Determination of the transmembrane topology of mammalian SLC11A2 by an epitope mapping approach

Maciej Czachorowski
Department of Biochemistry
McGill University
September 2008

A thesis submitted to McGill University in partial fulfillment of the requirements for the degree of Master of Science

© Maciej Czachorowski

TABLE OF CONTENTS

1.0 Abstract

The Slc11a family of integral, proton-coupled divalent metal transporters exhibits a high degree of conservation among phylogenetically distinct organisms and contributes to a variety of pleotropic effects in humans. The topology of mammalian Slc11a family members remains unclear and was investigated by insertion of hemagglutinin (HA) epitopes in the predicted hydrophilic segments of Slc11a2 isoform I, followed by cation transport assays to ensure proper protein function and targeting at the plasma membrane. Immunofluorescence, corroborated by a surface labeling assay, on stably transfected intact and permeabilized LLC-PK1 cells indicated that both termini and the intervening segments separating predicted transmembrane domains 4/5, 6/7, and 10/11 are intracellular, while those linking predicted TMDs 5/6, 7/8, and 11/12 correspond to extracellular regions. Epitope insertion in any of the first three predicted hydrophilic loops of the N-terminus abrogated cation transport activity. These results are consistent with a topological model for mammalian Slc11a2 having 12 TMDs and intracellular termini.

2.0 Résumé

La famille de transporteurs membranaires d'ions métalliques divalents couplés aux protons, Slc11a, présente un haut niveau de conservation parmi les organismes phylogénétiquement distincts et contribue à des effets pléiotropiques chez l'humain. La topologie des membres de la famille Slc11a de mammifère n'est pas clairement élucidée. Elle a été investiguée par l'insertion d'épitopes d'hémagglutinine (HA) dans les segments prédits hydrophiles de l'isoforme I de la protéine Slc11a2. Des essais de transport de cations ont été effectués afin de s'assurer du bon fonctionnement de la protéine et de sa localisation à la membrane plasmique. La technique d'immunofluorescence, effectuée sur des cellules intactes et perméabilisées LLC-PK1 exprimant de façon stable ces constructions, a indiqué que les deux extrémités de la protéine ainsi que les segments séparant les domaines transmembranaires prédits 4/5, 6/7 et 10/11 sont intracellulaires, alors que ceux séparant les domaines transmembranaires prédits 5/6, 7/8 et 11/12 correspondent à des régions extracellulaires. Données, qui ont été validées par un essai de marquage de surface. L'insertion d'épitopes dans l'une ou l'autre des trois premières boucles hydrophiles prédites de la région N-terminale a résulté en la perte d'activité de transport de cations de la protéine. Ces résultats sont en accord avec le modèle topologique existant pour la protéine Slc11a2 qui suggère 12 domaines transmembranaires et des extrémités intracellulaires.

3.0 Introduction

3.1 Membrane proteins

Integral membrane proteins (IMPs), have unique structural properties and organization in lipid bilayers that provide them with the capacity to selectively transport ions, metabolites and other macromolecules through different membranes [1]. The functional importance of integral membrane proteins in an array of diverse homeostatic roles is further highlighted by their high representation in the proteomes of phylogenetically distinct organisms. It is estimated that membrane proteins constitute nearly 1/3 of the proteins found in all proteomes [2], and, not surprisingly, their malfunction or absence is linked to a plethora of human diseases. For this reason, nearly half of all commercially available pharmaceutical agents target IMPs [3].

The development of the vast majority of pharmacologically active compounds targeting IMPs has resulted from at least a partial understanding of their structural/functional properties and those of their substrates. In this respect, the complete, high-resolution, three-dimensional structure of a protein can yield invaluable information detailing its mechanism of action in a normal state, and is essential for the design of subsequent structure-function analyses at the atomic level [4]. Although much headway has been made in structural techniques since the first high-resolution structure of myoglobin was determined [5], there remains a disparity between the number of known IMP structures and known soluble protein structures. Despite their biological importance, IMPs represent fewer than one percent of the more than ten thousand high-resolution complete protein structures in the Protein Data Bank [6]. This disparity arises primarily from the technical challenges associated with expressing, purifying and solubilizing large quantities of inherently hydrophobic IMPs and the difficulty in obtaining ordered arrays reviewed in [7]. However, despite the current challenges associated with obtaining IMP structures at atomic resolution, other lower resolution experimental techniques [8] combined with evolving algorithm based prediction methods [9] have provided valuable insight into the structure and membrane organization of IMPs. These techniques and prediction methods are explored further in section 3.6 of this thesis.

3.2 SLC11A transmembrane protein family

Among the long list of protein groups whose structural elucidation could enormously contribute to the current understanding of a multitude of human afflictions, is the Nramp (*Natural resistance associated macrophage protein*), or SLC11A (*Solute carrier 11a*) family of integral, proton-coupled divalent metal transporters. Divalent metals serve as essential elements for a multitude of chemical reactions and biochemical pathways, however they do not readily cross biological membranes and their cellular concentration must be tightly regulated via specific transporters reviewed in [10]. The SLC11A family has been implicated in the homeostatic control of divalent metals in phylogenetically distant organisms. Two Slc11a relatives have been identified in mammals. The first, SLC11A1, controls divalent metal concentration in specialized microbe-containing vacuoles formed in professional phagocytes (monocytes, macrophages), and plays an important role in innate resistance to infection [11-13]. The other, SLC11A2, plays an important role in both transferrin-independent and transferrin-dependent iron homeostasis, with mutations leading to severe anemia and hepatic iron overload in humans and in rodent models of microcytic anemia [14-16].

3.3 Basic membrane protein architecture

3.3.1 Secondary and tertiary structure

For proteins representing such a large proportion of the proteome and involved in an array of biological pathways, the fundamental structural characteristics of integral membrane proteins are relatively limited. Unlike, soluble proteins, that conform to a comparatively unordered globular mass whose backbone hydrogen-bonding requirements are met by the surrounding aqueous environment, the membrane spanning segments of IMPs are constrained by the highly hydrophobic environment of the lipid bilayer. To exist and function properly in this thermodynamically unfavorable milieu, protein transmembrane domains (TMD) satisfy their backbone hydrogen-bonding requirements by adopting one of two basic conformations in the lipid bilayer: α-helical bundles, or β-barrels. β-barrels consist of an even number of antiparallel β-sheets fixed in a cylindrical fashion through the membrane [17]. Despite being one of the two predominant IMP architectures, β-barrels are far less abundant than α-helical bundles and, so far, have only been found in the outer membranes of Gram-negative bacteria and some double-membrane bound organelles such as the mitochondria and chloroplasts [18]. In contrast, α-

helical bundles adopt a more versatile architecture, where, at the simplest level, they span the membrane perpendicularly and are fixed with their amino acid side chains slotted into the spaces between the side chains of apposing helices [19]. More complex α-helical architecture sees conformational changes, partial folding, parallel and obliquely positioned helices relative to the membrane face, and variations in helix packing density which all contribute to overall functional dynamics [19]. α-helical bundles are the most prevalent structural unit of IMPs [9] and appear to sustain significant diversity in function and transported substrates as exemplified by the Slc11a family of cation transporters [20].

3.4 The eukaryotic translocon and protein insertion into the membrane

3.4.1 The Sec61 translocon

Proteins of the secretory pathway, including IMPs are generated at the endoplasmic reticulum through the concerted actions of ribosomes and the translocon machinery [21,22]. However, unlike soluble proteins, the hydrophobic nature of IMPs requires that they be integrated into the equally hydrophobic membrane environmen. In eukaryotes this is accomplished by the tight seal that the ribosome creates with the Sec61 translocon complex and the subsequent cotranslational integration of the nascent peptide into the membrane [21]. Understanding the coordinated process of peptide transfer through the Sec61 heterotrimer and into the membrane could yield additional information about how IMP topology is determined. To this extent much information has been gleaned from the combined high-resolution crystal structure of the Sec61 archaebacterial homolog resolved by van den Berg and colleagues [23], and the lower resolution electron-microscopy structure of the *E. coli* translocon bound to a translating ribosome [21]. These models appear to suggest that recognition of a topogenic signal (i.e. the 20 or so hydrophobic amino acid residues making up a TMD of the nascent polypeptide - see sec. 3.5.2 below) initiates the opening of a lateral gate in the translocon through which the TMD passes to become integrated in the membrane [24]. Furthermore, a short extracytosolic helical plug prevents the thermodynamically unfavorable interaction of the translocating TMD with the reticular lumen [23].

Although there appears to be consensus on the basic functional architecture of the Sec61 heterotrimer, opinions differ as to the stoichiometry and functional organization of the active, ribosome bound complex. For example, cross-linking studies done by Sadlish

and colleagues (2005) seem to suggest a sequential model of polytopic protein insertion in which TMSs are displaced from a single Sec61 heterotrimer into the membrane one or two at a time by those further down the C-terminus [25]. However, Mitra et al (2005) suggest that two translocon heterotrimers may be positioned *en face* to create one large transport channel [21]. Most recently, Kida and colleagues added yet more complexity to existing theories by suggesting that functional translocons can accommodate up to two hydrophilic peptides separated by multiple TMD [26], in line with the previous proposition that TMDs do not integrate into the membrane in single file [27].

3.5 Membrane protein topology

3.5.1 Classification of membrane topology

Membrane protein topology can most simply be defined as the number of transmembrane segments found in a polytopic protein and their orientation in the membrane [8]. Although classification systems vary [28-31], the topology of helix-bundle proteins can be classified into three main groups depending on the number of membrane spanning α-helices they contain and the orientation of their termini relative to the membrane [31] (see figure sec. 5.1). The first two groups, I and II, represent bitopic IMPs having only a single transmembrane helix and an intracellularly located C- or N-terminus, respectively [29,31]. Group III membrane proteins are polytopic, consisting of two or more α-helical domains weaving through the membrane and linked by largely hydrophilic extramembranous polypeptide segments [31]. The termini of these proteins can have various orientations relative to the two faces of the membrane, and will be located in the same hydrophilic space should the number of protein TMDs be even, or on opposite sides of the membrane should the number be odd.

3.5.2 Topogenic signals contributing to protein topology

Aside from a few known exceptions in which interprotein interactions appear to have an impact on the type of topology a transmembrane protein ultimately adopts (discussed below), protein topology is largely determined by variations of two topogenic signals found within the amino acid sequence of the nascent peptide: stop-transfer and signal-anchor sequences [32]. Upon encountering a stop-transfer anchor sequence (characterized as a stretch of twenty or so hydrophobic amino acids), translocation ceases and the hydrophobic sequence is laterally incorporated into the membrane as an alpha

helical TMD [33]. Conversely, when a signal-anchor sequence is encountered, translocation of the nascent peptide is allowed to proceed through the translocon after the anchor sequence has been laterally integrated into the membrane as an alpha helical TMD [34]. Similarly to stop-transfer anchor sequences, signal anchor sequences are composed of about twenty hydrophobic amino acids, however they contain a high concentration of positively charged residues that are always cytoplasmically localized and thus dictate the orientation of the TMD within the membrane [35,36]. This positive charge bias, known as the "positive inside rule", also conveniently provides a means of predicting the topology of a novel protein based on its sequence and is used by some computational prediction programs (see sec. 3.6.1). The ultimate topology of a polytopic protein thus depends on the number of alternating signal sequences found along its length, and the positioning of positive charges with respect to the signal-anchor sequence.

3.5.3 Topological diversity

Until recently [8,37] IMP topology prediction was based on the assumption that TMDs span the membrane in a near perpendicular manner and are linked by hydrophilic segments alternating between the two compartments separated by the bilayer. This view, as exemplified by the crystal structure of bacteriorhodopsin [38], is fundamentally correct, as there exist only a limited number of ways in which a peptide can be inserted into the membrane given the functional limitations of the translocation machinery and the thermodynamic restrictions imposed by the hydrophobic bilayer. However, as more IMP structures become resolved, it is also becoming apparent that a great deal of topological heterogeneity exists and the manner in which IMPs span the membrane and become integrated into it, is more complicated than originally perceived.

For example the crystal structures of the chloride channel ClC [39] and the glutamate receptor [40] reveal not only the typical perpendicularly positioned TMDs, but also long helices that obliquely span the membrane, together with shorter discontinuous helices, also known as re-entrant loops, which break midway through the membrane and turn back to the compartment from which they originated. Similar discontinuous helix architecture has been observed in the Ca2+-ATPase [41] and, notably, in the leucine transporter Slc6 at TMDs 1 and 6 [42], with the primary difference being that the discontinuous helices in these proteins span the length of the membrane rather than

turning back in a hairpin fashion as observed with re-entrant loops [43]. Discontinuous helices are believed to function as the key structural elements involved in ion transport in these proteins, as the partial charges that their helix termini produce in the membrane core provide the appropriate electrostatic environment for substrate recognition and binding [43].

Aside from the irregularities in basic transmembrane helix architecture described, some proteins such as the prion protein (PrP) can take on multiple topologies attributed to different disease phenotypes depending on the efficiency of TMD integration into the membrane during translocation [44,45]. Furthermore, the ductin protein is believed to exhibit dual topology, characterized by two different topological forms present in a stoichiometry of 1:1 within a cell, depending on the particular macromolecular protein complex it may be associated with [46].

3.6 Protein topology prediction

As touched on earlier, the conformational constraints placed upon IMPs by the membrane environment limits their fold space and the number of discrete orientations they can undertake relative to soluble proteins. Various algorithmic programs, have been developed taking advantage of these structural limitations in an attempt to predict the topology of IMPs based on their amino acid sequence. However, the limited number of resolved IMP structures used as calibration or training sets, together with non-sequence specific influences on topology result in an accuracy of around 70 – 80% (based on subsequently resolved structures) for the most precise programs, facilitating the need for experimental validation or ideally a combined computational/experimental approach to topology prediction [9].

3.6.1 Computational topology prediction

The earliest IMP prediction methods were limited to predicting topology simply by determining whether a particular polypeptide segment had the propensity to be located within the membrane or within the hydrophilic environment by analyzing the hydrophobicity, or hydrophilicity, of its respective amino acids. For example, the Kyte and Doolittle [47] algorithm assigns an average hydrophobicity value to each amino acid within a specified segment size known as a window (generally the 20 or so amino acids of a membrane spanning helix). The length of the protein is analyzed in this manner ultimately providing a hydropathy plot whose positive peaks represent regions of high

hydrophobic nature, marking them as good TMD candidates (Figure 1A). Although this type of assessment provides a good initial estimate of protein structure it fails to account for the peculiarities in transmembrane protein structure discussed earlier (see sec. 3.6.3). To increase predictive power, modern computational topology programs build on this protein scanning approach of earlier methods but draw on a variety of sophisticated machine-learning algorithms such as hidden Markov models that search for givens or patterns within a calibration set of resolved proteins and apply them to the sequence in question [48]. Two such programs with high predictive accuracy [49] are TMHMM [2] and HMMTOP [50]. These programs look at several structural classes representing various architectural motifs of resolved IMPs, such as the collective properties of the amino acids making up the inside loop, inside helix and outside helix etc, and determine the likeliness of finding these same features in the target sequence [30]. Other popular programs include MEMSAT [51] which compares the target sequence to well characterized membrane proteins and generates a list of the most probable topologies, and TOPPRED [52,53] which determines the most likely topology based on a combination of hydrophobicity measurements and 'positive inside rule' predictions. Although an advancement over the simple hydropathy analyses, these prediction methods have their shortcomings: hydrophobic cores of soluble proteins can be mistaken for the TMDs, and short TMDs or those containing charged residues may be improperly identified [30].

To overcome some of these limitations, a combination of these programs has been used in unison thus increasing their predictive capacity by providing an overall topological consensus and discounting errors which would otherwise be considered false positives with a single program [48]. Another complementary approach is to include experimental data obtained on a protein or group of phylogenetically related proteins: in essence, constraining particular known segments and providing a greater foundation on which to base the final prediction [49,54]. This is especially useful for genome-wide topology scanning when at least one terminus of the protein set is known [54,55]

3.6.2 Experimental topology validation

Experimental topology validation methods generally involve the modification of the IMP under investigation by insertion of a reporter 'tag' whose presence or activity can be tested by various assays so as to ascertain its position relative to the membrane.

Various tagging approaches have been used to determine the topology of different proteins. The most commonly used techniques include: enzyme tagging and reporter fusions [56], glycosylation mapping [57], cysteine scanning [58], epitope mapping [59], and limited proteolysis [60]. Each method has its inherent advantages and shortcomings and must be chosen based on the known physical properties of the protein being studied and the system it is being expressed in.

Enzyme tagging [56] involves the genetic fusion of a reporter enzyme whose activity, or lack thereof, provides an indication of its location relative to the membrane . For example, reporter enzymes such as alkaline phosphatase (PhoA) [61] and β-lactamase (Blam) [62] have been used as periplasmic indicators because their respective activities (phosphate ester hydrolysis for PhoA; antibiotic resistance for Blam) are not apparent when they are intracellularly localized as a result of improper folding or substrate inaccessibility, respectively. Conversely, enzymes such as β-galactosidase (LacZ) [63] and chloramphenicol acetyltransferase (Cat) [64], are active only when localized in the cytoplasm and can be used to complement topological data gathered by using an extracellular/periplasmic reporter. Although reporter fusions have been used to predict the topology of various prokaryotic and eukaryotic proteins expressed in *E. coli* [65-68], the approach has important inherent limitations. Fusions involve protein modification by truncation and replacement of the endogenous c-terminal region with the reporter enzyme. The insertion of a large enzyme tag within a much smaller putative hydrophilic segment may ultimately affect function and/or topology of the protein under study [8].

Glycosylation mapping [57] takes advantage of the fact that many eukaryotic membrane proteins contain Asn-linked glycosylation sites and subsequently become glycosylated in the Golgi only on their luminal/extracellular side [69]. Thus, by introducing exogenous glycosylation sites within putative hydrophilic loops, their localization as either luminal/extracellular or intracellular can be determined based on the presence or absence of glycosylation, respectively [8]. The drawback of this method, however, is that all endogenous glycosylation sites must be removed which may have an effect on the native properties of the protein.

Several other experimental topology mapping techniques make use of the selectively permeable nature of the lipid bilayer in which all transmembrane proteins are

embedded. One such method, cysteine scanning mutagenesis [58], entails the removal of all endogenous cysteine residues within the protein of interest followed by the introduction of a single cysteine residue in each putative hydrophilic loop of the protein. These cysteine residues can then be exposed to various combinations of membrane im/permeable sulfhydryl reagents conjugated to an easily detectable marker such as a biotin tag, a fluorescent group, or a radiolabel [8]. These reagents will bind covalently, by way of disulphide bond formation, to the cysteine's endogenous sulfhydryl group, and, based on their membrane permeability, will indicate whether the exogenously introduced cysteine is found intra- or extracellularly. An advantage of this method over glycosylation mapping, epitope mapping and limited proteolysis (described below), is that the labeled sulfhydryl reagents used are relatively small and can more easily access residues in smaller loops otherwise inaccessible to the glycosylation machinery, larger epitope specific antibodies or proteases [70,71]. An obvious disadvantage is that removal of endogenous cysteines may cause misfolding, prevent membrane insertion and/or inactivate function.

