Available online at www.sciencedirect.com
ScienceDirect
Protein Expression and Purification 49 (2006) 228–234
Protein Expression & Purification
www.elsevier.com/locate/yprep

Expression and purification of the cytoplasmic N-terminal domain of the Na/HCO_3 cotransporter NBCe1-A: Structural insights from a generalized approach

Harindarpal S. Gill *, Walter F. Boron

Department of Cellular and Molecular Physiology, Yale University School of Medicine, 333 Cedar Street, New Haven, CT 06520, USA

Received 21 February 2006, and in revised form 31 March 2006
Available online 25 April 2006

Abstract

The cytoplasmic, N-terminal domain (Nt) of the electrogenic sodium/bicarbonate cotransporter—NBCe1—over-expresses in *Escherichia coli* and yields a large amount of soluble protein. A novel purification strategy, which involves a streptomycin precipitation, overcomes obstacles of instability and copurifying proteins, and leads to the first seen Nt-NBCe1 crystals. The purification procedure generally lends itself to the purification of Nts from other classes of the SLC4 family. Size-exclusion chromatography suggests that the Nt of NBCe1 as well as the Nt of other SLC4 members form dimers. A comparison of Nt-NBCe1 to SLC4 member Nt-AE1, based on purification properties and predicted secondary-structure sequence alignments, suggests a similar mechanism for dimer stabilization.

Keywords: Na^+-coupled HCO_3^- cotransporter; SLC4A4; Purification; Solubility; GroEL copurification; Secondary-structure; Estimated molecular mass

Na^+-coupled HCO_3^- transporters (SCBTs) are integral membrane proteins that play a vital role in transporting Na^+ and HCO_3^-—and Cl^- in at least one case—in tissues throughout the body [1–3]. They represent at least half the ten-membered SLC4 family of HCO_3^- transporters (Fig. 1). The SLC4 family can be divided into three major classes or groups based on their sequence homology. The first consists of the anion exchangers AE1–AE3. AE1, the founding member of the SLC4 family also known as band 3, was first cloned in 1985 from a spleen cDNA library by Kopito et al. [4]. One of the oldest-known transporters, AE1 is the Cl–HCO_3 exchanger of erythrocytes and plays a key role in the delivery of CO_2 from systemic tissues to the lung. The second group of SLC4 family members consists of three electroneutral, Na^+-coupled HCO_3^- transporters: NBCn1, NDCBE, and NCBE. The Na^+-driven Cl–HCO_3 exchanger NDCBE [5] plays a key role in the regulation of intracellular pH in neurons of the central nervous system. The third group of SLC4 members consists of the two electrogenic Na/HCO_3 cotransporters NBCe1 and NBCe2. NBCe1 is present at the basolateral membrane of the proximal tubule, which reabsorbs 80% of the filtered HCO_3^-. This transporter completes the process of reabsorbing HCO_3^- from lumen to blood, thereby playing a critical role in the regulation of blood pH. Patients with defective NBCe1 have been diagnosed with mutations in NBCe1 that cause autosomal-recessive disorders that may include severe proximal renal tubular acidosis (pRTA)[1], ocular abnormalities, and mental retardation [6,7].

The SLC4 family members have three distinct domains: (i) a large cytoplasmic N terminus, (ii) a transmembrane domain (TMD) that spans the membrane up to 14 times, and (iii) a small cytoplasmic C terminus. Not much is known about the TMD and the C terminus, but the N

* Corresponding author. Fax: +1 203 785 4951.
E-mail address: hs.gill@yale.edu (H.S. Gill).

[1] *Abbreviations used:* pRTA, proximal renal tubular acidosis; TMD, transmembrane domain; TCEP, Tris(2-carboxyethyl)phosphine hydrochloride solution; PMSF, phenylmethylsulponylfluoride; IPTG, isopropyl-β-D-thiogalactopyranoside; BSA, bovine serum albumin.

1046-5928/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.pep.2006.04.001

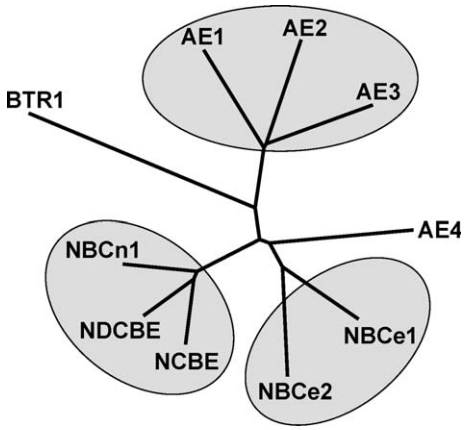

Fig. 1. Dendrogram of the cytoplasmic N-terminal domains of SLC4 family members. The percent identity between groups varies from 39–50% for the electroneutral Na/HCO_3 transporters vs the electrogenic NBCs, and 30–34% for the electroneutral Na/HCO_3 transporters vs the AE1–3. The dendrogram was made with the program TREEVIEW [16]. This dendrogram is similar to that of the full-length proteins. Both the grouping of the N termini and the grouping of the full-length proteins also reflect the physiology of these transporters. The two remaining family members, BTR1 and AE4, whose function are unknown or inconclusive, are less closely related phylogenetically to the others as well as between themselves.

