Differential Roles of Tryptophan Residues in the Functional Expression of Human Anion Exchanger 1

by

Yuka Okawa

A thesis submitted in conformity with the requirements for the degree of Master of Science

Department of Biochemistry
University of Toronto

© Copyright by Yuka Okawa 2012
Differential Roles of Tryptophan Residues in the Functional Expression of Human Anion Exchanger 1

Yuka Okawa
Master of Science
Graduate Department of Biochemistry
University of Toronto
2012

Abstract

Anion exchanger 1 (AE1) is a 95 kDa glycoprotein that facilitates Cl⁻/HCO₃⁻ exchange across the erythrocyte plasma membrane. Seven conserved tryptophan (Trp) residues are in the AE1 membrane domain; at the membrane interface (Trp⁶⁴⁸, Trp⁶⁶², and Trp⁷²³), in transmembrane segment (TM) 4 (Trp⁴⁹² and Trp⁴⁹⁶), and in hydrophilic loops (Trp⁸³¹, and Trp⁸⁴⁸). All 7 Trp residues were individually mutated into alanine (Ala) and phenylalanine (Phe) and transiently expressed in human embryonic kidney (HEK)-293 cells. The 7 Trp residues could be grouped into three classes according to the impact of the mutations on the functional expression of AE1: class 1, normal expression, class 2, expression decreased, and class 3, expression decreased by Ala substitution. These results indicate that Trp residues play differential roles in AE1 expression depending on their location in the protein and suggest that Trp mutants with a low expression are misfolded and retained in the ER.
Acknowledgments

Firstly, I would like to thank my supervisor, Dr. Reinhart Reithmeier for all his support and guidance through my project. I have acquired not only scientific knowledge but also his philosophy and wisdom. Thank you for providing me an opportunity to work in this laboratory. Deeply thank him for being much patient with me.

I would like to acknowledge my Supervisory Committee members, Dr. James Rini and Dr. David Williams for their expertise and advice toward polishing my project and for helping in my development of critical thinking.

I greatly thank Dr. Joseph Casey and his laboratory members from University of Alberta for contributing the Cl⁻/HCO₃⁻ transport assay as collaborators.

I would like to thank Battista Calvieri and Steven Doyle for the technical support for the confocal microscopy.

My large thanks to the present and past Reithmeier Lab members and the members in fifth floor Medical Science Building and the Department of Biochemistry for technical and moral supports. You all made the school so exciting place to come and learn. I was inspired by your hard work and responsibility toward your work. I also enjoyed gathering and experience outside of the laboratory.

I am thankful for Dr. Carolyn Kapron, Dr. Janet Yee, and Dr. Steven Rafferty in Trent University by encouraging me in my choice of study in a graduate school here. I thank for your care and support through my undergraduate study. I learned to have enjoyment in science.

I sincerely thank for Miyamae-sensei, Toyoshima-sensei, Ogawa-san, Tsuda-san, and Mimura-san for all advice and wisdom towards my project and career paths. I have been inspired by your enthusiasm for science and philosophy.

I deeply thank all friends for their support and encouragement. I thank them for being a friend with me, even though they know the worst of me. Without the friendships, I was not able to achieve my undergraduate and graduate studies and I could not go through the difficulties I had encountered. Your friendships are my fortune. I will not forget all the times, spaces, laughs, and tears, we spent together.

Finally, and the most tremendously, my gratitude is to my family for love and support, with my deepest appreciation and respect. Thank you for always being there for me.
# Table of Contents

Abstract ............................................................................................................................. ii  
Acknowledgments ............................................................................................................ iii  
Table of Contents .............................................................................................................. iv  
List of Tables ...................................................................................................................... vi  
List of Figures .................................................................................................................... vii  
List of Appendices .............................................................................................................. viii  
Abbreviations ................................................................................................................... ix  

Chapter 1 Introduction ..................................................................................................... 1  
1.1 Preamble ..................................................................................................................... 1  
1.2 Membrane proteins .................................................................................................... 3  
   1.2.1 Types of membrane proteins .............................................................................. 3  
   1.2.2 Properties of membrane proteins ...................................................................... 7  
   1.2.3 Structure and function of membrane proteins .................................................. 13  
   1.2.4 Tryptophan residues in membrane proteins ...................................................... 22  
1.3 Anion exchanger 1 (AE1) .......................................................................................... 25  
   1.3.1 Anion exchanger family ..................................................................................... 25  
   1.3.2 The structure and function ................................................................................. 26  
   1.3.3 Tertiary and quaternary structural domains ...................................................... 32  
   1.3.4 Topology ............................................................................................................ 36  
   1.3.5 Ion translocation ............................................................................................... 39  
   1.3.6 Inhibitor binding sites ....................................................................................... 40  
   1.3.7 Functionally important amino acid residues ..................................................... 41  
   1.3.8 Tryptophan residues .......................................................................................... 43  
1.4 Functional expression of AE1 .................................................................................... 45  
   1.4.1 Biosynthesis ...................................................................................................... 45  
   1.4.2 N-glycosylation ................................................................................................. 47  
   1.4.3 Folding and trafficking ...................................................................................... 49  
   1.4.4 Stability and degradation .................................................................................. 50  
   1.4.5 Expression systems .......................................................................................... 51  
1.5 Project focus and hypothesis ..................................................................................... 53  

Chapter 2 Materials and methods ............................................................................... 55  
2.1 Materials .................................................................................................................... 55  
2.2 Site-directed mutagenesis ........................................................................................ 55  
2.3 Transient transfection and expression in HEK-293 cells ........................................ 56  
2.4 Electrophoresis and immunoblotting ...................................................................... 56  
2.5 Immunofluorescence and microscopy .................................................................... 57  
2.6 Cell surface biotinylation ......................................................................................... 57  
2.7 Inhibitor (SITS) binding ........................................................................................... 58  
2.8 Transport assay ........................................................................................................ 59  
2.9 Purification ................................................................................................................ 59  
2.10 N-glycosylation analysis ......................................................................................... 60  

Chapter 3 Results ........................................................................................................... 61
3.1. Expression and stability of Trp mutants ................................................................. 61
3.2. Cell surface expression of Trp mutants ................................................................. 64
3.3. Inhibitor binding of Trp mutants ........................................................................... 70
3.4. Transport activity of Trp mutants ........................................................................ 73
3.5. Expression and trafficking of AE1, kAE1, and mdAE1 ........................................ 75
3.6. Oligosaccharide processing of Trp mutants of mdAE1 ....................................... 79
3.7. Cell surface expression of mdAE1 ....................................................................... 82
Chapter 4 Discussion..................................................................................................... 86
Chapter 5 Future Directions ...................................................................................... 95
References ..................................................................................................................... 99
Appendix 1 .................................................................................................................. 114
List of Tables