Lastly, epitope mapping [59] and limited proteolysis [60] are used in topology prediction by determining the accessibility of specific antibodies or proteases to exogenously introduced epitope or proteolytic cleavage sites under intact and membrane permeabilized conditions. Commonly used, and well characterized epitopes include FLAG [72], HA [73] and c-myc [74] (tags which are small (8, 9 and 10 residues respectively) and thus minimally disturb the native structure of the protein. Epitope specific, and membrane impermeable, antibodies are often conjugated to a fluorescent label allowing for signal detection by way of immunofluorescence. Proteolytic cleavage works in much the same way, however rather than conjugated antibodies, the protein of interest is exposed to proteases specific for carefully identified proteolytic cleavage sites. A substantial advantage of these methods is that they allow for protein topology to be studied in intact and transport-competent proteins, thereby reflecting proper folding and membrane insertion in a functional state.

3.7 SLC11A family of divalent metal transporters

Studies on the SLC11A family of proteins in prokaryotic and eukaryotic organisms did not begin in earnest until a connection was made between a locus on mouse chromosome 1, known at the time as *Bcg/Ity/Lsh*, that conferred resistance to several antigenically diverse pathogens in mice: *Mycobacterium, Salmonella* and *Leishmania* [13,75,76]. In a seminal study undertaken by Vidal and colleagues 15 years ago, the gene involved was identified by positional cloning and characterized as consisting of 15 exons spanning 11.5 Kb [13]. It was named *Natural resistance associated macrophage protein 1*, or Nramp1, coined after the phagocytic cells in which it was most highly expressed [13]. Since, the identification of Nramp1, a large family of structurally and presumably functionally homologous proteins has been discovered in organisms ranging from prokaryotes to eukaryotes and has since been reclassified as the SLC11A (*Solute carrier 11*) family of proteins [20]. Notably, aside from the SLC11A1 (i.e. Nramp1) homolog identified in mice, a second ubiquitously expressed paralog displaying 64% identity and 78% similarity [20] to SLC11A1, was later identified on mouse chromosome 15 and is now aptly known as SLC11A2, but also as DMT1 (*Divalent metal transporter 1*) and DCT1 (*Divalent cation transporter 1*) [77,78]. Murine Slc11a2 consists of 18 exons spanning more than 30 Kb. The structure:function relationships of these two mammalian paralogs, and their role in human health and disease will encompass the remainder of this introduction.

3.7.1 Structure of mammalian SLC11A paralogs

The high degree of conservation and sequence similarity observed between SLC11A1 and SLC11A2, off-handedly suggests that these two mammalian paralogs share substantial structural similarity. This is most certainly the case and applies not only to the mammalian paralogs but also a diverse spread of SLC11A members, with certain eukaryotic organisms sharing more sequence and structural similarity than others [20]. This disparity can be attributed to an ancestral gene duplication splitting eukaryotic SLC11A orthologs into prototypical and archetypical subclasses resulting differing in their number of TMDs [79]. SLC11A orthologs (such as SLC11A1/2) from multicellular organisms are predominantly classed into the archetypical group where, despite what the current nomenclature might suggest, SLC11A2 is believed to be an evolutionary precursor of

SLC11A1, diverging from a strictly homeostatic protein to one involved in innate resistance to infection while still maintaining its fundamental divalent metal transport role [80; see below]. This shared ancestral heritage underlies many of the shared characteristics of SLC11A1 and 2.

Both SLC11A paralogs constitute 90-100 kDa integral membrane phosphoglycoproteins, in which *N*-linked glycosylation accounts for nearly 50% of their observed molecular masses [81]. In SLC11A1, complex glycosylation of the mature protein does not appear to be critical for subcellular targeting [82]. On the contrary, glycosylation in SLC11A2 has been implicated in the polarized distribution of the protein in certain epithelial cell lines, and glycosylation site mutants become equally distributed at both the apical and basolateral membranes versus fully glycosylated proteins which are targeted strictly to the apical membrane [83].

The inherently hydrophobic nature of these proteins has undermined efforts to obtain high-resolution crystallographic information and structure prediction has been limited largely to computational prediction techniques such as hydropathy profiling, hydrophobic moment analysis and multiple sequence alignment [20]. These studies have suggested a 12 TMD domain architecture for both proteins, with intracellular termini [20] and sequence homology to other known membrane transporters [13]. Some effort by our group has been made to experimentally validate this predicted topology using epitope accessibility studies in intact and permeabilized cells, which have confirmed the polarity of both termini and the exofacial location of the 7th glycosylated loop [84-86]. This model has been given further credence by topology mapping studies carried out in the conserved bacterial homolog MntH by Courville and colleagues. This group used a targeted genetic fusion technique involving cytoplasmic and periplasmic reporters to provide a topology model consistent with 11 TMD and periplasmic c-terminal predictions for this prototypical ortholog [20,87]. The body of this thesis continues where the eukaryotic topology studies left off, and attempts to experimentally confirm the number, position and polarity of the remaining TMDs in mammalian SLC11A2 by an epitope mapping approach (see Results and Discussion).

An interesting observation made from the previously mentioned sequence alignment studies in SLC11A proteins, is the presence of a highly conserved hydrophobic

core encompassing the first 10 TMDs of these proteins [20]. The TMD domains found within this region are particularly well conserved and contain up to 9 invariant charged amino acid residues whose thermodynamically unfavorable location within the hydrophobic TMD environment immediately made them appealing targets for structure:function studies by several groups [88-90]. The majority of the work has been carried out on mammalian SLC11A2 with supporting data coming from experiments using the bacterial homolog MntH [91,92]. Lam-Yuk-Tseung and colleagues identified three negatively charged residues (D86, D192, E299) within TMD 1, 4 and 7 of SLC11A2 which they suggested contributed to a metal binding site or permeation pathway which interacts with oppositely charged metal substrates [88]. Moreover, in the same study, two conserved histidine residues (H267, H272) in TMD 6 were implicated as regulating pH dependence of metal transport, a theory subsequently backed by studies in *Xenopus* oocytes confirming that these two residues are required for proton-coupling and subsequent metal transport [89]. Most recently, mutagenesis studies done in MntH, implicated two conserved peptide motifs in TMDs 1 and 6 as important for proton-coupled symport of divalent metals [92].

The general consensus from these studies and others done on invariant TMD residues [93,94] and hydrophilic loops [95] within the conserved hydrophobic core of SLC11A proteins, is that these residues and regions collectively act in the recognition, binding and coupling of substrates for their eventual transport through the lipid bilayer, implicating the hydrophobic core as the key functional region in the SLC11A family of divalent metal transporters.

Lastly, a highly conserved structural feature observed in SLC11A orthologs and found in many transport proteins is the conserved transport motif (CTM) of the fourth intracellular loop. The exact function of the CTM is not completely understood; however its importance is emphasized in mutagenesis studies where changes to some conserved residues within this motif abrogate transport [93]. In bacterial periplasmic permeases the CTM is believed to be involved in orchestrating interactions between membrane components and peripheral ATP binding subunits of periplasmic permeases [96,97].

3.8 Function, localization and expression of mammalian SLC11A proteins

Although structurally very similar, SLC11A1 and SLC11A2 share different expression profiles and sub-cellular localization. Facilitated diffusion dependent on the membrane potential facilitates divalent metal movement across the membrane. Moreover, early characterization studies of mammalian SLC11A2 provided experimental evidence that the SLC11A family of proteins function as divalent metal transporters via facilitated diffusion coupled to a proton gradient see below; [78]. Consequently, the expression and functional characteristics of SLC11A2 will be examined first in the subsequent section.

3.8.1 Molecular function of SLC11A2

Although amino acid sequence analysis suggested SLC11A proteins may function as transmembrane transporters [13], determination of their specific molecular role was facilitated by characterization studies of orthologous SLC11A2 proteins [94] and their exogenous expression in *Xenopus* oocytes [78]. When overexpressed in yeast, the SLC11A2 yeast ortholog, SMF1p, was observed to complement a yeast mutant that could not grow in a low metal environment created by the addition of the metal chelator EGTA. Furthermore, mutation of this protein resulted in a reduction of Mn^{2+} uptake in yeast cells overexpressing the mutant, collectively suggesting a direct role for SLC11A2 in metal transport across the membrane [94]. Subsequent, studies in which SLC11A2 cDNA was expressed in *Xenopus* oocytes demonstrated a several hundred-fold increase in Fe^{2+} uptake compared with control cells. Further characterization of SLC11A2 uncovered a broad substrate range of divalent metals including Fe^{2+}, Mn^{2+}, Zn^{2+}, Co^{2+}, Cu^{2+}, Ni^{2+} and Pb^{2+}, whose transport was coupled to a proton gradient and dependent on the cell membrane potential [78]. The high level of conservation observed in the SLC11A family, suggested that the functional observations made for SLC11A in these studies could be applied to other known family members and labeled SLC11A proteins as proton-coupled, divalent metal transporters (see figure sec. 5.2).

3.8.2 SLC11A2 isoforms

Early characterization studies of the predicted amino acid sequences of Slc11a2 cDNA isolated from mouse [77], rat [78] and human [98], identified variation between the sequences of the rat and human C-terminal sequence when compared with that of the

mouse. It was later deduced that the apparent C-terminal variation observed in these orthologous protein sequences resulted from alternative gene splicing at two 3' exons, effectively generating two C-terminally variant isoforms (I +IRE and II -IRE) of SLC11A2 [99]. Isoform I is present as a 561 aa protein whose expression appears to be regulated by an iron responsive element (IRE) at its 3' untranslated region (UTR). In contrast to isoform I, isoform II is slightly larger at 568 aa and lacks a 3' IRE. In addition to these 3' modifications, alternate transcription site usage at the 5' end of Slc11a2 can incorporate an alternate exon (exon 1A) and result in a theoretical protein extended by 29-31 aa upstream of the original start codon identified in SLC11A2 isoforms I and II, which use exon 1B [100]. Collectively, these 5' and 3' modifications can generate four different *Slc11a2* transcripts varying at their 5' terminal end by the presence of either exon 1A or 1B, or the presence or absence of an IRE at the 3' end. These four potential transcripts, 1A-IRE, 1A+IRE, 1B-IRE and 1B+IRE, are schematically depicted in sec. 5.3.

The majority of studies done to date on SLC11A2 utilized the 1B isoforms resulting in a paucity of information regarding the inherent characteristics of 1A isoforms. However, a recent study done by Ludwiczek and colleagues in which COS cells were transiently transfected with each of the four known SLC11A2 isoforms and gauged for $^{59}Fe^{2+}$ uptake at increasing extracellular pH, showed that isoform 1A+IRE most effectively transports iron at pH levels between 5.5-6.5 with an abrupt decline in transport activity above pH 7.5. This was in contrast to 1B isoforms, which exhibited the most pronounced iron transport at neutral pH and only a gradual decline in activity with progressively increasing pH [101]. These observations led the group to suggest that the 1A isoforms could be responsible for metal transport from the highly acidic endosomal environment into the cytoplasm, while the 1B isoforms could be better adapted to metal transport across the plasma membrane into the cell from the relatively neutral extracellular milieu [101]. This interpretation is consistent with studies done by Lis *et al* which demonstrated that 1A isoforms are highly expressed under the acidic conditions associated with cerebral hypoxia, as evidenced in rat PC12 cells [102].

3.8.3 SLC11A2 expression

SLC11A2 is expressed ubiquitously with specific tissue and subcellular localization being correlated to the type of isoform expressed: isoform I is predominantly

expressed in epithelial cells while isoform II is observed in non-epithelial cells of peripheral tissues. However, the kidney, thymus and liver have shown the presence of transcripts belonging to both isoforms [78,83,103]. The epithelial expression of isoform I is most prevalently observed in the duodenum of the small intestine [104,105]. Here, it is expressed at the apical plasma membrane of enterocytes and involved in the uptake of dietary iron - being up-regulated in response to iron deprivation and repressed during iron overload [104,106,107]. Similarly to the apical expression observed in the duodenum, immunostaining studies in mouse kidney have detected the presence of isoform I on the apical surface of epithelial cells lining the proximal tubule [108], while studies in rat kidney showed similar expression in the distal convoluted tubules [109]. The presence of SLC11A2 at these sites points to its potential role in divalent metal absorption in the proximal tubules of the kidneys [110].

SLC11A2 isoform II expression has been primarily localized to peripheral tissues, and most prevalently to various blood cells such as reticulocyte precursors and phagocytic macrophages where it plays a role in iron uptake and recycling [111]. Reticulocyte precursors require large amounts of iron for hemoglobin production. This iron is co-internalized together with transferrin in recycling endosomes, where it is subsequently transported across the endosomal membrane by SLC11A2 into the cytosol. Thus, although isoform II exhibits plasma membrane expression in reticulocyte precursors, it is primarily found in transferrin-positive recycling endosomes [86]. Iron is also recycled by macrophages through the phagocytosis of senescent red blood cells - a rich source of heme iron. Perhaps not surprisingly then, immunofluorescence studies have detected isoform II localization at the phagosomal membrane of macrophages during phagocytosis, as well as in early endosomes with limited presence in late endosomes/lysosomes [111], thus facilitating iron transport into the cytosol.

Aside from the epithelial cells of the gut and kidney, and the peripheral cells of the blood, SLC11A2 has been found to be expressed to a variable degree in the hepatocytes of the liver, increasing in response to iron overload or decreasing under conditions of iron deficiency [106,112]. Furthermore, most neurons of the brain [78], and to a lesser extent cells in the thymus [78], testis, heart [78,113], and placenta [114] have also shown SLC11A2 expression.

3.8.4 SLC11A2 subcellular localization

As evidenced by the expression analyses described previously, both SLC11A2 isoforms are found at the plasma membrane of various cells and tissues, be it at the apical membrane of epithelial cells (isoform I) or at the cell surface of certain peripheral cells (isofrom II; see above). However, their subcellular targeting varies. Using various transfected cell lines, isoform II was shown to be expressed primarily at early and recycling endosomes, from where it cycled to and from the plasma membrane [86,111]. This plasma membrane cycling is largely attributed to a YXLXX motif at the c-terminus of isoform II [115], which also plays a role in the protein's targeting to the early endosomes, as evidenced by mutagenesis studies [83]. Isoform I, which lacks the YXLXX c-terminal targeting motif, exhibits a comparatively slower internalization rate from the plasma membrane, and is not recycled back after endocytosis but rather targeted to late endosomes and lysosomes [83,116]. The slower internalization kinetics of isoform I also contribute to the higher cell surface expression observed for this isoform relative to isoform II [116].

3.8.5 SLC11A1 expression, sub-cellular localization

Contrary to the ubiquitous expression of SLC11A2, SLC11A1 is predominantly expressed in the intracellular space of machrophages [117] and other phagocytic cells such as granulocytes [108] where it is localized to the membrane of late endosomes/lysosomes, and in the gelatinase positive tertiary granules during the steady state. However, upon pathogen induced phagocytosis, SLC11A1 is recruited to the phagosomal membrane and remains there until its maturation into the phagolysome [117]. Here it is believed that through the proton-coupled transport of physiologically important divalent metals such as Fe^{2+} and Mn^{2+}, SLC11A1 is capable of inhibiting pathogen maturation and thus acting in an antimicrobicidal manner. However, the precise manner by which divalent metal transport could have bacteriostatic effects remains somewhat controversial.

3.8.6 SLC11A1 molecular function

There are two schools of thought as to how SLC11A1 mediated metal transport acts to inhibit pathogen replication within the phagolysosome. The first suggests that SLC11A1 acts as a divalent metal influx pump which increases the concentration of redox active divalent metals such as Fe^{2+} and Mn^{2+} within the phagolysosome. A high

concentration of these divalent metals inititates the production of anti-pathogenic oxygen radicals by the Haber-Weiss or Fenton reactions, which in turn are detrimental to pathogen proliferation [118-122]. Moreover, the proposed mechanism for metal influx into the phagolysosome would suggest that SLC11A1 acts as an antiporter, transporting divalent metals from the cytoplasm against the proton gradient built up in the acidic environment of the phagolysosome. Several lines of evidence point to this method of transport including a study in which stimulated phagosomes demonstrated an increase of radiolabeled Fe^{2+} accumulation that could be attenuated by inhibiting SLC11A1 transport activity [122]. Similarly, a study by Goswami and colleagues using *Xenopus* oocytes expressing SLC11A1 showed an increase of Zn^{2+} transport into the phagosome from the cytosol after exposure of these cells to an alkaline external pH [119].

The second school of thought proposes that SLC11A1 acts as a proton-coupled symporter and confers bactericidal effects by acting in the efflux of divalent metals from the phagosome, in effect starving the phagocytosed pathogen of essential nutrients required for maturation [111,123] (see figure sec. 5.2A). The proponents of this model argue that the high degree of sequence conservation between SLC11A1 and SLC11A2, in which metal transport has been established to be pH-dependent [78], suggests that SLC11A1 would also share a similar functional mechanism. Strengthening this view is a study in which the insertion of a hemaglutinin epitope tag at the fourth exofacial loop of SLC11A1, resulted in the targeting of the protein to the plasma membrane, as observed naturally with SLC11A2. At the cell surface the epitope tagged protein transported Fe^{2+} and Mn^{2+} across the membrane in a pH-dependent manner mimicking SLC11A2 proton coupled symport [84].

Further strengthening the metal efflux model are observations that divalent metals play an essential role in the maturation of various pathogens under the control of SLC11A1. For example, *Mycobacterium tuberculosis* has been observed to up-regulate genes involved in metal uptake in activated macrophages, accounting for nearly half of all gene upregulation during pathogenesis [124]. Furthermore, mutations in divalent metal transporters found in *Salmonella,* or divalent metal chelation have been shown to reduce or abrogate its virulence [125]. Huynh and colleagues also showed that *Leishmania amozonensis* upregulates the expression of its own iron acquisition protein, LIT1, when

expressed in congenic mice containing a full complement of SLC11A1, but not in SLC11A1 null mice [126]. Lastly, a recent study done by Corbin et al, demonstrated that the protein calprotectin, expressed inside *S. aureus* induced abscesses in mice, acts as a critical factor in the innate immune response by chelating the divalent metals Mn^{2+} and Zn^{2+} thus reprogramming the bacterial transcriptome and inhibiting proliferation [127]. Collectively these studies appear to suggest that limiting divalent metal accessibility to invading pathogens within the host phagosome would be expected to negatively affect pathogen replication. Moreover, these studies strengthen the idea that the SLC11A1 induced bactericidal effects result from an efflux of divalent metals from the phagosome in a pH-dependent manner.

3.9 SLC11A in human health and disease

3.9.1 SLC11A1 and genetic susceptibility to infection

Several lines of evidence, stemming from mouse models, suggested early on that SLC11A1 played a role in innate resistance to infection. During the seminal study by Vidal and colleagues which identified Slc11a as the gene representing the *Bcg/Ity/Lsh* locus, it was observed that mice susceptible to mycobacterial infection all contained a non-conservative G169D mutation in SLC11A1 [13], leading to a misfolded and improperly targeted protein [81,82]. Furthermore, susceptible mouse strains harbouring this mutation became resistent to infection by *Mycobacterium,* and *Salmonella* when modified by the transgenic introduction of wild type Slc11a1 [128]. Conversely, knockout of Slc11a1 in mice originally resistant to infection resulted in susceptibility to *Mycobacterium, Salmonella,* and *Leishmania* [98].

