## Supplementary information

Cryo-EM Structure of the Sodium-driven Chloride/Bicarbonate Exchanger NDCBE

Wang et al.

**Supplementary Figures** 



**Supplementary Figure 1. Structure determination of NDCBE. a** Gel-filtration of NDCBE in a buffer containing Na<sup>+</sup>, Cl<sup>-</sup>, HCO<sub>3</sub><sup>-</sup>, Tris-HCl, with LMNG (pH 7.4) and Coomassie staining of purified NDCBE (n = 4 biologically independent experiments). Source data are provided as a Source Data file. **b** Representative cryo-EM micrograph of NDCBE in LMNG (n = 7,545 micrographs; scale bar = 20 nm). **c** Flowchart of image processing for NDCBE. **d** Density map of full-length NDCBE structure. **e** Angular distribution map of NDCBE. **f** Local resolution analysis of the NDCBE cryo-EM map using Resmap. **g** The Gold standard FSC curve of the final 3D reconstruction. **h** The FSC curves for cross-validation between the maps and the model.



Supplementary Figure 2. Cryo-EM density maps of NDCBE TMD regions superposed with atomic models. Representative maps at 14 TMs of NDCBE, H1 and NAG-Asn646 are shown.

	H1	ТМ	1	TM2		
rNDCBE hNDCBE hNBCe1 hAE1	GRLFGGLVLDVKRKAPWYWSDYRDALS GRLFGGLVLDIKRKAPWYWSDYRDALS GRFCGGLIKDIKRKAPFFASDFYDALM GQLFGGLVRDIRRRYPYYLSDITDAFS *:: ***: *::*: *:: *:: **	SLQCLASFLFLYCACM SLQCLASFLFLYCACM IIQALSAILFIYLATV PQVLAAVIFIYFAAL * *::*:* * :	SPVITFGGLLGEATEGF SPVITFGGLLGEATEGF INAITFGGLLGDATDNN SPAITFGGLLGEKTRN( ********	ISAIESLFGASMTGI ISAIESLFGASMTGI QGVLESFLGTAVSGA MGVSELLISTAVQGI	AYSLFAGQPLTIIGST AYSLFAGQALTIIGST IFCLFAGQPLTIISST LFALLGAQPLLVVGFS ** * :	538 538 485 465
	TM3	TM4	H2	TM5	H3	
rNDCBE hNDCBE hNBCe1 hAE1	GPVLVFEKILEKFCKDYALSYLSLRAC GPVLVFFKILEKFCKDYALSYLSLRAC GPVLVFERLENFSKDNFDYLEFRL GPLLVFEAFFSFCEINGLEYIVGRVW **:**** * * *	IGLWTAFLCIVLVAT IGLWTAFLCIVLVAT NGFWLILLVVLVAF *: * :* :::**	DASSLVCYITRFTEEAH DASSLVCYITRFTEEAH DASFLVQYFTRFTEEGI EGSFLVRFISRYTQEIH * ** :::*:*:* *	ASLICIIFIYEAIEA SSLICIIFIYEAIEK SSLISFIFIYDAFKK SFLISLIFIYEFFSK * **.:****	LIHLÆTYPIHMHSQL LIHLÆTYPIHMHSQL 4IKLÆDYYPINSNFKV LIKIFØDHPLQKTYNY *::::::::::	628 628 575 555
	Η3 β1	H4	H5 H6	β2	TM6	
rNDCBE hNDCBE	DHLSLYYOROALPENPNNHTL DHLSLYYOROTLPENPNNHTL	-QYWKEHSIPT -QYWKDHNIVT	ADVNWANLTVSEQQEMF AEVHWANLTVSEQQEMF	IGEFIGSACGHHGPYT IGEFMGSACGHHGPYT	PDVLFWSCILFFATFI PDVLFWSCILFFTTFI	707
hNBCe1 hAE1	GYNTLFSGIGVPPDPANISISNDTTLA NVLM	PEYLPTMSSTDMYHN	TTFDWAFLSKKE <mark>G</mark> SKYG	GNLVGNNCNFV	PDITLMSFILFLGTYT PNTALLSLVLMAGTFF	661 583
	. ' 2 *.		2		: : * :*: *:	
