	
No. reflections	52,199
$R_{\rm work} / R_{\rm free}$	0.2280/0.2913
No. atoms	12390
Protein	12263
Ligand	39
Sodium	4
<i>B</i> -factors	98.66
Protein	98.42
Ligand	127.02
Sodium	66.43
R.m.s. deviations	
Bond lengths (Å)	0.010
Angles (°)	1.557
Ramachandran statistics	
(%)	
Favored	88.1
Outliers	0.2

Supplementary Table 2. Crystallographic data collection and refinement statistics

Notes: \*: A single crystal was used for the structure.  $R_{merge} = \sum_{hkl} \sum_i |I_i^{hkl} - \langle I^{hkl} \rangle| / \sum_{hkl} \sum_i I_i^{hkl}$ , where I is the intensity of a reflection hkl and  $\langle I \rangle$  is the average over measurements of hkl.

Redundancy represents the ratio between the number of measurements and the number of unique reflections. *R* factor =  $\Sigma |F(obs) - F(cal)| / \Sigma F(obs)$ ; 5% of the data that were excluded from the refinement were used to calculate  $R_{\text{free}}$ . The average *B* factor was calculated for all non-hydrogen atoms. r.m.s.d. of bond is the root-mean-square deviation of the bond angle and length.

Numbers in parentheses are statistics of the highest resolution shell.

Coordinating atoms*	Distance to Na <sup>+</sup> (Å)
S146 carbonyl O	2.34
S146 hydroxyl O	2.53
S150 carbonyl O	3.33
N151 amide O	2.36
G199 carbonyl O	3.26

Supplementary Table 3. Distances between Na<sup>+</sup> ion and protein coordinating atoms

Notes: \*: Interatomic distances were measured using Chain D.



Supplementary Fig. 1. Pathway for fatty acid biosynthesis in liver and fat cells. The rate of fatty acid synthesis depends on the cytosolic citrate level. NaCT, Na<sup>+</sup>-dependent plasma membrane citrate transporter; TCA, citric acid cycle; CTP, mitochondrial citrate transporter;

ACC, acetyl CoA carboxylase; LDL, low-density lipoprotein.



Supplementary Fig. 2. Phylogenetic tree of the DASS family.

Indy_Vibrio	1	vulgismartavialiteren intrewelhensetviadvalfer yhfiefehnvulgismartavialter hvvvatuvvvatuvpvatuvffetferqaalnaansi
SdcS_Staphyl	1	MAYFNQHQSMISKRYLTFFSKSKKKKPFSAGQLIGIILGPLLFLTLIFFHPQDLPWKGVYVLAITIWUATWWITEATPIAATSLLPIVLLPIGHTLTPEQVSSEYGNDI
Indy_fly	1	MATETTKMIYTPPPLDIKMEIEIGEQPQPPVKCSNFFANHWKGLVVFLVPLLCLPVMLINEGAEFRGMYLLLVMAIFWVTEALPLYVTSMIPIVAEPIMGTMSEDQTCRLYFKDI
NaDC1_human	1	YCAYAIIIMALEWCTEALPLAWAINSYLIYFYYPIILLPPPILVESKEAYCAYAIIIMALEWCTEALPLAVTALEPLIFYMMGIVDASEVAVEYLKDS
NaDC1_mouse	1	YCAYSIILMALWAYRSYLIVICLPIFLLPIPLIVQTKEAYCAYSIILMALLWCTEALPLAVTALFPIILFPLMGIMEASKVCLEYFKDI
NaDC1_rabbit	1	YCAYAIILMALFWCTDALPLAVTALLPICLFPIQUEASEVGLEYLKDI
NaDC3_human	1	MAVWCTEALPLSVTALLPIVLFPFMGILPSNKVCPQYFLD
NaDC3_mouse	1	RCLYVILLMAVYWCTEALPLSVTALLPFMGILPSSKVCPQYFLDI
NaDC3_rat	1	RCLYVILLMAVYWCTEALPLSVTALLPIUSARRLLVLLLVPLALLPILFALPPKEGRCLYVILLMAVYWCTEALPLSVTALLPIILFPFMGILPSSKVCPQYFLDI
NaCT_human	1	RCAYVIILMAIWCTEVIPLILLETVILMPAKFVRCAYVIILMAIWCTEVIPLAVTSLMPVLLPPLFQILDSRQVCVQMKDI
NaCT_mouse	1	RCAYVIVIMAVYNCTDVIPVAVTSLLPPLLKVLDSKQVCIQIMKDI
NaCT_rat	1	RCAYVIIIMAIYMCTDVIPVALTSLLPVLLFPLLKVLDSKQVCVQIMTDI
NaS1_human	1	ECAYTLEVVATEWLTENDESUULEVATEWLTENDESUULEVATEWLTENDESUULEVATEWLTENDESUULEVATEWLTENDESUULEVATEWLTENDESUULEVA
NaS1_mouse	1	CCAYILFVTALFWITCHINYALVYRRFLLVWFTILVFLPPLIIRTKEACCAYILFVTALFWITEALPLSITALLPGLMFPMFGIMRSSCVASAYFKD
NaS1_rat	1	ECAYILFVIATFWITEALPLSITALLEGLMFPMFGIMSSTHWASAYFKD
NaS2_rat	1	ECAYILFVIATFWITEALPLSITALLEGLMFPMFGIMSSTHWASAYFKD