terminus of AE1 (Nt-AE1), at least, is known to serve as an anchorage site for components of the cytoskeleton, some glycolytic enzymes, and hemoglobin [8]. Very little is known about the N termini of the SCBTs, in part because of the difficulty of obtaining protein. Unlike AE1—which, at one million copies per red blood cell, is the dominant membrane protein in a cell with no organelles—SCBTs are expressed in much lower numbers, in membranes that also contain many other major proteins, in cells that also contain organellar membranes, in tissues that contain other cell types. However, it appears that the N terminus of NBCe1 is essential for the TMD to function as a transporter and that the two known N-terminal splice-variants modulate the HCO_3^--transport rate [9]. In addition, one of the naturally occurring missense mutations associated with pRTA in humans is located in the N terminus [7].

Zhang and co-workers in 2000 elucidated the crystal structure of a large portion of Nt-AE1 to 2.6 Å resolution [10], showing that two Nt-AE1 monomers interlock to form a dimer. To this interlocking mechanism, each monomer contributes an arm that is made up of several hydrophobic α-helices and one β-strand. Although the authors purified residues 1–379 of Nt-AE1, their structure actually lacked the first 54 residues at the N terminus as well as residues 202–211 and 350–379, either due to poorly defined electron-density maps or chemical cleavage during the crystallization process [10]. Prior to its crystallization, Nt-AE1 was also described as being unstable during its purification, requiring the presence of 30–50% glycerol with ultimate storage of the protein solution at −20 °C [11,12]. Despite very low protein concentrations, the purified protein led to small crystals and clusters of needles [13], which upon subsequent streak–seeding, ultimately yielded crystals that were suitable for X-ray diffraction.

In the present study, we developed a robust expression and purification system for the N terminus of NBCe1-A (Nt-NBCe1, residues 1–365). Our approach overcomes the difficulties with stability and solubility that were described for Nt-AE1 and that are apparently associated with the entire SLC4 family. Our procedure yields highly pure and homogenous N termini that formed the first seen crystals of Nt-NBCe1. Based upon our success with Nt-NBCe1, we show the usefulness of our approach for purifying the Nts of other Na^+-coupled HCO_3^- transporters. Our findings for the improved isolation of these N termini allow for early insights into the structures of SLC4 family members and for future X-ray crystallographic studies, thereby advancing of the field of HCO_3^- transport.

Methods

Materials

The following particulars were commercially obtained from the respective vendors: the expression vector pET15B and *Escherichia coli* strain Rossetta2™ (Invitrogen, Carlsbad, CA); Cloned-Pfu™ (Stratagene, La Jolla, CA); premixed dNTPs (Roche Applied Science, Indianapolis, IN); streptomycin-sulfate salt (#S6501, Sigma, St. Louis, MO); Tris(2-carboxyethyl)phosphine hydrochloride solution, TCEP (#646547, Sigma); chloramphenicol (#C0378, Sigma); ampicillin (#AB00115, American Bioanalytical, Natick, MA); nickel superflow resin (Qiagen Inc., Valencia, CA); protease-inhibitor cocktail tablets (#11836170001, Roche); phenylmethylsulponylfluoride, PMSF (#AB01620, American Bioanalytical); isopropyl-β-D-thiogalactopyranoside, IPTG (#AB00841, American Bioanalytical); lysozyme (#AB01178, American Bioanalytical); Superdex-200 HR 10/30 gel-filtration column and Akta Unit (Amersham Pharmacia), Amicon Ultra 30 kDa MW cut-off concentration units (#UFC903024, Millipore); bovine serum albumin, BSA (#A-2153, Sigma); and His-Tag Monoclonal Antibody (#70796-3, Novagen, San Diego, CA). Our buffer A solutions contained 50 mM Hepes (pH 7.5), 1 mM EDTA, 200 mM NaCl; buffer B solutions contained 50 mM Hepes (pH 7.5), 1 mM EDTA, 500 mM NaCl.

