



# The N-terminal cytoplasmic region of NCBE displays features of an intrinsic disordered structure and represents a novel target for specific drug screening

Kaare Bjerregaard-Andersen<sup>1,2</sup>, Harmonie Perdreau-Dahl<sup>1</sup>, Hanne Guldsten<sup>1</sup>, Jeppe Praetorius<sup>3</sup>, Jan K. Jensen<sup>2</sup> and Jens P. Morth<sup>1,4\*</sup>

<sup>1</sup> Norwegian Centre for Molecular Medicine, University of Oslo, Oslo, Norway

<sup>2</sup> Danish Chinese Centre for Cancer and Proteases, Department for Molecular Biology and Genetics, University of Aarhus, Aarhus, Denmark

<sup>3</sup> Department of Biomedicine, University of Aarhus, Aarhus, Denmark

<sup>4</sup> Institute for Experimental Medical Research, Oslo University Hospital, Oslo, Norway

## Edited by:

Ebbe Boedtkjer, Aarhus University, Denmark

## Reviewed by:

Seth L. Alper, Beth Israel Deaconess Medical Center, USA

Mark D. Parker, Case Western Reserve University, USA

## \*Correspondence:

Jens P. Morth, Membrane Transport Group, Nordic EMBL Partnership, Norwegian Centre for Molecular Medicine, University of Oslo, Gaustadalleen 21, PO Box 1137 Blindern, 0318 Oslo, Norway  
e-mail: j.p.morth@ncmm.uio.no

The sodium dependent bicarbonate transporter NCBE/NBCn2 is predominantly expressed in the central nervous system (CNS). The highest protein concentrations are found in the choroid plexus. The primary function of this integral plasma membrane transport protein is to regulate intracellular neuronal pH and also probably to maintain the pH homeostasis across the blood-cerebrospinal fluid barrier. NCBE is predicted to contain at least 10 transmembrane helices. The N- and C- termini are both cytoplasmic, with a large N-terminal domain (Nt-NCBE) and a relatively small C-terminal domain (Ct-NCBE). The Nt-NCBE is likely to be involved in bicarbonate recognition and transport and contains key areas of regulation involving pH sensing and protein-protein interactions. Intrinsic disordered protein regions (IDPRs) are defined as protein regions having no rigid three-dimensional structure under physiological conditions. They are believed to be involved in signaling networks in which specific, low affinity, protein-protein interactions play an important role. We predict that NCBE and other SoLute Carrier 4 (SLC4) family members have a high level of intrinsic disorder in their cytoplasmic regions. To provide biophysical evidence for the IDPRs predicted in Nt-NCBE, we produced pure (>99%), recombinant Nt-NCBE using *E. coli* as the expression host. The protein was used to perform differential scanning fluorescence spectroscopy (DSF), in order to search for small molecules that would induce secondary or tertiary structure in the IDPRs. We expect this to assist the development of selective pharmaceutical compounds against individual SLC4 family members. We have also determined a low resolution (4 Å) X-ray crystal structure of the N-terminal core domain. The N-terminal cytoplasmic domain (cdb3) of anion exchanger 1 (AE1) shares a similar fold with the N-terminal core domain of NCBE. Crystallization conditions for the full-length N-terminal domain have been sought, but only the core domain yields diffracting crystals.

**Keywords:** SLC4, intrinsic disorder, drug screen, NCBE, IDP, bicarbonate

## INTRODUCTION

In general, regulation of intra- and extracellular pH is crucial for cellular function, because most metabolic enzymes have optimal functionality within the narrow pH range of 6.8–7.4. The main buffering system is based on carbonic acid and bicarbonate, which effectively minimizes the impact of short-term pH deviations (Cordat and Casey, 2009). Most mammalian cells depend on selective transport of bicarbonate across the plasma membrane by specific integral membrane proteins belonging to the SoLute Carrier 4 (SLC4) protein family. Ten members of this family are involved in maintaining pH homeostasis (Parker and Boron, 2013). The SLC4 protein family members are divided into three major functional classes; (i) Na<sup>+</sup>-independent Cl<sup>-</sup>/HCO<sub>3</sub><sup>-</sup> exchangers, (ii) Na<sup>+</sup>, HCO<sub>3</sub><sup>-</sup> cotransporters and, (iii) Na<sup>+</sup> driven Cl<sup>-</sup>/HCO<sub>3</sub><sup>-</sup> exchangers (Boron et al., 2009). The SLC4 family also exhibits differences associated

with electrogenic/electroneutral transport characteristics and stilbene derivative sensitivity. Stilbene derivatives, DIDS and SITS, are the classic inhibitors that target the Sodium Bicarbonate Cotransporter (NBC) family. Other potent inhibitors are diBAC oxonol (Liu et al., 2007), S0859 (Ch'en et al., 2008), tenidap and benzamil (Ducoudret et al., 2001). Members of NBC family have earlier been highlighted as potential therapeutic targets of ischemic reperfusion. Inhibition of the electrogenic Na<sup>+</sup>/HCO<sub>3</sub><sup>-</sup> cotransporter (NBCe1-b), found in the heart, showed a decrease in ischemic injury (Khandoudi et al., 2001). The family may also be targets to reduce tumor progression in, for example, breast cancer (Boedtkjer et al., 2012). NCBE (NBCn2/SLC4A10) is an electroneutral Na<sup>+</sup>-dependent HCO<sub>3</sub><sup>-</sup> transporter (Wang et al., 2000; Giffard et al., 2003; Parker et al., 2008; Damkier et al., 2010) found predominantly in the brain (Liu et al., 2010) with particular high expression levels found in the choroid plexus (Praetorius

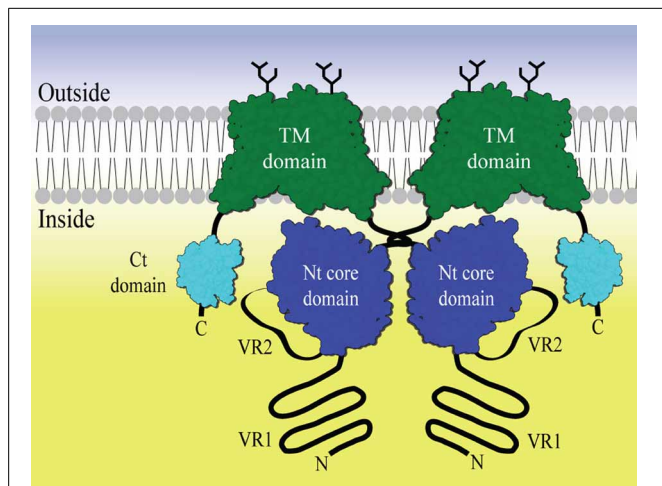
et al., 2004). NCBE is DIDS-sensitive (Wang et al., 2000). A recent report indicates that NCBE is a potential neuronal drug target. A study performed in mice showed that disruption of the *SLC4A10* gene could prevent fatal epileptic seizures (Jacobs et al., 2008) indicating drugs counteracting this particular NBC might be usable in anticonvulsive therapy. To date no unique inhibitor of NCBE transport activity is known. The development of specific inhibitors against NCBE, or of any SLC4 members in general, would benefit our understanding of the sodium and bicarbonate transport mechanism and perhaps lead to novel drugs that selectively target specific SLC4 family members.