In humans 12 SLC11A1 gene polymorphisms [80] have been observed in different populations and, as observed in the early mouse studies, have been associated with increased susceptibility to various infectious diseases including leprosy [129], non-tuberculous mycobacterial (NTM) lung disease [130], and human immunodeficiency virus HIV; [131]. Studies in the Chinese Han population have found several polymorphisms in Slc11a1 directly associated with susceptibility to tuberculous pleurisy including a single nucleotide change in intron 4 and a TGTG deletion in the 3' untranslated region of the protein [132]. Furthermore, a non-conservative amino acid substitution at position 543 (D543N) within the same population was also believed to be substantial enough to alter

protein function and contribute to infection susceptibility [132]. Aside from acting as a site-specific guard against infection at the phagosomal membrane where it sequesters divalent metals from the invading pathogen, Slc11a1 also contains a response element for the inflammatory cytokine interferon gamma, along with binding sites for transcription factors such as NF-κB, amongst others, implicating the protein in multiple pleotropic effects linked to launching a systemic immune response [133-135]. Interestingly, a 5' polymorphism in the promoter region of Slc11a1 associated with susceptibility to certain infections is also associated with protection against certain autoimmune disorders such as diabetes [136]. In fact studies by Kissler and colleagues on non-obese diabetic (NOD) mouse models, showed a decrease in susceptiblity to diabetes onset when SLC11A1 levels were reduced in the macrophages of these animals [137]. It was suggested that a reduction in SLC11A1 in the NOD mice led to a decrease in various proinflammatory signal transduction cascades which would otherwise favor autoimmune pathology [137]. Several other inflammatory diseases associated with SLC11A1 polymorphism include rheumatoid arthiritis [138], inflammatory bowel diseases [139] and multiple sclerosis [140].

3.9.2 SLC11A2 and iron homeostasis

Just as mouse models pointed to a possible role for SLC11A1 in resistance to infection, mutations in mice and rats suggested a role for SLC11A2 in systemic iron homeostasis. Mice bearing a SLC11A2 mutation in TMD 4 (G185R), known as the *mk* mouse [141], suffer from hypochromic microcytic anemia due to decreased intestinal iron absorbtion and impaired reticulocyte iron uptake [142,143]. Interestingly, an identical mutation in the *Belgrade* rat [144], causes the same defect in iron metabolism [145,146].

Several SLC11A2 mutations have been identified and characterized in humans which altered the function of the protein and, as evidenced in the *mk* mouse and *Belgrade* rat, conferred a severe anemic phenotype and hepatic iron overload attributed to aberrant iron homeostasis. The first Slc11a2 mutation contributing to hypochromic microcytic anemia and hepatic iron overload was discovered as a G→C mutation (G1285C) in exon 12 which impaired splicing and introduced a E399D substitution within the CTM of SLC11A2 [16,147]. Characterization of the mutated protein by Lam Yuk-Tseung and colleagues suggested that the disease phenotype was not the result of an aberrant protein *per se*, but rather a reduction in *Slc11a2* mRNA as a consequence of aberrant splicing [148].

Moreover, two deletions in a compound heterozygote patient encompassing a 3 base pair deletion in intron 4 which partially impaired splicing, and a C1246T mutation in exon 13 which resulted in an R416C substitution were also identified by Iolascon and colleagues [15]. R416 is a highly conserved residue in TMD 9 whose non-conservative substitution with cysteine results in a complete loss of SLC11A2 function and retention of the protein in the endoplasmic reticulum [149]. Lastly, two mutations in Slc11a2 exons 5 and 8, were found to generate the in-frame deletion of V114 in TMD 2 and a G→V substitution at position 212 in TMD 5, resulting in similar anemic phenotypes to those observed in the previous two patients described [14]. These mutations have yet to be characterized.

3.10 Experimental rationale

Despite the growing body of data accumulated from experiments in both prokaryotic and eukaryotic SLC11A orthologs [80,150], the hydrophobic nature of the Slc11a family makes high resolution structural data difficult to obtain. Such information would be beneficial in deciphering the mechanism of metal/proton symport in SLC11A proteins, and identifying substrate binding sites. Ultimately, this could aid in the development of novel therapeutic interventions for disease states linked to Slc11a polymorphisms and help clarify the mechanism by which intracellular pathogenesis and iron homeostasis are controlled in humans. To this end, the body of this thesis details the steps taken by our group to advance the known structural information regarding eukaryotic SLC11A proteins through an experimental epitope mapping approach aimed at elucidating the topology of this family of important transmembrane proteins. Aside from providing direct information concerning the two-dimensional structure of the SLC11A family of proteins, this approach allows for the elucidation of functionally and structurally important sites, as every epitope tagged construct is tested for metal transport activity across the membrane.

4.0 Materials and Methods

4.1 Materials

Reagent-grade chemicals were purchased from Sigma Chemical (St. Louis, MO). Genetycin (G418) was obtained from Invitrogen. Monoclonal mouse antibody (HA.11) directed against the influenza hemagglutin epitope (HA) was purchased from Covance (Berkeley, CA). Calcein-acetoxymethlyester (calcein-AM) was purchased from Invitrogen Molecular Probes (Eugene, OR). Cy3-conjugated goat anti-mouse secondary antibodies and peroxidase-coupled donkey anti-mouse and goat anti-rabbit antisera were purchased from Jackson ImmunoResearch Laboratories (West Grove, PA). All restriction enzymes and Vent DNA polymerase were obtained from New England Biolabs (Ipswich, MA).

4.2 Plasmids and Site-Directed Mutagenesis

The full-length cDNA for murine Slc11a2 isoform II (-IRE/1B) cloned into the *HindIII* site of pCB6 (pN2-2myc) was described elsewhere [88], as was the Slc11a2 isoform I (+IRE/1B) full-length cDNA in pCB6 containing a hemagglutinin [73] (HA) tag at the fourth extracellular loop (pN2-HAIRE) [116]. Slc11a2 isoform II was excised from pN2-2myc by *HindIII* digestion and cloned into pBluescript KS+ (Stratagene, La Jolla, CA) modified by the elimination of the *SacI* and *XbaI* sites from the polylinker to generate pN2myc-B/S-KO.

To improve protein expression in transfected cells the full-length Slc11a2 was modified at its 5' end by the addition of a GCCACC Kozak sequence, preceded by *EcoRI* and *MluI* sites, as well as a *HindIII* site to enable subsequent cloning into the mammalian expression vector pCB6 [59]. To achieve this, PCR mutagenesis was undertaken as previously described [85] using oligonucleotides EMHK-NT F (5'–CAGAATTCACGCGTAAGCTT***GCCACC***ATGGTGTTGGATCCTAAAGAAAAGATG-3') and N2 501 R (5'-GCCAATGACTTCCTGCATGTC-3' annealing just upstream of an endogenous Slc11a2 *BstE II* site [381 bp downstream of AUG]), and pN2-HAIRE as a template. The 3' end of Slc11a2 in pN2myc-B/S-KO was also modified to remove the isoform II specific nucleotides and replace them with those found in isoform I. To accomplish this, the same technique was used as for the 5' end, but oligonucleotides N2 1051 F (5'- ACTCTGGCTGTGGACATCTAC-3' annealing just downstream of an

endogenous Slc11a2 *Sac I* site [1152 bp downstream of AUG]) and N2-I-CT-HX R (5'-CAGTCTCGAGAAGCTTTTACTTAATGTTGCCACCGCTGG-3'), were used with pN2-HAIRE acting as the template. These modified 5' (N-terminal) and 3' (C-terminal) fragments of Slc11a2 were then digested and subcloned as *EcoRI*/*BstEII* and *SacI/XhoI* fragments, respectively, into corresponding sites of pN2myc-B/S-KO to generate the construct pN2IRE-B/S-KO.

HA epitopes were inserted as 10 amino acid peptides (YPYDVPDYAS) in murine Slc11a2 (+IRE/1B) [116] by using a recombinant PCR approach previously described [151]. To facilitate the potential insertion of additional HA epitopes, oligonucleotides were designed with an *NheI* restriction site at the 3' end of each HA sequence, accounting for the additional S residue found c-terminal to the endogenous sequence of each tag. Epitope insertion sites relative to the protein amino acid sequence are shown in Figure 1B. pN2-HAIRE was used as a template for all HA insertional mutatgenesis reactions, together with the oligonucleotides listed in Table 1. The one exception was in the generation of construct 6 for which pN2-2myc was used as a template due to the presence of a previously inserted HA-tag within the same cloning cassette. The modified Slc11a2 products were then digested and subcloned either as *EcoRI/BstE II* (constructs NT and 1), *BstE II/Xba I* (constructs 2, 3, 4, and 5), *XbaI/SacI* (construct 6) or *SacI/XhoI* (constructs 8, 9, 10, 11 and CT) cassettes into pN2IRE-B/S-KO. All constructs were sequenced for integrity prior to subcloning of each HA-tagged full-length Slc11a2 fragment into the corresponding *HindIII* site of pCB6 for transfection. In the case of constructs 5b, 8b and 11b, the HA epitope was inserted by cloning a short double-stranded fragment composed of two complementary oligonucleotides into the engineered *NheI* sites of the previously inserted HA epitope sequences [152].

4.3 Cell Culture, Transfection and Western blotting

LLC-PK1 cells [153] were grown in Dulbecco's modified Eagle's medium (Invitrogen) supplemented with 10% fetal bovine serum and 100 units/ml penicillin/streptomycin at 37°C, and 5% ambient CO_2. Cells were transfected with Slc11a-HA/pCB6 vectors using Lipofectamine2000 (Invitrogen) as per manufacturer's instructions. Selection of stably transfected clones was done using 1.4 mg/mL of G418 for approximately 14 days. Individual colonies were then isolated and expanded. LLC-

PK1 cells transfected with the variably HA-tagged constructs, were lysed in buffer (50 mM Tris-HCL pH 8.0, 137 mM NaCl, 1% Nonidet P-40, 10% glycerol) supplemented with protease inhibitors (1 mM phenylmethylsulfonyl fluoride [PMSF], 1 μM pepstatin, 0.3 μM aprotinin, 1 μM leupeptin). These whole cell extracts were subsequently prepared and separated by SDS-PAGE and clones exhibiting high levels of SlC11A2-HA expression were identified by immunoblotting as previously described [85]. Antibodies were used as follows: affinity purified rabbit polyclonal anti-mouse Nramp2 (1:1000) or mouse anti-HA epitope (1:2000). Anti-rabbit and anti-mouse secondary antibodies conjugated to hoseradish peroxidase were used at 1:20,000.

4.4 Divalent Metal Transport Assay

Measurement of metal transport by Slc11a2-HA transfected LLC-PK1 cells was performed as previously described by our group [148]. Calcein-AM was prepared as a 500 μM stock solution in dimethyl sulfoxide (DMSO), while ferrous ammonium sulfate and cobalt chloride were dissolved in deionized water to obtain 2 mM stock solutions of Fe^{2+} and Co^{2+}. Fluorescence quenching curves, derived to calculate the initial rates of metal transport in stably transfected cells, were obtained through a fluorescence quenching assay used previously by our group [85].

4.5 Immunofluorescence

We investigated several immunofluorescence protocols from the literature [85,111,152,154,155] that dealt with various cell lines under intact and permeabilized conditions, but were met with inconsistent and ambiguous results surrounding the membrane polarity of our expressed HA tags (data not shown). In our opinion, the thick carbohydrate glycocalyx enveloping LLC-PK1 cells [153], and the natural internalization of Slc11a2 between the plasma membrane and intracellular environment [116] contributed to this limited success. Respectively, these two factors curtailed antibody accessibility to the plasma membrane and the amount of HA-tagged SlC11A2 protein present at the cell surface. To overcome these two issues a variation on a cell surface labeling protocol developed by Lam-Yuk-Tseung and Gros was used [116]. All stable transfectants were grown to confluency on glass coverslips for two days, with empty vector transfected cells serving as a negative control. To localize the HA epitopes on the cell surface under non-permeabilized conditions, LLC-PK1 cells were incubated with anti-HA antibody (1:200

dilution) in 2% nonfat milk/DMEM for 2 hours at 37°C prior to fixation with 4% paraformaldehyde in phosphate buffered saline (PBS) supplemented with 1 mM MgCl2 and 0.1 mM CaCl2 (PBS++). Cells were then incubated with goat anti-mouse rhodamine-conjugated secondary antibody (1:1000 dilution) in 5% nonfat milk/PBS for 1 hr at room temperature. In experiments with permeabilized cells, prior to fixation with 4% paraformaldehyde, cells were incubated similarly to intact cells however in a solution devoid of any primary antibody. Following fixation these cells were permeabilized with 0.1% Triton X-100 for 30 min at room temperature. The cells were then incubated in primary antibody in 5% nonfat milk/PBS for 1 hr at room temperature before exposure to secondary antibody as described above. All cells, intact and permeabilized, were incubated in DAPI/PBS (1:10,000 dilution) for 5 min at room temperature prior to being mounted on glass microscope slides. Cells were visualized using an Axiovert 200M epifluorescence microscope with a 63x oil immersion objective. Digital images were acquired with a Zeiss AxioCam HRm camera operated with AxioVision 4.3. Images were cropped, assembled and labeled using Adobe Photoshop 7.0, and Illustrator 10.0 software.

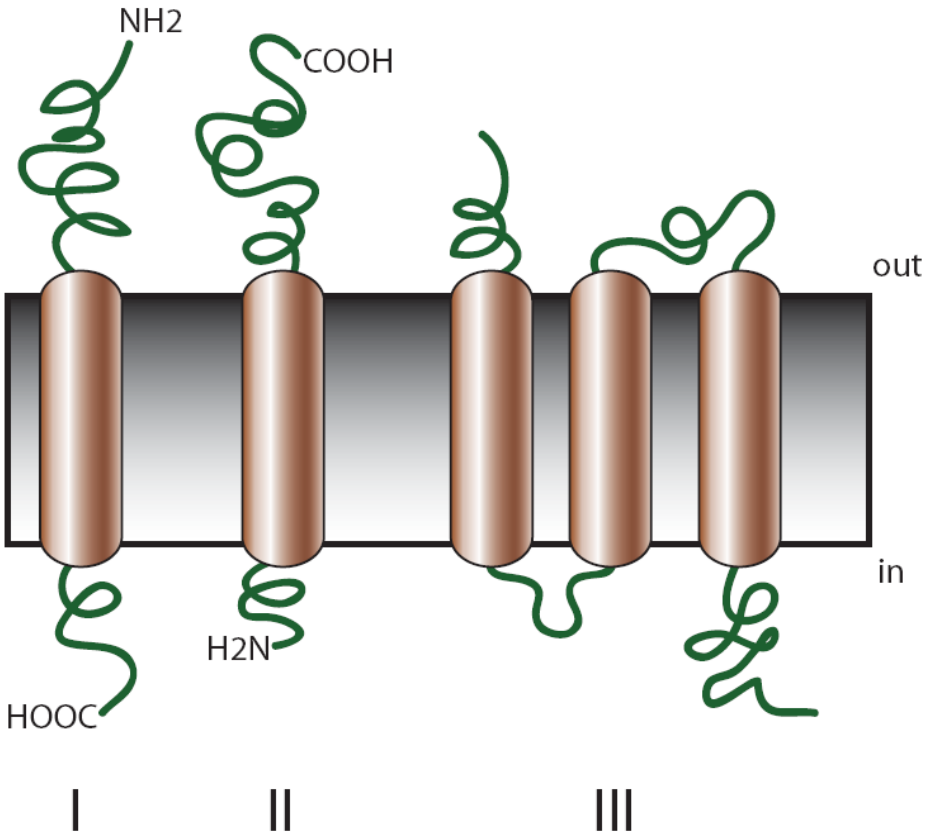
4.6 Measurement of Relative Slc11a2-HA Surface Expression

As a means of verifying and quantifying the immunofluorescence derived membrane polarity of HA epitope tags, the surface protein expression (non-permeabilized conditions) was measured as a percentage of total protein (permeabilized conditions) in transfectants, using an enzyme based colorimetric assay previously described with slight changes [115]. Cells were treated exactly as previously described in preparation for immunofluorescence studies with the exclusion of DAPI staining (see above). Moreover, 2.0×10^4 cells/well were seeded in 48-well tissue culture plates and allowed to reach confluency over 2 days. Secondary antibody incubations were done using horseradish peroxidase (HRP) conjugated Abs (1:4000 dilution) in 5% nonfat milk/PBS for 1 hr at room temperature. Following secondary Ab incubation, the HRP substrate (0.4 mg/mL *o*-phenylenediamine dihydrochloride [OPD], Sigma *FAST* OPD; 550 µL/well) was added to each well according to manufacturer's instructions. Absorbance readings (492 nm) were taken in an ELISA plate reader (Bio-Rad Model 450) and background absorbance

readings from i.) nonspecific binding of secondary Ab and ii.) nonspecific binding of primary Ab to vector-transfected cells were subtracted for each sample.