rNDCBE	IM6 VSSTLKTEKTSBYEPTBVBSTVSDEAV	I <b>M/</b> FLTTFTMVILDFLIG	V-PSPKLOVPSVFKPT-	BDDRGWFTSPTGP	IM8	792
hNDCBE	LSSTLKTFKTSRYFPTRVRSMVSDFAV	FLTIFTMVIIDFLIG	V-PSPKLQVPSVFKPT-	RDDRGWIINPIGP	-NPWWTVIAAIIPALL	792
hAE1	FAMMLRKFKNSSYFPGKLRRVIGDFGV	PISILIMVLVDFFIQ	DTYTQKLSVPDGFKVSN	ISSARGWVIHPLGLRSI	EFPIWMMFASALPALL	673
	тм8	тм9		410	 TM11	
rNDCBE	CTILIFMD2Q1TAVIINRKEHKLKKGC	GYHLDLLVVAIMLGV	CSLMGLPWFVAATVLSI	THVNSLKLESECSAP	GEQPKFLGIREQRVTG	882
hNDCBE hNBCe1	CTILIFMDQQITAVIINRKEHKLKKGO VTILIFMDQQITAVIVNRKEHKLKKGA	GYHLDLLMVAIMLGV GYHLDLFWVAILMVI	CSIMGLPWFVAATVLSJ CSLMALPWYVAATVISJ	THVNSLKLESECSAP( AHIDSLKMETETSAP(	EQPKFLGIREQRVTG EQPKFLGVREQRVTG	882 836
hAE1	VFILIFLEBQITILIVSKPERKMVKGS	GFHLDLLLVVGMGGV	AALFGMPWL <mark>SATTVR</mark> S\ ** ***	THANALTVMGKASTP	GAAAQIQEVKEQRISG *:** :: ::***:	763 :*
	<u>TM11 H7 T</u>	V12	H8	H9	TM13	
rNDCBE	LMIFVLMGCSVFMTAVLKFIPMPVLYG	VFLYMGVSSLQGIQF VFLYMGVSSLOGIOF	FDRLKLFGMPAKHQPDE FDRLKLFGMPAKHOPDE	'IYLRHVPLRKVHLFTI 'IYLBHVPLRKVHLFTI	LVQLTCLVLLWVIKAS	972 972
hNBCe1	TLVFILTGLSVFMAPILKFIPMPVLYG	VFLYMGVASLNGVQF	MDRLKLLLMPLKHQPD	IYLRHVPLRRVHLFT	FLQVLCLALLWIIKST	926
1111111	:: :* * * *::* :*. **: **:*	*:*****:**.*:*	:**: *: * *::* *	· *:::* ::****	:*: **.:**::*::	000
	TM14		H10			
rNDCBE hNDCBE	PAAIVFPMMVLALVFVRKV-MDLCFSK PAAIVFPMMVLALVFVRKV-MDLCFSK	RELSWLDDLMPESKK RELSWLDDLMPESKK	KKLDDAKKKE-EEEEAE KKLDDAKKKAKEEEEAE	KMLDIGGDKFPLESR KMLEIGGDKFPLESR	KLLSSPGKNNSFRCDP KLLSSPGKNISCRCDP	$\begin{array}{c} 1060 \\ 1061 \end{array}$
hNBCe1 hAE1	VAAIIFPVMILALVAVRKG-MDYLFSÇ PASLALPFVLILTVPLRRVLLPLIFRN	HDLSFLDDVIPEKDK IVELQCLDAD	KKKEDEKKKKKKKKGSLI DAKATFDEEEGRI	SDNDDSDCPYSEK EYDEVAMPV	VPSIKIPMDIME	1009 911
	*:: :*.::: * :*: : * : :	* * *	* * . :: :.	: *		
rNDCBE	SEINISDEMPKTTVWKALSINSGNTKF	KSPFN 10	92			
hNDCBE	SEINISDEMPKTTVWKALSMNSGNAKE	KSLFN 10	93			
hAE1	2DKEKS	91	.1			

**Supplementary Figure 3. Sequence alignment of NDCBE, AE1 and NBCe1.** The Clustal Omega program<sup>62</sup> was used for the alignment. Secondary structure assignments are based on the cryo-EM structure of NDCBE. The sequence alignment of the N-terminal domain is not shown. Red boxes and numbers mark the cysteine residues that form disulfide bonds and green boxes show the glycosylation sites of the asparagine residues in EL3. Blue boxes highlight the residues from the binding pockets of AE1, NBCe1 and NDCBE (presented in Figs. 2,3,5). Brown boxes show the most pronounced differences in the charged residues lining TMs 3 and 5 of the ion permeation pathways (presented in Fig. 3a).