The general topology of the transporters includes a large amino-terminal (Nt) cytoplasmic domain, a transmembrane domain including extracellular regions of which certain loops can be glycosylated and a smaller carboxy-terminal (Ct) cytoplasmic domain (Figure 1). In general, very little structural data is available for members of this protein family. Two homology models of the anion exchanger 1 (AE1) transmembrane domain based on the *E. coli* chloride-proton channel (ClC) (Bonar et al., 2013) and the uracil transporter (UraA) (Barneaud-Rocca et al., 2013) have been proposed, although there is no significant similarity of AE1 primary structure to these. Historically, the large N-terminal cytoplasmic domain (Nt) has been the target for functional and structural characterization. It is known to be involved in the regulation of transport activity. For example an auto inhibitory domain is found in NBCe1-B/C (McAlear et al., 2006; Lee et al., 2012). Protein–protein and metabolite–protein interactions have also been reported, for example glyceraldehyde-3-phosphate dehydrogenase (GAPDH) with erythrocyte AE1 (Chu and Low, 2006) and inositol-1,4,5-trisphosphate (IP3) receptors binding protein released with IP3 (IRBIT) with NBCe1-B (Shirakabe et al., 2006). However, knowledge of the molecular basis for NBC function and

regulation is still limited. The cytoplasmic domain of human AE1 (cdb3) has been crystallized and the structure determined at 2.6 Å resolution. Only residues 55–356 out of the 379 residues available in the crystallized construct were visible in the electron density. The structure forms a dimer, with a dimerization arm formed by residues 317–356. A central elongated  $\beta$ -sheet is present in both monomers surrounded by several  $\alpha$ -helices (Zhang et al., 2000). The cytoplasmic Nt domain from NBCe1-A, an electrogenic  $\text{Na}^+$ ,  $\text{HCO}_3^-$  co-transporter, has been expressed and purified in *E. coli* (Gill and Boron, 2006), but the structure determination is problematic (Gill et al., 2013). The SLC4 family proteins are, like the ClC channels, believed to be homodimers. Chemical cross-linking followed by SDS-PAGE indicates that membrane bound full-length NBCe1-A forms dimers and traces of tetramers (Kao et al., 2008). AE1 is believed to form both dimers and tetramers (Jiang et al., 2013). Nt-NBCe1-A is able to form monomers, dimers and tetramers in solution (Gill, 2012). It has been observed that the oligomeric state is pH dependent and it is postulated that conformational changes occur within the monomers (Zhang et al., 2000; Gill, 2012). It is therefore possible that the activity of the SLC4 family proteins is related to changes in the oligomeric state and that this happens as a consequence of changes in pH.

Protein function is traditionally thought to depend primarily on the chemical environment created by the tertiary structure of the macromolecule, and it is believed that the composition of the primary structure supports certain secondary and tertiary structures (or folds) of the macromolecule. In recent years, attention has turned to proteins containing non-structured regions—and, in particular, to the function of these regions. These intrinsically disordered protein regions (IDPRs) (Uversky, 2013) are abundant in protein sequences of all organisms, and the number of IDPRs increases as the complexity of the organism increases. In general, the highest percentage of the predicted disorder is found in eukaryotes (Xue et al., 2012) probably because regulation needs to be tighter and is intrinsically more complex than in other domains of life. Experimental identification of IDPRs is tedious and large scale analyses rely on *in silico* predictions from the primary structures. Today disordered regions can be predicted with between 80 and 90% accuracy (He et al., 2009). IDPRs are especially common as mediators of protein–protein interactions. Many are found in proteins related to signaling and are therefore important in understanding diseases such as neurodegenerative disease and cardiovascular disease (Metallo, 2010). For these reasons they represent a class of high-value pharmaceutical targets but due to their poorly characterized properties they have not yet been investigated to any great extent.

In polarized epithelia, the SLC members can reside either in the apical or in the basolateral plasma membranes depending on the primary structure and perhaps the cellular membrane sorting, or membrane protein retaining machineries. To date it is unknown which specific intracellular motifs are involved in the membrane targeting of the SLC4 proteins. Motifs in the Nt or Ct most probably carry a signal for the polarized insertion of the transporters into the plasma membrane (Toye et al., 2004), but extracellular loop glycosylation or even NBC-binding adaptor proteins may also control membrane targeting, as is the case for the  $\text{Na}^+$ ,  $\text{K}^+$ -ATPase (Mellman and Nelson, 2008).



**FIGURE 1 | The general domain structure of a SLC4 bicarbonate transporter represented as a dimer.** The Transmembrane (TM) domain is shown in green. Dimerization is driven by dimerization arms close to the Nt core domains (blue) in the cytoplasm. Variable region 2 (VR2) is drawn as an extended loop from the core domain. Variable region 1 (VR1) is shown as an N-terminal extension from the core domain. The C-terminal region often contains a small PDZ domain (Kennedy, 1995) (cyan). Glycosylations are shown as black sticks.

The function of the Nt cytoplasmic domain, in concert with the transmembrane domain, has been studied in AE1 and NBCe1. Unlike AE1 (Groves and Tanner, 1992), NBCe1 seems unable to function without this domain (Espiritu et al., 2006; McAlear et al., 2006). The structure of the link between the cytoplasmic domain and transmembrane domain may therefore be very class specific. Specifically, for NBCe1, a substrate channel formed by the Nt cytoplasmic domain has been suggested (Chang et al., 2008). The Nt cytoplasmic domain, located close to the membrane, consists of two constant regions (CR1 and CR2) separated by variable region 2 (VR2) (Boron et al., 2009). The core domain is N-terminally flanked by variable region 1 [VR1 or Nt-appendage (Parker and Boron, 2013)] and is conserved in all members of the SLC4 family. In this report we have focused on VR1 and VR2 and show that they are likely to be intrinsically disordered in all the SLC4 family members, but do, however, share distinct recognizable signatures. We believe it is this intrinsic disorder that has prevented thorough structural characterization of the cytoplasmic domain. However, we hypothesize that the intrinsically disordered regions found in VR1 and VR2 may be used as specific drug targets to distinguish between the individual SLC4 members and that together they form a distinct structural “fingerprint” for each individual SLC4 transporter. We have identified buffer conditions and a purification protocol to produce pure and monodispersed solutions of both the complete cytoplasmic domain of NCBE and what we refer to as the core domain, without VR1. We also present a novel screening approach, developed to identify small interacting molecules. The screening procedure is potentially applicable as an initial screen to identify leads targeting the VR1 and/or the VR2 regions which are present in the cytoplasmic domain of NCBE and other SLC4 members.

## MATERIALS AND METHODS

### CLONING AND EXPRESSION

The gene encoding *SLC4A10* isoform 3 from *Rattus norvegicus* (rb2NCBE) was purchased from GenScript. Primers were designed for ligation independent cloning (LIC) into the pET-46 Ek/LIC vector (Novagen) for the production of an N-terminal hexa-histidine tagged fusion protein. For construct Nt-NCBE residues 2–421 and the construct Nt-NCBE- $\Delta$ D1 residues 96–396. A tobacco etch virus (TEV) protease recognition site was included in the primer located between the hexa-histidine and protein sequence of interest.