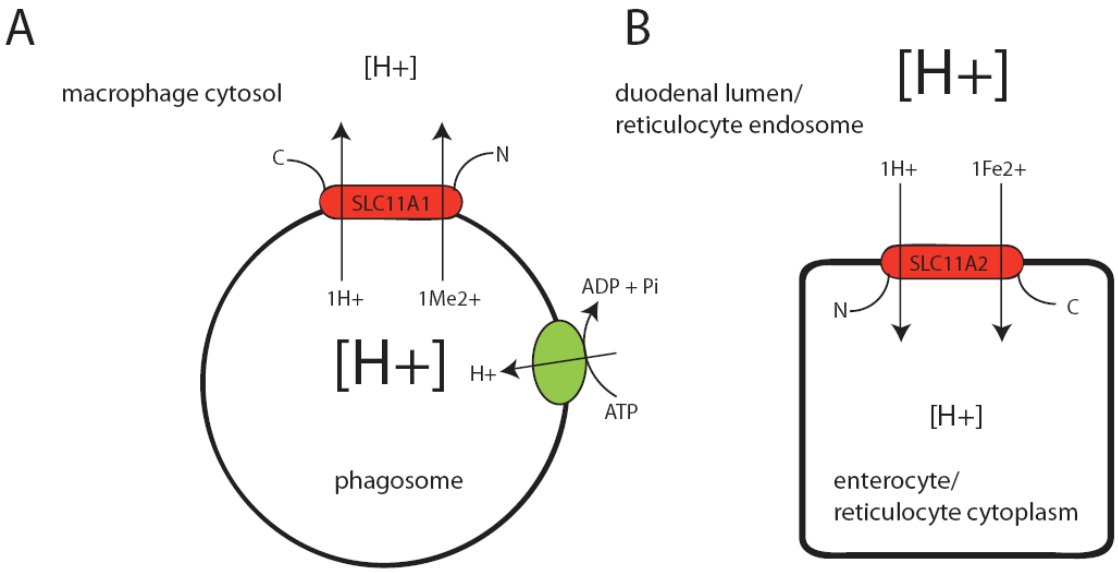
5.0 Figures

5.1 Transmembrane protein orientation in the lipid bilayer



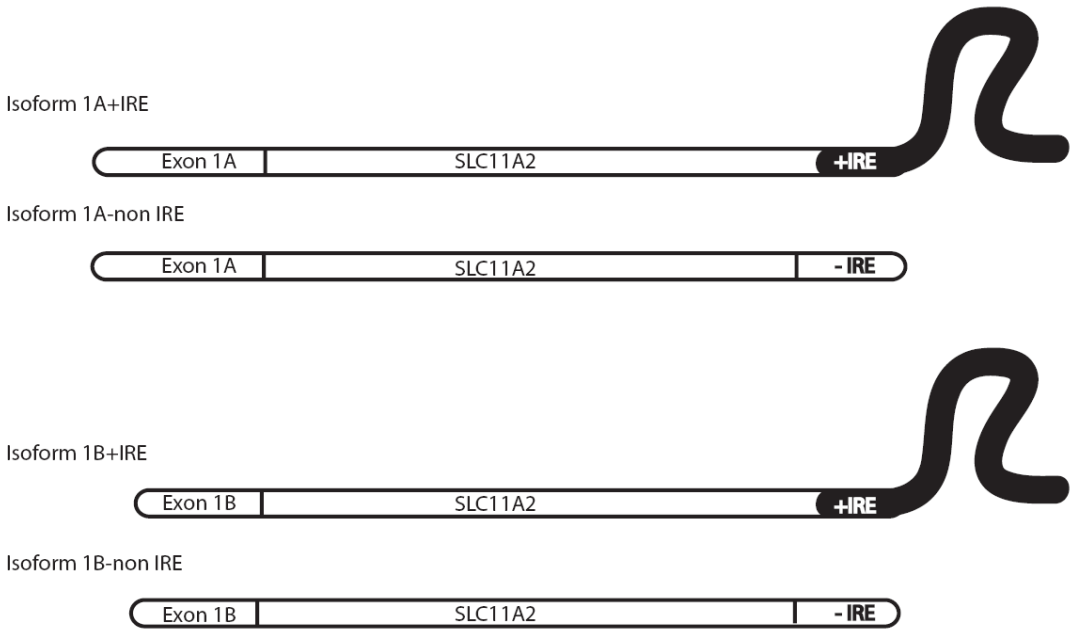
Integral membrane proteins are classified into three classes based on the orientation of their terminal ends relative to the membrane, and the number of transmembrane segments they contain. Type I and II integral membrane proteins have single membrane-spanning alpha helices and differ with respect to the orientation of their termini relative to the membrane. Type III proteins can have several membrane-spanning alpha helices, whose presence, in even or odd numbers, places their termini on the same or opposite faces of the membrane, respectively.

5.2 Proton-coupled divalent metal transport by mammalian SLC11A isoforms



Under physiological conditions, divalent metal transport across biological membranes by mammalian SLC11A isoforms (shown in red) is coupled to proton transport with a stoichiometry of 1:1. *A)* SLC11A1 can transport a range of divalent metals (Me^{2+}) across the phagosomal membrane and into the cytosol of macrophages. This transport is facilitated by a proton gradient across the membrane maintained by V-ATPases (shown in green). *B)* The transport of dietary iron (Fe^{2+}) by SLC11A2 from the duodenum into the cytosol of enterocytes lining the gut, is coupled to a proton gradient created by the acidic conditions present in the duodenum. Similarly, in reticulocytes, iron is transported by SLC11A2 across the endosomal membrane into the cytosol.

5.3 Theoretical SLC11A2 isoforms resulting from alternative splicing of terminal ends



Two potential transcription initiation sites contribute to diversity at the 5' end of Slc11a2, resulting in the incorporation of either exon 1A or 1B. At the 3' end alternative splicing may generate a transcript with (+) or without (-) an iron responsive element (IRE). Collectively, this terminal diversity generate four potential SLC11A2 isoforms.

5.4 Table 1

Table 1. Oligonucleotides used for epitope insertion by site-directed mutagenesis

		M1 [b]			**V2**		
NT		CA GAA TTC ACG CGT AAG CTT GCC ACC ATG	**(HA)** [c]		GTG TTG GAT CCT AAA GAA AAG ATG	**(27)**
		A98			**V99**		
1	**(271)** [a]	GAA TCT GAT TTG CAG TCT GGA GCA	**(HA)**		GTG GCT GGA TTT AAG CTG CTC TG	**(317)**
		L131			**H132**		
2	**(393)**	TT GGA GTG GTC ACC GGC TTG	**(HA)**		CAT CTT GCT GAA GTA TGT CAC CG	**(416)**
		L175			**S176**		
3	**(524)**	CA GCC ATC GCC ATC AAT CTG CTG	**(HA)**		TCT GCA GGA AGG GTC CCC CTG TG	**(548)**
		K201			**Y202**		
4	**(603)**	GTG TTT CTT TTT TTG GAC AAA	**(HA)**		TAT GGC TTG CGG AAG CTG GAA GC	**(626)**
		P243			**S244**		
5	**(729)**	CTC AGG GGC ATG TTC GTG CCG	**(HA)**		TCC TGT CCA GGG TGC CGC ACC	**(750)**
5b			**(HA)**	x2			
		A248			**N249**		
6	**(852)**	C AAG TCT AGA CAG GTG AAT CGG GCC	**(HA)**		AAT AAG CAG GAA GTG CGG GAA GC	**(875)**
		L344			**F345**		
7			**(HA)**				
		N403			**L404**		
8	**(1209)**	GTC ATG GAG GGA TTC CTG AAC	**(HA)**		CTA AAA TGG TCG CGC TTT GCC	**(1230)**
8b			**(HA)**	x2			
		V432			**E433**		
9	**(1296)**	C GTC GCT GTC TTC CAG GAT GTG	**(HA)**		GAG CAC CTA ACG GGG ATG AAT G	**(1318)**
		S468			**E469**		
10	**(1404)**	CA AGC CTG CGG CCA GTG ATG AGT	**(HA)**		GAG TTT TCC AAT GGA ATA GGC TGG	**(1421)**
		H504			**V505**		
11	**(1512)**	GTT TAT GTC CAG GAG CTA GGG CAT	**(HA)**		GTG GCA CTC TAT GTG GTG GCT GC	**(1532)**
11b			**(HA)**	x2			
		K561					
CT	**(1660)**	GAT ACC AGC GGT GGC AAC ATT AAG	**(HA)**		TAA AAG CTT CTC GAG ACT G	

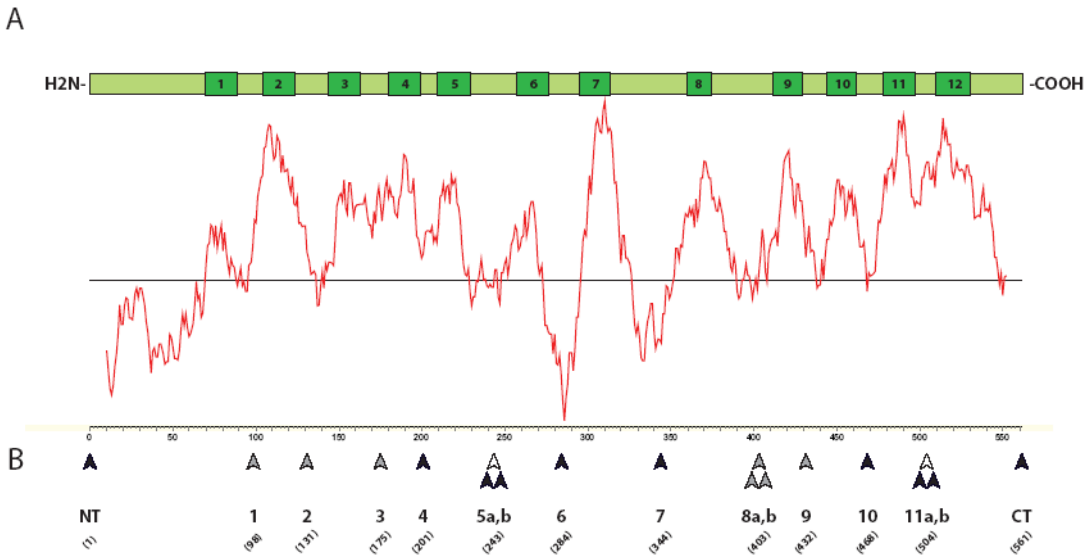
(HA)[c] : **YPYDVPDYAS**

[a] Nucleotide position in sequence of SLC11A2 isoform I

[b] Amino acid residue immediately preceeding the site of insertion of the HA epitope

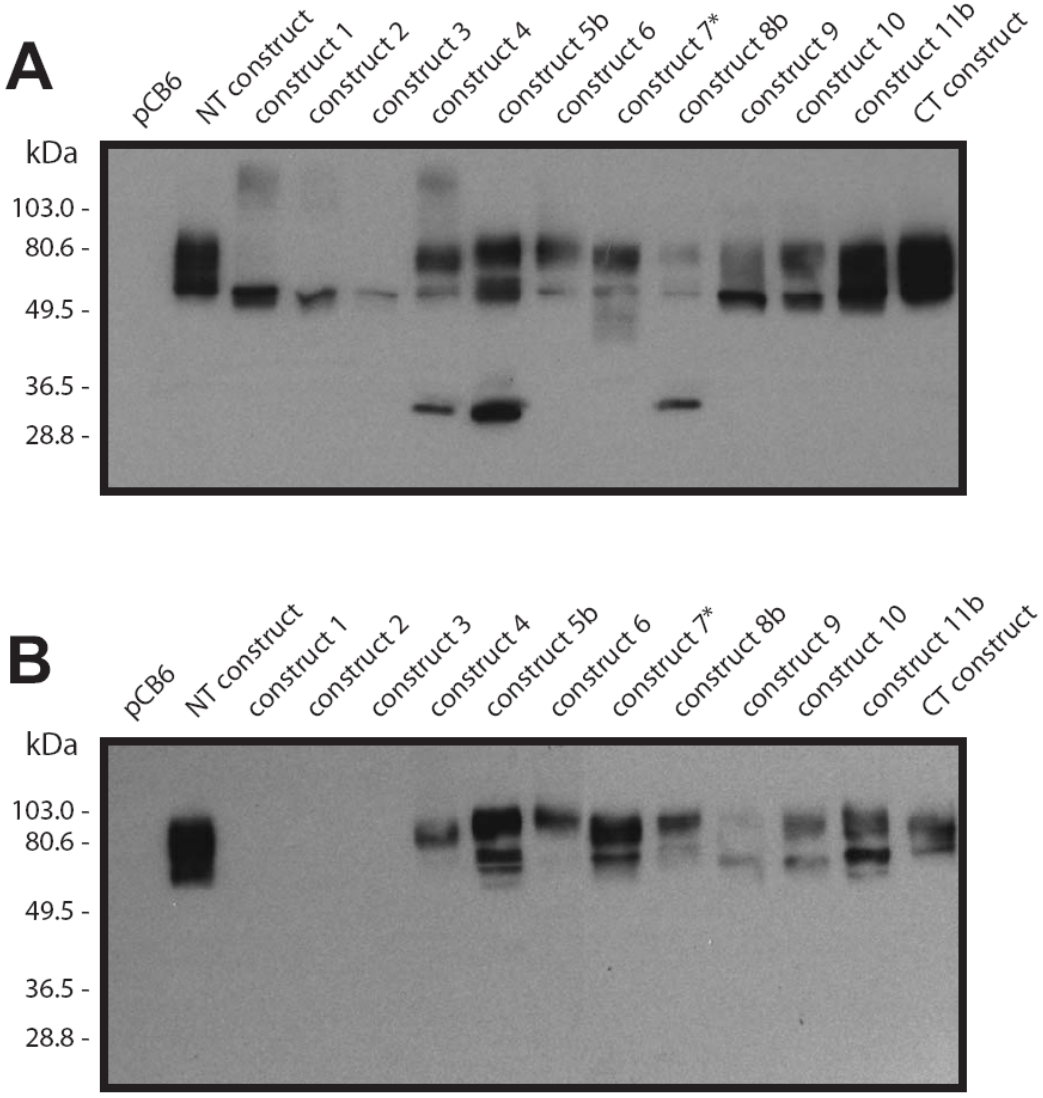
[c] HA epitope sequence

5.5 Epitope insertion in Slc11a2 (Figure 1)



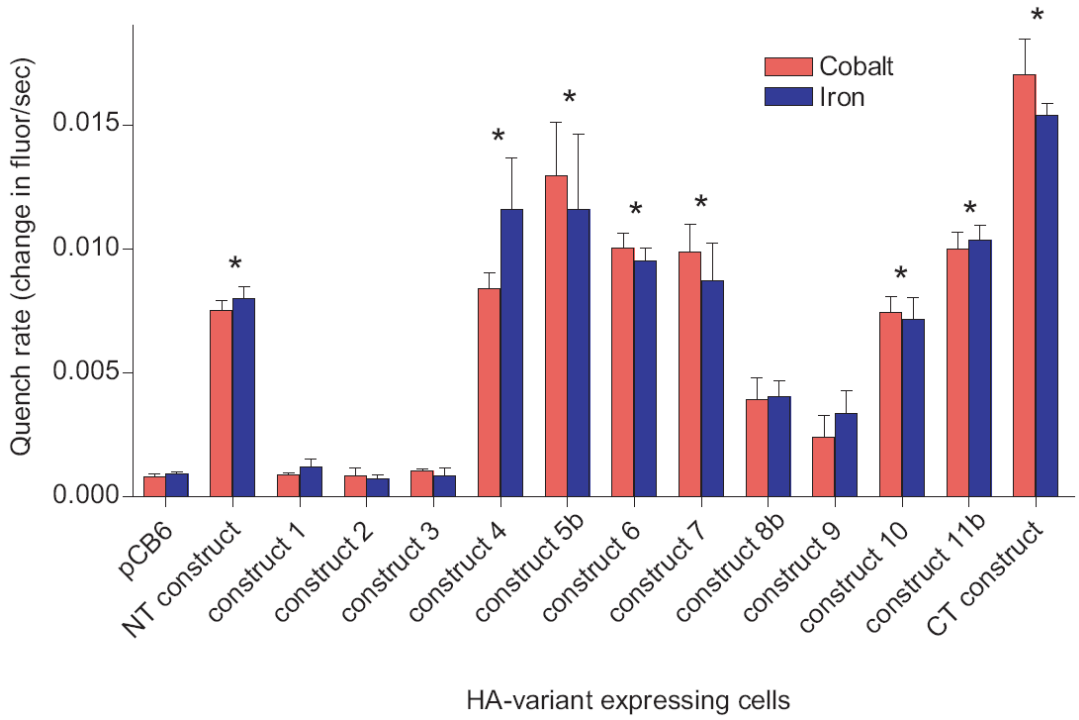
A, schematic representation of SLC11A2 isoform I (amino acid positions 1-561) derived from hydropathy analysis based on the algorithm of Kyte and Doolittle using a sliding window of 19 amino acids. Numbered dark green boxes represent potential transmembrane domains spanning the protein from the N-terminus, consisting of at least 19 amino acids and exhibiting hydrophobic properties. *B*, insertion sites for the YPYDVPDYA HA epitopes. Modified proteins are identified numerically corresponding to the putative hydrophilic loop into which epitope insertion was made, or as NT or CT if epitope insertion was made in the N- or C- terminus, respectively. The amino acid residue preceding the epitope insertion site is identified in *parentheses*. *Black arrowheads* indicate insertion sites that produced a functional protein in which the tag could be mapped by immunofluorescence. *White arrowheads* indicate insertion sites that could not be mapped by immunofluorescence despite producing a functional protein, while *grey arrowheads* represent insertions which inactivated the protein. *Single arrowheads* represent single epitope tags while *double arrowheads* represent two epitope tags inserted in tandem.

5.6 Expression analysis of epitope-tagged SLC11A2 constructs in LLC-PK1 cells (Figure 2)



LLC-PK1 cells were stably transfected with *Slc11a2 isoform I* cDNAs containing HA epitopes, or with empty vector, pCB6, and selected for stable expression of tagged protein in 1.4 mg/mL G418. Western blots of whole cell lysates were incubated with anti-Slc11a2 N-terminal specific polyclonal antibody, *A,* or with the monoclonal anti-HA epitope antibody 16B12, *B.* Five μg of protein was loaded in each lane, except for the *NT* and *CT constructs* in *B*, where $1/10^{th}$ of this value was loaded to prevent excessive blot saturation. Lysates were resolved by 10% SDS-polyacrylamide gel electrophoresis. The positions of the molecular mass markers are indicated on the *left* of the figure in kDa. The *asterisk* (*) denotes that *construct 7* was inherited from previous studies (Lam Yuk-Tseung *et al.*, 2005) *and* is being used as a positive control.