Construction of the Nt-NBCe1 expression vector

Two primers were synthesized to construct a plasmid that encodes a six-histidine leader sequence fused to the N terminus of residues 1–365 of NBCe1. The forward primer contained an *Nco*I restriction site (underscored), which merged immediately into a glycine codon plus a six-histidine codon stretch that preceded the first post-Met codon of NBCe1: 5′-CATGCCATGGGA-CATCAT CATCATCATCAT-TCCACTGAAAATGTGG-3′. The reverse primer contained a stop-codon after the 365th

codon of NBCe1 and a *Xho*I site (underscored) thereafter: 5′-CCGCTCGAGTTAACCTGAGTACATATTC-3′. Reaction mixtures for the polymerase chain reaction were made according to the manufacture's protocol for Cloned-Pfu^{TM} and contained 1× supplied buffer, 100 ng template, 0.8 μM of each primer, and 0.4 mM of each dNTP in a 50 μl final volume. The program of the thermal-cycler consisted of: (i) 1 cycle at 98 °C for 5 min, 84 °C for 30 s; (ii) 28 cycles at 94 °C for 1 min, 55 °C for 1 min, 72 °C for 10 min; and (iii) 1 cycle at 72 °C for 10 min and 4 °C thereafter. DNA products were then subcloned into the *Nco*I and *Xho*I restriction sites of the pET15B bacterial expression vector, thereby removing the company's cleavable His-Tag sequence within the linker region of the vector. We refer to this construction as pET15b/Δ*slc4a4*. Constructions for other SLC4 members were similarly made.

Expression of Nt-NBCe1

Nt-NBCe1 was expressed in *Escherichia coli* strain Rossetta2^{TM}, which contains an additional plasmid that co-expresses seven eukaryotic tRNAs to overcome species-related codon biases and thus promotes the over-expression of foreign genes. Cells were grown in 18 L of LB media, supplemented with 50 μg/ml ampicillin and 34 μg/ml chloramphenicol, at 37 °C in a shaker until OD_{600} reached 1.0, at which point the temperature was lowered to 15 °C. Cells were induced with 0.5 mM IPTG, harvested after 20–24 h, and visually checked for over-expression of a ~41-kDa product on a gel, using SDS–polyacrylamide gel electrophoresis (SDS–PAGE) to separate proteins by size. Roughly 60 g of cells were divided into 4 × 15 g aliquots and frozen at −80 °C for long-term storage.

Preparation of cell extracts

Wet *E. coli* cells (15 g) that were over-expressing Nt-NBCe1 were resuspended in a volume of 5–6 times the wet-cell gram weight in the presence of 0.2% β-mercaptoethanol, five protease-inhibitor cocktail tablets, and 1 mM PMSF. Lysozyme was added to a final concentration of 1 mg/ml and incubated for 1 h at 4 °C. After the resuspended cells were placed in an ice-bath, they were further disrupted by 6 × 15 s bursts of sonication, and then spun-down at 10,000*g* for 15 min. The pellet was removed. The protein concentration of the supernatant (S1) was ~20 mg/ml.

Purification of Nt-NBCe1

Streptomycin-sulfate precipitation

Streptomycin-sulfate salt, 10% by volume of a 10% solution, was slowly added to S1 over the course of 15 min at 4 °C. The solution was stirred for 1 h and then centrifuged as before. The resulting white pellet was discarded, and the new supernatant (S2) retained.

DEAE–sepharose filtration (optional)

The S2 supernatant was then applied over a 25-ml bed-volume of DEAE–sepharose by gravity and washed with 1 column-volume of buffer A (see Materials). The flow-though (S3) was retained. The DEAE–resin was subsequently regenerated by washing with buffer B.

Ni^{2+}-column affinity chromatography

The S3 flow-through was then applied by gravity onto 3–4 Ni^{2+}-resin columns each with a maximum bed-volume of 5 ml (1 ml bed-volume per 10–12 mg Nt-NBCe1) that were pre-equilibrated with buffer B. The resins were washed with buffer B supplemented with 20 mM imidazole, pH 8, and the protein was eluted with the buffer B supplemented 300 mM imidazole, pH 8, in a total volume of ~25 ml per 5 ml bed-resin. To each pooled elution, 30% by volume of a saturated (96% w/v) ammonium sulfate solution was slowly added over 15 min and continuously stirred for an hour. After centrifugation, the precipitate was saved and was stored overnight at 4 °C.

Gel-filtration chromatography

Each ammonium sulfate pellet was resuspended in a volume of 7 ml in buffer A. Further concentration to a volume 0.5 ml was achieved using an Amicon concentration unit with a 30-kDa MW cut-off. The concentrated sample was applied over a Superdex-200 HR 10/30 FPLC gel-filtration column, equilibrated with 20 mM Tris (pH 7.5), 150 mM NaCl, 1 mM EDTA, 1 mM TCEP, 0.02% NaN_3. The flow rate was 0.5 ml/min and the fraction size, 0.75 ml. The peak fractions having a retention volume around 13 ml were pooled. Samples were either stored at room temperature or at −20 °C.

SDS–PAGE, Western blots, protein estimation, and crystallization

SDS–PAGE was performed by the method of Laemmli [14], the gel stained with Coomassie brilliant blue R-250. Purity was checked by visual inspection from SDS–PAGE and by Western blotting using a commercially obtained monoclonal antibody directed at the engineered N-terminal 6 × His-Tag. The Western blotting was done according to the manufacture's protocol, using chemiluminescent detection. To further check the purity and homogeneity, crystallization trails were performed using the hanging-drop, vapor-diffusion method with 40% (v/v) saturated ammonium sulfate solution as a precipitant at pH 7. Each drop had a total volume of 4 μl, consisting of equal parts of well solution and protein. Protein concentration was determined by the method of Bradford using BSA as a standard [15].