HP <sub>in</sub>							
		TM4b H4c H4c H4c					
Indy Vibrio	97	IFUFLCCFALAAAHHQCFDKVLADKVLADKVLADKMSVAUFMLFCVTALLSMVLSNVATAAMMUPUVLCVTSKUDADKQ					
SdcS Staphyl	111	IFLFLGEFILAIAMERWILHTRVALTIINLICASTSKILLGEMVATCFLSMFVSNTAAVMIMIPIGLAIIKEAHDLQEAN					
Indy fly	116	LVMEMGCINVALAVEYCNLHKRUALRVIQIVCCSPRRIHFCLIMVIMFLSMWISNAACTAMMCPIICAVLEEIQAQGVCKINHEPQYQIVG					
NaDC1_human	87	NLEFGGLVAIAVEHWNLHKRTALRVLIVGVRPAPLTLGFMLVTAFLSMWISNTATSAMMVPIAHAVLDOHSSQAS-SNVEEGSNNPTFETOPPSPOKEVTKLD					
NaDC1_mouse	87	NILFVGGLMVA <mark>IAVEHWNLHKRIAL</mark> GVLLIIGV <mark>RPALLILGFMIVTAFLSMWISNTATTAMMIPIGYAVLE</mark> OLQGSQKDVEEGNSNPSFELQBASPQKBETKLD					
NaDC1_rabbit	87	NVLFIGGLLIA <mark>TAVEHWNLHKRIALRVLLITGVRPALLILGFMVVTAFLSMWISNTASTAMMVPI</mark> AH <mark>AVL</mark> OELNNTQSNVEEGSDNPTFELQBPSPQKETSKVDEKD					
NaDC3_human	43	NELFISCIIMASATEEWNIHRRIALKIIMIVGVOPARIIIGMMVTTSEISMWISNTASTAMMIPIANATIKSIEGOKEVRKDPSQESEENTAAVRRNGIHTVPTEMQEIA					
NaDC3_mouse	90	NELFLS <mark>GIIMASAIEEWNLHRRIALKVLMIVGVO</mark> PARLIIGMMVTT <mark>SFLSMWLSNTASTAMMIPI</mark> AS <mark>AIL</mark> KSLFGOREARKDLPREGDESTAAVQGNGLRTVPTEMOFLA					
NaDC3_rat	90	NELFLS <mark>GIMASATEERNIHRRIALKVIMIVGVO</mark> PARLIIGMMVTT <mark>SFISMWISNTASTAMMIPI</mark> AS <mark>AII</mark> KSIFGORDARKDIPREGEDSTAAVRGNGIRTVPTEMOFIA					
NaCT_human	87	NYLFLGGLIVAVAVERWNLHKRIALRTLLWVCAKPARLMLGFMCVTALLSMWISNTATTAMMVPIVEAILOOMEATSAATEAGLEVVKGKAKELPGSQ					
NaCT_mouse	87	NMLFLGSLIVAVAVERNKLHKRVALRMLLFVCTKPSRLMLGFMFVTAFLSMWISNTAATAMMIPIVEAMLOOMIAANTAVEASLGTLELLKNKTSELPGSQ					
NaCT_rat	87	NMLFLGSLIVATAVERNELHKRIALRMLLFVGTKPSRLMLGFMFVTAFLSMWISNTATTAMMIPIVEAMLEOMVATNVAVDASORTMEILOKNKASELPGSQ					
NaS1_human	86	h LIIGVICLATSIEKWNLHKRIALKMVMMVGVNPANLTLGFMSSTAFLSMWLSNTSTAAMVMPIAEAVVOOIINAEAEVEATQMTYFNGSTNHGLEIDESVNGHDINERKEKTKPVPGY					
NaS1_mouse	86	h li <mark>g</mark> u chaistekwilikrialrwymygyn <mark>pawd tlgfm</mark> ss <b>taflswu snistaam</b> ympi ve <mark>avaqo</mark> ti saeaeaeatomtyf iesaahgldide <mark>i</mark> vigoetnekkekikpapgs					
NaS1_rat	86	h li <mark>g</mark> u chaistekwilikrialrwymygynpawd t <mark>lgfm</mark> ss <b>taflswu sni</b> staamwpiveavaqottsaeaeaeatomtyfnesaaqglevdeni i goetnerkektkpalgs					
NaS2 rat	86	HILLGVICTANSIEKWNLHKRIALRMVMWVGVNPAWLTLGFMSSTAFLSMWISNTSTAAMVMPIVEAVAQQITSAEAEAEATQMTYFNESAAQGLEVDETIIGQETNERKEKIKPALGS					