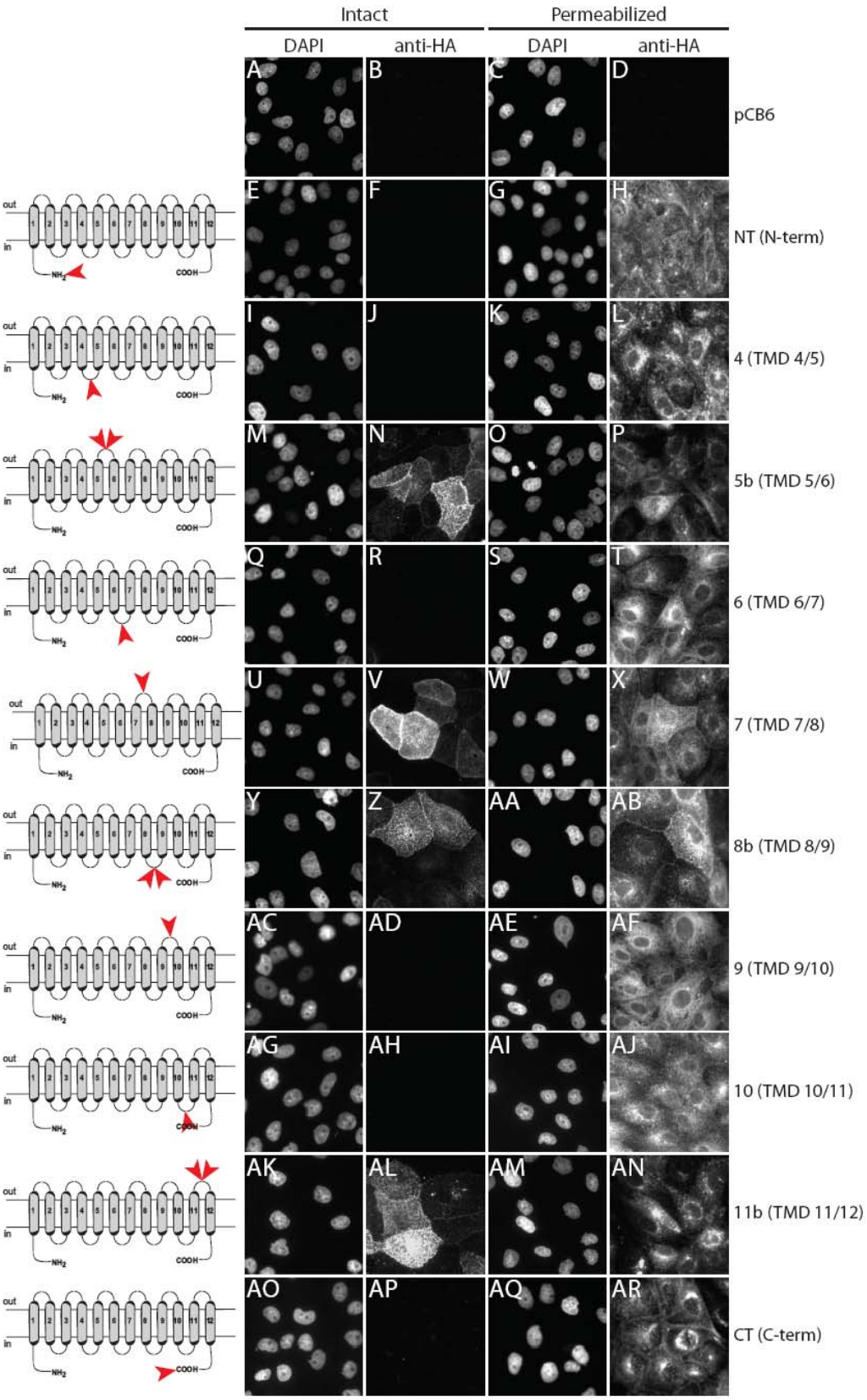
5.7 Metal transport activity of Slc11a2-HA constructs in LLC-PK1 cells (Figure 3)



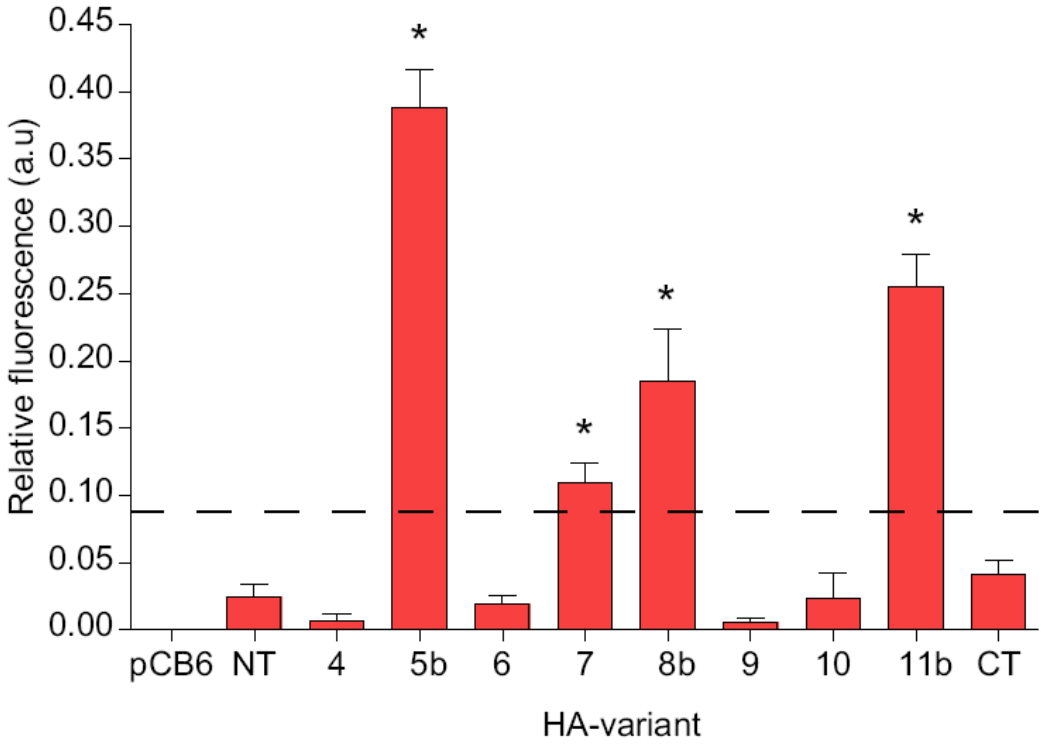
Metal-transport activity of HA-tagged variants was tested using a fluorescence-quenching assay. Briefly, $1X10^6$ cells were loaded with the divalent metal-sensitive fluorescent dye calcein and incubated with 20 μM of Co^{2+} or Fe^{2+} in acidic (pH 5.0) buffer, allowing the rate of fluorescence quenching to be measured over time (2 min) with a spectrofluoremeter (excitation wavelength, 488 nm; emission wavelength 517 nm). The slopes of the initial fluorescence quenching curves were calculated, allowing relative cation transport activity of the variant HA-constructs to be gauged. The results are shown as the initial rates of fluorescence quenching with error bars representing the standard error of the means of three to five independent experiments. * One-way ANOVA followed by a Dunnett's *post-hoc* test indicated no significant difference (P>0.05) in quench rate between the empty vector control (pCB6) and *constructs 1-3, 8 and 9*. Furthermore, no significant difference (P=0.95; two-way ANOVA) was observed in quench rate between the two metals used.

5.8 Detection of epitope-tagged Slc11a2 proteins by immunofluorescence microscopy (Figure 4)

LLC-PK1 transfectants stably expressing empty pCB6 vector or Slc11a2-HA variants, were exposed to mouse monoclonal anti-HA epitope antibody 16B12 without pretreatment (intact *two leftmost columns*) or with pretreatment (permeabilized *two rightmost columns*) with 0.1% Triton X-100. Cells were then incubated with secondary goat anti-mouse antibody conjugated to rhodamine (Cy3) and images were acquired by epifluorescence microscopy. The HA epitopes in constructs 5b (M-P), 7 (U-X), 8b (Y-AB) and 11b (AK-AN) were detected in intact and permeabilized cells, while the HA epitopes in constructs NT (E-H), 4 (I-L), 6 (Q-T), 9 (AC-AF), 10 (AG-AJ) and CT (AO-AR) were detectable only in permeabilized cells. Nuclear staining (DAPI *columns 1 and 3*) was also done as a control in the absence of a rhodamine signal. Exposure times were different for the DAPI and rhodamine images, but were kept the same within each of these two groups with the exception of permeabilized cells for constructs NT (H), 5b (P) and CT (AR) which were photographed at one third of the previous exposure time. A schematic of Slc11a2 derived from the data presented in *Fig. 1* was added to facilitate conceptualization of the fluorescence data with *red arrowheads* representing the inserted HA-epitopes.

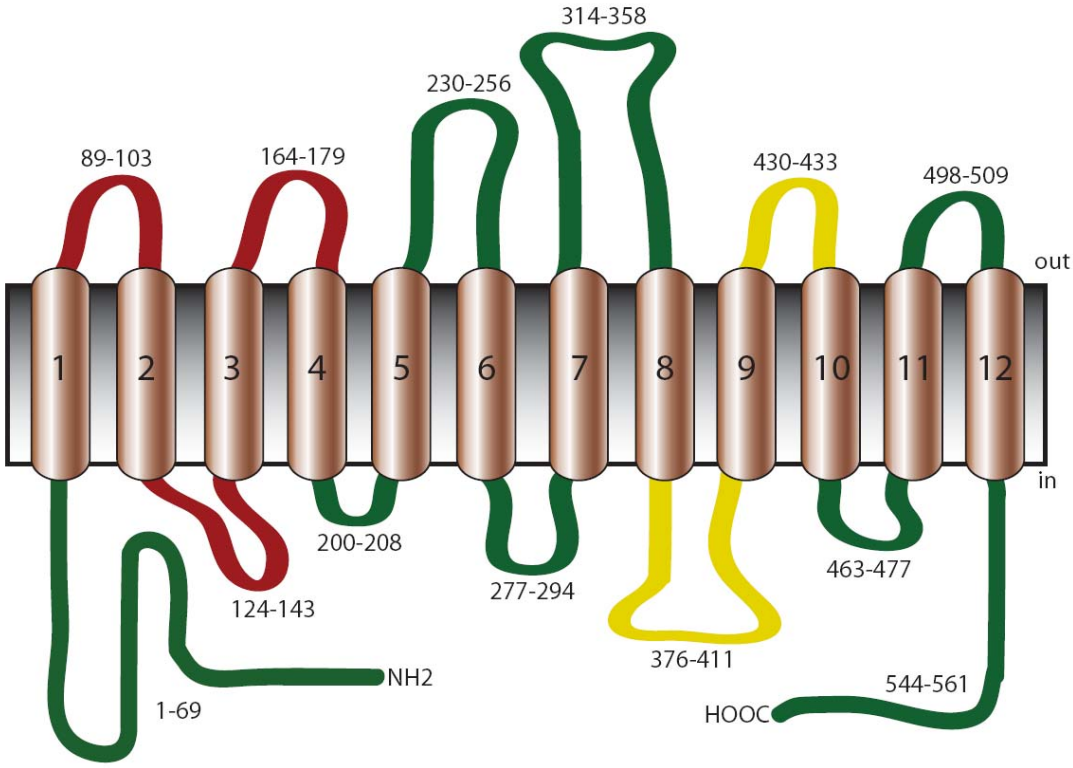

Intact
Permeabilized
DAPI
anti-HA
DAPI
anti-HA
pCB6
NT (N-term)
4 (TMD 4/5)
5b (TMD 5/6)
6 (TMD 6/7)
7 (TMD 7/8)
8b (TMD 8/9)
9 (TMD 9/10)
10 (TMD 10/11)
11b (TMD 11/12)
CT (C-term)

5.9 Corroboration and quantification of immunofluorescence data by cell surface labeling assay (Figure 5)



Stably transfected LLC-PK1 cells were fixed and incubated with primary anti-HA antibody with or without prior detergent permeabilization. Cells were then incubated with a Horse Radish Peroxidase (HRP)-coupled secondary antibody, and the amount of bound primary antibody present was determined for both conditions by a colorimetric reaction using *o*-phenylenediamine dihydrochloride (OPD) followed by spectrometry (492 nm). The presence of Slc11a2-HA expressed at the cell surface (in intact cells) is shown as a fraction of total protein expression (in permeabilized cells) normalized for nonspecific binding of primary and secondary antibodies in non-transfected LLC-PK1 cells. Error bars represent the standard error of the means of two independent experiments done in duplicate. * One-way ANOVA followed by a Dunnett's *post-hoc* test indicated a significant difference ($P<0.01$) in cell-surface fluorescence between the empty vector control (pCB6) and *constructs 5b, 7, 8b, and 11b*. Dashed line represents the minimum relative fluorescence value required to suggest that a clone is expressing an SLC11A2 construct containing an exofacially localized HA tag.

5.10 Predicted transmembrane topology of SLC11A2 (Figure 6)



Hydrophilic loops coloured in green represent those whose membrane orientation was experimentaly confirmed or reconfirmed in this study. The orientation relative to the membrane of the hydrophilic loops coloured in red could not be determined by epitope mapping in this study, as they resulted in improperly processed, non-functional proteins that did not produce a signal in immunofluorescence experiments. Hydrophilic loops coloured in yellow could be mapped, but provided ambiguous data as a result of their drastically reduced metal transport activity relative to the positive control. Amino acid residues constituting each hydrophilic segment are indicated as a range above or below each respective loop or terminus.

6.0 Results

6.1 Construction and Expression of HA-tagged SLC11A2

The SlC11 protein family shares a high degree of conservation among distant phylogenetic members, however differences in C-terminal SLC11 protein topology between some prokaryotic and eukaryotic variants has been suggested [20,87]. Hydropathy analysis (Figure 1A) suggests a 12 transmembrane domain topology for murine SLC11A2 as does a consensus topology prediction approach [48] using an assortment of strongly performing and algorithmically varied topology prediction programs which cumulatively suggested an intracellular location for both termini (data not shown). To experimentally validate the currently held topological model of eukaryotic SLC11A we used an epitope insertion technique previously adopted by our group to investigate the topology of human multidrug resistance protein 1 MRP1; [151]. Aside from allowing for topology mapping in the full-length protein, a corollary of this approach is that it provides a functional assessment of the consequences of disrupting the selected insertion sites. The antigenic epitope tag YPYDVPDYAS (hemagglutinin, HA) was inserted within the putative hydrophilic segments of SLC11A2 isoform I as represented by the most hydrophilic peaks in Figure 1A. Furthermore the specific sites of epitope insertion were chosen as those having the least conservation among various known SLC11 orthologs based on sequence alignment analysis (data not shown), suggesting areas of limited functional/structural importance. Sites of epitope insertion are shown in Figure 1B and, from the N- to C-terminus, are found directly after amino acids 1 (NT construct), 98 (construct 1), 131 (construct 2), 175 (construct 3), 201 (construct 4), 243 (construct 5ab), 284 (construct 6), 344 (construct 7), 403 (construct 8ab), 432 (construct 9), 468 (construct 10), 504 (construct 11ab) and 561 (CT construct). Due to the limited signal that a single epitope exhibited at positions 243, 403 and 504 in immunofluorescence experiments (see below), a second epitope was added at these positions directly following the first with the goal of improving epitope accessibility and amplifying the signal as done in other studies (Kast 1996). The resulting doubly tagged constructs were differentiated from the singly tagged HA constructs by the addition of a ‘b’ after the original name (Figure 1B).

SLC11A2 isoform I HA-tagged cDNAs were cloned into the mammalian expression vector pCB6 and transfected into porcine kidney cells (Hull 1976) that well

replicate the cellular environment in which SLC11A2 isoform I is expressed *in vivo* [110]. Our decision to use SLC11A2 isoform I as a representative template for topological studies of the SLC11 transporter family stems from its high surface expression and slow internalization kinetics from the plasma membrane in comparison to SLC11A2 isoform II [116], or SlC11A1 which is intracellularly localized. This membrane localization facilitates the determination of topology in intact/permeabilized cells (see below). Stable transfectants were selected in G418 and expanded, at which point expression of HA-tagged protein was analyzed by Western blotting (Figure 2AB). SLC11A2 has been previously detected by Western blotting as a pair of bands approximately 60 kDa and 90 kDa, accounting for core and complex glycosylated forms of the protein, respectively [86]. Whole cell lysates from control, vector transfected LLC-PK1 cells and HA-tagged protein expressing cells (constructs NT, 1 through 11 and CT) were separated by SDS-PAGE, transferred to a blotting membrane and analyzed using an N-terminal specific anti-SLC11A2 Ab (Figure 2A), and an anti-HA epitope Ab (Figure 2B). The negative vector-transfected control (pCB6) did not display any antibody specific banding pattern, while the positive control (construct 7) showed the expected immunoreactive species representing the core and complex glycosylated protein in both the epitope and SLC11A2 specific blots (Figures 2AB). Construct 7 was previously generated and used by our group in SLC11A2 isoform characterization studies [149] and shown not to exhibit variably different expression, transport activity, or subcellular localization when compared with untagged SLC11A2 [86]. Nearly all whole cell extracts exposed to anti-SLC11A2 Ab showed the presence of two immunoreactive bands corresponding to the two variably glycosylated protein forms, with the exception of constructs 1-3 which presented as a single 60 kDa core glycosylated species (Figure 2A). Furthermore, extracts incubated with anti-HA Ab displayed a similar double banding pattern as previously observed with the protein specific Ab, but with the mature, complex glycosylated protein appearing more prominently in constructs 4 and 6 (Figure 2B). Moreover, no banding pattern was observed for constructs 1, 2 or 3 exposed to the anti-HA Ab (Figure 2B). These ambiguous banding patterns could potentially be explained by the inaccessibility of the anti-HA Ab to the epitope in the incompletely processed core glycosylated protein, resulting in the presence of only the Ab accessible mature protein (constructs 4 and 6;

Figure 2B), or no banding pattern at all in transfectants where only the core glycosylated form was produced to begin with (constructs 1, 2, and 3; Figures 2AB). Further support for this explanation is drawn from constructs with HA-epitopes inserted at the terminal ends of SLC11A2 (constructs NT and CT). Despite showing a slightly higher level of expression in LLC-PK1 cells compared with other constructs (Figure 2A), a disproportionate reduction in protein loading ($1/10^{th}$ that of non terminally-tagged constructs) was undertaken to prevent excessive over saturation of the blot exposed to anti-HA Ab (Figure 2B). The terminal ends of SLC11A2 isoform I are relatively large (69 and 28 amino acids for N- and C-termini, respectively) and unordered, and consequently their epitopes may be more accessible than the epitopes inserted in the putative loops (even after SDS treatment). Species of approximately 33 kDa were observed in the whole cell extracts from several constructs (Figure 2A) and attributed to proteolytic cleavage despite all efforts to minimize proteolysis by using protease inhibitors in the lysis buffer.

Epitope insertion at selected sites within SLC11A2 did not appear to adversely affect protein processing or native structure in the majority of constructs. However, in constructs 1, 2, 3, 8ab, and 9, it appears that epitope insertion may have disrupted sites required for the structural/functional integrity of the mature protein or its proper processing into a fully glycosylated form.

6.2 Functional Analysis of Insertion Mutants

To verify that epitope insertion did not compromise the topological architecture of our mutant constructs and validate that they were properly targeted to the plasma membrane of LLC-PK1 cells, we assessed the divalent metal transport activity of each construct by using a fluorescence quenching assay [85,88]. We attributed metal transport activity to a properly processed, membrane-targeted and topologically unaltered protein. *In vivo*, mammalian SLC11A paralogs couple their metal transport to a proton gradient and share distinct but overlapping substrate specificity encompassing various physiologically vital divalent metals [78,123]. To account for both of these factors, the transport of two essential redox active metals, Fe^{2+} and Co^{2+}, was assessed in an acidic environment (pH 5.0). Calcein-AM, is a metal-sensitive, acetoxymethlyester (AM) conjugated, fluorescent dye that fluoresces upon cleavage of the ester group in the cytosol

of the cell. The fluorescence quench rate of calcein loaded LLC-PK1 transfectants was measured in response to the extracellular addition of the divalent metals and compared to the quench rate observed for the vector transfected and WT (construct 7) controls (Figure 3). Divalent metal uptake was not significantly different ($P>0.05$) between the empty vector control and constructs 1, 2, and 3. Moreover, constructs 8b and 9 displayed an intermediate level of transport activity roughly 4 and 3 fold higher, respectively, than the empty vector transfected cells, but ~2 and 3 fold lower, respectively, than the positive control (Figure 3). As expected, the metal transport activity of these same constructs (1, 2, 3, 8b and 9) was significantly ($P<0.01$) less than that observed for the WT control but, surprisingly, the CT construct harboring an HA-epitope at the C-terminus of SLC11A2 exhibited a significant ($P<0.01$) increase in transport activity over the positive control of roughly 1.7 fold (Figure 3). These results suggest that epitope insertion in the SLC11A2 protein sites selected in constructs 1, 2, 3, 8ab and 9 render the protein inactive, preventing the use of any epitope localization data obtained from these constructs in the validation of SLC11A2 topology. However, insertion sites in constructs 1, 2 and 3 appear to have a greater negative impact on transport activity than do 8 and 9, suggesting that the altered regions in the former three constructs could be more critical to metal transport or protein targeting to the plasma membrane. The transport activity of all other constructs was not noticeably affected by epitope insertion, suggesting that these mutant proteins were properly processed, functional, and efficiently targeted to the plasma membrane of LLC-PK1 cells.