Results

Early attempts at purification

Although large amounts of Nt-NBCe1 (~85 mg/L) were expressed in *E. coli*, our early attempts to purify Nt-NBCe1

using DEAE–sepharose chromatography—as previously used in a purification of Nt-AE1 [10]—to bind Nt-NBCe1 were ineffective. Moreover, the standard Ni^{2+}-affinity chromatography approach previously used to purify 6 × His-tagged Nt-AE1 [12] yielded Nt-NBCe1 protein that was only ~70% pure and quickly precipitated. In an attempt to stabilize Nt-NBCe1, we added large amounts of glycerol, explored various salt conditions, and stored the protein at low temperatures. Regardless, Nt-NBCe1 continued to fall out-of solution slowly. Another problem with the Nt-NBCe1 purification was the apparent co-expression of the *E. coli* chaperonin GroEL, a 60-kDa protein that normally functions to help fold or sequester misfolded proteins. Despite the best available measures to eliminate GroEL, including lowering growth temperatures, ATP–Mg^{2+}-washes, and repeated size-exclusion chromatography, small amounts of GroEL remained tightly bound to Nt-NBCe1 as judged by SDS–PAGE and size-exclusion chromatography. Lowering growth temperatures after induction reduced GroEL contamination and improved solubility. Nevertheless, ~10% of the Nt-NBCe1 still copurified with GroEL and the protein continued to precipitate. Although we attempted to remove the precipitate by centrifugation, the protein remaining in solution was low in concentration (1 mg/ml) but nevertheless formed small composite crystals (i.e., they did not grow uniformly and contained defects).

Detergent washes

We believe that, aside from the GroEL discussed above, other contaminating *E. coli* proteins non-specifically bound to Nt-NBCe1. Our evidence is: (i) after applying the cleared lysate to Ni^{2+} beads and subjecting the beads to SDS–PAGE, we observed numerous bands of low molecular weight (MW) and (ii) after eluting Nt-NBCe1 from the Ni^{2+} beads and applying it to a gel-filtration column, we observed overlapping broad and irregular, high-MW peaks, all of which contained Nt-NBCe1, but we observed no low-MW peaks. We could not remove the contaminants that aggregated with Nt-NBCe1 simply by detergent washes during the initial Ni^{2+}-affinity chromatography step. Rather, we found it necessary—after Ni^{2+}-affinity chromatography, gel filtration, and pooling of the major Nt-NBCe1 peak—to apply the material to the Ni^{2+} column a second time, and then wash with detergent. Apparently, it was necessary to remove some of the GroEL before the detergent wash could break the non-specific interactions. We did not further pursue this approach because it: (i) was cumbersome, (ii) yielded sub-optimal amounts of protein for crystallization trials, (iii) did not totally remove bound GroEL in the sample, and (iv) introduced trace amounts of detergent that may have adversely affected crystallization conditions.

Improved purification of Nt-NBCe1

To overcome the obstacles evident in our early purification attempts, we used the new approach detailed in Materials and methods—streptomycin-sulfate precipitation, DEAE–sepharose filtration, Ni^{2+}-column affinity chromatography, gel-filtration chromatography—which resulted in an extremely stable Nt-NBCe1. The yield was ~108 mg protein per 15 g *E. coli* cell paste. The resulting protein had a purity of >98% and could be concentrated to 50 mg/ml. Fig. 2A illustrates the relative purity of Nt-NBCe1 after each step, and Table 1 lists the recovery after each step. The final gel-filtration chromatographic step yielded a single major peak (Fig. 2B), the retention volume of which corresponds to a molecular mass (~81 kDa) that is about twice that of an Nt-NBCe1 monomer (i.e., 41 kDa). Thus, Nt-NBCe1 appears to migrate as a dimer. The purified Nt-NBCe1 formed single, well-shaped crystals. Fig. 3 compares the morphological differences between crystals grown using the early purification methods (Fig. 3A) and those grown using the improved methods presented here (Fig. 3B). Both groups of crystals diffracted X-rays to ~10 Å resolution at the X29 beamline at the National Synchrotron Light Source, Brookhaven,