Table 1. The IUBMB classification of transporters................................. 6
Table 2. Sequence identity of human AE1 homologues................................. 27
Table 3. Location of Trp residues in the membrane domain of human AE1....... 27
Table 4. Membrane domain homology of AE1 in paralogues and orthologues...... 27
Table 5. Percent complex oligosaccharide in AE1 variants.............................. 77
Table 6. Summary of the present study...................................................... 88
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1</td>
<td>Folding motifs of bitopic and polytopic proteins</td>
<td>4</td>
</tr>
<tr>
<td>Figure 2</td>
<td>Three-dimensional structure of an aquaporin 1 (AQP1)</td>
<td>14</td>
</tr>
<tr>
<td>Figure 3</td>
<td>Three-dimensional structure of KcsA K⁺ channel</td>
<td>16</td>
</tr>
<tr>
<td>Figure 4</td>
<td>A schematic GlpT structure</td>
<td>19</td>
</tr>
<tr>
<td>Figure 5</td>
<td>Three-dimensional structure of LacY</td>
<td>20</td>
</tr>
<tr>
<td>Figure 6</td>
<td>Structure of the prokaryotic chloride channel (CIC)</td>
<td>21</td>
</tr>
<tr>
<td>Figure 7</td>
<td>Domain organization of erythrocyte AE1, kidney (k)AE1, and membrane domain (md)AE1</td>
<td>28</td>
</tr>
<tr>
<td>Figure 8</td>
<td>Topology of human anion exchanger 1 (AE1)</td>
<td>30</td>
</tr>
<tr>
<td>Figure 9</td>
<td>Anion exchanger 1 (AE1) in the erythrocyte and kidney</td>
<td>31</td>
</tr>
<tr>
<td>Figure 10</td>
<td>2.6 Å crystal structure of the cytosolic domain of the human AE1 (cdAE1)</td>
<td>33</td>
</tr>
<tr>
<td>Figure 11</td>
<td>Three dimensional map of the membrane domain of the human AE1 (mdAE1)</td>
<td>34</td>
</tr>
<tr>
<td>Figure 12</td>
<td>Comparison of the AE1 membrane domain with the CIC chloride channel</td>
<td>35</td>
</tr>
<tr>
<td>Figure 13</td>
<td>Predicted folding model of the C-terminal membrane domain of human AE1</td>
<td>38</td>
</tr>
<tr>
<td>Figure 14</td>
<td>N-glycosylation processing pathway in mammalian cells</td>
<td>48</td>
</tr>
<tr>
<td>Figure 15A</td>
<td>Immunoblot analysis of C-terminal His₆-tagged AE1</td>
<td>62</td>
</tr>
<tr>
<td>Figure 15B</td>
<td>Immunoblot analysis of C-terminal His₆-tagged Trp mutants</td>
<td>62</td>
</tr>
<tr>
<td>Figure 15C</td>
<td>Protein expression of C-terminal His₆-tagged AE1 and Trp mutants</td>
<td>62</td>
</tr>
<tr>
<td>Figure 16A</td>
<td>Immuno-localization of C-terminal His₆-tagged AE1</td>
<td>66</td>
</tr>
<tr>
<td>Figure 16B</td>
<td>Immuno-localization of C-terminal His₆-tagged Trp mutants</td>
<td>66</td>
</tr>
<tr>
<td>Figure 17A</td>
<td>Cell surface biotinylation of C-terminal His₆-tagged AE1</td>
<td>68</td>
</tr>
<tr>
<td>Figure 17B</td>
<td>Cell surface biotinylation of C-terminal His₆-tagged Trp mutants</td>
<td>68</td>
</tr>
<tr>
<td>Figure 17C</td>
<td>Percent cell surface biotinylation of C-terminal His₆-tagged AE1 and Trp mutants</td>
<td>68</td>
</tr>
<tr>
<td>Figure 18A</td>
<td>SITS-Affi-Gel binding of C-terminal His₆-tagged AE1</td>
<td>71</td>
</tr>
<tr>
<td>Figure 18B</td>
<td>SITS-Affi-Gel binding of C-terminal His₆-tagged Trp mutants</td>
<td>71</td>
</tr>
<tr>
<td>Figure 18C</td>
<td>Percent SITS-Affi-Gel binding of C-terminal His₆-tagged AE1 and Trp mutants</td>
<td>71</td>
</tr>
<tr>
<td>Figure 19</td>
<td>Anion exchange rates of C-terminal His₆-tagged AE1, mdAE1, and Trp mutants of eAE1 expressed in HEK-293 cells</td>
<td>74</td>
</tr>
<tr>
<td>Figure 20</td>
<td>Immuno-blots and N-glycosylation analysis of AE1 variants</td>
<td>77</td>
</tr>
<tr>
<td>Figure 21</td>
<td>Immuno-localization of AE1 variants</td>
<td>78</td>
</tr>
<tr>
<td>Figure 22A</td>
<td>Immunoblot analysis of C-terminal His₆-tagged mdAE1</td>
<td>80</td>
</tr>
<tr>
<td>Figure 22B</td>
<td>Immunoblot analysis of C-terminal His₆-tagged mdAE1 Trp mutants</td>
<td>80</td>
</tr>
<tr>
<td>Figure 22C</td>
<td>Protein expression of C-terminal His₆-tagged mdAE1 and mdAE1 Trp mutants</td>
<td>80</td>
</tr>
<tr>
<td>Figure 23A</td>
<td>Immuno-localization of C-terminal His₆-tagged mdAE1</td>
<td>83</td>
</tr>
<tr>
<td>Figure 23B</td>
<td>Immuno-localization of C-terminal His₆-tagged mdAE1 Trp mutants</td>
<td>83</td>
</tr>
</tbody>
</table>
List of Appendices

Table of Trp mutant primers........................................................................................................................................ 114
Abbreviations

AE, anion exchanger
AQP, aquaporin
ABC, ATP-binding cassette
ATP, adenosine-5'-triphosphate
BADS, 4-benzamido-4’aminostilbene-2-2’-disulfonate
βAR, β-adrenergic receptor
BBF, β-barrel finder
BIDS, 4-benzamido-4’isothiocyanostilbene-2-2’-disulfonate
BR, bacteriorhodopsin
BTR, bicarbonate-transport related protein
CA, carbonic anhydrase
cdAE1, cytosolic domain of AE1
CNX, calnexin
COX, cytochrome c oxidase
CRT, calreticulin
DBDS, 4,4’-benzamidostilbene-2-2’-disulfonate
DIDS, 4,4’-isothiocyanostilbene-2,2’-disulfonate
DNFB, 2,4-dinitrofluorobenzene
DNDS, 4,4’-dinitrostilbene-2-2’-disulfonate
dRTA, distal renal tubular acidosis
cAE1, erythrocyte AE1
EC, extracellular loop
EndoH, endoglycosidase H
EPR, electron paramagnetic resonance
EMA, eosin-5-maleimide
ER, endoplasmic reticulum
GAPDH, glyceraldehyde-3-phosphate dehydrogenase
GI, Glucosidase I
GII, Glucosidase II
GES, Goldman-Engelman-Steitz
GalNAc, N-acetylgalactosamine
GlcNAc, N-acetylglucosamine
GPA, glycoporphin A
GPCR, G-protein-coupled receptors
GT, glucosyltransferase
GTP, guanosine triphosphate
HA, haemagglutinin
HEK, embryonic kidney cells
HLB, hydrophilic lipophilic balance
HS, hereditary spherocytosis
IUBMB, International Union for Biochemistry and Molecular Biology
kAE1, kidney AE1
KcsA, K⁺ channel from Streptomyces lividans
KvAP, voltage-dependent K⁺ channel
mdAE1, membrane domain AE1
MDCK, Madin-Darby Canine Kidney
M₉, molecular weight
MthK, calcium-activated K⁺ channel
NBC, sodium-bicarbonate co-transporter
NDCBE, sodium-driven chloride/bicarbonate exchanger
OST, oligosaccharyl transferase
PBS, phosphate buffered saline
PDI, protein disulfide isomerase
PEP, phosphoenolpyruvate
Pmf, proton motive force
PNGase F, N-glycanase F
RC, photosynthetic reaction center
SAO, Southeast Asian ovalocytosis
SD, standard deviation
SDS, sodium dodecyl sulfate
SDS-PAGE, SDS polyacrylamide gel electrophoresis
SERCA, sarcoplasmic reticulum
Ca²⁺ATPase
SITS, 4-acetamido-4’isothiocyanostilbene-2-2’-disulfonate
SLC4A1, solute carrier family 4, anion exchanger, member 1
SLC4A2, solute carrier family 4, anion exchanger, member 2
SLC4A3, solute carrier family 4, anion exchanger, member 3
Smf, sodium motive force
SRP, signal-recognition particle
TM, transmembrane
Trp, tryptophan
Wt, wild type
WW, Winley-White
Chapter 1
Introduction

1 Introduction

1.1 Preamble

Membrane proteins are specialized macromolecules adapted to associate with biological membranes. They play a diverse set of functions in cells and are important in health and disease. Functionally, membrane proteins are crucial in transporting specific molecules across membranes, catalyzing membrane-associated reactions, providing structural linkage to cytoskeletal proteins, in mediating extracellular interactions between adjacent cells, and in detecting and transducing chemical signals (Alberts, 2008). Approximately 30% of genomes encode membrane proteins in both eukaryotes and prokaryotes (Stevens and Arkin, 2000; Fagerberg et al., 2010), highlighting the importance of membrane proteins in living systems. Mutations in the genes encoding membrane proteins are linked to many different diseases and membrane proteins are therapeutic targets for almost 60% of the pharmaceuticals on the market today (Fagerberg et al., 2010). Despite the importance of membrane proteins, relatively few membrane proteins are represented by the 65,000 structures in the RCSB Protein Data Bank (website, http://www.pdb.org/pdb/home/home.do) (Baker, 2011). Currently available high-resolution structures of membrane proteins are on the website of the Stephen White Laboratory (http://blanco.biomol.uci.edu/Membrane_proteins_xtal.html).