			TM5	a -	TM5b	)	_	TM6		
Indy Vibrio	175	RSTYVI	VLLGV	AYSASIGGI	ATLVGSPPNAIA	AEVG		KFGLPTAMMMLPM.	AIAILYFLLKP	TLNGMFELDRAPVNW
SdcS_Staphyl	191	TNQTSIQKFEK	LVLAI	GYAGTIGGL	GTLIGTPPLIIL	GQYMQHEGHE	-ISFAKWM	IVG <mark>IPT</mark> VIVL <mark>L</mark> GI	IWLYLRYVAFR	HDLKYLPGGQTLIKQ
Indy_fly	207	GNKKNNEDEPPYPTKITLC	YYLGI	AYAS <mark>SLGG</mark> C	GTI IGTATNLTF	GIYEARFKNSTE	QMDFPTFMI	FYSVPSMLVYTLL	IFVF <b>LQ</b> W <mark>H</mark> FMG	LWRPKSKEAQEVQRG
NaDC1_human	193	NGQALPVTSASSEGRAHLSQKHLHLTQC	MSLCV	C <mark>YSA</mark> SIGGI	ATLTGTAPNLVL	QGQINS <mark>LFP</mark> QNGN	VVNFASWF	SFAFPTMVILLLL	AWLWLQ <mark>I</mark> LFLG	FNFRKN-FGIGEKMQ
NaDC1_mouse	191	NGQAVSVSSEPRAQKTKEHHRFSQ	LS <mark>LCI</mark>	C <mark>YSASIGG</mark> I.	ATLTGT TPNLVL	QQVNSIFPENSN	VVNFASWF	GFAFPTMVILLLL	AWLWLQ <mark>V</mark> LFLG	VNFRKN-FGFGEGEE
NaDC1_rabbit	194	NGQAQPLPAVPLESGEHMTQEQLRFSQ	MSLCV	C <mark>YSASIGG</mark> I	ATLTGT TPNLVL	QGQMTSLFPQNPN	VVNFASWF	GFAFPIMVILLLL	SWLWLQILFLG	INFRKN-FGIREQEH
NaDC3_human	153	STEAKDHPGETEVPLDLPADSRKEDEYRRNIWK	FLISI	PYSASIGGT	ATLTGTAPNLIL	.GQLKSFFPQC-D	VVNFGSWF	IFAFPLMLLFLLA	GWLWISFLYG <mark>G</mark>	LSFRGWRKNKSEIRT
NaDC3_mouse	200	SSEG-GHTEDAEAPMELPDDS-KEEEHRRNIWK0	FLISI	PYSASIGGT	ATLTGTAPNLIL	GQLKSFFPQC-D	VVNFGSWF	IFAFPLMLLFLLV	GWLWI <mark>SFL</mark> YG <mark>G</mark>	MSWRSWRKKKSKIRA