**Supplementary Figure 4. The dimeric assembly of NDCBE.** Both EL3 and TMs in TMD are involved in dimerization.



**Supplementary Figure 5.** Ion time series for the 1µs MD trajectories. a Aggregated ion time series for the four 1 µs MD trajectories (Trials 1-4 in Table S1). The time series for  $CO_3^{2^-}$ , Na<sup>+</sup>, and Cl<sup>-</sup> are shown in red, blue, and green color respectively. The S1-ion distance is evaluated as the minimum distance between the center of an ion of a certain type (Na<sup>+</sup>, Cl<sup>-</sup>, HCO<sub>3</sub><sup>-</sup>, CO<sub>3</sub><sup>2^-</sup>) and the C<sub>a</sub> atom of residue Ala846 from site S1. **b** Same as (**a**), however only the first 200 ns of the trajectories are shown. The relative position of sites S1 and S2, as well as residues of importance (Glu608, Glu611) and the cavity exit with respect to the C<sub>a</sub> atom of residue Ala846 from site S1 are also indicated in the figure. All trajectories start from pre-equilibrated MD simulations with a Na<sup>+</sup>-CO<sub>3</sub><sup>2-</sup> pair in the S1/S2 binding pocket. The initial position of the Na<sup>+</sup>-CO<sub>3</sub><sup>2-</sup> differ in each trajectory. The CO<sub>3</sub><sup>2-</sup> and Na<sup>+</sup> from the ion pair eventually dissociate together from the binding pocket and exit the OF cavity. Afterwards Na<sup>+</sup> and Cl<sup>-</sup> ions from the solution occasionally permeate the cavity and reside briefly in the areas of site S2 (Cl<sup>-</sup>) or in the vicinity of two acidic residues on TM5, Glu608 and Glu611 (Na<sup>+</sup>).



Supplementary Figure 6. Correlation of initially bound Na<sup>+</sup> and CO<sub>3</sub><sup>2-</sup>. Aggregated ion correlation graph for the four 1 µs MD trajectories (Trials 1-4 in Table S1). S1-Na<sup>+</sup> and S1-CO<sub>3</sub><sup>2-</sup> are the distances of the initially bound Na<sup>+</sup> and CO<sub>3</sub><sup>2-</sup> ions to the C<sub>a</sub> atom of residue Ala846 from site S1, respectively. Na<sup>+</sup>-CO<sub>3</sub><sup>2-</sup> is the distance between the initially bound Na<sup>+</sup> and CO<sub>3</sub><sup>2-</sup> ions. As the linear dependence and the small Na<sup>+</sup>-CO<sub>3</sub><sup>2-</sup> distance in the OF cavity (the area enclosed by 20 Å S1-Na<sup>+</sup> x 20 Å S1-CO<sub>3</sub><sup>2-</sup>) suggest, the two ions from the initially bound Na<sup>+</sup>-CO<sub>3</sub><sup>2-</sup> dissociate together from the binding pocket.



Supplementary Figure 7. Ion time series for the 300 ns MD trajectories with different ion loads presented in Table S1. a Trial 5 and b Trial 6 from Table S1 for NDCBE loaded with a Na<sup>+</sup>-CO<sub>3</sub><sup>2-</sup> pair in the area of site S1. c NDCBE loaded with a single Cl<sup>-</sup> ion. d NDCBE loaded with a single CO<sub>3</sub><sup>2-</sup> ion. e NDCBE loaded with a single HCO<sub>3</sub><sup>-</sup> ion. f NDCBE loaded with a Na<sup>+</sup>-Cl<sup>-</sup> pair. g NDCBE loaded with a Na<sup>+</sup>-HCO<sub>3</sub><sup>-</sup> pair. h NDCBE loaded with a single Na<sup>+</sup> ion. The time series for CO<sub>3</sub><sup>2-</sup> and HCO<sub>3</sub><sup>-</sup> are shown in red while the time series for Na<sup>+</sup>, and Cl<sup>-</sup> are shown in blue and green, respectively. The S1-ion distance is evaluated as the minimum distance between the center of an ion of a certain type (Na<sup>+</sup>, Cl<sup>-</sup>, HCO<sub>3</sub><sup>-</sup>, CO<sub>3</sub><sup>2-</sup>) and the C<sub>a</sub> atom of residue Ala846 from site S1. The relative position of sites S1 and S2, as well as residues of importance (Glu608, Glu611) and the cavity exit with respect to the C<sub>a</sub> atom of residue A846 from site S1 are also indicated in the figure.