The red blood cell anion exchanger 1 (AE1), commonly known as Band 3, is the founding member of the solute carrier family 4 of anion exchangers (SLC4A1) and is the main subject of this thesis. AE1 is the predominant integral membrane glycoprotein of the red blood cell, occupying 50% of total erythroid membrane protein and is responsible for the electroneutral exchange of bicarbonate and chloride across the plasma membrane. AE1 exists in two forms, the full-length protein expressed in erythrocytes (AE1) and a truncated form expressed in the kidney (kAE1). Erythrocyte AE1 is essential in the respiratory system by enhancing the carbon dioxide-carrying capacity of the blood. In the tissues, carbon dioxide diffuses into the red cell and is converted to bicarbonate by the
action of cytosolic carbonic anhydrase II (CAII). The bicarbonate is then transported out of the cell in exchange for chloride by AE1, thus increasing the carbon dioxide carrying capacity of the blood as soluble plasma bicarbonate. In the lungs, bicarbonate is transported into the red cell in exchange for chloride via AE1, where it is dehydrated by CAII to form carbon dioxide, which diffuses out of the red cell and is expired by the lungs. Kidney AE1 is expressed in α-intercalated cells where it mediates bicarbonate reabsorption into the blood and facilitates acid secretion into the urine; the bicarbonate and protons originating from the action of intracellular CAII. AE1 is the best-characterized member of the SLC4 family of anion transporters. It consists of two major domains; an N-terminal cytosolic domain that mediates interactions with the cytoskeleton and other proteins and a C-terminal membrane domain responsible for its anion transport function. While a high-resolution structure of the cytosolic domain is now available, a high-resolution structure of the membrane domain is not, nor has the molecular basis of its mechanism of action been completely established.

The topology and location of important amino acids in the AE1 membrane domain have been reported in numerous studies. However, the role of the seven tryptophan (Trp) residues in the functional expression of AE1, the subject of this thesis, has not been studied in a systematic fashion. Human AE1 has seven highly conserved endogenous Trp residues (492, 496, 648, 662, 723, 831, and 848) in its membrane domain. As commonly found at the boundary of transmembrane (TM) segments, these Trp residues may play an important role in the positioning of TM segments during biosynthesis. Trp residues in AE1 may also function in mediating TM helix-helix interactions, in forming the inhibitor binding site found in AE1, and in the ion translocation pathway of AE1. This thesis is focused on the roles of Trp residues in the membrane domain on the functional expression of AE1 in transfected HEK-293 cells. This was achieved by mutating each individual Trp residue to Phe, a bulky aromatic residue, or Ala, a small residue. We were able to distinguish three classes of Trp residues based on the effect of the functional expression of AE1: class 1, Trp residues that can be mutated to Phe or Ala; class 2, Trp residues that cannot be mutated to Phe nor Ala; and class 3, Trp residues that be substituted by Phe but not Ala. Thus, Trp residues
play differential roles in the functional expression of AE1 depending upon their particular location in the protein.

1.2 Membrane proteins

1.2.1 Types of membrane proteins

Membrane proteins can be categorized based on the nature of their association with biological membranes. The Fluid Mosaic Model (Singer, 1972) divides membrane proteins into extrinsic (peripheral) and intrinsic (integral) types. Peripheral membrane proteins are weakly associated with membrane lipids or integral membrane proteins without extending into the hydrophobic core of the lipid bilayer. Peripheral membrane proteins may be present at either face of the membrane bilayer via noncovalent or electrostatic interactions with lipids or proteins. Thus peripheral membrane proteins can be dissociated from the lipid bilayer by elevating the ionic strength or by changing the pH of the medium. Lipid-anchored extrinsic membrane proteins are the exception in that they are covalently modified with lipids, that allow them to associated tightly with the lipid bilayer requiring detergents for their extraction (Luckey, 2008). The lipid modifications include an amide linkage of N-terminal glycine with a myristoyl group and covalent bonding of palmitoyl or other lipid groups with serine or cysteine side chains via ester or thioester bonds respectively.

Integral membrane proteins, frequently simply referred as “membrane proteins”, are embedded in the lipid bilayer by hydrophobic interactions, and detergent is required to extract them from the membrane and solubilize them (Singer, 1972). They can be further classified with respect to their folding motifs (Reithmeier, 1996a; Luckey, 2008) (see Figure 1). Monotopic proteins are inserted into the lipid bilayer without spanning it, such as prostaglandin H2 synthase, the target for Aspirin™. Bitopic proteins span the bilayer once. The orientation of their single TM segments may have its N-terminal facing outside (type I) or facing inside (type II). Polytopic proteins span the membrane more than once and their hydrophobic TM segments are connected by hydrophilic turns, loops or larger domains. Their orientations can be both C- and N-termini extracellular (type III), or intracellular (type IV), and C- and N-termini facing different sides (type V: N-terminal in C-terminal out, type VI: N-terminal out C-terminal in). All types exist in
Figure 1. Folding motifs of bitopic and polytopic proteins. Bitopic proteins span the bilayer once with the orientation of their single TM segments having its N-terminal facing outside (type I) or facing inside (type II). Polytopic proteins span the membrane more than once and their hydrophobic TM segments are connected by hydrophilic turns, loops or larger domains. Their orientations can be both C- and N-termini facing - extracellular (type III), intracellular (type IV), and C- and N-termini facing different sides (type V: N-terminal in C-terminal out, type VI: N-terminal out C-terminal in).
nature with 7-span G-protein coupled receptors (Type VI) and 6- or 12-span transporters (Type IV) making up major classes (Liu et al., 2002).

Membrane transport proteins are also classified according to a functional/phylogenetic system by the International Union for Biochemistry and Molecular Biology (IUBMB), in which 400 families of transport proteins are grouped into five classes (Table 1)(Busch and Saier, 2002). The IUBMB transporter classification system (TC) (Busch and Saier, 2002) is summarized as follows: Class 1 includes channels and pores in which the substrate passes through the membrane by passive or facilitated diffusion, and is subdivided into α–helical protein channels, β-barrel protein porins, toxin channels, nonribosomally synthesized channels, and holins. Porins are used as survival strategies by some pathogenic bacteria in which pore-forming TMs enable rapid diffusion of their solutes (Achouak et al., 2001) and are found in the outer membranes of Gram-negative bacteria, mitochondria, and chloroplasts (Schirmer, 1998). Channel-forming toxins made from plants and bacteria include the family of AB toxins and colicins can enter the cell by penetrating the plasma membrane or exploiting translocation systems present in the cell, leading to toxic effects on cell physiology (Falnes and Sandvig, 2000). Nonribosomally synthesized channels are formed from secondary metabolites including ionophores such as alamethicin and gramicidin, which create pores in the membrane and cause ions to leak from the cell. Holins are small bacteriophage-encoded proteins that cause lysis of the host cell by forming a lesion that permeabilizes the membrane (Grundling et al., 2001).