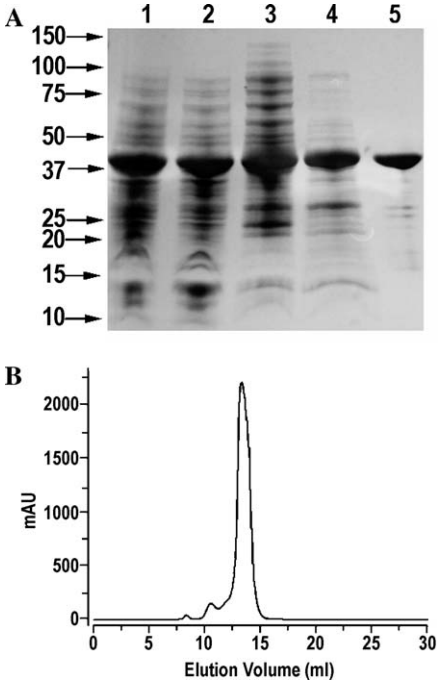

Fig. 2. Purification of Nt-NBCe1. (A) A 4–20% gradient SDS–PAGE shows the relative purification of Nt-NBCe1 during each fractionation step. Lane 1 (S0), crude extract of induced *E. coli* Rossetta2™ transformed with pET15b/Δslc4a4; lane 2 (S1), cleared extract or supernatant of S0; lane 3 (S2), streptomycin-cut of S1; lane 4 (S3), DEAE-filtration of S2; lane 5, Ni^{2+} beads showing bound Nt-NBCe1 after application of S2 and washing with 20 mM imidazole. The protein amount in each lane was adjusted to contain ~20 μg Nt-NBCe1. (B) Nt-NBCe1 was subsequently eluted from the Ni^{2+} beads, concentrated and applied to an HR S-200 analytical gel-filtration column. The column separated some minor, aggregated products (which also contained Nt-NBCe1) with molecular masses at the column's size-exclusion limit (<10.5 ml). The molecular-mass protein standards and their respective retention volumes used in estimating molecular mass were thyroglobulin, 600 kDa, 8.5 ml; bovine γ-globulin, 158 kDa, 11.5 ml; chicken ovalbumin, 44 kDa, 14.3 ml; and equine myoglobin, 15 kDa, 16.5 ml.

Table 1
Steps of Nt-NBCe1 purification

	Total protein (mg)	Amount of Nt-NBCe1 (mg)	Apparent purity (%)	Recovery (%)
Cleared lysate	2000	450	22.5	75
Streptomycin cut	1500	250	16.5	55
DEAE–sepharose filtration	226	193	85	77
Ni^{2+}-affinity chromatography (per 5-ml bed)	50	47	94	97.5
Gel-filtration chromatography	27.3	27	98.5	58

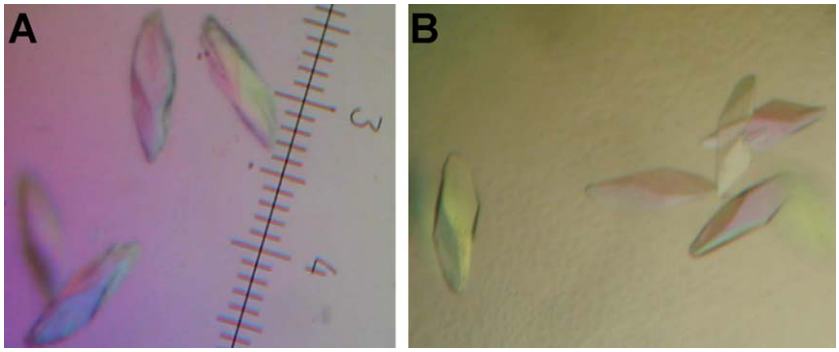

Fig. 3. Preliminary crystals of Nt-NBCe1. Nt-NBCe1 crystals formed in the presence of ammonium-sulfate. Their dimensions are $0.08 \times 0.08 \times 0.2$ mm. The crystals are shown colored by a polarization filter with each small, ruler tick mark separated by 25 μm. Crystallization conditions and protein concentrations were identical in the two panels. (A) The crystals were grown as part of our early purification attempts. Note the defects in the crystals. (B) The crystals were grown as part of our improved purification method. Note the morphological improvement.

```
Nt-AE1    1 ..MEELQDDY EDMMEENLEQ EEYEDPDIPE SQMEEPAAHD TEATATDYHT TS....HPG THKVYVELQE   63
Nt-NBC    1 MSTENVEGKP SNLGERGRAR SSTFLRVVQP MFNHSIFTSA VSPAAERIRF ILGEEDDSPA PPQLFTELDE   70

Nt-AE1   64 LVMDEKNQEL RWMEAARWVQ LEENLGENG. AWGRPHLSHL TFWSLLELRR VFTKGTVLLD LQETSLAGVA  132
Nt-NBC   71 LLAVD.GQEM EWKETARWIK FEEKVEQGGE RWSKPHVATL SLHSLFELRT CMEKGSIMLD REASSLPQLV  139

Nt-AE1  133 NQLLDRFIFE DQIRPQDREE LLRALLLKHS HAGELEALGG VKPAVLTRSG DPSQPLLP.. .....QHSSL  195
Nt-NBC  140 EMIVDHQIET GLLKPELKDK VTYTLLRKHR HQTKKSNLRS LADIGKTVSS ASRMFTNPDN GSPAMTHRNL  209

Nt-AE1  196 ETQLFCEQGD GGTEGHSPSG ILEKIPPDSE ATLVLVGRAD FLEQPVLGFV RLQEAAELE. AVELPVPIRF  264
Nt-NBC  210 TSSSLNDISD KPEKDQLKNK FMKKLPRDAE ASNVLVGEVD FLDTPFIAFV RLQQAVMLGA LTEVPVPTRF  279

Nt-AE1  265 LFVLLGPEAP HIDYTQLGRA AATLMSERVF RIDAYMAQSR GELLHSLEGF LDCSLVLPPT DAPSEQALLS  334
Nt-NBC  280 LFILLGPKGK AKSYHEIGRA IATLMSDEVF HDIAYKAKDR HDLIAGIDEF LDEVIVLPPG EWDPAIRIEP  349

Nt-AE1  335 LVPVQRELLR RRYQSS 350
Nt-NBC  350 PKSLPSSDKR KNMYSG 365
```