NaDC3_rat	200	SSEG-GHAEDVEAPLELPDDS-KEEEHRRNIWK	FLISI	PYSASIGGT	ATLTGTAPNLIL	GQLKSFFPQC-D	VVNFGSWF	IFAFP <mark>LMLLF</mark> LLV	GWLWISFLYG <mark>G</mark>	MSWRGWRKKNSKLRD
NaCT_human	186	VIFEGPTLGQQEDQERKRLCK/	MTLCI	ICYAASIGGT	ATLTGTGPNVVL	. <mark>GQ</mark> MN <mark>EL</mark> FPDSKD	LVNFASWF	AFAFPNMLVMLLF	A <mark>WLWLQ</mark> FVYM <mark>R</mark>	FNFKKS-WGCGLESK
NaCT_mouse	189	VVFEDPNVQEQEDEETKNMYK/	MHLCV	CYSASIGGT	ATLTGTGPNVVL	. <mark>GQ</mark> MQ <mark>EL</mark> FPDSKD	VLNYASWF	G <b>FAFP<mark>NM</mark>VMMLVL</b>	AWLWLQC <mark>L</mark> YM <mark>R</mark>	HNLKKTCICCGEKKR
NaCT_rat	189	VVFEDPSVQKQEDEETKNMYK/	MNLCV	CYAASIGGT	ATLTGTGPNVVLI	. <mark>GQ</mark> MQ <mark>EL</mark> FPDSKD	VMNFASWF	AFALPNMLLMLVM	AWLWLLCFYMR	PNLKKTCICCGRKKK
NaS1_human	206	NNDTGKISSKVELEKNSGMRTKYRTKKGHVTRK	TCLCI	AYSSTIGGL	TTITGTSTNLIF?	AEYFNTRYPDC-R	CLNFGSWF	F <mark>SFPAA</mark> LIILLL	SWIWLQWLFLG	F <mark>NFK</mark> EMFKCGKTKTV
NaS1_mouse	206	SHDKGKVSRKMETEKNAVTGAKYRSRKDHMMCK1	MCLSV	/ <mark>AYS</mark> STIGGL	TTITGTSTNLIFS	SEHFNTRYPDC-R	CLNFGSWF	LF <mark>SFPVA</mark> LILLL	SWIWLQWLYLG	FD <mark>FK</mark> -MFKCGKTKTL
NaS1_rat	206	SNDKGKVSSKMETEKNTVTGAKYRSKKDHMMCK1	MCLCI	AYSSTIGGL	TTITGTSTNLIFS	SEHFNTRYPDC-R	CLNFGSWF	LFSFPVAVILLLL	SWIWLQWLFLG	F <mark>NFK</mark> EMFKCGKTKTL
NaS2_rat	206	SNDKGKVSSKMETEKNTVTGAKYRSKKDHMMCK		IAYSSTIGGL	TTITGTSTNLIFS	SEHFNTRYPDC-R	CLNFGSWF	LFSFPVAVILLLL	SWIWLQWLFLG	FNFKEMFKCGKTKTL