Class 2 is composed of electrochemical potential-driven transporter proteins, also called secondary carrier-type facilitators, that use energy from the proton motive force (pmf) or the sodium motive force (smf) to drive transport processes. Protein porters can transport a single species (uniport), transport two or more species in the same direction (symport), or transport two or more species in opposite directions (antiport). Nonribosomally synthesized porters are mostly produced by bacterial and fungi, including the ionophores such as valinomycin and nigericin. Class 2 proteins are also involved in energy driven bacterial flagellar rotation and active transfer in the outer membranes of Gram-negative bacteria, such as the TonB family.
Table 1. The IUBMB classification of transporters (Busch and Saier, 2002)

<table>
<thead>
<tr>
<th>Transporter Classification (TC) System</th>
<th>Subclass</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Channels/Pores</td>
<td>α-Helical protein channels</td>
</tr>
<tr>
<td></td>
<td>β-Barrel protein porins</td>
</tr>
<tr>
<td></td>
<td>Toxin channels</td>
</tr>
<tr>
<td></td>
<td>Nonribosomally synthesized channels</td>
</tr>
<tr>
<td></td>
<td>Holins</td>
</tr>
<tr>
<td>2. Electrochemical Potential-Driven Transporters</td>
<td>Porters</td>
</tr>
<tr>
<td></td>
<td>Nonribosomally synthesized porters</td>
</tr>
<tr>
<td></td>
<td>Ion-gradient-driven energizers</td>
</tr>
<tr>
<td>3. Primary Active Transporters</td>
<td>P-P-bond hydrolysis-driven systems</td>
</tr>
<tr>
<td></td>
<td>Decarboxylation-driven systems</td>
</tr>
<tr>
<td></td>
<td>Methyltransfer-driven systems</td>
</tr>
<tr>
<td></td>
<td>Oxidoreduction-driven systems</td>
</tr>
<tr>
<td></td>
<td>Light-absorption-driven systems</td>
</tr>
<tr>
<td>4. Group Translocators</td>
<td>Phosphotransfer-driven systems</td>
</tr>
<tr>
<td>5. Transmembrane Electron Carriers</td>
<td>Two-electron transfer carriers</td>
</tr>
<tr>
<td></td>
<td>One-electron transfer carriers</td>
</tr>
</tbody>
</table>
Class 3 is composed of primary active transporters including phosphate-bond hydrolysis driven transporters (including ATPases and the ABC superfamily), decarboxylation-driven transporters, methyltransfer-driven transporters, oxidoreduction-driven transporters, and light absorption-driven transporters (such as the microbial rhodopsin family and the photosynthetic reaction center family). ATPases are further categorized into P-type, (such as Ca$^{2+}$ pump and the Na$^+\text{-K}^+$ ATPase), F-type (such as the ATPase synthases found in the mitochondria and bacteria), V-type (transporting proton across organelles such as vacuoles, lysosomes, and endosomes), and A-type (mainly found in Archaea and transporting anions). The ABC transporter superfamily, which is named for their nucleotide-binding domain (ATP-binding cassettes), is found in various species from *Escherichia. coli* (*E. coli*) to human. ABC transporters operate either as uptake or efflux pumps and may have either specific or broad substrate preference.

Class 4 is a group of translocators using metabolic energy to translocate a substrate, which is chemically modified during transport, into the cytoplasm. The bacterial phosphoenolpyruvate (PEP): sugar phosphotransferase system is the best known example of this group. The transport process is tightly coupled with the catalytic reaction of extracellular PEP into sugar-phosphate, which is phosphorylated during the transport process.

Class 5 is composed of transmembrane electron carriers, including two-electron transfer carriers (such as the disulfide bond oxidoreductase, DsbD and DsbB, families) and one-electron transfer carriers (the phagocyte NADPH oxidase family).

### 1.2.2 Properties of membrane proteins

Intrinsic membrane proteins are composed mainly of two hydrogen-bonded secondary structures, $\alpha$-helical bundles and $\beta$-barrels (Basyn *et al.*, 2003). The formation of main-chain hydrogen bonds is a strong driving force in the hydrophobic environment of the lipid bilayer (White and Wimley, 1999). Due to the scarcity of water molecules in the hydrocarbon core of the membrane, a fully hydrogen-bonded main chain can help compensate for the energy cost of partitioning from the high dielectric constant ($\epsilon$) of water phase ($\epsilon = 78$) to the low dielectric constant apolar phase in the bilayer ($\epsilon = \sim 2$) (Luckey, 2008).
Transmembrane $\alpha$-helices are the dominant secondary structure of the membrane-embedded portion of membrane proteins. They are typically composed of approximately 20 amino acids sufficient to span a 30Å thick hydrophobic phase when arranged as an $\alpha$-helix, showing intrahelical hydrogen bonds between the $i$ and $i + 4$ residues with a rise of 1.5Å per residue (Ubarretxena-Belandia and Engelman, 2001). Apolar amino acids, such as Leu, Ile, Val, Phe are abundant in TM helices (~ 57%) with their side chains facing the hydrophobic interior of the bilayer (Reithmeier and Deber, 1992). Small residues, Ala, Gly, Ser, and Thr, are also frequently found in TM segments where they mediate helix-helix interaction (Javadpour et al., 1999). Charged residues are not abundant in TM segments, with the frequency less than 5%, compared with 22% in water-soluble regions (Ubarretxena-Belandia and Engelman, 2001). Due to the low dielectric environment of the lipid bilayer, acidic and basic amino acids are predominantly uncharged or form ion pairs. However, in some case they remain charged, facing water-filled channels, playing special roles, such as transporting a substrate, or binding a metal or cofactor (Luckey, 2008). Charged residues like Lys can also position their side chains at the interface of TM segments by “snorkeling” with the terminal charged group exposed to water (Killian and von Heijne, 2000).

Interhelical hydrogen bonds are also found in TM segments. Asp, for example, in the center of TM segments can lead to the formation and stabilization of TM helix homooligomerization (Zhou et al., 2000; Ubarretxena-Belandia and Engelman, 2001). Gly and Ala in helical segments can minimize interhelical distances allowing the close packing of helices (Eilers et al., 2000). Pro, a rare amino acid in $\alpha$-helices of globular proteins, when present in the center of TM helices causes a kink in the helices and alters their directions (Reithmeier and Deber, 1992). Ser, Thr, and Asn can be hydrogen-bonded to the backbone carbonyls or amide nitrogen providing caps at the N-terminus or C-terminus end of helices. Gly, also commonly found as a C-cap residue, is a predominant helix termination at the C-terminus. Aromatic amino acids, Trp, Tyr, and Phe, are found mainly at the interface of the hydrophilic phase and hydrophobic core of the bilayer in both $\alpha$-helices and $\beta$-barrels (von Heijne, 1999), suggesting that they are important in positioning membrane proteins in the bilayer.
The α-helical TM segments of integral membrane proteins can be predicted from the amino acid sequences by recognizing stretches of ~20 hydrophobic residues favorably partitioning into the bilayer (Reithmeier and Deber, 1992). Twenty amino acids arranged in an α-helix are able to span the 30 Å hydrophobic phase of the lipid bilayer. Hydropathy plots are graphical displays of hydrophobicity as a function of sequence and allow the identification of potential TM segments and thereby provide a two-dimensional topology model. The Kyte-Doolittle scale (Jayasinghe et al., 2001) is a frequently used method (Reithmeier and Deber, 1992). Hydrophobicity scales of amino acids are determined by free energy changes occurring by transferring each amino acid from a polar to an apolar phase. The other frequently used hydrophobicity scales are the Goldman-Engelman-Steitz (GES) (Eisenberg et al., 1984) and the Wimley-White (WW) (Wimley and White, 1996) scales. The GES, similar to the von Heijne and Blomberg scale, calculates free energy transfer from polar to apolar phases of amino acids in α-helical peptides, the hydrophobic surface area, and charge neutralization based on pKa in charged groups (Reithmeier and Deber, 1992). The WW scale is an empirical method based on the solubility of small random-coil peptides in water and octanol phases (Luckey, 2008). Statistical analysis, such as TMHMM and HMMTOP using hidden Markov models can predict TM helical region with accuracy up to 85% (Rost et al., 1996). MEMSAT and TOPPRED use dynamic programming and show less accuracy (Cuthbertson et al., 2005). Combining these four methods and by building a neural network using PHD (Rost et al., 1996) can provide a high successful prediction rate with 95% accuracy (Cuthbertson et al., 2005).