Fig. 4. Primary and predicted secondary-structure alignment of Nt-NBCe1 with Nt-AE1. Primary amino-acid alignment was performed using the program CLUSTALW [17] and yielded a 37% identity of Nt-NBCe1 to Nt-AE1 over the indicated regions. The secondary-structure of Nt-NBCe1 was predicted by the program (PHD), which has accuracy of >73% for helices [18,19]. The secondary-structure of Nt-AE1, starting at residue 55, was taken from the Protein Databank Bank with code 1HYN [10] and judged by looking at the structure. Large blocks of color suggest conserved elements between Nt-AE1 and Nt-NBCe1. Blue indicates helices; green, conserved residues. If a conserved residue falls within a helix, the residue is colored the darker blue. The helical regions with leucine clusters at the dimer interface of Nt-AE1 [10] are underlined. Note that Nt-NBCe1 is predicted to have these same helical regions and contain conserved, hydrophobic (leucine and isoleucine) residues, suggesting a similar mechanism of stabilization for its putative dimer interface.

Table 2
Expression and purification of the cytoplasmic domain of various Na/HCO_3 SLC4 family members

SLC4 family member	Species	N-terminal residue no.	Calc. MW	Calc. p*I*	Total expression (mg)/15 g cells	Purity (%)	Est. M_r (kDa)	Calc. no. of subunits
NBCe1	Human	1–365	41	6.4	603	>98	81	1.97
NBCe1	Human	63–365	35	6.5	300	>98	67	1.91
NBCe2	Human	1–407	46	6.5	360	93	86	1.89
NBCn1	Human	1–572[b]	64	5.9	36	90	124	1.93
NBCn1	Rat	1–124	15	6.0	18	>95	97	>6
NDCBE	Human	1–116[a]	13	5.9	162	>95	121	>9
NDCBE	Squid	1–498	57	6.2	90	94	108	1.89
NCBE	Human	1–124	15	6.3	40	>95	97	>6

[a] Cleaved from a GST-fusion; substituted nickel with glutathione resin in the purification.
[b] Purified without DEAE-filtration, Western showed some degradation products.

LI. The diffraction pattern from the crystals in B appears less mosaic than those in A.

Finally, we have tested our improved protocol in the expression and purification of cytoplasmic Nts of all five human Na^+-coupled HCO_3^- transporters, as well as a rat and squid transporter. Three of these proteins are from electrogenic and five are from electroneutral SLC4 family members. In all eight cases, our early attempts at expressing and purifying the proteins suffered from the same sorts of difficulties that we outlined above for Nt-NBCe1. Also in all eight cases, our improved protocol yielded favorable results. Table 2 summarizes several of the properties of the purified proteins. All five of the large Nt fragments appear to be dimers. All three of the smaller Nt fragments—consisting of the ~120 most unique residues that precede the signature ETARWIKFEE motif of the SLC4 family—appear to form higher order oligomers.

Discussion

Distinctive forms of Nt-NBCe1

To obtain a stable Nt-NBCe1, and in preparation for growing high-quality crystals, we needed to separate three distinctive forms of Nt-NBCe1 in the cell lysate:

1. a less soluble form of Nt-NBCe1 that contains bound bacterial proteins;
2. a highly soluble form of Nt-NBCe1 that appears to be extremely aggregated; and
3. a highly soluble dimerized form of the Nt-NBCe1.

Using streptomycin sulfate, we easily forced the first unwanted form described above (1) to precipitate early in the purification. Note the increased background in Fig. 2 (lane 3) that is, the reduced ratio of total Nt-NBCe1 to other proteins. As summarized in Table 1 (columns 3 and 4), the streptomycin cut removed nearly half of the total Nt-NBCe1 and reduced the apparent purity. However, by removing the poorly soluble form of Nt-NBCe1, this purification step markedly improved the homogeneity of our sample. We believe that the Nt-NBCe1 removed in this step has a low solubility because it contains a mixture of (i) GroEL complexed with Nt-NBCe1 and bacterial proteins and/or (ii) Nt-NBCe1 non-specifically bound to *E. coli* proteins other than GroEL.