Nt-NCBE and Nt-NCBE- $\Delta$ D1 were expressed in *E. coli* strain Rosetta2 (Novagen) expressing seven additional tRNAs to compensate for the codon bias in the mammalian gene during prokaryotic expression. An O/N culture of 100 mL LB media with chloramphenicol (50  $\mu$ g/mL) and ampicillin (34  $\mu$ g/mL) from a single colony of transformed Rosetta2 cells was made for inoculation of fermentation in 1.5 l of LB media (antibiotics added +1 mL of polyethylene glycol). At optical density measured at 600 nm ( $OD_{600}$ ) of 1.0, the media was cooled to 18°C and induced with 1 mM (IPTG). The cells were harvested after 16 h yielding ~10 g of wet-weight cells. Cell cultures were grown using a LEX HT Bioreactor (Harbinger). Expression levels were visualized by SDS-PAGE with samples taken before and after induction. Protein identity was confirmed using tandem mass-spectrometry.

### PROTEIN PURIFICATION

For each gram of frozen cell pellet 10 mL lysis buffer (50 mM NaCl, 100 mM Tris-HCl pH 8.0, 1  $\mu$ g/mL DNaseI, 1 mM PMSF, 1x EDTA-free inhibitor cocktail tablet (Roche), 5 mM  $\beta$ -ME and 40 mM imidazole) was used for resuspension. Resuspended cells were lysed by forcing the suspension three times through high-pressure homogenizer (Emulsiflex C3, Avestin). Cell lysate was kept on ice. The cell lysate was ultracentrifuged for 1 h at 200,000 g to remove cell debris and membrane fractions. The cleared lysate was loaded onto a Ni-NTA column (GE) in equilibration buffer (50 mM NaCl, 100 mM Tris-HCl pH 8.0, 40 mM imidazole and 5 mM  $\beta$ -ME) using an Äkta Prime FPLC system. The column was washed with 10 column volumes (CV) of equilibration buffer. An additional wash with 10 CV of high-salt buffer (500 mM NaCl, 100 mM Tris-HCl pH 8.0, 40 mM imidazole and 5 mM  $\beta$ -ME) was performed. The protein was released with elution buffer (50 mM NaCl, 100 mM Tris-HCl pH 8.0, 200 mM imidazole and 5 mM  $\beta$ -ME). The elution peak fractions were collected and checked by SDS-PAGE analysis. The concentration of the eluted protein was measured using a NanoDrop Spectrophotometer (Thermo Scientific). 1:50 molar ratio of GFP-tagged TEV-protease was added and the suspension was dialyzed O/N against dialysis buffer (50 mM NaCl, 100 mM Tris pH 8.0, 5 mM  $\beta$ -ME and 40 mM imidazole). To remove uncleaved protein and the TEV protease, the suspension was passed over the pre-equilibrated Ni-NTA column and the flow-through containing the cleaved target protein was collected.

The protein was then concentrated using a 10 MWCO Viva spin filter (GE Healthcare). The concentrated sample was loaded on a Superdex 200 HiLoad 10/600 prep-grade column as a final purification step and for determination of homogeneity using an Äkta Purifier HPLC system. Fractions containing pure target protein were collected and kept at 4°C for further analysis. Purity was assessed using SDS-PAGE (Supplementary Figure S1).

### CRYSTALLIZATION AND STRUCTURE DETERMINATION

Crystallization conditions were screened using commercially available kits. Suitable crystals were only identified for the construct NCBE- $\Delta$ D1 in the condition 100 mM potassium thiocyanate (KSCN) and 20% PEG3350. This condition produced showers of crystals of limited size (50  $\times$  50  $\times$  50  $\mu$ m). Crystals were flash cooled to 100 K in the well solution supplemented with 10% ethylene glycol. Diffraction data was collected at the European Synchrotron Radiation Facility (ESRF) ID29 beamline. Data collection strategy was determined using iMOSFLM (Battye et al., 2011) and the data were processed at the beamline using XDS (Kabsch, 2010) and PHENIX (Adams et al., 2002). A molecular replacement solution was found with the search model for Nt-AE1 [PDB ID 1HYN, (Zhang et al., 2000)]. A homology model of the NCBE- $\Delta$ D1 based on 1HYN was created with Modeler (Eswar et al., 2007) and used to perform rigid body refinement against the experimental data. Data and refinement statistics are collected in Supplementary Table S2.

### DIFFERENTIAL SCANNING FLUORESCENCE (DSF)

Differential scanning fluorescence (DSF) measurements were carried out using a 7900 HT Fast Real-Time PCR system (Applied



Biosystems) using  $\lambda_{\text{ex}} = 462 \text{ nm}$  and  $\lambda_{\text{em}} = 569 \text{ nm}$  (Hawe et al., 2008). The protein stock was 1 mg/mL and contained 100 mM NaCl, 20 mM HEPES pH 7.2, and 0.5 mM TCEP. A prescreening of dye (Sypro Orange, Novagen) and protein was done in order to find the best dye and protein concentration. This was found to be 0.5 mg/mL protein and 100x Sypro Orange. Measurements were carried out in 96 well plate format sealed to prevent evaporation during experiment. Temperature increments of 1°C/min and a temperature range of 20–90°C were used (Niesen et al., 2007). Small molecules were supplied by the Silver Bullets Screen (Hampton Research) and were diluted five-fold for use. Sypro Orange was added as last component. Data treatment was performed using MS Excel and GraphPad Prism 6. A data sheet containing information on individual screen conditions, with hit conditions highlighted, can be found in the Supplementary Material.

### PREDICTION OF DISORDER IN THE SLC4 FAMILY

Predictions of SLC4 protein disorder were made using the PONDR-FIT metapredictor (Xue et al., 2010). The predictor is the newest development of the PONDR (Predictors Of Natural Disordered Regions) series and shows good accuracy. Query sequences from human SLC4 family proteins were obtained from the UNIPROT sequence database. All isoforms that have been identified at the protein level were chosen for analysis. The average disorder of the whole family, subfamilies and individual isoforms were calculated using GraphPad Prism 6.