(Supplementary Fig. 3. Continued)

			- <b>TM</b> 7		TM8	TM9a
Indy Vibrio	258	DKG	KVVTLGIFGLTVFLWIFSSPIN	AALGGBK	FDTLVALGATIMLSFA	RVVH <mark>WK</mark> EIQKTAD <mark>W</mark> G
SdcS Staphyl	286	KLDELGKMKYE	EKVVQTIFVLASLLWITREFLLI	KKWEVTSSVA	DGTIAIFISILLFIIP	AKNTEKHRRIIDWEVAK-ELPWG
Indy_fly	312	REG-ADVAKKVIDQRYKDLGPMSIH	EIQVMILFIFMVVMYFTR <mark>KPG</mark> II	FL <mark>GW</mark> AD <mark>L</mark> LNSKDIF	NSMPTIFVVVMCFMLP	ANYAFLRYCTRRGGPVPTG-PTPSLITWKFIQTKVPWG
NaDC1_human	306	EQQQAAYCVIQTEHRLLGPMTFA	E <mark>KAISILF</mark> VILVL <mark>LWF</mark> TREPGF1	fl <mark>gw</mark> gn <mark>l</mark> afpnakgesm <mark>vs</mark>	DGTVAI <mark>F</mark> IGIIMFIIP	SKFPGLTQDPENPGKLKAPLGLLDWKTVNQKMPWN
NaDC1_mouse	301	ERKQAAFQVIKTQHRLLGPMSFA	E <mark>KA</mark> VTFLFVLLVVLWFTREPGF	FP <mark>GW</mark> GDTAFANKKGQSM <mark>VS</mark>	DGTVAI <mark>FIS</mark> LIMFIIP	SKIPGLTEDPKKPGKLKAPPAILTWKTVNDKMPWN
NaDC1_rabbit	307	EQQRKQAAYRVIQTQYRLLGPMSEA	E <mark>KAV</mark> FILFVILVL <mark>LWFTREPGF</mark> I	fh <b>gw</b> gn <b>l</b> vfsdasgrvm <mark>v</mark> s	DGS <mark>AS</mark> ILIGV <mark>FLFMV</mark> P	SKIPGLTQDPDNPGRLKAPPALLNWKLVNKKMPWN
NaDC3_human	272	NAEDRARAVIREEYQNLGPIK		<u></u> FLS	DAVTGVAIVTILFFFP	SQRPSLKWWFDFKAPNTETEPLLTWKKAQETVPWN
NaDC3_mouse	317	DAEDQAKAVIQEEFQNLGPIKFA	EQAVFILFCTFAILLFSRDPKF.	IP <mark>GW</mark> ASLFAPGFVS	DAVTGVAIVTILFFFP	SQKPSLKWWFDFKAPNSETEPLLSWKKAQETVPWN
NaDC3_rat	317	VAEDKAKAVIQEEFQNLGPIKFA	EQAVFILFCLFAILLFSRDPKF.	IPGWASLFAPGFVS	DAVTGVAIVTILFFFP	SQKPSLKWWFDFKAPNSETEPLLSWKKAQETVPWN
NaCT_human	293	KNEKAALKVLQEEYRKLGPLSFA	E <mark>IN<mark>VLIC</mark>FFLLVI<b>LWFSRDPGF</b>I</mark>	MP <mark>GW</mark> LTVAWVEG-ETKYVS	DATVAI FVA <mark>T</mark> LLFIVP	SQKPKFNFRSQTEEERKTPFYPPPLLDWKVTQEKVPWG
NaCT_mouse	297	DTEKIAYKVLNEEYQKLGSLSYP	E <mark>CNVLFC</mark> FTLLVILWFSRDPGF1	P <mark>GW</mark> LSFAWVEG-NTVHII	DATVAI <mark>F</mark> VAILLFIIP	SQKPKFNFSSQTEEERKTPFYPPALLDWKVAQEKVPWD
NaCT_rat	297	DTEKIASKVLYEEYRKLGPLSYA	E <mark>CNVLFC</mark> FGLLIILWFSRDPGF1	PGWLSIAWIEG-NTKH <mark>V</mark> I	DATVAI FVAILLFIVP	SQKPKFNFSRQTEEERKTPFYPPPLLNWKVTQEKVPWG
NaS1_human	325	QQKACAEVIKQEYQKLGPIRYQ	E <mark>IVT</mark> LVLFIIM <mark>A</mark> LLWFSRDPGF	PGWSALFSEYPGFAT	DSTVALLIGLLFFLIP	AKTLTKTTPTGEIVAFDYSPLITWKEFQSFMPWD
NaS1_mouse	324	KEKACAKVIKQEYEKLGPMRYQ	E <mark>IVT</mark> LVIFIVM <mark>A</mark> LLWFSRDPGF	/T <mark>GW</mark> SVLFSEYPGYVI	D <mark>STVALVAGIL</mark> FFLIP	AKKVTKMTSAGEIIAFDYTPLITWKEFQSFMPWD
NaS1_rat	325	KEKACAEVIKQEYEKLGPMRYQ	EIVTLVIFIVMALLWFSRDPGF	TGWSVIIFSEYPGYVI	DSTVALVAGILFFLIP	AKKLTKMTSTGDIIAFDYSPLITWKEFQSFMPWD
NaS2_rat	325	KEKACAEVIKQEYEKLGPMRYQ	EIVTLVIFIVMALLWFSRDPGF	TGWSVIFSEYPGYVI	DSTVALVAGILFFLIP	AKKLTKMTSTGDIIAFDYSPLITWKEFOSFMPWD