Fig. 4 illustrates that Nt-NBCe1 has hydrophobic-rich (L, I, F) regions that are not only predicted to be helices, but also align with the helical arms (underscored in Fig. 4) that stabilize the dimer interface of Nt-AE1 [10]. If the hydrophobic arms of monomers are exposed to solvent during overexpression, they may be responsible for non-specifically binding *E. coli* proteins, aggregating, and becoming a target for GroEL. The low solubility of this aggregated form of Nt-NBCe1 may have been the form that continuously precipitated in our early attempts at purification. Copurification with GroEL or other contaminants is a common problem in *E. coli* recombinant expression systems. Our approach for eliminating GroEL contamination may have more general applicability than just the SLC4 family.

The remaining, soluble forms (2) and (3) of Nt-NBCe1 were stable at room temperature and without the use of glycerol. To separate the aggregated but soluble Nt-NBCe1 (2) from the Nt-NBCe1 dimers (3), we used gel filtration. The highly aggregated form elutes in the void volume (>600 kDa), and appears in Fig. 2 as the smaller peaks with peak retention volumes less than 10.5. Finally, preliminary work suggests that use of the Source Q (Pharmacia) anion-exchange resin, which is a quicker approach than the cumbersome gel-filtration step, yields comparable results in terms of separating (2) and (3).

Comparison to Nt-AE1 purification

Similarities

Nt-AE1 and Nt-NBCe1 share several properties. First, as was our experience in our early attempts at the purification of Nt-NBCe1, Nt-AE1 was apparently not stable by itself. Purified Nt-AE1 required at least 30% glycerol and freezing temperatures for storage of up to 6 months [11,12]. Second, as for Nt-NBCe1, Nt-AE1 expressed in *E. coli* appears to copurify with a 60-kDa bacterial protein (e.g., GroEL), as judged from the apparent mobility in SDS–PAGE [11]. A precipitation cut with streptomycin might thus improve the purification of Nt-AE1. Finally, as we found for Nt-NBCe1, Nt-AE1 was more soluble upon induction when grown at low temperatures [11], possibly because more of the protein folded properly and did not complex with GroEL.

Differences

Although Nt-NBCe1 is similar to Nt-AE1 in many respects, Nt-NBCe1 has several distinctive properties that warrant a novel purification strategy. First, the conditions that appear to stabilize Nt-AE1 (glycerol and low temperature) at best only slow the precipitation of Nt-NBCe1. Moreover, high concentrations of glycerol prevent crystallization of Nt-NBCe1.

Second, unlike Nt-AE1, Nt-NBCe1 does not effectively bind to DEAE–sepharose. Although DEAE–sepharose is pivotal for Nt-AE1 purification [11], we optionally employ the resin here simply as a filter to lower the background before applying the Nt-NBCe1 to a Ni^{2+}-resin. We probably could have achieved the same result for Nt-NBCe1 using the Ni^{2+}-resin alone; but the DEAE–sepharose was still helpful in the other cases (e.g., squid NDCBE, Table 2), where expression levels were not so high. Alternatively, it should also be noted that because the DEAE–sepharose yielded 85% pure Nt-NBCe1, we probably could have applied the sample directly to the gel-filtration column, thereby eliminating the Ni^{2+}-resin and the need for a His-tag construct.

Finally, a third distinction between Nt-AE1 and Nt-NBCe1 is that, in the Nt-AE1 crystals, the first 54

residues were not visible in electron-density maps, either because these residues had an extremely high atomic displacement or temperature factor, or possibly because of spontaneous proteolysis. In any case, these 54 residues are unimportant for crystal contacts to form. Our initial attempt was to engineer a truncated version of Nt-NBCe1 (residues 63–365, Table 2) that, based on sequence alignments (see Fig. 4), was similar to the truncated version of Nt-AE1. Yet, unlike the longer version of Nt-NBCe1 (residues 1–365), this shorter version of Nt-NBCe1 has not crystallized.

Summary

(i) The expression and purification procedures that led to crystals of Nt-NBCe1 are new. Moreover, they an important improvement over those described for Nt-AE1 because the protein is now stable, homogenous, and free from bacterial contaminating proteins. Purified Nt-NBCe1 protein can now keep for indefinite periods of time, without glycerol, at higher temperature, and at higher protein concentration. Moreover, our procedures can be applied to other SLC4 family members and the isolated protein can now be effectively used for a variety of biochemical experiments.

(ii) The Nt-NBCe1 crystals themselves are new. Our improved purification scheme yields crystals that now have more uniformity and have distinctive edges compared to crystals from early purification attempts. Moreover, the improved purification led to Nt-NBCe1 crystals that are a significant improvement over the small crystals and clusters of needles that formed directly from the purification of Nt-AE1. With our improved purification scheme, which allows for higher protein concentration, we should be able to grow bigger crystals and thereby increase the resolution for future crystallographic studies. It will be interesting to see in future structure determination whether the extreme N terminus of Nt-NBCe1, the comparable portion of which was absent in the crystals of Nt-AE1, might be stabilized and present in Nt-NBCe1 crystals.