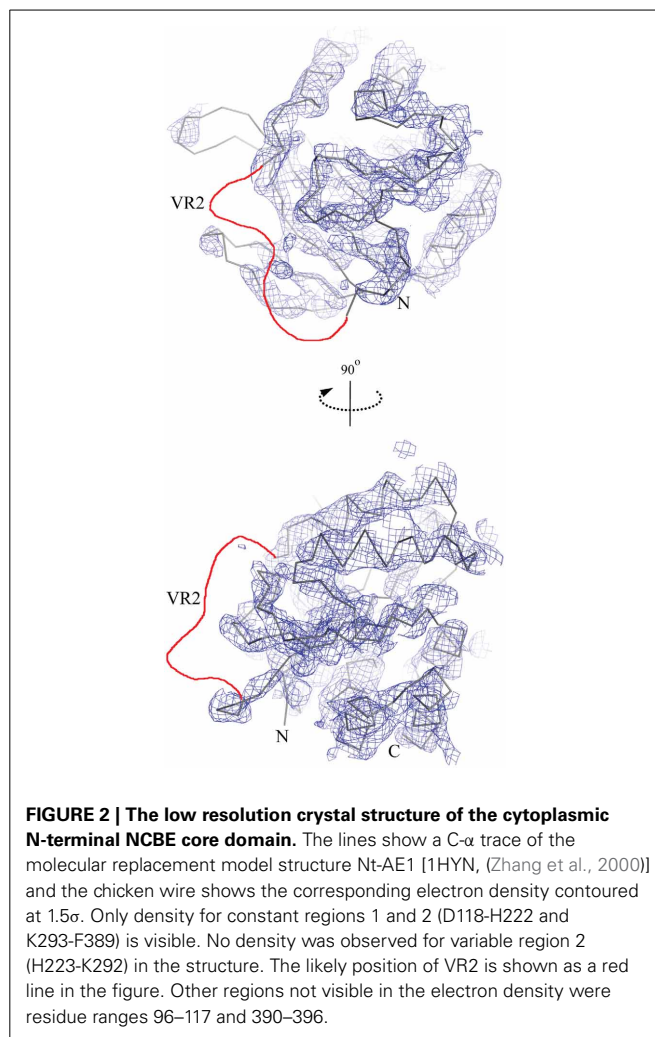
## RESULTS

### CRYSTAL STRUCTURE OF N-TERMINAL NCBE CORE DOMAIN

A crystallization condition was identified for NCBE- $\Delta$ D1, and the crystal structure of the Nt-NCBE core domain was determined at 4.0 Å resolution (Figure 2). At this resolution, ion binding sites and side-chain positions cannot be resolved, only the backbone density is visible for residues 116–222 and 292–389 of the NCBE- $\Delta$ D1 construct containing residues 96–396. No electron density is observed for variable region 2 (VR2). There is only one single molecule in the asymmetric unit and no clear dimer with a symmetry related molecule could be found in the crystal lattice. Extensive crystallization screening was undertaken for Nt-NCBE, but no crystallization condition has been identified even after more than 5000 unique conditions have been screened. We suspect that the intrinsic disorder of the Nt-NCBE construct may have prevented crystallization.

### PREDICTION OF DISORDERED REGIONS

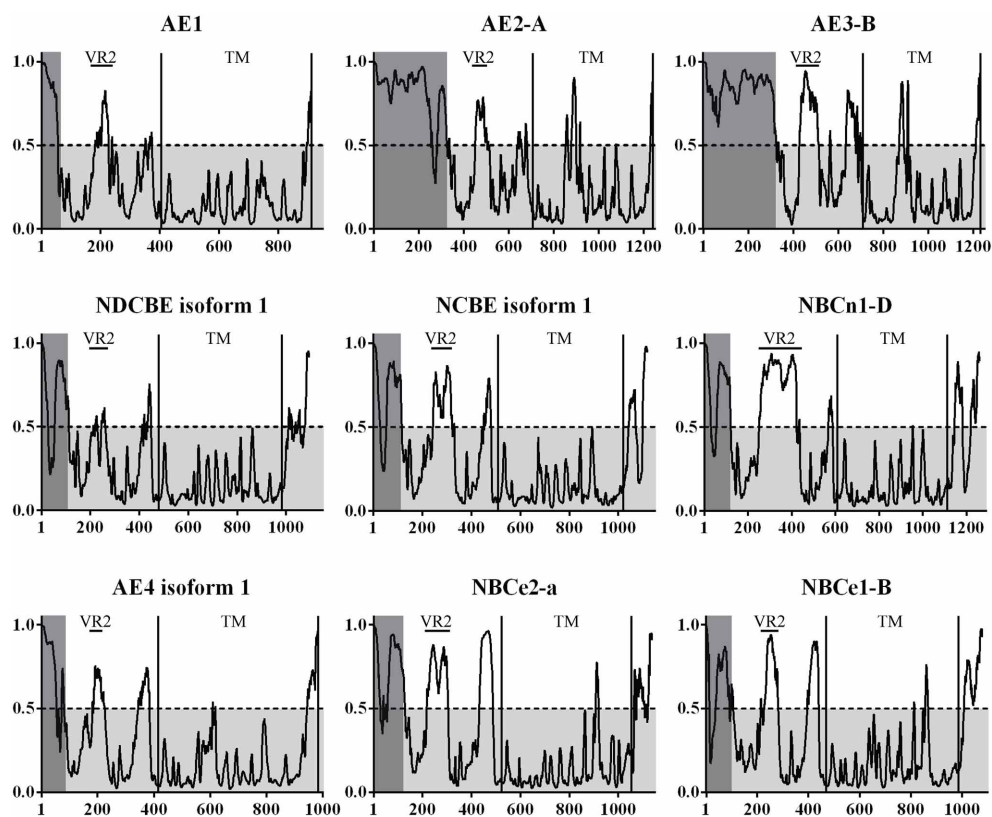
A complete and large scale disorder prediction of the SLC4 family was undertaken, as the crystallization trials and partial structure indicated that large disordered regions could be present. All isoforms confirmed at protein level for SLC4A 1-5 and SLC4A 6-11 were analyzed. Selected isoforms are presented in Figure 3. The transmembrane (TM) domain are the most conserved part of the SLC4 family and sequences were aligned to the starting residue of transmembrane helix 1 (TMH1). A high probability of disorder was found for variable region 1 (VR1) of all SLC4 family proteins. Likewise, a rather high probability of disorder was found for variable region 2 [VR2, (Boron et al., 2009)] of the core domain and for the flexible linker region occurring before the TMH1.



**FIGURE 2 | The low resolution crystal structure of the cytoplasmic N-terminal NCBE core domain.** The lines show a C- $\alpha$  trace of the molecular replacement model structure Nt-AE1 [1HYN, (Zhang et al., 2000)] and the chicken wire shows the corresponding electron density contoured at 1.5 $\sigma$ . Only density for constant regions 1 and 2 (D118-H222 and K293-F389) is visible. No density was observed for variable region 2 (H223-K292) in the structure. The likely position of VR2 is shown as a red line in the figure. Other regions not visible in the electron density were residue ranges 96–117 and 390–396.

### STABILIZATION OF DISORDERED REGION OF EXTREME N-TERMINUS OF NCBE BY SMALL MOLECULES

DSF spectroscopy, commonly known as the thermofluor method, is well documented for its use in determining optimal buffer composition and compounds for the stabilization of proteins (Ericsson et al., 2006; Vedadi et al., 2006). Briefly, a hydrophobic molecule (e.g., Sypro Orange) is added to the protein and, optionally, a stabilizing compound. The solution is then heated—as tertiary structure elements reach their melting point, they fall apart and expose their hydrophobic cores which interact with Sypro Orange. The increasing hydrophobic environment enhances the quantum yield of the Sypro Orange chromophore resulting in a dramatic increase of the fluorescent signal. The method is therefore suited for measuring the possible induction of a tertiary structure by small molecules and for high-throughput screening (Pantoliano et al., 2001). Assuming that VR1 and VR2 of NCBE are IDPRs, we would like, if possible, to introduce some tertiary structure to these regions for two reasons. Firstly, to identify compounds that would stabilize a more rigid fold and in turn aid the crystallization process. Secondly, this technique might result in a method for finding leads against pharmacologically



**FIGURE 3 | The prediction of disorder in the bicarbonate transporters of the SLC4 family.** Probability plots of the longest isoforms of SLC4A1-5 and SLC4A7-10 are presented here. A total of 35 isoforms were analyzed. On average, the highest probability of disorder is found in two regions of the N-terminal cytoplasmic domain in the SLC4 family. A high level of disorder is predicted for the variable regions 1 and 2 [VR1, VR2 (Boron et al., 2009)]. VR1 is highlighted in gray. This region was not visible in the Nt-AE1 structure and no experimental structural information exists for it. VR2 is marked on the plots and was modeled in the Nt-AE1 structure. In that structure the high