6.3 Localization of HA-epitope Tags

Determination of the exofacial position of inserted HA epitopes in SLC11A2 was performed using the anti-HA monoclonal antibody 16B12 on transfectants expressing singly HA-tagged SLC11A2 protein. Immunofluorescence was performed on both intact cells and those permeabilized with a low concentration of Triton X-100 allowing for the detection of extracellularly and intracellularly localized tags respectively (Figure 4: columns 2 and 4). Vector transfected LLC-PK1 cells serving as a negative control for antibody binding did not display a noticeable fluorescent signal in localization studies (Figure 4B,D). To ensure that absence of an epitope induced signal was not the result of an absence of cells in the field of view, DAPI nuclear staining was performed indicating

ample cell numbers for each cell line under both conditions tested (Figure 4: columns 1 and 3). Strong fluorescent signals were observed under both intact and permeabilized conditions for constructs 5b (HA between TMD 5/6), 7 (HA between TMD 7/8) and 11b (HA between TMD 11/12) (Figure 4: N,P; V,X; AL, AN). This suggested that the regions into which these tags were inserted in SLC11A2 were all found in the extracellular compartment (Figure 4: left panel schematic) and confirmed the previously reported exofacial localization of the fourth extracellular loop construct 7; [85]. Conversely, cells transfected with constructs NT (HA at N-terminus), 4 (HA between TMD 4/5), 6 (HA between TMD 6/7), 9 (HA between TMD 9/10), 10 (HA between TMD 10/11) and CT (HA at C-terminus) (Figure 4: F,H; J,L; R,T; AH,AJ; AP,AR), only produced a fluorescent signal under permeabilized conditions, suggesting the intracellular location of the HA epitope in these regions of SLC11A2 and confirming the previously reported cytosolic orientation of both termini constructs NT and CT; [86].

SLC11A2 isoform I is targeted to the apical surface of the plasma membrane in epithelial cells, ultimately becoming internalized in the late endosomes/lysosomes of the cytosol [83,104,149]. This cellular distribution was observed in the immunofluorescence panels in Figure 4 where apical membrane localization was seen in intact cells, whereas a punctate and perinuclear distribution was observed in permeabilized cells.

Constructs 8b and 9 will be dealt with separately from the others as they were not shown to be functionally active in metal transport experiments (Figure 3) and consequently it is questionable whether they conform to native SLC11A2 topology. Despite this caveat, cell lines expressing these constructs were examined by immunoflurescence like the others and fluorescence was monitored. Fluorescent signals were observed in construct 8b (HA between TMD 8/9) under both intact and permeabilized conditions, whereas fluorescence was only observed in permeabilized cells expressing construct 9 (HA between TMD 9/10) suggesting extracellular and intracellular locations, respectively (Figure 4: Z,AB; AD,AF).

Transfectants expressing mutants 5b, 8b and 11b contained two HA tags in tandem. All three constructs exhibited fluorescence under both intact and permeabilized conditions (Figure 4: 5b; 8b; 11b) suggesting an exofacial location for their inserted epitope tags. It should also be noted that variants of construct 9 containing two and three

epitope tags in tandem, were also created and exhibited the same epitope localization characteristics as the singly tagged construct (data not shown).

6.4 Corroboration of HA Epitope Localization in SLC11A2

The immunofluorescence results provided a qualitative assessment of epitope localization which could be used to propose a topological model for the SLC11A2 protein. To verify and strengthen this experimental data, we used an enzyme-based colorimetric assay, which provided a highly sensitive measure of HA epitope present at the LLC-PK1 plasma membrane relative to the total expressed as SLC11A2-HA in transfectants. To this end, cells were prepared similarly to the way they were during immunofluorescence studies but relative surface expression levels were measured using a secondary anti-mouse antibody coupled to horseradish peroxidase allowing for colorimetric analysis using a fluoremeter. LLC-PK1 cells stably transfected with constructs 5b, 7, 8b and 11b were found to have a significantly greater proportion of accessible HA-tagged SLC11A2 at their surface than the remaining constructs (NT, 4, 6, 9, 10 and CT; Figure 5). This was in accordance with the immunofluorescence studies which indicated the same division of constructs between the extra- and intracellular compartments (Figure 3). Interestingly, transfectants expressing constructs 5b, 8b and 11b, exhibited substantially more surface protein than the positive control (construct 7), possibly arising from the improved accessibility of their two tandem HA-epitopes in the extracellular space compared with the single epitope found in construct 7 (Figure 5).

7.0 Discussion

7.1 SLC11A topology model

SLC11A defines a large, highly conserved family of integral, proton-coupled, divalent metal transporters of medical relevance in human health and disease. A representative high-resolution crystal structure for these proteins has yet to be resolved. Through this study we have attempted to contribute to the growing base of information regarding the structure:function relationships of the SLC11A proteins with the immediate aim of experimentally validating their topology. In the absence of high resolution structural data, deciphering the topology of a membrane protein provides important information relevant to the design of subsequent structure:function studies. Previous computationally [20] and experimentally [85-87] derived data favors a 12 transmembrane topology with cytosolic termini for eukaryotic SLC11A proteins. The results obtained herein from our epitope mapping approach are consistent with this model, and experimentally validate the currently held topology model. Moreover, a corollary of this approach is that it allowed testing of structurally or functionally important regions of SLC11A2 in parallel, corroborating information collected by others implicating the N-terminal exofacial loop regions of SLC11A2 as of high functional importance [95].

Current topological models for eukaryotic SLC11A members are based predominantly on hydropathy profiling, hydrophobic-moment analysis and multiple sequence alignments of various orthologs [20]. These models suggest a cytosolic N-terminus followed by a core of ten highly conserved hydrophobic TMDs, and a variable C-terminal region which can include two additional less conserved TMDs and a cytosolic terminus in higher eukaryotes, or, a single TMD necessitating an exofacial terminal end in yeast and bacteria [20,87]. This model has been validated experimentally in the SLC11A2 bacterial ortholog MntH by random and targeted genetic fusions of cyto- and periplasmic reporters in the predicted hydrophilic loops of the protein, yielding an 11 TMD model with a periplasmic C-terminus as suggested initially by hydropathy analysis [87]. Despite their distant phylogenetic ancestry, MntH and eukaryotic SLC11A share a high degree of conservation and functional homology, with the former believed to be an evolutionary precursor of the latter [156], indicating that the prokaryotic topological model could apply to eukaryotes as well, especially in the conserved ten TMD core of the proteins. Further

experimental evidence to support the accuracy of the existing hydropathy-based topology predictions has been performed by immunofluorescence experiments in eukaryotic SLC11A by our group [84,157] and others [83]. By using antisera targeting a specific N-terminal region of the proteins, together with strategic insertion of epitope tags and construction of fusion proteins expressed in intact and permeabilized cells, it was demonstrated that both SLC11A1/2 terminal ends are intracellular, while the hydrophilic loop linking predicted TMDs 7 and 8 is extracellular [81,84,85]. These studies are consistent with the 2-dimensional topology predicted by the current, predominantly hydropathy-based models. Furthermore, limited 3-dimensional information has also been put forward. NMR and CD analyses have shown that synthetic peptides of various lengths [158,159] representing SLC11A2 TMD 4 and its N-terminal loop, adopt an alpha-helical secondary structure in membrane mimetic environments as would be expected of a TMD in an alpha-helical bundle protein. Moreover, by using computer modeling programs and holding experimentally derived and TMD localized, functional residues from *E. coli* MntH as constraints, Courville and colleagues suggested that SLC11A members may share a similar general fold as the recently crystallized 12 TMD SLC6 leucine transporter, LeuT [42,92]. This model suggests a pseudo-two fold symmetry for LeuT, in which TMDs 1 and 6 contain membrane spanning discontinuous helices involved in substrate recognition and ion binding. Furthermore, the LeuT crystal structure reveals a protein of rich topological diversity with features such as long obliquely tilted helical TMDs, one of which penetrates deep into the cytosol, along with secondary structural features and potential re-entrant behaviour in some of the hydrophilic loops [42]. This model has been used to elucidate several of the structural and functional intricacies observed in the SLC11A family [92].

Using an HA-epitope mapping technique, it was our goal to experimentally identify the number and membrane polarity of the predicted TMDs of mouse SLC11A2 in hopes of validating the C-terminal differences between eukaryotic and prokaryotic SLC11A proteins, and providing an accurate model on which to base future structure:function experiments. This topology mapping approach provides several benefits over existing experimental techniques such as limited proteolysis or gene fusions [56,60], in that the protein under study is maintained in its full-length, biologically active

state without the need for truncation. Furthermore, the tagged protein can be expressed in mammalian cells, where proper translocation and processing of the nascent protein can take place [8]. Several methods such as cysteine and glycosylation scanning mutagenesis also use the full-length protein for analysis, however they require that endogenous cysteines and glycosylation sites be removed [57,58]. Murine SLC11A2 has 10 cysteine residues, several of which are highly conserved even in prokaryotic MntH, suggesting that removal of these residues may compromise the structural or functional integrity of the protein. The creation of a cysteine-less SLC11A2 mutant by our group resulted in a non-functional protein (data not presented). Moreover, studies in which the glycosylation signals in the fourth extracellular loop of SLC11A2 were disrupted, demonstrated that the protein's apical membrane distribution is affected in MDCK epithelial cells, thus circumventing the use of glycosylation scanning to explore topology [83]. Using HA tags also has advantages over other similarly sized and well-characterized epitope sequences such as c-Myc [74], and FLAG [72]. The major advantage involves the low intrinsic charge of HA (-2) in comparison with c-Myc and FLAG (net charge of –3 each), which becomes particularly important in topology mapping studies where charge distribution can have an effect on the orientation of a protein in the membrane [35]. This approach was previously used by our group to determine the topology of P-glycoprotein [152] and the multidrug resistance-associated protein MRP [151], and more recently by others to map the topology of the human sodium-dependent bile acid transporter SLC10A2; [154].

7.2 Experimentally validated hydrophilic loops: constructs NT, 4-7, 10, 11 and CT

Epitope insertion sites in murine SLC11A2 (Figure 1B) were selected as those sites exhibiting limited conservation among orthologs and found within highly hydrophilic segments believed to be inter-TMD loops, as predicted by hydropathy analysis (Figure 1A). Modified proteins were expressed in mammalian cells enabling the monitoring of any deleterious effects on SLC11A2 function resulting from epitope insertion, and epitope accessibility was determined in intact and permeabilized cells to establish its intra- or extracellular location. Using this approach, we have unambiguously localized HA epitopes inserted at positions 243, 344 and 504 to the extracellular side of the membrane (Figures 1B and 4). Furthermore, HA epitopes inserted at positions 1, 201, 284, 468 and 561 were detected only in permeabilized cells and consequently assigned an

intracellular location (Figures 1B and 4). Expression analysis in these stable transfectants (Figure 2AB) identified the presence of two possible isoforms of SLC11A2 at 60 kDa and 90 kDa. A similar expression pattern has been observed in other studies for fully functional SLC11A2 and attributed to the immature core, and mature complex *N*-glycosylated forms of the protein respectively [86,115,149]. This qualitative data was further reiterated by a quantitative assessment gauging the relative presence of HA epitopes present at the cell surface compared with intracellularly distributed epitopes in each construct (Figure 5). Collectively, the epifluorescence microscopy data (Figure 4), taken together with hydropathy profiling (Figure 1B), computer modeling (data not shown) and cell surface labeling (Figure 5) for these hydrophilic loops, are strongly consistent with a 12 transmembrane domain topology model for SLC11A2 with intracellular termini (Figure 6; green segments).

7.3 Putative hydrophilic loops which could not be experimentally validated

The presence of a series of highly conserved, charged amino acid residues in the first nine TMDs of the hydrophobic core in SLC11A proteins has implicated these domains as important in functional or structural roles, an implication that has been verified by mutagenesis studies [88,89]. However, several of the relatively less conserved, intervening hydrophilic loops linking the core TMD may also serve an important functional role. In a seminal sequence-based analysis of SLC11A family members, Cellier and colleagues found that the intermittent hydrophilic loops between predicted TMDs 1-3 and 8-10 all share a conserved length, while sequence conservation was especially high in loops between predicted TMDs 1-4 and 8-9 [20]. The group noted that conservation of the amino acids constituting the hydrophilic loops of integral membrane proteins is a characteristic observed in ion channels, where these conserved hydrophilic segments contribute to a common channel architecture in orthologs [20,160]. It is perhaps not coincidental then, that all of the functionally inactive transfectants (constructs 1, 2, 3), and functionally reduced transfectants (constructs 8b and 9) in this study, have epitope insertions made in precisely these highly sequence and size conserved hydrophilic loops, warranting a more detailed discussion.

7.4 Transport null mutants: constructs 1-3

Constructs having epitope insertions at positions 98, 131 and 175 (constructs 1, 2 and 3 respectively; Figure 1B) did not exhibit any metal transport activity, demonstrating calcein quench rates similar to the empty vector control (Figure 3). Interestingly, expression analyses in these three transfectants, showed a complete preponderance of the immature 60 kDa SLC11A2 isoform, and no detectable presence of the mature and highly glycosylated 90 kDa protein (Figure 2A). The predominant 60 kDa species observed in these transfectants and their metal transport defects, appear to suggest that there exists a correlation between these two states resulting from the insertion of an epitope tag within either of the first three putative hydrophilic loops of SLC11A2 (Figure 1AB). For example, epitope insertion within these constructs may directly result in inefficient protein maturation, protein retention in the endoplasmic reticulum, diminished protein targeting to the plasma membrane, and/or a reduced half-life resulting in the observed expression pattern (Figure 2A). These properties would subsequently contribute to the observed lack of metal transport activity (Figure 3) by conferring a protein that was improperly folded or retained at the plasma membrane in a non-functional state. These properties, together with an absence of any detectable signal in the immunofluorescence epitope localization experiment (Figure 4; data not shown), prevented the mapping of epitopes inserted at these sites, thus curtailing the validation of membrane polarity for the putative hydrophilic loops into which they were inserted (loops 1-3; Figure 6).

In a mutagenesis study of the first predicted exofacial loop in SLC11A2, Cohen and colleagues changed several unconserved residues and expressed the mutated protein in yeast strains lacking the endogenous SLC11A ortholog SMF1p, as well as in *Xenopus* oocytes which do not express an endogenous ortholog of the protein [95]. Nearly every mutation made (particularly to G119, D124 and Q126) influenced uptake activity, substrate specificity and proton affinity in SLC11A2, leading the authors to conclude that the first exofacial loop of the transporter is involved in metal ion binding and proton coupling [95,161]. If single point mutations in this loop had such a deleterious effect on transport function, then perhaps it is not surprising that insertion of an entire HA epitope tag may completely abrogate function and even affect proper folding of the protein, generating the expression and transport profiles observed in this study (Figures 2AB, 3).

This analogy could also be put forward to predicted intra- and extracellular loops 2 and 3 which exhibited similar expression and metal transport defects to that of loop 1 (Figures 2AB, 3). The leucine transporter, LeuT, which has recently been suggested to share a similar fold to SLC11 proteins [92], bears substantial secondary structure in the hydrophilic loops between TMDs 2-4 and 7-8, loops implicated in conformational changes during substrate transport [42]. The presence of small α-helices or even β-strands in the hydrophilic regions defining loops 2 and 3 of SLC11A2 (Figure 1A), presumably playing a functional or structural role, could explain why epitope insertion at these sites (Figure 1B) results in the expression pattern (Figure 2AB) and deleterious effects (Figure 3) observed in our study. This may especially apply to predicted extracellular loop 3, which is positioned between putative TMD 3/4 – an area of particular predictive ambiguity as exemplified by hydropathy profiling (Figure 1A).

7.5 Ambiguous mutants: constructs 8b and 9

Constructs containing epitope insertions at positions 403 and 432 (constructs 8b and 9 respectively; Figure 1B), provided comparatively ambiguous data to that collected for the previously discussed functional and non-functional constructs. These constructs harboured epitopes which were accessible to antibody, as evidenced by the immunofluorescence localization study (Figure 4), yet suggested a membrane orientation for hydrophilic loops 8 and 9 which contradicted the proposed topology model (Figure 6). Interestingly, these two constructs also displayed marginally more activity than the transport null mutants (construct 1-3; Figure 3), and appeared to express both core and complex glycosylated SLC11A2 isoforms (Figure 2AB). The fraction of core/complex glycosylated protein for construct 9, however, appeared modestly shifted towards the core glycosylated protein (Figure 2AB), lending an explanation as to why construct 9 (predicted to be extracellular by hydropathy analysis; Figure 1 AB) only displayed a fluorescent epitope-induced signal under permeabilized conditions and not in intact cells as would be expected (Figure 4, AD and AF). This epitope-modified protein may be affected by the same processing/targeting defects as constructs 1, 2 and 3 (discussed above), and similarly may not be targeted to the plasma membrane or may be improperly folded at the bilayer, thus hindering epitope accessibility. Notably, construct 8b (Figure 2AB), which also displayed a significant loss of transport acitivity (on a level similar to

construct 9), appeared to be expressed in the form of both glycosylated variants, indicating that it was properly processed and targeted to the plasma membrane. This would suggest that the site of epitope insertion in this construct (see Figure 1B) may lead to functional inactivity resulting not from targeting, folding or maturation defects in the protein *per se*, but rather from a change in the ability to bind and transport divalent metals.

The fourth predicted intracellular loop of SLC11A2 is highly conserved among orthologs and contains a 21 residue long segment known as the conserved transport motif (CTM). The CTM shows several invariant residues and mutagenesis studies demonstrate that changes to key amino acids within this site abrogate transport function [91,93]. An epitope insertion within this putative loop was made at the C-terminal periphery of the CTM (Figure 1B) to avoid any deleterious effects possibly associated with compromising the integrity of this vital segment. Best effort aside, construct 8b was not functional (Figure 3) despite being properly processed as evidenced by expression analysis (Figure 2AB). A previous attempt at HA epitope insertion within the same homologous loop in the yeast ortholog SMF1 was similarly shown to abrogate function [162], further highlighting the functional importance of the region. More recently, in a yeast complementation study involving SLC11A1/2 chimeras, Techau *et al* identified four amino acid residues potentially involved in mediating metal/proton symport in a region bordered by TMD 8 and the C-terminus of SLC11A2 [163]. One of these candidate residues was N403, whose carbonyl oxygens provide it with metal binding properties and, which directly precedes the epitope insertion site in our study Figure 1B; [163]. Consequently it is enticing to suggest that insertion of the HA epitope apposing N403 may interfere with the metal binding affinity of this residue, or that of other highly charged proximal residues such as K405 and/or R408, which in turn could account for the lack of transport activity observed in construct 8b transfectants (Figure 3).