				I	1P <sub>out</sub>		
		TM9b	- H9c -	-		TM10a	
Indy_Vibrio	322	VLLLFGGGLCLSNVLKQT	GTSVFLANALSDMVSHM	GIFVVILVVATF <b>V</b> VFL	FASNTASAALLIPVFATVAE	AFGMSPVLLSVLIAVAASCAFM	LPVATPPNAIVFASGHIKQSEMMRV
SdcS_Staphyl	367	VLILFGGGLALAKGISES	GLAKWLGEQLKSLNGVS	PILIVIVITIFVLFLT	VTSNTATATMILPILATLSV	AVGVHPLLLMAPAAMAANCAYM	LPVGTPPNAIIFGSGKISIKQMASV
Indy_fly	425	LVFLLGGGFALAEGSKQS	<u>GMAKLIGNALIGLKVLP</u>	-NSVLLEVVIEVAVFLE	FSSNVAIANIIIPVLAEMSL	AIEIHPLYLILPAGLACSMAFH	LPVSTPPNALVAGYANIRTKDMAIA
NaDC1_human	421	IVLLLGGGYALAKGSERS	<b>GLSEWLG</b> NKLTPLQSVP	-APAIAIIISLIVATET	CTSNVATTTIFLPILASMAQ	AICLHPLYVMLPCTLA <mark>T</mark> SLAFM	LPVATPPNAIVFSFGDLKVLDMARA
NaDC1_mouse	416	ILILLGGGFALAKGSEES	<b>GLSKWLG</b> DK <mark>LTPL</mark> QH <b>VP</b>	-PSATVLILSLLVAIFT	CTSNVATTIFLPILASMAQ	AICLHPLYVMLPCTLAASLAFM	LPVATPPNAIVFSFG <mark>GLKV</mark> SDM <mark>A</mark> RA
NaDC1_rabbit	424	IVLLLGGGYALAKGSEES	GLSQWLGNKLMPLQHVP	PPATVFIICLLVATFT	CTSNAATTTLLLPILASMAQ	AICLHPLYVMLPCTLASSLAFM	LPVATPPNAIVFSFG <mark>GLRVS</mark> DM <mark>A</mark> RA
NaDC3_human	347	IILLLGGGFAMAKGCEES	GLSVWIGGQLHPLENVP	PALAVLLITVVIAFFT	FASNTATIIIFLPVLAELAI	RLRVHPLYLMIPGTVGCSFAFM	lpvstppnsiafasghllvkdmvrt
NaDC3_mouse	427	IILLLGGGFAMAKGCE <mark>E</mark> S	GLS <mark>AWIG</mark> GQLHPLEHVP	-PLLAVLLITVVIAFFT	FASNTATIIFLPVLAELAI	RL <mark>HVHPLYLMIP</mark> GTVGCSYAFM	LPVSTPPNSIAFSTGHLLVKDMVRT
NaDC3_rat	427	IILLGGGFAMAKGCEES	GLSAWIGGQLHPLEHVP	-PLLAVLLITVVIAFFTI	FASNTATIIIFLPVLAELAI	RLHVHPLYLMI PGTVSCSYAFM	LPVSTPPNSIAFSTGHLLVKDMVRT
NaCT_human	410	IVLLLGGGFALAKGSEAS	GLSVWMGKQMEPLHAVP	PAAITLILSLLVAVFT	CTSNVATTILFLPIFASMSR:	SIGLNPLY IMLPCTLSASFAFM	LPVATPPNAIVFTYGHLKVADMVKT
NaCT_mouse	414	IVLLLGGGFAMAKGCETS	GLS <mark>KWMA</mark> AQMEPLRLVK	-PAVITLILSCLVAMTTI	CTSNVATTLFLPIFASMAR.	SIGIHPLYVMIPCTMSASLAFM	LPVATPPNAIVF <mark>A</mark> YGHLRV <mark>V</mark> DMMK <mark>T</mark>
NaCT_rat	414	IVLLLGGGFAMAKGCETS	GLS <mark>EWMA</mark> RQME <mark>PL</mark> SSVR·	-PAIITLILSCIVAMTTI	CTSNVATTLFLPIFASMAR.	SIGIHPLYVMIPCTLSASLAFM	LPVATPPNAIVF <mark>A</mark> YGHLKV <mark>I</mark> DMVKT
NaS1 human	435	IAILVGGGFALADGCEES	GLSKWIGNKLSPLGSIP	-AWLIILISSLMVTSLT	VASNPATITLFLPILSPLAE	AIHVNPLYILIP <mark>STLCT</mark> SFAFL	LPVANPPNAIVFSYGHLKVIDMVKA
NaS1 mouse	434	IAILVGGGFALADGCQVS	GLSNWIGSKLSPLGSLP	-VWLIILISSLIVTSLT	VASNPATITILFPILSPLAE	AIQVNPLQILLPSTLCTSFAFL	LPVA <mark>N</mark> PPNAIVFSYGHLKV <mark>I</mark> DMVKA
NaS1_rat	435	IAILVGGGFALADGCQVS	GLSSWIGSKLSPLGSLP	-VWLIILISSLIVTSLT	VASNPATITILFPILSPLAE	AIHVNPLHILLPSTLCTSFAFL	LPVA <mark>N</mark> PPNAIVFSYGHLKV <mark>I</mark> DMVKA
NaS2 rat	435	IAILVGGGFALADGCQVS	GLSSWIGSKLSPLGSLP	-VWLIILISSLIVTSLT	VASNPATITILFPILSPLAE	AIHVNPLHILLPSTLCTSFAFL	LPVA <mark>N</mark> PPNAIVFSYGHLKVIDMVKA