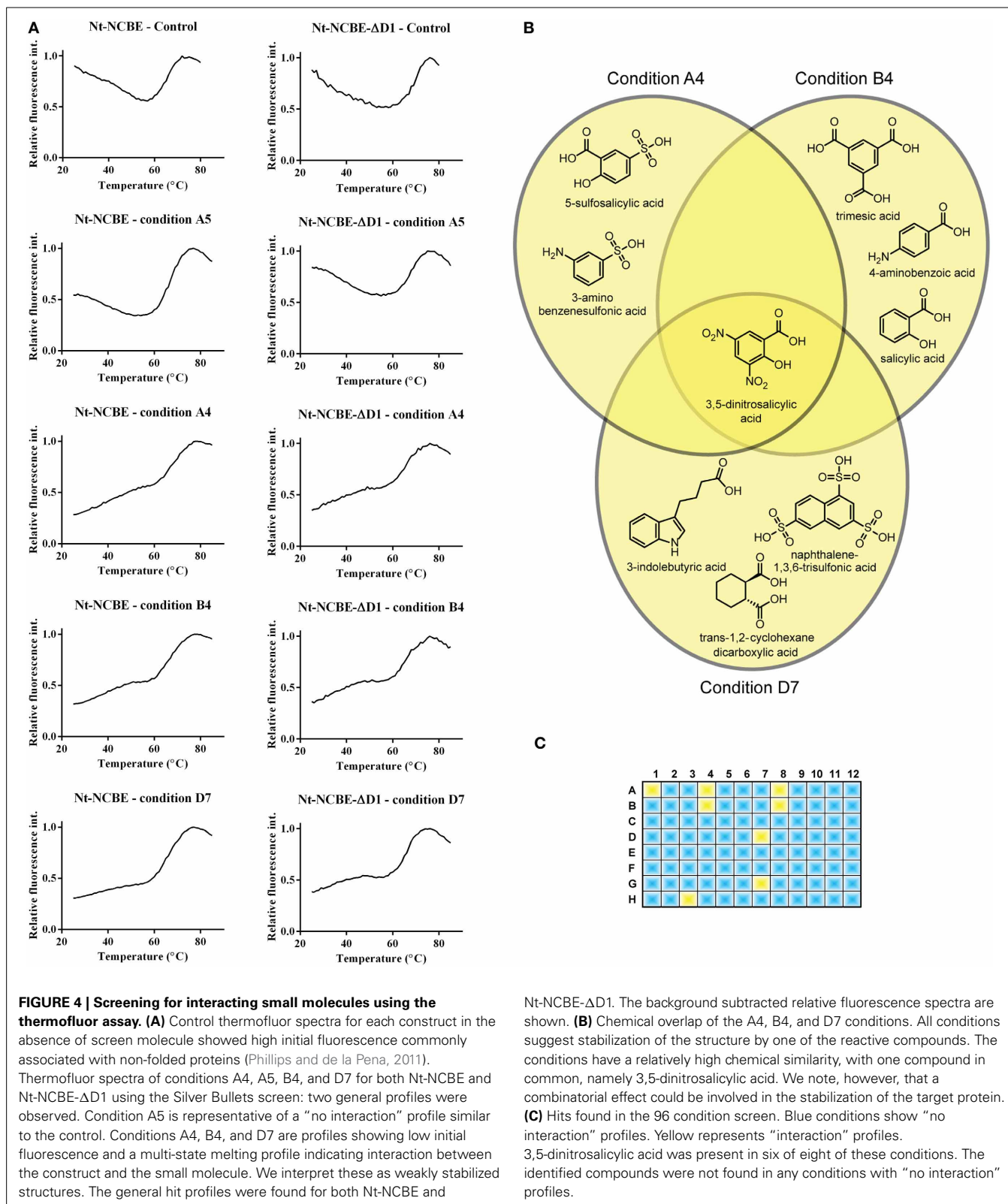
temperature factors and the random coil structure of the region (residues 166–213) support the disorder prediction. VR1 and VR2 are likely involved in protein-protein interactions and regions could constitute a structural “fingerprint” of each individual SLC4 transporter. UniProt accession codes for presented plots are: AE1 (SLC4A 1, P02730), AE2-A (SLC4A 2, P04920-1), AE3-B (SLC4A 3, P48751-1), NDCBE isoform 1 (SLC4A 8, Q2Y0W8-1), NCBE isoform 1 (SLC4A 10, Q6U841-1), NBCn1-D (SLC4A 7, Q9Y6M7-7), AE4 isoform 1 (SLC4A 9, Q96Q91-1), NBCe2-a (SLC4A 5, Q9BY07-1), NBCe1-B (SLC4A 4, Q9Y6R1-1).

interesting IDPRs. The 96 condition Silver Bullets combinatorial library (HR2-096, Hampton Research) provides ~400 unique molecules for this purpose. Overlapping groups of similar compounds are tested in each experiment. In general, a relatively high initial fluorescence was observed for both NCBE constructs in the absence of any screen molecules as well as in the presence of the small molecules in the screen. A high initial fluorescence is often seen for non-globular folded proteins and exposed hydrophobic residues (Phillips and de la Pena, 2011) are anyway expected for IDPRs. The control sample contained 100 mM NaCl and 20 mM HEPES pH 6.8. This melting profile, exemplified by condition A5 in **Figure 4A**, is considered a non-hit melting profile as we do not see a difference in overall profile from adding small molecules to the protein. A significant change in profile was observed for a number of conditions, for example by conditions A4, B4, and D7 in **Figure 4A** (third, fourth, and fifth rows). The initial fluorescence starts lower than in the control and the spectra indicate a more complex transition model. We regarded such a profile as indicating an interaction (a hit). Fluorescence backgrounds for the dye and screen molecules alone were subtracted from the

spectra. From chemical overlap of conditions we can identify the molecule that is likely to cause the effect (**Figure 4B**). In total eight hit-profiles were identified from the screen (**Figure 4C**). Six of the eight hit conditions containing 3,5-dinitrosalicylic acid. Three of the eight hit conditions contained 4-aminobenzoic acid. We, therefore, regard these molecules as candidates for further investigation. The hit molecules were present in concentrations of 0.05% w/v. The remaining spectra and chemical alignments are shown in Supplementary Figure S3.

## DISCUSSION

In general, the predicted levels of structural disorder are high among SLC4 members in comparison to other eukaryotic membrane proteins (Supplementary Figure S4), and therefore we wish to further discuss the IDPRs of that family. The Na<sup>+</sup>-K<sup>+</sup> ATPase  $\alpha$ -1 subunit (Supplementary Figure S4A) and the Cystic Fibrosis Transmembrane conductance Regulator (CFTR) (Supplementary Figure S4B) exhibit disorder profiles that show short spans of disorder (e.g., loop regions) and no extended disordered regions. When analyzing the crystal structure of the



Na<sup>+</sup>-K<sup>+</sup> ATPase α-1 subunit an important loop region can be identified in the disorder profile. The C-terminal of the CIC-channel, however, (Supplementary Figure S4C), exhibits extended disordered regions. This may relate directly to the CIC

channel ball-and-chain mechanism of regulation (Grunder et al., 1992). A similar pattern to that of the CIC channel is found for SLC26A6 (Supplementary Figure S4D), a transporter co-occurring with CFTR (He et al., 2010). However, no evidence

for a ball-and-chain mechanism for SLC26A6 has been reported. NCBE isoform 1 and NBCn1-D (Figure 3) are shown to illustrate the extent of the disordered regions found in the N-termini of these proteins. The predicted degree of disorder seems to correlate with the fact that many protein–protein interactions have been established for the intracellular domains of SLC4 family proteins. Description of a disordered “fingerprint” formed by VR1 and VR2 of the N-terminal domain could enhance our understanding of the nature of these interactions. We believe that this fingerprint will constitute a unique identifier for each family member. Large variations in the predicted disorder levels are also predicted between isoforms of SLC4 family members (e.g., NBCn1), and it seems that IDPRs may be present at sites of alternative splicing.