Lastly, although the lack of significant transport activity (Figure 3) in constructs 8b and 9 suggests that the native topology of SLC11A2 in these transfectants may be compromised - thus invalidating epitope localization studies done with these mutants (Figure 4Z, AB and AD, AF) - it is nonetheless worthwhile to use this information to gain insight into the native topology of the respective TMDs. For example, our hydropathy

predictions (Figure 1A) were largely based on the presumption that most TMDs spanning the lipid bilayer lie perpendicularly to its face and consequently consist of roughly 19 or 20 amino acids or so. However, this assumption may have led us to underestimate the actual size of several TMDs and their orientation within the membrane, resulting in epitope insertion sites that were actually made within *de facto* TMDs rather than hydrophilic loops. For example, in the LeuT transporter, several TMDs, including TMD 8 coincidently, consist of 30 or more amino acids and span the membrane at oblique angles [42]. Any epitope insertion within the hydrophobic core of a TMD would presumably greatly compromise native topology and function. Conversely, several transporters such as GLT-1 [164] and NXC1 [165] display re-entrant loops in several larger hydrophilic segments which may weave in and out of the membrane environment. The putative loop between TMDs 8/9 of SLC11A2 is estimated to be the second largest in the protein, after the loop between TMDs 7/8, suggesting that it too may demonstrate this re-entrant behaviour. Either of these two scenarios (large TMD or re-entrant characteristics) could account for the observation that the insertion of two HA epitope tags within loop 8 of SLC11A2 leads to a fluoresent signal in intact cells despite hydropathy predictions which would indicate the contrary (Figures 1AB, 4Z, AB). In other words, a partial transmembrane architecture in this loop could potentially cause a large double epitope tag insertion (as in loop 8) to bud out on the opposite side of the membrane resulting in a positive signal by immunofluorescence and the observed decrease in transport activity.

8.0 Conclusion

Mammalian SLC11A2 proteins have long been predicted to conform to a 12 transmembrane domain topology model that places their termini in the cytoplasm [20]. Recently, the topologically diverse leucine transporter LeuT was proposed to share a similar fold to SLC11A proteins [42,92]. The topological information collected herein is consistent with both these models and provides the first experimental validation (to our knowledge) of the orientation of the hydrophilic loops separating TMDs 4/5, 5/6, 6/7, 10/11 and 11/12 (Figure 6). Our inability to unambiguously map the putative hydrophilic loops separating TMDs 1/2, 2/3, 3/4, 8/9 and 9/10 (Figure 6) suggests that epitope insertion within these sites may compromise the functional and/or structural properties of SLC11A2. These sites should be central in future structure:function studies and an attempt at mapping their topology may involve repeating the same epitope mapping approach but making insertions at alternate positions within these putative loops, or using an alternate experimental approach. The data collected herein is also consistent with experimentally derived topology models made in prokaryotic SLC11 orthologs [87], and consequently suggests that it will be equally applicable to other eukaryotic SLC11A proteins.

9.0 Acknowledgements

Foremost, I would like to thank my supervisor, Philippe Gros, for his guidance, patience and encouragement during the course of this project and the writing of this thesis. It was greatly appreciated. I would also like to thank Steven Lam, a past member of the Gros lab from whom I took over the SLC11A project. Steven assisted with primer design and provided invaluable technical support and training during the early stages of this project, introducing me to many of the techniques used during my graduate studies. I am also grateful to other past and present members of the Gros lab for providing a fantastic working atmosphere and for their encouragement and friendship over the years. I am also grateful to Anne Fortier for providing the French translation of the abstract to this thesis and her contagious '*lâches pas'* attitude.

Last, but definitely not least, I thank my family for their unrelenting support, encouragement and unconditional love during the course of all of my life's projects.

10.0 Abbreviations

aa: amino acids
HA: hemagglutinin
IMP: integral membrane protein
IRE: iron responsive element
NRAMP: Natural resistance associated macrophage protein
PCR: polymerase chain reaction
SDS-PAGE: sodium dodecyl sulfate polyacrilamide gel electrophoresis
SLC11A1/2: Solute carrier 11 member 1/2
UTR: untranslated region

References

1. von Heijne, G. The membrane protein universe: what's out there and why bother? *J Intern Med* **261**, 543-57 (2007).
2. Krogh, A., Larsson, B., von Heijne, G. & Sonnhammer, E.L. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* **305**, 567-80 (2001).
3. Hopkins, A.L. & Groom, C.R. The druggable genome. *Nat Rev Drug Discov* **1**, 727-30 (2002).
4. Granseth, E., Seppala, S., Rapp, M., Daley, D.O. & Von Heijne, G. Membrane protein structural biology--how far can the bugs take us? *Mol Membr Biol* **24**, 329-32 (2007).
5. Kendrew, J.C. et al. A three-dimensional model of the myoglobin molecule obtained by x-ray analysis. *Nature* **181**, 662-6 (1958).
6. Berman, H.M. et al. The Protein Data Bank. *Nucleic Acids Res* **28**, 235-42 (2000).
7. Wiener, M.C. A pedestrian guide to membrane protein crystallization. *Methods* **34**, 364-72 (2004).
8. van Geest, M. & Lolkema, J.S. Membrane topology and insertion of membrane proteins: search for topogenic signals. *Microbiol Mol Biol Rev* **64**, 13-33 (2000).
9. Elofsson, A. & von Heijne, G. Membrane protein structure: prediction versus reality. *Annu Rev Biochem* **76**, 125-40 (2007).
10. Dunn, L.L., Rahmanto, Y.S. & Richardson, D.R. Iron uptake and metabolism in the new millennium. *Trends Cell Biol* **17**, 93-100 (2007).
11. Bassuny, W.M. et al. Association study of the NRAMP1 gene promoter polymorphism and early-onset type 1 diabetes. *Immunogenetics* **54**, 282-5 (2002).
12. John, S. et al. Linkage and association studies of the natural resistance associated macrophage protein 1 (NRAMP1) locus in rheumatoid arthritis. *J Rheumatol* **24**, 452-7 (1997).
13. Vidal, S.M., Malo, D., Vogan, K., Skamene, E. & Gros, P. Natural resistance to infection with intracellular parasites: isolation of a candidate for Bcg. *Cell* **73**, 469-85 (1993).
14. Beaumont, C. et al. Two new human DMT1 gene mutations in a patient with microcytic anemia, low ferritinemia, and liver iron overload. *Blood* **107**, 4168-70 (2006).
15. Iolascon, A. et al. Microcytic anemia and hepatic iron overload in a child with compound heterozygous mutations in DMT1 (SCL11A2). *Blood* **107**, 349-54 (2006).
16. Mims, M.P. et al. Identification of a human mutation of DMT1 in a patient with microcytic anemia and iron overload. *Blood* **105**, 1337-42 (2005).
17. Galdiero, S., Galdiero, M. & Pedone, C. beta-Barrel membrane bacterial proteins: structure, function, assembly and interaction with lipids. *Curr Protein Pept Sci* **8**, 63-82 (2007).
18. von Heijne, G. Principles of membrane protein assembly and structure. *Prog Biophys Mol Biol* **66**, 113-39 (1996).
19. Gruber, M. & Lupas, A.N. Historical review: another 50th anniversary--new periodicities in coiled coils. *Trends Biochem Sci* **28**, 679-85 (2003).

20. Cellier, M. et al. Nramp defines a family of membrane proteins. *Proc Natl Acad Sci U S A* **92**, 10089-93 (1995).
21. Mitra, K. et al. Structure of the E. coli protein-conducting channel bound to a translating ribosome. *Nature* **438**, 318-24 (2005).
22. Skach, W.R. The expanding role of the ER translocon in membrane protein folding. *J Cell Biol* **179**, 1333-5 (2007).
23. Van den Berg, B. et al. X-ray structure of a protein-conducting channel. *Nature* **427**, 36-44 (2004).
24. Higy, M., Junne, T. & Spiess, M. Topogenesis of membrane proteins at the endoplasmic reticulum. *Biochemistry* **43**, 12716-22 (2004).
25. Sadlish, H., Pitonzo, D., Johnson, A.E. & Skach, W.R. Sequential triage of transmembrane segments by Sec61alpha during biogenesis of a native multispanning membrane protein. *Nat Struct Mol Biol* **12**, 870-8 (2005).
26. Kida, Y., Morimoto, F. & Sakaguchi, M. Two translocating hydrophilic segments of a nascent chain span the ER membrane during multispanning protein topogenesis. *J Cell Biol* **179**, 1441-52 (2007).
27. Borel, A.C. & Simon, S.M. Biogenesis of polytopic membrane proteins: membrane segments assemble within translocation channels prior to membrane integration. *Cell* **85**, 379-89 (1996).
28. Brown, C.A. & Black, S.D. Membrane topology of mammalian cytochromes P-450 from liver endoplasmic reticulum. Determination by trypsinolysis of phenobarbital-treated microsomes. *J Biol Chem* **264**, 4442-9 (1989).
29. Hartmann, E., Rapoport, T.A. & Lodish, H.F. Predicting the orientation of eukaryotic membrane-spanning proteins. *Proc Natl Acad Sci U S A* **86**, 5786-90 (1989).
30. Ott, C.M. & Lingappa, V.R. Integral membrane protein biosynthesis: why topology is hard to predict. *J Cell Sci* **115**, 2003-9 (2002).
31. Singer, S.J., Maher, P.A. & Yaffe, M.P. On the transfer of integral proteins into membranes. *Proc Natl Acad Sci U S A* **84**, 1960-4 (1987).
32. Bayle, D. et al. In vitro translation analysis of integral membrane proteins. *J Recept Signal Transduct Res* **17**, 29-56 (1997).
33. Mothes, W., Jungnickel, B., Brunner, J. & Rapoport, T.A. Signal sequence recognition in cotranslational translocation by protein components of the endoplasmic reticulum membrane. *J Cell Biol* **142**, 355-64 (1998).
34. Do, H., Falcone, D., Lin, J., Andrews, D.W. & Johnson, A.E. The cotranslational integration of membrane proteins into the phospholipid bilayer is a multistep process. *Cell* **85**, 369-78 (1996).
35. von Heijne, G. Control of topology and mode of assembly of a polytopic membrane protein by positively charged residues. *Nature* **341**, 456-8 (1989).
36. von Heijne, G. & Gavel, Y. Topogenic signals in integral membrane proteins. *Eur J Biochem* **174**, 671-8 (1988).
37. von Heijne, G. Recent advances in the understanding of membrane protein assembly and structure. *Q Rev Biophys* **32**, 285-307 (1999).
38. Grigorieff, N., Ceska, T.A., Downing, K.H., Baldwin, J.M. & Henderson, R. Electron-crystallographic refinement of the structure of bacteriorhodopsin. *J Mol Biol* **259**, 393-421 (1996).

39. Dutzler, R., Campbell, E.B., Cadene, M., Chait, B.T. & MacKinnon, R. X-ray structure of a ClC chloride channel at 3.0 A reveals the molecular basis of anion selectivity. *Nature* **415**, 287-94 (2002).
40. Yernool, D., Boudker, O., Jin, Y. & Gouaux, E. Structure of a glutamate transporter homologue from Pyrococcus horikoshii. *Nature* **431**, 811-8 (2004).
41. Toyoshima, C., Nakasako, M., Nomura, H. & Ogawa, H. Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 A resolution. *Nature* **405**, 647-55 (2000).
42. Yamashita, A., Singh, S.K., Kawate, T., Jin, Y. & Gouaux, E. Crystal structure of a bacterial homologue of Na+/Cl--dependent neurotransmitter transporters. *Nature* **437**, 215-23 (2005).
43. Screpanti, E. & Hunte, C. Discontinuous membrane helices in transport proteins and their correlation with function. *J Struct Biol* **159**, 261-7 (2007).
44. Hegde, R.S. et al. A transmembrane form of the prion protein in neurodegenerative disease. *Science* **279**, 827-34 (1998).
45. Kim, S.J., Rahbar, R. & Hegde, R.S. Combinatorial control of prion protein biogenesis by the signal sequence and transmembrane domain. *J Biol Chem* **276**, 26132-40 (2001).
46. Dunlop, J., Jones, P.C. & Finbow, M.E. Membrane insertion and assembly of ductin: a polytopic channel with dual orientations. *EMBO J* **14**, 3609-16 (1995).
47. Kyte, J. & Doolittle, R.F. A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**, 105-32 (1982).
48. Punta, M. et al. Membrane protein prediction methods. *Methods* **41**, 460-74 (2007).
49. Melen, K., Krogh, A. & von Heijne, G. Reliability measures for membrane protein topology prediction algorithms. *J Mol Biol* **327**, 735-44 (2003).
50. Tusnady, G.E. & Simon, I. Principles governing amino acid composition of integral membrane proteins: application to topology prediction. *J Mol Biol* **283**, 489-506 (1998).
51. Jones, D.T., Taylor, W.R. & Thornton, J.M. A mutation data matrix for transmembrane proteins. *FEBS Lett* **339**, 269-75 (1994).
52. Claros, M.G. & von Heijne, G. TopPred II: an improved software for membrane protein structure predictions. *Comput Appl Biosci* **10**, 685-6 (1994).
53. von Heijne, G. Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. *J Mol Biol* **225**, 487-94 (1992).
54. Nilsson, J., Persson, B. & von Heijne, G. Consensus predictions of membrane protein topology. *FEBS Lett* **486**, 267-9 (2000).
55. Daley, D.O. et al. Global topology analysis of the Escherichia coli inner membrane proteome. *Science* **308**, 1321-3 (2005).
56. Manoil, C. Analysis of membrane protein topology using alkaline phosphatase and beta-galactosidase gene fusions. *Methods Cell Biol* **34**, 61-75 (1991).
57. Chang, X.B., Hou, Y.X., Jensen, T.J. & Riordan, J.R. Mapping of cystic fibrosis transmembrane conductance regulator membrane topology by glycosylation site insertion. *J Biol Chem* **269**, 18572-5 (1994).
58. Kimura, T., Ohnuma, M., Sawai, T. & Yamaguchi, A. Membrane topology of the transposon 10-encoded metal-tetracycline/H+ antiporter as studied by site-directed chemical labeling. *J Biol Chem* **272**, 580-5 (1997).

59. Canfield, V.A. & Levenson, R. Transmembrane organization of the Na,K-ATPase determined by epitope addition. *Biochemistry* **32**, 13782-6 (1993).
60. Wilkinson, B.M., Critchley, A.J. & Stirling, C.J. Determination of the transmembrane topology of yeast Sec61p, an essential component of the endoplasmic reticulum translocation complex. *J Biol Chem* **271**, 25590-7 (1996).
61. Manoil, C. & Beckwith, J. A genetic approach to analyzing membrane protein topology. *Science* **233**, 1403-8 (1986).
62. Broome-Smith, J.K., Tadayyon, M. & Zhang, Y. Beta-lactamase as a probe of membrane protein assembly and protein export. *Mol Microbiol* **4**, 1637-44 (1990).
63. Silhavy, T.J., Shuman, H.A., Beckwith, J. & Schwartz, M. Use of gene fusions to study outer membrane protein localization in Escherichia coli. *Proc Natl Acad Sci U S A* **74**, 5411-5 (1977).
64. Zelazny, A. & Bibi, E. Biogenesis and topology of integral membrane proteins: characterization of lactose permease-chloramphenicol acetyltransferase hybrids. *Biochemistry* **35**, 10872-8 (1996).
65. Danielsen, S., Boyd, D. & Neuhard, J. Membrane topology analysis of the Escherichia coli cytosine permease. *Microbiology* **141 (Pt 11)**, 2905-13 (1995).
66. Bibi, E. & Beja, O. Membrane topology of multidrug resistance protein expressed in Escherichia coli. N-terminal domain. *J Biol Chem* **269**, 19910-5 (1994).
67. Geller, D. et al. Comparative topology studies in Saccharomyces cerevisiae and in Escherichia coli. The N-terminal half of the yeast ABC protein Ste6. *J Biol Chem* **271**, 13746-53 (1996).
68. Eitinger, T. & Friedrich, B. A topological model for the high-affinity nickel transporter of Alcaligenes eutrophus. *Mol Microbiol* **12**, 1025-32 (1994).
69. Welply, J.K., Shenbagamurthi, P., Lennarz, W.J. & Naider, F. Substrate recognition by oligosaccharyltransferase. Studies on glycosylation of modified Asn-X-Thr/Ser tripeptides. *J Biol Chem* **258**, 11856-63 (1983).
70. Chen, J.G., Liu-Chen, S. & Rudnick, G. Determination of external loop topology in the serotonin transporter by site-directed chemical labeling. *J Biol Chem* **273**, 12675-81 (1998).
71. Zhou, J., Fazzio, R.T. & Blair, D.F. Membrane topology of the MotA protein of Escherichia coli. *J Mol Biol* **251**, 237-42 (1995).
72. Einhauer, A. & Jungbauer, A. The FLAG peptide, a versatile fusion tag for the purification of recombinant proteins. *J Biochem Biophys Methods* **49**, 455-65 (2001).
73. Field, J. et al. Purification of a RAS-responsive adenylyl cyclase complex from Saccharomyces cerevisiae by use of an epitope addition method. *Mol Cell Biol* **8**, 2159-65 (1988).
74. Evan, G.I., Lewis, G.K., Ramsay, G. & Bishop, J.M. Isolation of monoclonal antibodies specific for human c-myc proto-oncogene product. *Mol Cell Biol* **5**, 3610-6 (1985).
75. Bradley, D.J. Letter: Genetic control of natural resistance to Leishmania donovani. *Nature* **250**, 353-4 (1974).
76. Plant, J. & Glynn, A.A. Genetics of resistance to infection with Salmonella typhimurium in mice. *J Infect Dis* **133**, 72-8 (1976).
77. Gruenheid, S., Cellier, M., Vidal, S. & Gros, P. Identification and characterization of a second mouse Nramp gene. *Genomics* **25**, 514-25 (1995).