The orientation of TM segments that make up an integral membrane protein in the bilayer is another component to understanding its properties. The topology of membrane proteins has been extensively studied and is commonly used to create folding models (Manoil et al., 1988; Calamia and Manoil, 1990; von Heijne, 1999). Membrane protein orientation can be predicted according to the sequences in the loop regions. By statistically analyzing the amino acid residues, the cytoplasmic loops of TM segments possess 2 to 4 times higher proportion of positively charged amino acids, such as Lys and Arg, than the external loops (von Heijne, 1992). This “positive-inside rule” can be applied to most integral membrane proteins in the three kingdoms of life (von Heijne,
By combining hydrophobicity plots with the positive-inside rule, the prediction of integral membrane protein topology can be achieved with highly reliability (von Heijne, 1999). These predictions are useful in building membrane protein folding models, yet biochemical or structural studies are required to confirm the prediction.

Biochemical methods, such as fusion with the coding regions of reporter enzymes, like alkaline phosphatase, β-lactamase, and β-galactosidase, can aid in determining the orientation of integral membrane proteins recombinantly expressed in E. coli (Manoil et al., 1988; Calamia and Manoil, 1990; Manoil, 1990; Derman and Beckwith, 1991). Inserting an N-glycosylation site is a robust method utilizing the eukaryotic cell translational system and can identify the extra-cytosolic regions of membrane glycoproteins exposed to the ER lumen during biosynthesis (Popov et al., 1999). Insertion of an epitope tag such as hemagglutinin (HA), c-myc, or FLAG, in a predicted extracellular site followed by detection on the surface of transfected cells via antibodies is another technique to identify surface exposed loops (LaVallie and McCoy, 1995; Kaltwasser et al., 2002). Proteolytic analysis can also be carried out for endogenous or introduced protease cleavage sites to investigate membrane protein topology (Hamaskaki et al., 1997). Non-penetrating chemical reagents reactive with amino acid side chains can be used to determine membrane protein topology (Jennings and Nicknish, 1984; Fujinaga et al., 1999). Scanning cysteine mutagenesis, whereby cysteine residues are introduced at various sites throughout a constructed Cys-less membrane protein, is a powerful method for examining membrane protein topology since residues within hydrophobic regions are unreactive, while external sites can be labeled in intact cells expressing the membrane protein (Fujinaga et al., 1999).

β-barrel proteins make up only 2-3% of the proteomes of Gram-negative bacteria and are also found in outer membranes of mitochondria and chloroplasts (Wimley, 2003). TM segments of β-barrels are more extended and thinner than α-helices with ~ 0.5 nm diameters, composed of anti-parallel β-sheets and short turns (Schulz, 2000; Luckey, 2008). β-strands spanning the membrane are 8-9 residues long if perpendicular to the plane of bilayer. Yet, many β-sheets are twisted in the bilayer and typical TM β-strands possess 9 to 11 residues and are likely to be tiled from 20-45° to the plane of the membrane (Wimley, 2003). The first and the last strands are hydrogen bonded to each
other forming a closed barrel to satisfy the main-chain hydrogen bonds of the carbonyl and amino groups. The interstrand hydrogen bonds allow β-barrels to form rigid and stable structures within the bilayer (Wimley, 2003). Non-polar amino acid residues are found on the exterior surface of the β-barrels facing the hydrophobic region of the membrane while the interior is lined with hydrophilic residues forming selective channels or pores (Wimley, 2002, 2003). Aromatic residues are found at the interface region of β-barrels.

In contrast to α-helices, β-barrels are harder to predict from the sequence and physical principles (Wimley, 2002; Waldispohl et al., 2006). The program, β-barrel finder (BBF), can effectively identify an integral membrane β-barrel and screen whole genomes for proteins predicted to be present in the outer membrane (Zhai and Saier, 2002). The program predicts β-barrel proteins based on the secondary structure, hydrophobicity, and amphipathicity after screening for an N-terminal signal sequence of outer membrane protein precursors. The program examines hydrophobicity and amphipathicity peaks in possible TM β-strands based on hydrophobicity and the alternation of polar and apolar residues in the TM region. The final step searches for homologous sequences of known β-barrel proteins via the Basic Local Alignment Search Tool (BLAST) (Wimley, 2003).

Interactions with annular lipids are important in stabilizing integral membrane proteins (Lee, 2003). Indeed, the crystal structures of many membrane proteins have revealed the presence of tightly bound lipids or detergent molecules (Lee, 2003; Marsh and Pali, 2004). The annular lipids surrounding a membrane protein are equivalent to the solvent around a soluble protein (Lee, 2003). Annular lipids can be separated from the other bulk lipids in the bilayer, and their interactions with integral membrane proteins have been extensively studied using electron paramagnetic resonance (EPR), fluorescence quenching, energy transfer, and molecular dynamic simulations (Marsh and Horvath, 1998; Marsh and Pali, 2004). The heterogeneous and rough surface of TM domains with shallow grooves and protrusions, leads surrounding lipids to distort their fatty acyl chains and results in a decrease in trans-gauche isomerisation to enable tight packing of the proteins and maintain the permeability barrier of the bilayer (Lee, 2003). Many membrane proteins prefer lipids in liquid crystalline phase to gel phase with
respect to free volume and effective van der Waals contacts, but they show relatively little selectivity in binding annular lipids (Lee, 2003). Rhodopsin, for example, has no selectivity, while the majority of intrinsic membrane proteins show a small selectivity for anionic phospholipids (Lee, 2003). The stoichiometry of annular lipids is correlated with the shape of proteins (Lee, 2003). With ~0.5 nm lipid diameter, 10 lipids surround ~1 nm TM helix in a bitopic protein, and the number of annular lipids depends proportionally on the number of TM helices and the geometric arrangement in polytopic proteins (Lee, 2004). Composed of a much more extended structure than an α-helix, β-barrels have the same number of annular lipids with a β-strand and twice more with 60° tilt (Marsh and Horvath, 1998). The thickness of the hydrophobic regions of TM segments needs to match the thickness of the hydrophobic regions of the lipid bilayer (Lee, 2003, 2004). The annular lipids can be thinner or thicker than the bilayer to match the hydrophobic portions of the membrane proteins but membrane proteins and the surrounding lipids can also result in a hydrophobic mismatch (Lee, 2003, 2004).