IDPRs have previously been thought of as assemblies fluctuating between a 3D structure, maybe statically disordered, and a dynamically disordered state (Tsvetkov et al., 2008). Inhibitors of these fluctuations could be other proteins, metal ions, small molecules and post-translational modifications. The seeding of secondary and tertiary structures can be measured using a range of biophysical methods e.g., circular dichroism (CD), isothermal titration calorimetry (ITC), electron paramagnetic resonance (EPR), and nuclear magnetic resonance (NMR). These are methods applied on a sample to sample basis and are not suited for high-throughput screening (HTS). We show that the thermofluor method provides a high-throughput tool for screening of tertiary structure induction of multiple samples. This method of screening for small molecule interactions with IDPRs could potentially be useful, in conjunction with orthogonal methods, as a first approach to identify pharmacological leads. Our data do not and cannot provide a molecular description of the interaction. However, a lower initial relative fluorescence and a multi-state transition, indicating a stabilization of the structure, provide a selection criterion for HTS methods. Molecules identified by chemical overlap of components of the screen will be candidates for studies using serial methods (e.g., CD or ITC) for further biophysical characterization of the interaction. As with all screening, false positives may occur and many pharmacologically active molecules contain chemical moieties that can absorb light or fluoresce. For this reason we believe it is important to subtract the spectra of the screen molecules and the dye. Fluorescence from screen molecules may interfere with the spectrum if the emission peak is close to that of the used dye. Absorbance may dampen the signal, but we assume this effect is rather small due to the very short light path in the experimental setup and the signal strength usually gained from the dye.

The low-resolution crystal structure represents a first step toward a structural understanding of NCBE. It is also the first crystal structure of a SLC4 protein to be reported since 2006. We show that the fold of the N-terminal core domain is conserved between Nt-NCBE- $\Delta$ D1 and Nt-AE1. In AE1 no electron density was observed for residues 1–55 corresponding to VR1 in the crystal structure. VR2 was modeled, but found to be devoid of secondary structure and to have high temperature factors, indicating a large degree of mobility (Zhang et al., 2000) (Supplementary Figure S5). The structure of the fingerprint region remains to be determined, however, the nature of this region as presented in

this paper requires methods for studying protein dynamics rather than a crystallographic approach.

Conservation of biological function is often translated directly from sequence conservation at the level of the individual residues, i.e., as changes determined by the chemistry available to the side chain in a defined environment. At the same time the evolutionary selection pressure may work on other levels to conserve protein function. Selection pressure on IDPRs may also work at the biophysical level and to keep a balance between retaining an ability to form a defined structure in the presence of another protein or a metabolite or slipping into a state that cannot be ordered. It could therefore be beneficial to preserve the disorder to prevent order—that is to prevent the formation of secondary structural elements that would hinder the recognition process. We believe there must be such selection pressure to keep the disorder, to prevent the energetically favorable secondary or tertiary structures from forming and thus prevent or induce biological function, and that this is preserved in the fingerprint region. To study this, the experimental approach must be designed to accommodate the specific biophysical characteristics of IDPRs. A common approach to probe protein function is by mutation of key residues—often to alanine to remove amino acid side chain properties. However, in the case of IDPRs one should consider the propensity of alanine to form  $\alpha$ -helices when designing a mutational experiment.

The SLC4 family of membrane transporters has proven to be a challenging group of proteins to target for structural and biophysical studies. We hope to elucidate biophysical properties of this group of proteins in order to target individual members selectively with novel inhibitors. The development of these could greatly benefit the field and become valuable tools for further *in vivo* investigations.

## AUTHOR CONTRIBUTIONS

Kaare Bjerregaard-Andersen performed the crystallization experiments and structure determination, performed IDPR predictions, DSF screening, analyzed the data and wrote the paper. Harmonie Perdreau-Dahl and Hanne Guldsten assisted with cloning and purification and analyzed data. Jeppe Praetorius assisted with the data analysis and general physiological comparison. Jan K. Jensen assisted with the biophysical analysis of the DSF data. Jens P. Morth supervised the project, designed experiments and analyzed the data and wrote the paper. All authors commented on the paper.

## ACKNOWLEDGMENTS

The crystallographic data collection experiments were performed on the ID-29 beam line at the European Synchrotron Radiation Facility (ESRF), Grenoble, France. We are grateful to Drs. Daniele de Sanctis and Christoph Mueller-Dieckmann at ESRF for providing assistance in using beam line ID-29. This study was supported by a grant from Lundbeckfonden (Kaare Bjerregaard-Andersen and Jens P. Morth) and the Norwegian research council (Harmonie Perdreau-Dahl and Jens P. Morth). We are grateful to Dr. Paul A. Tucker, for valuable comments and proofreading of the manuscript.



## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fphys.2013.00320/abstract>