(iii) The tendency of the low-solubility form of Nt-NBCe1 to attract GroEL and non-specifically bind *E. coli* proteins are consistent with its conserved hydrophobic-rich residues in regions of Nt-NBCe1 that align with exposed helical regions of Nt-AE1. In Nt-NBCe1, these leucine and isoleucine residues presumably stabilize the dimerization.

Acknowledgments

We thank Drs. Peter Piermarini and Mark Parker for their helpful discussions, Drs. Liming Chen and Michelle Kelly for their help with the immunological detection methods. We also thank Prof. Joseph Schlessinger and Dr. Yuande Yang for assistance with the Source Q column. H.G. was supported by NIH Postdoctoral Training Grant 5T32 DK007259. This work was supported by NIH Grants R37 DK30344 and R01 NS18400.

References

[1] W.F. Boron, Regulation of intracellular pH, Adv. Physiol. Educ. 28 (2004) 160.
[2] M.F. Romero, M.A. Hediger, E.L. Boulpaep, W.F. Boron, Expression cloning and characterization of a renal electrogenic Na^+/HCO_3^- cotransporter, Nature 387 (1997) 409.
[3] W.F. Boron, E.L. Boulpaep, Intracellular pH regulation in the renal proximal tubule of the salamander: basolateral HCO_3^- transport, J. Gen. Physiol. 81 (1983) 53.
[4] R.R. Kopito, H.F. Lodish, Primary structure and transmembrane orientation of the murine anion exchange protein, Nature 316 (1985) 234.
[5] I.I. Grichtchenko, I. Choi, X. Zhong, P. Bray-Ward, J.M. Russell, W.F. Boron, Cloning, characterization, and chromosomal mapping of a human electroneutral Na^+-driven Cl–HCO_3 exchanger, J. Biol. Chem. 276 (2001) 8358.
[6] C.M. Laing, A.M. Toye, G. Capasso, R.J. Unwin, Renal tubular acidosis: developments in our understanding of the molecular basis, Int. J. Biochem. Cell Biol. 37 (2005) 1151.
[7] T. Igarashi, T. Sekine, J. Inatomi, G. Seki, Unraveling the molecular pathogenesis of isolated proximal renal tubular acidosis, J. Am. Soc. Nephrol. 13 (2002) 2171.
[8] Q. Zhu, D.W. Lee, J.R. Casey, Novel topology in C-terminal region of the human plasma membrane anion exchanger, AE1, J. Biol. Chem. 278 (2003) 3112.
[9] X.F. Liu, S.D. McAlear, M.O. Bevensee, Functional consequences of truncating Na/bicarbonate cotransporters (NBCe1s) at the cytoplasmic N termini, FASEB J. 19 (2005) A142.
[10] D. Zhang, A. Kiyatkin, J.T. Bolin, P.S. Low, Crystallographic structure and functional interpretation of the cytoplasmic domain of erythrocyte membrane band 3, Blood 96 (2000) 2925.
[11] C.C. Wang, J.A. Badylak, S.E. Lux, R. Moriyama, J.E. Dixon, P.S. Low, Expression, purification, and characterization of the functional dimeric cytoplasmic domain of human erythrocyte band 3 in *Escherichia coli*, Protein Sci. 1 (1992) 1206.
[12] Y. Ding, W. Jiang, Y. Su, H. Zhou, Z. Zhang, Expression and purification of recombinant cytoplasmic domain of human erythrocyte band 3 with hexahistidine tag or chitin-binding tag in *Escherichia coli*, Protein Expr. Purif. 34 (2004) 167.
[13] A.B. Kiyatkin, P. Natarajan, S. Munshi, W. Minor, J.E. Johnson, P.S. Low, Crystallization and preliminary X-ray analysis of the cytoplasmic domain of human erythrocyte band 3, Proteins 22 (1995) 293.
[14] U.K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4, Nature 227 (1970) 2680.
[15] M.M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, Anal. Biochem. 72 (1976) 248.
[16] R.D. Page, TreeView: an application to display phylogenetic trees on personal computers, Comput. Appl. Biosci. 12 (1996) 357.
[17] R. Chenna, H. Sugawara, T. Koike, R. Lopez, T.J. Gibson, D.G. Higgins, J.D. Thompson, Multiple sequence alignment with the Clustal series of programs, Nucleic Acids Res. 31 (2003) 3497.
[18] B. Rost, C. Sander, Prediction of protein secondary structure at better than 70% accuracy, J. Mol. Biol. 232 (1993) 584.
[19] D. Przybylski, B. Rost, Alignments grow, secondary structure prediction improves, Proteins 46 (2002) 197.