Detergents that disrupt biological membranes can be used to replace the bulk lipids surrounding the membrane proteins and solubilize hydrophobic membrane proteins in aqueous solutions. Ideal detergents for membrane proteins keep the protein structure intact often with their native quaternary structures (Gennis, 1988). An optimal detergent is challenging to find as some may irreversibly denature proteins while others may be too weak to effectively solubilize the membrane protein. A systematic parameter, the hydrophilic lipophilic balance (HLB) is frequently used for selecting a proper detergent (Gennis, 1988). HLB values in the range from 12.5 to 14.5 effectively solubilize membrane proteins and can be chosen as an initial factor (Gennis, 1988). Triton X-100, a commonly used detergent, has an HLB of 13.4 (Egan, 1976). Other than maximizing solubilization of a desired protein in a desired form with its function intact, screening a proper detergent also requires considering its compatibility with the preparative protocol (Gennis, 1988). While functional assays of receptors can utilize ligand binding, transport proteins must be reconstituted into sealed lipid vesicles in order to measure their activity. Detergent compatibility includes length of the hydrophobic chain, size of the head group, charge, critical micelle concentration, micelle size, and the optical absorbance of the detergent (Gennis, 1988).
1.2.3 Structure and function of membrane proteins

Despite occupying 30% of the human genome, high-resolution structures of membrane proteins are challenging to acquire, and much fewer structures (263 compared to over 70000 of soluble proteins: from http://blanco.biomol.uci.edu/Membrane_Proteins_xtal.html) have been reported to 2010 (Baker, 2011). A few membrane proteins are naturally abundant, and they have been purified from natural sources, such as mammalian and bacterial rhodopsins, aquaporins, respiratory complexes, ATPases, photosynthetic complexes, reaction centers, and light harvesting proteins (Bill et al., 2011). More generally, recombinant production by engineering vectors to over-express eukaryotic membrane proteins in various cell systems has been attempted, but has had a low success rate. Purifying these proteins in stable forms is another bottleneck especially for proteins from higher eukaryotes. Instability often results from the loss of essential lipids during the purification procedure resulting in protein unfolding and aggregation. As a result, the majority of membrane protein structures are resolved from bacteria or Archaea sources with an emphasis on mammalian homologs (Bill et al., 2011). High-resolution structural information is, however, often necessary to comprehend membrane protein function, including their mechanism of action and regulation.

Structures of membrane proteins have been reported in an accelerated manner since the late 90’s. Peter Agre and Roderick MacKinnon were awarded the Nobel Prize in 2003 for solving the X-ray structures of aquaporin (AQP) and the potassium channel and unveiling their mechanisms of action (Luckey, 2008). AQPs are abundant channels and selectively facilitate the permeation of water and small amphipathic solutes such as glycerol across the cell membrane (Stroud et al., 2003). They conduct water at a high rate, close to the diffusion limit, but are remarkably selective excluding hydroxide and hydronium ions (Alberts, 2008). The water channel contains a narrow pore, which is too small for any hydrated ions to pass through (Figure 2) (Sui et al., 2001; Stroud et al., 2003). The channel is also impermeable to protons by placing two Asn residues in the center of the pore (Sui et al., 2001), tethering water molecules in the file and hindering a line of water molecules from conducting protons via the Grotthuss mechanism (Stroud et al., 2003).
Figure 2. Three-dimensional structure of an aquaporin 1(AQP1) monomer in a side view (PDB code: 1J4N) (Sui et al., 2001). Backbone is shown in ribbon diagram in blue (yellow spheres: water molecules in the pore, red sticks: Asn$^{78}$ and Asn$^{194}$). The monomer contains an independent pore with a length about 20 Å, displayed in blue mesh. The narrowest pore has a diameter ~2.8 Å. Both C- and N-termini face the cytoplasmic side (Type IV). The figure was generated using PyMOL.
The first high-resolution K+ channel structure resolved was the KcsA channel that uncovered the paradox of their ability to combine high selectivity and K+ conductance 10,000-fold higher than Na+ (Doyle et al., 1998). The K+ channel is a tetramer with 4-fold symmetry with a central pore (Figure 3). Each subunit contains two TM α-helices forming a cone shape with a water-filled pore facing the cytoplasm (Doyle et al., 1998). The two TM helices connect to a short α-helix (the pore helix) and a loop creating a selectivity filter tilting carbonyl oxygen atoms of the polypeptide backbone into the pore. The cytosolic entrance of the pore is rich in negatively charged residues, concentrating cations, including Na+ and K+ in the pore. The selectivity filter of the K+ channel shows K+ ions remain mostly hydrated throughout the internal pore and cavity but must be dehydrated when entering the 12 Å narrow filter (Doyle et al., 1998). The dehydrated K+ ions fit in precise coordination with the carbonyl oxygen atoms from the backbone of Val76, Gly77, Tyr78, and Gly79, and hydroxyl oxygen of Thr75, compensating for the energetic cost of dehydration, while the smaller radius of Na+ ions cannot fully interact with these carbonyl oxygen atoms, thereby favouring remaining in aqueous solution (Doyle et al., 1998). The authors also propose that the rapid conductance of the channels is achieved by close proximity of two K+ ions repelling each other in the selectivity filter.

Recent progress in high-resolution membrane structures include G-protein-coupled receptors (GPCR), P-type ATPases, secondary active transporters, ABC transporters, and ion channels (Bill et al., 2011). The structures of the best characterized two groups from the GPCR family, rhodopsin (Murakami and Kouyama, 2008; Park et al., 2008; Scheerer et al., 2008) and β-adrenergic receptor (βAR) (Cherezov et al., 2007; Rasmussen et al., 2007; Rosenbaum et al., 2007; Warne et al., 2008), were solved in the last few years. Rhodopsins, sufficiently purified from bovine or squid tissues, were extremely stable in detergent and adopted a single inactive state until activated by addition of retinal (Murakami and Kouyama, 2008; Park et al., 2008; Scheerer et al., 2008). βAR activation is, however, more complex than the bimodal switch between inactive and active states observed in rhodopsin which possesses little basal (agonist-independent) activity and reaches maximal activation from single photon absorption (Kobilka and Deupi, 2007). The breakthroughs enabling the crystallization of βAR
Figure 3. The three dimensional KcsA $\text{K}^+$ channel shown as a tetramer viewed from the extracellular side (A) and as a dimer (B) viewed from the side (PDB code: 1K4C) (Zhou et al., 2001). The ion conduction pore is $\sim$10 Å in the diameter in the open state. The selectivity filter residues, Thr$^{75}$, Val$^{76}$, Gly$^{77}$, Tyr$^{78}$, and Gly$^{79}$ are shown in ball and stick representation (red: oxygen, blue: nitrogen). Four of seven $\text{K}^+$ ions (yellow) are dehydrated and coordinated by carbonyl and side chain oxygens (B). The figures were generated using PyMOL.
include using the binding of a Fab antibody fragment (Rasmussen et al., 2007) or inserting T4 lysozyme into the third intracellular loop (Cherezov et al., 2007; Rosenbaum et al., 2007) of β2AR, and using a thermostabilized mutant (m23) of β1AR (Warne et al., 2008). The β2AR-Fab complex, β1AR-m23, and β2AR-T4 lysozyme were successfully crystallized, showing that the nature of the ligand binding pockets are very consistent with one another (Cherezov et al., 2007; Rosenbaum et al., 2007; Warne et al., 2008) and also with biochemical data (Yao et al., 2006; Kobilka and Deupi, 2007), providing insight into the structural basis of pharmacological selectivity between closely related receptors (Cherezov et al., 2007).

Recently solved P-type ATPases include the sarcoplasmic reticulum Ca2+ ATPase (SERCA) with the ion pathway open (Olesen et al., 2007), the sodium-potassium (Na+/K+) pump (subfamily II) (Morth et al., 2007; Shinoda et al., 2009), and the proton (H+) pump (subfamily III) (Pedersen et al., 2007). The SERCA structure was determined in the presence of a phosphate analog, BeF3−, caught in the act of releasing Ca2+ ions by splaying the TM helices out into the sarcoplasmic reticulum (SR) lumen, forming a funnel-shaped pathway (Olesen et al., 2007). The solved structures of Na+/K+ and H+ pumps, surprisingly, share the same TM helix arrangement with the same relative position of the cytosolic domain with that of SERCA regardless of their distinct size, charge, or number of transported ions (Morth et al., 2007; Pedersen et al., 2007). The latest structure of the Na+/K+ pump with 2K+ · MgF42− (2K+ · Pi analog) revealed that one of the subunits absent in the SERCA extracellular domain and a single spanning TM helix (β-subunit), which stabilizes a TM helix involved in ion binding and cholesterol by hydrogen bonding from clusters of aromatic residues, are required for K+ binding (Shinoda et al., 2009).