## REFERENCES

- Adams, P. D., Afonine, P. V., Grosse-Kunstleve, R. W., Moriarty, N. W., Sauter, N. K., Zwart, P. H., et al. (2002). Automated structure determination with phenix. *Acta Crystallogr. D Biol. Crystallogr.* 58(Pt 11), 1948–1954. doi: 10.1107/S090744902016657
- Barneaud-Rocca, D., Etchebest, C., and Guizouarn, H. (2013). Structural model of the anion exchanger 1 (SLC4A1) and identification of transmembrane segments forming the transport site. *J. Biol. Chem.* 288, 26372–26384. doi: 10.1074/jbc.M113.465989
- Battye, T. G. G., Kontogiannis, L., Johnson, O., Powell, H. R., and Leslie, A. G. W. (2011). iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr. D Biol. Crystallogr.* 67(Pt 4), 271–281. doi: 10.1107/S0907444910048675
- Boedtker, E., Bunch, L., and Pedersen, S. F. (2012). Physiology, pharmacology and pathophysiology of the pH regulatory transport proteins NHE1 and NBCn1: similarities, differences, and implications for cancer therapy. *Curr. Pharm. Des.* 18, 1345–1371. doi: 10.2174/138161212799504830
- Bonar, P., Schneider, H. P., Becker, H. M., Deitmer, J. W., and Casey, J. R. (2013). Three-dimensional model for the human Cl<sup>-</sup>/HCO<sub>3</sub><sup>-</sup> exchanger, AE1, by homology to the *E. coli* CIC protein. *J. Mol. Biol.* 425, 2591–2608. doi: 10.1016/j.jmb.2013.04.005
- Boron, W. F., Chen, L., and Parker, M. D. (2009). Modular structure of sodium-coupled bicarbonate transporters. *J. Exp. Biol.* 212, 1697–1706. doi: 10.1242/jeb.028563
- Chang, M. H., Dipiero, J., Sonnichsen, F. D., and Romero, M. F. (2008). Entry to “formula tunnel” revealed by SLC4A4 human mutation and structural model. *J. Biol. Chem.* 283, 18402–18410. doi: 10.1074/jbc.M709819200
- Ch'en, F. F., Villafuerte, F. C., Swietach, P., Cobden, P. M., and Vaughan-Jones, R. D. (2008). S0859, an N-cyanosulphonamide inhibitor of sodium-bicarbonate cotransport in the heart. *Br. J. Pharmacol.* 153, 972–982. doi: 10.1038/sj.bjpp.0707667
- Chu, H., and Low, P. S. (2006). Mapping of glycolytic enzyme-binding sites on human erythrocyte band 3. *Biochem. J.* 400, 143–151. doi: 10.1042/BJ20060792
- Cordat, E., and Casey, J. R. (2009). Bicarbonate transport in cell physiology and disease. *Biochem. J.* 417, 423–439. doi: 10.1042/BJ20081634
- Damkier, H. H., Aalkjaer, C., and Praetorius, J. (2010). Na<sup>+</sup>-dependent HCO<sub>3</sub><sup>-</sup> import by the slc4a10 gene product involves Cl<sup>-</sup> export. *J. Biol. Chem.* 285, 26998–27007. doi: 10.1074/jbc.M110.108712
- Diederichs, K., and Karplus, P. A. (1997). Improved R-factors for diffraction data analysis in macromolecular crystallography. *Nat. Struct. Biol.* 4, 269–275. doi: 10.1038/nsb0497-269
- Ducoudret, O., Diakov, A., Muller-Berger, S., Romero, M. F., and Fromter, E. (2001). The renal Na-HCO<sub>3</sub>-cotransporter expressed in *Xenopus laevis* oocytes: inhibition by tenidap and benzamil and effect of temperature on transport rate and stoichiometry. *Pflugers Arch.* 442, 709–717. doi: 10.1007/s004240100594
- Ericsson, U. B., Hallberg, B. M., Detitta, G. T., Dekker, N., and Nordlund, P. (2006). Thermofluor-based high-throughput stability optimization of proteins for structural studies. *Anal. Biochem.* 357, 289–298. doi: 10.1016/j.ab.2006.07.027
- Espirito, D. J., Bernardo, A. A., and Arruda, J. A. (2006). Role of NH(2) and COOH termini in targeting, stability, and activity of sodium bicarbonate cotransporter 1. *Am. J. Physiol. Renal Physiol.* 291, F588–F596. doi: 10.1152/ajprenal.00361.2005
- Eswar, N., Webb, B., Marti-Renom, M. A., Madhusudhan, M. S., Eramian, D., Shen, M. Y., et al. (2007). Comparative protein structure modeling using Modeller. *Curr. Protoc. Protein Sci.* Chapter 2:Unit 2.9. doi: 10.1002/0471140864.ps0209s50
- Giffard, R. G., Lee, Y. S., Ouyang, Y. B., Murphy, S. L., and Monyer, H. (2003). Two variants of the rat brain sodium-driven chloride bicarbonate exchanger (NCBE): developmental expression and addition of a PDZ motif. *Eur. J. Neurosci.* 18, 2935–2945. doi: 10.1046/j.1460-9568.2003.03053.x
- Gill, H. S. (2012). pH-sensitive self-associations of the N-terminal domain of NBCe1-A suggest a compact conformation under acidic intracellular conditions. *Protein Pept. Lett.* 19, 1054–1063. doi: 10.2174/092986612802762642
- Gill, H. S., and Boron, W. F. (2006). Expression and purification of the cytoplasmic N-terminal domain of the Na/HCO<sub>3</sub> cotransporter NBCe1-A: structural insights from a generalized approach. *Protein Expr. Purif.* 49, 228–234. doi: 10.1016/j.pep.2006.04.001
- Gill, H. S., Dutcher, L., Boron, W. F., Patel, S., and Guay-Woodford, L. M. (2013). X-ray diffraction studies on merohedrally twinned Delta-1-62NtNBCe1-A crystals of the sodium/bicarbonate cotransporter. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 69(Pt 7), 796–799. doi: 10.1107/S1744309113016710
- Groves, J. D., and Tanner, M. J. (1992). Glycophorin A facilitates the expression of human band 3-mediated anion transport in *Xenopus oocytes*. *J. Biol. Chem.* 267, 22163–22170.
- Grunder, S., Thiemann, A., Pusch, M., and Jentsch, T. J. (1992). Regions involved in the opening of CIC-2 chloride channel by voltage and cell volume. *Nature* 360, 759–762. doi: 10.1038/360759a0
- Hawe, A., Sutter, M., and Jiskoot, W. (2008). Extrinsic fluorescent dyes as tools for protein characterization. *Pharm. Res.* 25, 1487–1499. doi: 10.1007/s11095-007-9516-9
- He, B., Wang, K., Liu, Y., Xue, B., Uversky, V. N., and Dunker, A. K. (2009). Predicting intrinsic disorder in proteins: an overview. *Cell Res.* 19, 929–949. doi: 10.1038/cr.2009.87
- He, Q., Chen, H., Wong, C. H., Tsang, L. L., and Chan, H. C. (2010). Regulatory mechanism underlying cyclic changes in mouse uterine bicarbonate secretion: role of estrogen. *Reproduction* 140, 903–910. doi: 10.1530/REP-10-0178
- Jacobs, S., Ruusuvoori, E., Sipila, S. T., Haapanen, A., Damkier, H. H., Kurth, I., et al. (2008). Mice with targeted Slc4a10 gene disruption have small brain ventricles and show reduced neuronal excitability. *Proc. Natl. Acad. Sci. U.S.A.* 105, 311–316. doi: 10.1073/pnas.0705487105
- Jiang, J., Magilnick, N., Tsurulnikov, K., Abuladze, N., Atanasov, I., Ge, P., et al. (2013). Single particle electron microscopy analysis of the bovine anion exchanger 1 reveals a flexible linker connecting the cytoplasmic and membrane domains. *PLoS ONE* 8:e55408. doi: 10.1371/journal.pone.0055408
- Kabsch, W. (2010). Xds. *Acta Crystallogr. D Biol. Crystallogr.* 66(Pt 2), 125–132. doi: 10.1107/S0907444909047337
- Kao, L., Sassani, P., Azimov, R., Pushkin, A., Abuladze, N., Peti-Peterdi, J., et al. (2008). Oligomeric structure and minimal functional unit of the electrogenic sodium bicarbonate cotransporter NBCe1-A. *J. Biol. Chem.* 283, 26782–26794. doi: 10.1074/jbc.M804006200
- Kennedy, M. B. (1995). Origin of PDZ (DHR, GLGF) domains. *Trends Biochem. Sci.* 20, 350. doi: 10.1016/S0968-0004(00)89074-X
- Khandoudi, N., Albadine, J., Robert, P., Krief, S., Berrebi-Bertrand, I., Martin, X., et al. (2001). Inhibition of the cardiac electrogenic sodium bicarbonate cotransporter reduces ischemic injury. *Cardiovasc. Res.* 52, 387–396. doi: 10.1016/S0008-6363(01)00430-8
- Lee, S. K., Boron, W. F., and Parker, M. D. (2012). Relief of autoinhibition of the electrogenic Na-HCO<sub>3</sub> [corrected] cotransporter NBCe1-B: role of IRBIT vs. amino-terminal truncation. *Am. J. Physiol. Cell Physiol.* 302, C518–C526. doi: 10.1152/ajpcell.00352.2011
- Liu, X., Williams, J. B., Sumpter, B. R., and Bevensee, M. O. (2007). Inhibition of the Na/bicarbonate cotransporter NBCe1-A by diBAC oxonol dyes relative to niflumic acid and a stilbene. *J. Membr. Biol.* 215, 195–204. doi: 10.1007/s00232-007-9018-z
- Liu, Y., Xu, K., Chen, L. M., Sun, X., Parker, M. D., Kelly, M. L., et al. (2010). Distribution of NBCn2 (SLC4A10) splice variants in mouse brain. *Neuroscience* 169, 951–964. doi: 10.1016/j.neuroscience.2010.06.005
- McAlear, S. D., Liu, X., Williams, J. B., McNicholas-Bevensee, C. M., and Bevensee, M. O. (2006). Electrogenic Na/HCO<sub>3</sub> cotransporter (NBCe1) variants expressed in *Xenopus oocytes*: functional comparison and roles of the amino and carboxy termini. *J. Gen. Physiol.* 127, 639–658. doi: 10.1085/jgp.200609520
- Mellman, I., and Nelson, W. J. (2008). Coordinated protein sorting, targeting and distribution in polarized cells. *Nat. Rev. Mol. Cell Biol.* 9, 833–845. doi: 10.1038/nrm2525
- Metallo, S. J. (2010). Intrinsically disordered proteins are potential drug targets. *Curr. Opin. Chem. Biol.* 14, 481–488. doi: 10.1016/j.cbpa.2010.06.169
- Niesen, F. H., Berglund, H., and Vedadi, M. (2007). The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* 2, 2212–2221. doi: 10.1038/nprot.2007.321
- Pantoliano, M. W., Petrella, E. C., Kwasnoski, J. D., Lobanov, V. S., Myslik, J., Graf, E., et al. (2001). High-density miniaturized thermal shift assays as