78. Gunshin, H. et al. Cloning and characterization of a mammalian proton-coupled metal-ion transporter. *Nature* **388**, 482-8 (1997).
79. Richer, E., Courville, P., Bergevin, I. & Cellier, M.F. Horizontal gene transfer of "prototype" Nramp in bacteria. *J Mol Evol* **57**, 363-76 (2003).
80. Courville, P., Chaloupka, R. & Cellier, M.F. Recent progress in structure-function analyses of Nramp proton-dependent metal-ion transporters. *Biochem Cell Biol* **84**, 960-78 (2006).
81. Vidal, S.M., Pinner, E., Lepage, P., Gauthier, S. & Gros, P. Natural resistance to intracellular infections: Nramp1 encodes a membrane phosphoglycoprotein absent in macrophages from susceptible (Nramp1 D169) mouse strains. *J Immunol* **157**, 3559-68 (1996).
82. White, J.K., Stewart, A., Popoff, J.F., Wilson, S. & Blackwell, J.M. Incomplete glycosylation and defective intracellular targeting of mutant solute carrier family 11 member 1 (Slc11a1). *Biochem J* **382**, 811-9 (2004).
83. Tabuchi, M., Tanaka, N., Nishida-Kitayama, J., Ohno, H. & Kishi, F. Alternative splicing regulates the subcellular localization of divalent metal transporter 1 isoforms. *Mol Biol Cell* **13**, 4371-87 (2002).
84. Forbes, J.R. & Gros, P. Iron, manganese, and cobalt transport by Nramp1 (Slc11a1) and Nramp2 (Slc11a2) expressed at the plasma membrane. *Blood* **102**, 1884-92 (2003).
85. Picard, V., Govoni, G., Jabado, N. & Gros, P. Nramp 2 (DCT1/DMT1) expressed at the plasma membrane transports iron and other divalent cations into a calcein-accessible cytoplasmic pool. *J Biol Chem* **275**, 35738-45 (2000).
86. Touret, N., Furuya, W., Forbes, J., Gros, P. & Grinstein, S. Dynamic traffic through the recycling compartment couples the metal transporter Nramp2 (DMT1) with the transferrin receptor. *J Biol Chem* **278**, 25548-57 (2003).
87. Courville, P., Chaloupka, R., Veyrier, F. & Cellier, M.F. Determination of transmembrane topology of the Escherichia coli natural resistance-associated macrophage protein (Nramp) ortholog. *J Biol Chem* **279**, 3318-26 (2004).
88. Lam-Yuk-Tseung, S., Govoni, G., Forbes, J. & Gros, P. Iron transport by Nramp2/DMT1: pH regulation of transport by 2 histidines in transmembrane domain 6. *Blood* **101**, 3699-707 (2003).
89. Mackenzie, B., Ujwal, M.L., Chang, M.H., Romero, M.F. & Hediger, M.A. Divalent metal-ion transporter DMT1 mediates both H+ -coupled Fe2+ transport and uncoupled fluxes. *Pflugers Arch* **451**, 544-58 (2006).
90. Nevo, Y. & Nelson, N. The mutation F227I increases the coupling of metal ion transport in DCT1. *J Biol Chem* **279**, 53056-61 (2004).
91. Chaloupka, R. et al. Identification of functional amino acids in the Nramp family by a combination of evolutionary analysis and biophysical studies of metal and proton cotransport in vivo. *Biochemistry* **44**, 726-33 (2005).
92. Courville, P. et al. Solute carrier 11 cation symport requires distinct residues in transmembrane helices 1 and 6. *J Biol Chem* **283**, 9651-8 (2008).
93. Pinner, E., Gruenheid, S., Raymond, M. & Gros, P. Functional complementation of the yeast divalent cation transporter family SMF by NRAMP2, a member of the mammalian natural resistance-associated macrophage protein family. *J Biol Chem* **272**, 28933-8 (1997).

94. Supek, F., Supekova, L., Nelson, H. & Nelson, N. A yeast manganese transporter related to the macrophage protein involved in conferring resistance to mycobacteria. *Proc Natl Acad Sci U S A* **93**, 5105-10 (1996).
95. Cohen, A., Nevo, Y. & Nelson, N. The first external loop of the metal ion transporter DCT1 is involved in metal ion binding and specificity. *Proc Natl Acad Sci U S A* **100**, 10694-9 (2003).
96. Dassa, E. & Hofnung, M. Sequence of gene malG in E. coli K12: homologies between integral membrane components from binding protein-dependent transport systems. *EMBO J* **4**, 2287-93 (1985).
97. Kerppola, R.E. & Ames, G.F. Topology of the hydrophobic membrane-bound components of the histidine periplasmic permease. Comparison with other members of the family. *J Biol Chem* **267**, 2329-36 (1992).
98. Vidal, S. et al. The Ity/Lsh/Bcg locus: natural resistance to infection with intracellular parasites is abrogated by disruption of the Nramp1 gene. *J Exp Med* **182**, 655-66 (1995).
99. Lee, P.L., Gelbart, T., West, C., Halloran, C. & Beutler, E. The human Nramp2 gene: characterization of the gene structure, alternative splicing, promoter region and polymorphisms. *Blood Cells Mol Dis* **24**, 199-215 (1998).
100. Hubert, N. & Hentze, M.W. Previously uncharacterized isoforms of divalent metal transporter (DMT)-1: implications for regulation and cellular function. *Proc Natl Acad Sci U S A* **99**, 12345-50 (2002).
101. Ludwiczek, S. et al. Ca2+ channel blockers reverse iron overload by a new mechanism via divalent metal transporter-1. *Nat Med* **13**, 448-54 (2007).
102. Lis, A. et al. Hypoxia induces changes in expression of isoforms of the divalent metal transporter (DMT1) in rat pheochromocytoma (PC12) cells. *Biochem Pharmacol* **69**, 1647-55 (2005).
103. Johnson, D.M., Yamaji, S., Tennant, J., Srai, S.K. & Sharp, P.A. Regulation of divalent metal transporter expression in human intestinal epithelial cells following exposure to non-haem iron. *FEBS Lett* **579**, 1923-9 (2005).
104. Canonne-Hergaux, F., Gruenheid, S., Ponka, P. & Gros, P. Cellular and subcellular localization of the Nramp2 iron transporter in the intestinal brush border and regulation by dietary iron. *Blood* **93**, 4406-17 (1999).
105. Griffiths, W.J., Kelly, A.L., Smith, S.J. & Cox, T.M. Localization of iron transport and regulatory proteins in human cells. *QJM* **93**, 575-87 (2000).
106. Trinder, D., Oates, P.S., Thomas, C., Sadleir, J. & Morgan, E.H. Localisation of divalent metal transporter 1 (DMT1) to the microvillus membrane of rat duodenal enterocytes in iron deficiency, but to hepatocytes in iron overload. *Gut* **46**, 270-6 (2000).
107. Yeh, K.Y., Yeh, M., Watkins, J.A., Rodriguez-Paris, J. & Glass, J. Dietary iron induces rapid changes in rat intestinal divalent metal transporter expression. *Am J Physiol Gastrointest Liver Physiol* **279**, G1070-9 (2000).
108. Canonne-Hergaux, F. et al. Expression and subcellular localization of NRAMP1 in human neutrophil granules. *Blood* **100**, 268-75 (2002).
109. Ferguson, C.J. et al. Cellular localization of divalent metal transporter DMT-1 in rat kidney. *Am J Physiol Renal Physiol* **280**, F803-14 (2001).
110. Canonne-Hergaux, F. & Gros, P. Expression of the iron transporter DMT1 in kidney from normal and anemic mk mice. *Kidney Int* **62**, 147-56 (2002).

111. Gruenheid, S. et al. The iron transport protein NRAMP2 is an integral membrane glycoprotein that colocalizes with transferrin in recycling endosomes. *J Exp Med* **189**, 831-41 (1999).
112. Zhang, A.S., Xiong, S., Tsukamoto, H. & Enns, C.A. Localization of iron metabolism-related mRNAs in rat liver indicate that HFE is expressed predominantly in hepatocytes. *Blood* **103**, 1509-14 (2004).
113. Jabado, N., Canonne-Hergaux, F., Gruenheid, S., Picard, V. & Gros, P. Iron transporter Nramp2/DMT-1 is associated with the membrane of phagosomes in macrophages and Sertoli cells. *Blood* **100**, 2617-22 (2002).
114. Georgieff, M.K., Wobken, J.K., Welle, J., Burdo, J.R. & Connor, J.R. Identification and localization of divalent metal transporter-1 (DMT-1) in term human placenta. *Placenta* **21**, 799-804 (2000).
115. Lam-Yuk-Tseung, S., Touret, N., Grinstein, S. & Gros, P. Carboxyl-terminus determinants of the iron transporter DMT1/SLC11A2 isoform II (-IRE/1B) mediate internalization from the plasma membrane into recycling endosomes. *Biochemistry* **44**, 12149-59 (2005).
116. Lam-Yuk-Tseung, S. & Gros, P. Distinct targeting and recycling properties of two isoforms of the iron transporter DMT1 (NRAMP2, Slc11A2). *Biochemistry* **45**, 2294-301 (2006).
117. Gruenheid, S., Pinner, E., Desjardins, M. & Gros, P. Natural resistance to infection with intracellular pathogens: the Nramp1 protein is recruited to the membrane of the phagosome. *J Exp Med* **185**, 717-30 (1997).
118. Blackwell, J.M. et al. SLC11A1 (formerly NRAMP1) and disease resistance. *Cell Microbiol* **3**, 773-84 (2001).
119. Goswami, T. et al. Natural-resistance-associated macrophage protein 1 is an H+/bivalent cation antiporter. *Biochem J* **354**, 511-9 (2001).
120. Kuhn, D.E., Baker, B.D., Lafuse, W.P. & Zwilling, B.S. Differential iron transport into phagosomes isolated from the RAW264.7 macrophage cell lines transfected with Nramp1Gly169 or Nramp1Asp169. *J Leukoc Biol* **66**, 113-9 (1999).
121. Kuhn, D.E., Lafuse, W.P. & Zwilling, B.S. Iron transport into mycobacterium avium-containing phagosomes from an Nramp1(Gly169)-transfected RAW264.7 macrophage cell line. *J Leukoc Biol* **69**, 43-9 (2001).
122. Zwilling, B.S., Kuhn, D.E., Wikoff, L., Brown, D. & Lafuse, W. Role of iron in Nramp1-mediated inhibition of mycobacterial growth. *Infect Immun* **67**, 1386-92 (1999).
123. Forbes, J.R. & Gros, P. Divalent-metal transport by NRAMP proteins at the interface of host-pathogen interactions. *Trends Microbiol* **9**, 397-403 (2001).
124. Schnappinger, D. et al. Transcriptional Adaptation of Mycobacterium tuberculosis within Macrophages: Insights into the Phagosomal Environment. *J Exp Med* **198**, 693-704 (2003).
125. Boyer, E., Bergevin, I., Malo, D., Gros, P. & Cellier, M.F. Acquisition of Mn(II) in addition to Fe(II) is required for full virulence of Salmonella enterica serovar Typhimurium. *Infect Immun* **70**, 6032-42 (2002).
126. Huynh, C., Sacks, D.L. & Andrews, N.W. A Leishmania amazonensis ZIP family iron transporter is essential for parasite replication within macrophage phagolysosomes. *J Exp Med* **203**, 2363-75 (2006).

127. Corbin, B.D. et al. Metal chelation and inhibition of bacterial growth in tissue abscesses. *Science* **319**, 962-5 (2008).
128. Govoni, G. et al. The Bcg/Ity/Lsh locus: genetic transfer of resistance to infections in C57BL/6J mice transgenic for the Nramp1 Gly169 allele. *Infect Immun* **64**, 2923-9 (1996).
129. Abel, L. et al. Susceptibility to leprosy is linked to the human NRAMP1 gene. *J Infect Dis* **177**, 133-45 (1998).
130. Koh, W.J. et al. NRAMP1 gene polymorphism and susceptibility to nontuberculous mycobacterial lung diseases. *Chest* **128**, 94-101 (2005).
131. Marquet, S. et al. Variants of the human NRAMP1 gene and altered human immunodeficiency virus infection susceptibility. *J Infect Dis* **180**, 1521-5 (1999).
132. Kim, J.H. et al. NRAMP1 genetic polymorphisms as a risk factor of tuberculous pleurisy. *Int J Tuberc Lung Dis* **7**, 370-5 (2003).
133. Blackwell, J.M. et al. Genomic organization and sequence of the human NRAMP gene: identification and mapping of a promoter region polymorphism. *Mol Med* **1**, 194-205 (1995).
134. Cellier, M. et al. Human natural resistance-associated macrophage protein: cDNA cloning, chromosomal mapping, genomic organization, and tissue-specific expression. *J Exp Med* **180**, 1741-52 (1994).
135. Kishi, F., Tanizawa, Y. & Nobumoto, M. Structural analysis of human natural resistance-associated macrophage protein 1 promoter. *Mol Immunol* **33**, 265-8 (1996).
136. Blackwell, J.M., Searle, S., Mohamed, H. & White, J.K. Divalent cation transport and susceptibility to infectious and autoimmune disease: continuation of the Ity/Lsh/Bcg/Nramp1/Slc11a1 gene story. *Immunol Lett* **85**, 197-203 (2003).
137. Kissler, S. et al. In vivo RNA interference demonstrates a role for Nramp1 in modifying susceptibility to type 1 diabetes. *Nat Genet* **38**, 479-83 (2006).
138. Shaw, M.A. et al. Linkage of rheumatoid arthritis to the candidate gene NRAMP1 on 2q35. *J Med Genet* **33**, 672-7 (1996).
139. Hofmeister, A., Neibergs, H.L., Pokorny, R.M. & Galandiuk, S. The natural resistance-associated macrophage protein gene is associated with Crohn's disease. *Surgery* **122**, 173-8; discussion 178-9 (1997).
140. Kotze, M.J. et al. Analysis of the NRAMP1 gene implicated in iron transport: association with multiple sclerosis and age effects. *Blood Cells Mol Dis* **27**, 44-53 (2001).
141. Fleming, M.D. et al. Microcytic anaemia mice have a mutation in Nramp2, a candidate iron transporter gene. *Nat Genet* **16**, 383-6 (1997).
142. Edwards, J.A. & Hoke, J.E. Defect of intestinal mucosal iron uptake in mice with hereditary microcytic anemia. *Proc Soc Exp Biol Med* **141**, 81-4 (1972).
143. Edwards, J.A. & Hoke, J.E. Red cell iron uptake in hereditary microcytic anemia. *Blood* **46**, 381-8 (1975).
144. Fleming, M.D. et al. Nramp2 is mutated in the anemic Belgrade (b) rat: evidence of a role for Nramp2 in endosomal iron transport. *Proc Natl Acad Sci U S A* **95**, 1148-53 (1998).
145. Edwards, J.A., Garrick, L.M. & Hoke, J.E. Defective iron uptake and globin synthesis by erythroid cells in the anemia of the Belgrade laboratory rat. *Blood* **51**, 347-57 (1978).

146. Edwards, J.A., Sullivan, A.L. & Hoke, J.E. Defective delivery of iron to the developing red cell of the Belgrade laboratory rat. *Blood* **55**, 645-8 (1980).
147. Priwitzerova, M. et al. Severe hypochromic microcytic anemia caused by a congenital defect of the iron transport pathway in erythroid cells. *Blood* **103**, 3991-2 (2004).
148. Lam-Yuk-Tseung, S., Mathieu, M. & Gros, P. Functional characterization of the E399D DMT1/NRAMP2/SLC11A2 protein produced by an exon 12 mutation in a patient with microcytic anemia and iron overload. *Blood Cells Mol Dis* **35**, 212-6 (2005).
149. Lam-Yuk-Tseung, S., Camaschella, C., Iolascon, A. & Gros, P. A novel R416C mutation in human DMT1 (SLC11A2) displays pleiotropic effects on function and causes microcytic anemia and hepatic iron overload. *Blood Cells Mol Dis* **36**, 347-54 (2006).
150. Cellier, M.F., Courville, P. & Campion, C. Nramp1 phagocyte intracellular metal withdrawal defense. *Microbes Infect* **9**, 1662-70 (2007).
151. Kast, C. & Gros, P. Epitope insertion favors a six transmembrane domain model for the carboxy-terminal portion of the multidrug resistance-associated protein. *Biochemistry* **37**, 2305-13 (1998).
152. Kast, C., Canfield, V., Levenson, R. & Gros, P. Transmembrane organization of mouse P-glycoprotein determined by epitope insertion and immunofluorescence. *J Biol Chem* **271**, 9240-8 (1996).
153. Hull, R.N., Cherry, W.R. & Weaver, G.W. The origin and characteristics of a pig kidney cell strain, LLC-PK. *In Vitro* **12**, 670-7 (1976).
154. Banerjee, A. & Swaan, P.W. Membrane topology of human ASBT (SLC10A2) determined by dual label epitope insertion scanning mutagenesis. New evidence for seven transmembrane domains. *Biochemistry* **45**, 943-53 (2006).
155. Reinhardt, J., Grishin, A.V., Oberleithner, H. & Caplan, M.J. Differential localization of human nongastric H(+)-K(+)-ATPase ATP1AL1 in polarized renal epithelial cells. *Am J Physiol Renal Physiol* **279**, F417-25 (2000).
156. Cellier, M.F., Bergevin, I., Boyer, E. & Richer, E. Polyphyletic origins of bacterial Nramp transporters. *Trends Genet* **17**, 365-70 (2001).
157. Picard, V., Epsztejn, S., Santambrogio, P., Cabantchik, Z.I. & Beaumont, C. Role of ferritin in the control of the labile iron pool in murine erythroleukemia cells. *J Biol Chem* **273**, 15382-6 (1998).
158. Li, H., Gu, J.D. & Sun, H. Structure, topology and assembly of a 32-mer peptide corresponding to the loop 3 and transmembrane domain 4 of divalent metal transporter (DMT1) in membrane-mimetic environments. *J Inorg Biochem* **102**, 1257-66 (2008).
159. Li, H., Li, F., Qian, Z.M. & Sun, H. Structure and topology of the transmembrane domain 4 of the divalent metal transporter in membrane-mimetic environments. *Eur J Biochem* **271**, 1938-51 (2004).
160. Ellinor, P.T., Zhang, J.F., Horne, W.A. & Tsien, R.W. Structural determinants of the blockade of N-type calcium channels by a peptide neurotoxin. *Nature* **372**, 272-5 (1994).
161. Nevo, Y. & Nelson, N. The NRAMP family of metal-ion transporters. *Biochim Biophys Acta* **1763**, 609-20 (2006).

162. West, A.H. et al. Two related genes encoding extremely hydrophobic proteins suppress a lethal mutation in the yeast mitochondrial processing enhancing protein. *J Biol Chem* **267**, 24625-33 (1992).
163. Techau, M.E. et al. Evolution of differences in transport function in Slc11a family members. *J Biol Chem* **282**, 35646-56 (2007).
164. Brocke, L., Bendahan, A., Grunewald, M. & Kanner, B.I. Proximity of two oppositely oriented reentrant loops in the glutamate transporter GLT-1 identified by paired cysteine mutagenesis. *J Biol Chem* **277**, 3985-92 (2002).
165. Iwamoto, T., Uehara, A., Imanaga, I. & Shigekawa, M. The Na+/Ca2+ exchanger NCX1 has oppositely oriented reentrant loop domains that contain conserved aspartic acids whose mutation alters its apparent Ca2+ affinity. *J Biol Chem* **275**, 38571-80 (2000).

12.0 APPENDIX

McGill University
Environmental Health and Safety

THIS IS TO CERTIFY THAT

Maciej Czachorowski

Department of Biochemistry

SUCCESSFULLY COMPLETED CORE TRAINING IN

Workplace Hazardous Materials Information System (W.H.M.I.S.)

ON

Wednesday, June 14, 2006

Pietro Gasparrini
Environmental Health & Safety Officer

Wayne Wood
Manager, Environmental Health & Safety

Valid Until Sunday, June 14, 2009