Several structures of secondary active transporters have been reported that represent the alternating access mechanism, which is the well-developed typical substrate-translocation mechanism in this family (Weyand et al., 2011). In this mechanism, the motion of molecules is coupled with downhill movement of ions along an electrochemical gradient by a reciprocal alternation of two conformations, inward-facing and outward-facing (Weyand et al., 2011). Two structures of the MFS, major facilitator superfamily, including GlpT, glycerol-3-phosphate (G3P) transporter (Huang
et al., 2003) and LacY, lactose permease (Abramson et al., 2003) from *E. coli* were published in 2003. Their overall structures (Figure 4 and Figure 5) are similar as the functional monomers form an hourglass shape with each N- and C-terminal 6 TM helices halves connected by a long central loop, showing pseudo two-fold symmetry perpendicular to the membrane plane (Abramson et al., 2003; Huang et al., 2003). The substrate binding sites of GlpT are located at the central pore at the interface of the two halves, and this site is enriched in aromatic amino acids, which regulate opening of the pore (Huang et al., 2003). Two Trp residues, Trp138 and Trp161, at the N-terminus of GlpT, protruding into the C-terminus to separate the inside and outside of the pore, presumably help closing the cytosolic pore in its outer-facing conformation (Huang et al., 2003). A “rocker-switch” in the substrate binding site may be achieved by tilting Arg45 and Arg269 at the central pore upon substrate binding (Huang et al., 2003). The substrate binding site of LacY is also in the central cavity, and Trp151 is involved, as commonly seen in galactoside binding proteins (Abramson et al., 2003). Similar to the rocker-switch in GlpT, LacY forms a salt-bridge between Arg144 and Glu269 upon substrate binding (Abramson et al., 2003). The alternating access mechanism is also observed in other recently solved structures of secondary active transporters (Faham et al., 2008; Shimamura et al., 2008; Weyand et al., 2008; Gao et al., 2009; Weyand et al., 2011), and may serve as a general model for the action of facilitated transporters.

Recent structures from chloride channel (ClC) families of *S. typhimurium* and *E. coli* were reported at 3.0, 3.5, and 2.5 Å resolution (Dutzler et al., 2002; Dutzler et al., 2003) (Figure 6). The structures form a homodimer representing a double hourglass shape, related by a pseudo two-fold axis perpendicular to the membrane plane (Dutzler et al., 2002). Each subunit in the dimer holds its own translocation pore as a channel forming unit and displays an antiparallel architecture (Dutzler et al., 2002). The selectivity filter is placed at the pseudo two-fold symmetry axis, with the positive N-termini of two helices stabilizing the Cl\(^-\) ion in the membrane (Dutzler et al., 2002). The Cl\(^-\) ion does not directly contact a full positive charge from Lys or Arg side chains but interacts with partial positive charges derived from the helix dipole, a main chain nitrogen, and side chain oxygen atoms (Dutzler et al., 2002; Huang et al., 2003). Glu148,
Figure 4. A schematic GlpT structure generated using PyMOL. The original three-dimensional structure resolved from X-ray diffraction data is in the PDB (code 1PW4) (Huang et al., 2003). GlpT adopts a pseudo two-fold symmetry with 6 helices arranged into each half. α-helices are shown as straight cylinders with the number of each of the 12 TM segments indicated. Arg^{45} and Arg^{269} in TM1 and TM7, respectively, are proposed to be involved in the substrate binding. The TM1 and TM7 are enriched in aromatic residues, presumably filling the space in the central pore and helping to close the pore. Trp^{138} and Trp^{161} located at the half of the N-terminal side may be involved in closing the cytosolic pore into the outer-facing conformation during the alternating access mechanism.
Figure 5. Three dimensional structure of LacY in a side view (PDB code 1PV6) (Abramson et al., 2003). The overall structure is similar to GlpT with a pseudo two-fold symmetry of six helices forming two bundles. Both N- and C-termini face the cytoplasmic side. Trp$^{151}$, Arg$^{144}$, and Glu$^{269}$ are shown as sticks and balls. The figure was generated from PyMOL.
Figure 6. Structures of a prokaryotic chloride channel (ClC) from Salmonella typhimurium (S. typhimurium). (PDB code: 1KPL) (Dutzler et al., 2002) (A) The ClC forms a homodimer. Each monomer (pink and skyblue) has its own translocation pore as a channel forming unit and displays antiparallel architecture. (Cl⁻ ion: green sphere) (B) The selectivity filter has positive N-termini (blue cylinder helices) of two helices projecting and stabilizing the Cl⁻ ion in the membrane (C-termini: red cylinder helices, Cl⁻ ion: green sphere). Glu¹⁴⁸, nearby the selectivity filter, protrudes into the pore, presumably functioning as an opening gate for Cl⁻ (green). (C) The Cl⁻ ion (green sphere) is coordinated by a main chain nitrogen atoms (Ile⁳⁵⁶ and Phe⁴⁵⁷), and side chain oxygen atoms (Ser¹⁰⁷ and Tyr⁴⁴⁵) (black dashed lines: polar interactions and green dashed lines: hydrophobic contacts). The figure was generated using PyMOL.
nearby the selectivity filter, protrudes into the pore, presumably functioning as a gate for Cl\(^-\) (Dutzler et al., 2002; Dutzler et al., 2003). This Glu\(^{148}\) gating mechanism is confirmed by structures of ClC mutants, E148A and E148Q, whereby Glu\(^{148}\) obstructs the pore by its carboxyl group acting as a Cl\(^-\) analog (Dutzler et al., 2003). This gating mechanism deviates from typical channels and may exhibit features of a protonated outward-facing conformation in a coupling exchange cycle (Dutzler et al., 2003; Gouaux and Mackinnon, 2005). While showing a relatively slow turnover rate (\(10^5\) s\(^{-1}\)) as a channel but a fast turnover rate as a transporter, the conformational changes upon chloride ion exchange may be small (Accardi and Miller, 2004). A recent electrophysiology analysis directly observed that the bacterial homologue of ClC is a Cl\(^-\)/H\(^+\) exchanger, rather than a Cl\(^-\) channel as suggested by accumulated data (Accardi and Miller, 2004). Moreover, some of the ClC family members also function as Cl\(^-\)/H\(^+\) exchangers (Accardi and Miller, 2004; Gouaux and Mackinnon, 2005). Interestingly, the ClC family may be an intermediate class lying between a channel and transporter (Accardi and Miller, 2004; Gouaux and Mackinnon, 2005).