- a general strategy for drug discovery. *J. Biomol. Screen.* 6, 429–440. doi: 10.1177/108705710100600609
- Parker, M. D., and Boron, W. F. (2013). The divergence, actions, roles, and relatives of sodium-coupled bicarbonate transporters. *Physiol. Rev.* 93, 803–959. doi: 10.1152/physrev.00023.2012
- Parker, M. D., Musa-Aziz, R., Rojas, J. D., Choi, I., Daly, C. M., and Boron, W. F. (2008). Characterization of human SLC4A10 as an electroneutral Na/HCO<sub>3</sub> cotransporter (NBCn2) with Cl<sup>-</sup> self-exchange activity. *J. Biol. Chem.* 283, 12777–12788. doi: 10.1074/jbc.M707829200
- Phillips, K., and de la Pena, A. H. (2011). The combined use of the ThermoFluor assay and ThermoQ analytical software for the determination of protein stability and buffer optimization as an aid in protein crystallization. *Curr. Protoc. Mol. Biol.* Chapter 10:Unit 10.28. doi: 10.1002/0471142727.mb1028s94
- Praetorius, J., Nejsum, L. N., and Nielsen, S. (2004). A SCL4A10 gene product maps selectively to the basolateral plasma membrane of choroid plexus epithelial cells. *Am. J. Physiol. Cell Physiol.* 286, C601–C610. doi: 10.1152/ajpcell.00240.2003
- Shirakabe, K., Priori, G., Yamada, H., Ando, H., Horita, S., Fujita, T., et al. (2006). IRBIT, an inositol 1,4,5-trisphosphate receptor-binding protein, specifically binds to and activates pancreas-type Na<sup>+</sup>/HCO<sub>3</sub><sup>-</sup> cotransporter 1 (pNBC1). *Proc. Natl. Acad. Sci. U.S.A.* 103, 9542–9547. doi: 10.1073/pnas.0602250103
- Toye, A. M., Banting, G., and Tanner, M. J. (2004). Regions of human kidney anion exchanger 1 (kAE1) required for basolateral targeting of kAE1 in polarised kidney cells: mis-targeting explains dominant renal tubular acidosis (dRTA). *J. Cell. Sci.* 117(Pt 8), 1399–1410. doi: 10.1242/jcs.00974
- Tsvetkov, P., Asher, G., Paz, A., Reuven, N., Sussman, J. L., Silman, I., et al. (2008). Operational definition of intrinsically unstructured protein sequences based on susceptibility to the 20S proteasome. *Proteins* 70, 1357–1366. doi: 10.1002/prot.21614
- Uversky, V. N. (2013). The most important thing is the tail: multitudinous functionalities of intrinsically disordered protein termini. *FEBS Lett.* 587, 1891–1901. doi: 10.1016/j.febslet.2013.04.042
- Vedadi, M., Niesen, F. H., Allali-Hassani, A., Fedorov, O. Y., Finerty, P. J. Jr., Wasney, G. A., et al. (2006). Chemical screening methods to identify ligands that promote protein stability, protein crystallization, and structure determination. *Proc. Natl. Acad. Sci. U.S.A.* 103, 15835–15840. doi: 10.1073/pnas.0605224103
- Wang, C. Z., Yano, H., Nagashima, K., and Seino, S. (2000). The Na<sup>+</sup>-driven Cl<sup>-</sup>/HCO<sub>3</sub><sup>-</sup> exchanger. Cloning, tissue distribution, and functional characterization. *J. Biol. Chem.* 275, 35486–35490. doi: 10.1074/jbc.C000456200
- Xue, B., Dunbrack, R. L., Williams, R. W., Dunker, A. K., and Uversky, V. N. (2010). PONDR-FIT: a meta-predictor of intrinsically disordered amino acids. *Biochim. Biophys. Acta* 1804, 996–1010. doi: 10.1016/j.bbapap.2010.01.011
- Xue, B., Dunker, A. K., and Uversky, V. N. (2012). Orderly order in protein intrinsic disorder distribution: disorder in 3500 proteomes from viruses and the three domains of Life. *J. Biomol. Struct. Dyn.* 30, 137–149. doi: 10.1080/07391102.2012.675145
- Zhang, D., Kiyatkin, A., Bolin, J. T., and Low, P. S. (2000). Crystallographic structure and functional interpretation of the cytoplasmic domain of erythrocyte membrane band 3. *Blood* 96, 2925–2933.

**Conflict of Interest Statement:** The Authors, Editor and Chief Editor declare that while the author J. Praetorius and the editor E. Bødtkjer are currently employed by the same institution (University of Aarhus, Denmark) there has been no conflict of interest during the review and handling of this manuscript.

Received: 23 August 2013; paper pending published: 02 September 2013; accepted: 15 October 2013; published online: 07 November 2013.

Citation: Bjerregaard-Andersen K, Perdreau-Dahl H, Guldsten H, Praetorius J, Jensen JK and Morth JP (2013) The N-terminal cytoplasmic region of NCBE displays features of an intrinsic disordered structure and represents a novel target for specific drug screening. *Front. Physiol.* 4:320. doi: 10.3389/fphys.2013.00320

This article was submitted to *Membrane Physiology and Membrane Biophysics*, a section of the journal *Frontiers in Physiology*.

Copyright © 2013 Bjerregaard-Andersen, Perdreau-Dahl, Guldsten, Praetorius, Jensen and Morth. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.