Supplementary Fig. 3. Amino acid sequence alignment of vcINDY and its homologs. Positions of secondary structures of the protein as observed in its crystal structure are indicated, and they are colored using the same rainbow schemes as in Figs. 1-4. The red dots indicate the two SNT motifs for carboxyl group binding.



Supplementary Fig. 4. Analytical size-exclusion chromatography trace of purified vcINDY in dodecylmatoside detergent. vcINDY ran as a single, monodisperse peak at 16.951 mins. The retention times of the monomeric glycerol-3-phosphate transporter (GlpT, 18.654 mins) and the dimeric tetracycline transporter (TetL, 16.133 mins), two membrane transporter proteins with a similar molecular weight as vcINDY, are indicated.



Time (min)

Supplementary Fig. 5. Effects of various salts on the thermostability of purified vcINDY characterized by analytical size-exclusion chromatography. vcINDY purified in DDM detergent and 100 mM NaCl ran as a sharp peak (without NaCl, the protein would aggregate). Upon being heated to 44 °C for 10 minutes, the peak height dropped by ~50%. The presence of succinate or malate during incubation was able to stabilize the protein. The largest effect in thermostabilization was observed for citrate, in which the peak height was completely recovered, indicating a specific interaction between the transporter and citrate. As a control, MgCl<sub>2</sub> had no effect. The void was at ~11 mins.