**Supplementary Figure 8. Membrane expression of NDCBE mutants.** a,d Representative experiments showing immunoblot analysis of cell-surface and cell-lysate expression of wt and mutant NDCBE proteins (n = 4 biologically independent experiments). The positions of molecular weight size markers (kDa) are shown on the left. Blot splicing is indicated with a vertical white line. Source data are provided as a Source Data file. **b,c,e,f** Densitometry analysis of cell-surface expression and the ratio of cell-surface to cell-lysate intensity of NDCBE protein expression (n = 4 biologically independent experiments). One-way ANOVA and Dunnett's test were used to compare multiple study group means with wt-NDCBE. Mutant NDCBE data was not statistically different from wt-NDCBE. Results are depicted as mean  $\pm$  SEM. Open circles represent individual data points. Source data are provided as a Source Data file.



Supplementary Figure 9. C $\alpha$  RMSD values calculated for the TMD portion of NDCBE obtained from the four 1  $\mu$ s MD simulations of NDCBE. (Trials 1-4 in Table S1).

## **Supplementary Tables**

**Supplementary Table 1.** Ion residence times (in ns) in the area of sites S1/S2 of NDCBE with different ion loads obtained from 1  $\mu$ s and 300 ns MD simulations. The ion time series for the trajectories listed in Table S1 are presented in Supplementary Figures 5 and 7.

lon load	MD length	Anion residence	Na⁺ residence	
		time, ns	time, ns	
Na⁺ + CO <sub>3</sub> ²⁻ (trial 1)	1 µs	48.5	48.5	
Na <sup>+</sup> + CO <sub>3</sub> <sup>2–</sup> (trial 2)	1 µs	67	66	
Na <sup>+</sup> + CO <sub>3</sub> <sup>2–</sup> (trial 3)	1 µs	13	13	
Na <sup>+</sup> + CO <sub>3</sub> <sup>2–</sup> (trial 4)	1 µs	152.5	152.5	
Na <sup>+</sup> + CO <sub>3</sub> <sup>2–</sup> (trial 5)	300 ns	153	153	
Na <sup>+</sup> + CO <sub>3</sub> <sup>2–</sup> (trial 6)	300 ns	191	191	
CΓ	300 ns	3	-	
CO32-	300 ns	74*	-	
HCO₃⁻	300 ns	9	-	
Na⁺ + Cl⁻	300 ns	142	142	
Na⁺ + HCO₃ <sup>−</sup>	300 ns	256	296**	
Na <sup>+</sup>	300 ns	-	12	

\*A portion of this simulation is presented in Supplementary Movie 2. The prolonged anion residence time here is the result of Na<sup>+</sup> from the solution binding to Asp800 (see Supplementary Figure 7d) and stabilizing the  $CO_3^{2^-}$  ion in analogy to the other Na<sup>+</sup>-  $CO_3^{2^-}$  simulations (Trials 1-6).

\*\* Upon HCO<sub>3</sub><sup>-</sup> departure from the binding pocket, Cl<sup>-</sup> ions from the solution permeate to area of binding sites S1/S2, attracted by the Na<sup>+</sup> ion which remains there (see Supplementary Figure 7g).

Supplementary Table 2. Free energy values (in kcal/mol) evaluated for different ion loads in site S1 of NDCBE.  $G_{hydr}$  is the absolute free energy of hydration for the individual anions.  $G_{site}$ is the absolute free energy of the interaction between the anion and the protein matrix (the Na<sup>+</sup> in the Na<sup>+</sup> bound systems is included in the protein matrix).  $\Delta G_{bind}$  is the difference between the free energies of hydration and the free energies of interaction with the protein matrix. More negative  $\Delta G_{bind}$  implies stronger ion binding to site S1 of NDCBE.

	Na⁺- Cl⁻	Na⁺- HCO <sub>3</sub> <sup>−</sup>	Na <sup>+</sup> -CO <sub>3</sub> <sup>2-</sup>	CI⁻	HCO₃ <sup>−</sup>	CO32-
G <sub>hydr</sub>	-	-	-	-81.3	-86.5	-272.5
G <sub>site</sub>	-83.1	-89.8	-291.3	-75.1	-86.1	-282.4
$\Delta G_{\text{bind}}$	-1.8	-3.3	-19.8	6.2	-0.4	-9.9