Supplementary Fig. 6. Quality of vcINDY electron density maps. The  $2F_o - F_c$  map (contoured at 1.5 $\sigma$ ) is superimposed with the anomalous difference Fourier map (contoured at 3.7 $\sigma$ ) of the SeMet crystal, which was used to identify the position of 22 out of the 23 methionine residues of the vcINDY protein. The crystallographic asymmetric unit cell contains four vcINDY protomers.



Supplementary Fig. 7. Periplasmic view of the vcINDY dimer. Structure of the protein is viewed from the extracellular space. The cross section of the protein dimer in the membrane plane measures about 80 Å by 55 Å, whereas the height of the protein along the membrane normal is 60 Å. Many residues at the interface are conserved from bacteria to mammals, including Trp320, which forms an inter-protomer  $\pi$ - $\pi$  interaction with the Trp320 of the other protomer.



Supplementary Fig. 8. Structure of the vcINDY protomer. **a**, Viewed from within the membrane. **b**, Viewed from the cytosol.



Supplementary Fig. 9. Comparison of the N-terminal and C-terminal halves in vcINDY. **a**. Overlay of the N- and C-terminal halves, with their helical bundles superimposed. **b**. Overlay of the N- and C-terminal helical bundles.



Supplementary Fig. 10. Comparison of the N-terminal and C-terminal halves of the vcINDY structure. Overlay of the N- and C-terminal halves, with their thumbs superimposed.



Supplementary Fig. 11. Electron density  $(2F_o - F_c \text{ map}, \text{ contoured at } 2.5\sigma \text{ and } 3.5\sigma)$  for the bound Na<sup>+</sup> ion at the Na1 site. The electron density between the hairpin HP<sub>in</sub> tip and the loop L5ab is too large for a Li<sup>+</sup> ion, and its ligand coordination and the conservation of the coordinating amino acid sequence do not support a bound water molecule. Therefore, we infer that the density belongs to a bound Na<sup>+</sup> ion.



Supplementary Fig. 12. Western blot analysis of expression levels of vcINDY mutants in *E. coli* for uptake experiments. Most of the mutants expressed at comparable levels to the wild-type protein, indicating that any decrease observed in transport activity was not due to reduced protein expression.



Supplementary Fig. 13. Overlay of Na1 and Na2 site structures and comparison of distances between the corresponding hairpin tip and capping loop. The distances between the  $HP_{out}$  and L10ab in the C-terminal Na2 clamshell (pink) is significantly larger than the corresponding distances between the  $HP_{in}$  and L5ab in the N-terminal Na1 clamshell (green), supporting the hypothesis that a Na<sup>+</sup> ion has been released from the Na2 site. s.c. denotes side chain.



Supplementary Fig. 14. Electron density  $(2F_o - F_c \text{ map}, \text{ contoured at } 1.0\sigma)$  for the bound citrate molecule.



Supplementary Fig. 15. Succinate transport activity of wild-type and mutant vcINDY in the presence of sulfate. N = 3.



Supplementary Fig. 16. Comparison of the vcINDY protomer structure with that of the concentrative nucleotide transporter (CNT) from *Vibrio cholerae*. vcINDY is colored orange and yellow, whereas CNT is colored green and cyan. **a**, Periplasmic view. **b**, Viewed from within the membrane. The r.m.s.d between C $\alpha$  atoms of these most conserved parts, the helical hairpins and the two Transmembrane helices that follow them, is 4.9 Å for those 243 amino acid residues.



Supplementary Fig. 17. Proposed transport mechanism of vcINDY. In the outward-facing  $C_o$  conformation, the N-terminal half adopts a U-shape, whereas the C-terminal half adopts a V-shape. Following Na<sup>+</sup> and substrate binding, the N- and C-terminal halves change to a V- and a U-shape, respectively, resulting in the inward-facing  $C_i$ -S-Na<sup>+</sup> conformation. After the release of Na<sup>+</sup> and substrate to the cytosol, the transporter returns to the  $C_o$  state, completing a substrate translocation cycle.