### 1.2.4 Tryptophan residues in membrane proteins

Trp is the largest amino acid residue bearing an indole ring, though the residue is the least abundant building block in proteins, accounting for a frequency of 1.2% from known protein sequences in the SWISS-PROT database (Wallace B. A., 1999). Trp is known for its manifold roles in protein architecture and functions (Lee et al., 2009). The secondary structural propensity of Trp via Ramachandran \(\Phi-\Psi\) plots of 84 high-resolution structures shows that Trp can adopt both \(\alpha\)-helical and \(\beta\)-sheet conformations (Topham et al., 1993). Trp residues are unevenly distributed in proteins with two thirds of them present in the interior region of 46 monomeric proteins (Miller et al., 1987). From the structure-based distribution, the free energy of transfer of Trp to the lipid bilayer was calculated as 0.4 kcal/mol (Miller et al., 1987), suggesting Trp is non-polar. The hydrophobicity of Trp is controversial as it is designated to be very hydrophobic or moderately hydrophilic depending on the type of scales (Miller et al., 1987; Radzicka, 1988). Indeed, Trp may be considered the ideal interface residue due to the amphipathic nature of its side chain and the ability of the indole nitrogen atom to participate in hydrogen bonds.
In membrane proteins, Trp is enriched in the boundary region of TM segments. Trp residues in the ~12 Å interface of the lipid bilayer, known as the aromatic belt (Yau et al., 1998; von Heijne, 1999), are found in the photosynthetic reaction center (RC), bacteriorhodopsin (BR), porins, cytochrome c oxidase (COX), and the bacterial potassium channel (Schiffer et al., 1992; Lee, 2003). They function to stabilize a membrane protein in the bilayer by localizing close to the glycerol backbone and donating hydrogen bonds with ester carbonyls of membrane lipids (Sanchez et al.). With a comparable volume of 228 Å³ with phospholipid headgroups, Trp residues can disrupt the packing of the glycerol backbone, allowing nearby charged residues like Lys to be exposed on the membrane surface (Lee, 2003). In single-spanning membrane proteins, Trp residues show significant preference for presenting at the outer interface (Ridder et al., 2000), suggesting that Trp residues may help anchor TM helices at the membrane-water interface (trans-side of the membrane) (Clark et al., 2003). The anchoring effect of Trp was also observed in RC by enrichment of Trp on the periplasmic side of the bacterial membrane (Schiffer et al., 1992). Trp residues at the center of a TM segment tend to be buried and facing away from the lipid (Pilpel et al., 1999; Lee, 2003).

The ability of Trp to function at a protein-protein interaction interface is established in soluble proteins by the concept of “hot-spots” (Bogan and Thorn, 1998). Trp residues in membrane proteins may also play a key role in helix-helix interactions, functioning as a driving force of tertiary and quaternary structure in membrane protein packing (Eilers et al., 2000). Adamian and Liang (Adamian and Liang, 2001) investigated the pairwise interaction propensities of amino acid residues participating in interhelical interactions in TM helices. They reported that Trp residues have the highest non-bonded atomic contacts in TM regions reflecting van der Waals and hydrogen bonding interactions (Adamian and Liang, 2001). The interhelical interaction propensity of Trp is the second highest (1.1) after Met (1.17) in 14 α-helical structures and the highest value in soluble proteins (1.32) as determined by computational geometry, which measures the nearest atomic contact (Adamian and Liang, 2001). Trp is likely to interact with Leu in both soluble and membrane proteins and with Phe, Ile, and Val in membrane proteins, accounting for 40% of all pairwise contacts (Adamian and Liang, 2001). The highest propensity of helical interfacial pairwise contact with Trp is Arg (2.5) followed
by His (1.9), Phe (1.9), Gln (1.3), Met (1.2), and Pro (1.2). Trp residues have the highest interhelical contact by non-polar-non-polar atomic contact, accounting for 69% of all atomic pairwise contacts, with His (2.7), Gly (2.4), and Phe (1.8). The frequently observed Trp-Gly interaction at interhelical packing in membrane proteins is rare in soluble proteins. Many fewer self-pairs of Trp were observed in membrane proteins, which show high propensity in soluble proteins (2.2). Evidence of Trp functioning in TMS (TM segment)-TMS interactions was also provided by some other studies (Langosch and Heringa, 1998; Ridder et al., 2005). Trp residues are over-represented within the TM-helix-helix interface of the heptad repeat in BR, RC, and COX (Langosch and Heringa, 1998). Trp is the most abundant amino acid found at the position of TMS interaction sites, with a four times higher frequency of occurrence at this position (Ridder et al., 2005). Braun and Heijne also demonstrated that an introduced poly-Trp segment can act as an efficient stop-transfer sequence across the endoplasmic reticulum (ER) membrane (Braun and von Heijne, 1999). Taken together, Trp may extensively interact with adjacent helices and interhelical regions and function in a stop-transfer sequence during the translocation, allowing for the proper folding and assembly of membrane proteins (Braun and von Heijne, 1999; Clark et al., 2003).

Analysis of 14 α-helical membrane proteins by Adamian and Liang revealed that BR, halorhodopsin (HR), and RD are tightly packed, while other membrane proteins such as mechanosensitive ion channel contain extensive voids and pockets in their TM regions (Adamian and Liang, 2001). Adamian and Liang also noted that Trp residues are the most likely to be found in a pocket or void among all amino acid residues for both soluble and membrane proteins. Since voids and pockets are often occupied by ligands or prosthetic groups, this suggests an important role of Trp residues in the formation of binding sites for ligands or prosthetic groups (Adamian and Liang, 2001). Trp residues in active sites are observed in many soluble proteins, for example acetylcholinesterase (Ordentlich et al., 1995), streptavidin (Katz, 1995), myosin (Yengo et al., 1998), and chitin-binding proteins (Zeltins and Schrempf, 1997). Trp residues are also found in the active sites of many saccharidases (Ozaki et al., 1991; Rao et al., 1996), suggesting an essential role in carbohydrate binding and hydrolysis (Samanta et al., 2000).
The aromaticity of Trp enables it to behave as a chromophore in fluorescence spectrometry with an absorption maximum at 295 nm and emission maximum at 350 nm. The fluorescence of Trp is very sensitive to its environment, being enhanced under hydrophobic conditions. Furthermore, fluorescence energy transfer from Trp residues to suitable bound acceptors can be used to measure relative distances of 20-50 Å in proteins (Sun et al., 2001). These properties make Trp a very useful intrinsic probe of the environment and dynamics of localized regions of proteins.

1.3 Anion exchanger 1 (AE1)

1.3.1 Anion exchanger family

The anion exchanger (AE) gene family is a subfamily of the SLC4 family of bicarbonate transporters, encoding integral membrane proteins with highly conserved sequence similarity especially in the membrane domain (Bruce and Tanner, 1999). With distinct modes of ion transport, all ten human genes of the SLC4 family can be sorted into three groups as: 1) chloride/bicarbonate anion exchangers (AEs), 2) sodium-bicarbonate co-transporters (NBCs), and 3) sodium-driven chloride/bicarbonate exchangers (NDCBEs) (Romero et al., 2004). All of the SLC4 members transport their substrates electroneutrally, except for two electrogenic transport systems of NBCs and perhaps AE4, sodium-coupled bicarbonate transporter (NCBE), and bicarbonate-transporter related protein 1 (BTR1) with unknown functions (Romero et al., 2004).

The human anion exchanger (AE) family contains three genes: AE1, AE2, and AE3, mapped to chromosome 17q21-qter, 7q35-7q36, and 2q36, respectively (Reithmeier, 1996b). The genes are transcribed at multiple promoters, leading to a range of AE proteins with differences in the N-terminal regions of their cytosolic domains. AE1 (SLC4A1) is a single-copy gene of ~18 kb containing 20 exons (Alper, 1991). This gene produces two transcripts for AE1 forms expressed in the erythrocyte plasma membrane and the basolateral membrane of α-intercalated cells of the kidney distal and collecting tubules (Alper, 1991). Under the control of tissue specific promoters, mRNA products are alternatively spliced immediately upstream of exon 1, giving rise to erythrocyte AE1, and within intron 3, starting at Met66, resulting in a truncated kidney (k)AE1 form that is missing the first 65 residues in AE1 (Alper, 1991). With the presence of three
transcription start sites, the human AE2 (SLC4A2) gene yields five variants (Alper, 1991). AE2 contains a longer N-terminus, which can extend up to 300 additional residues depending on the transcript. AE2 also contains an insertion in the third extracellular (EC) loop (Z-loop) that contains three N-glycosylation acceptor sites and no acceptor sites in the EC loop, the site of N-glycosylation in AE1 (Reithmeier, 1996b). AE2 proteins are ubiquitously expressed in the basolateral membrane of most epithelial cells (Romero et al., 2004). Human AE3 (SLC4A3) possesses two alternative promoters generating two variants found mostly in excitable neural tissues and some epithelial tissues (Romero et al., 2004). These variants again are longer than AE1 by about 100 to 300 residues, arising from the N-terminal extension and the Z-loop insertion (Reithmeier, 1996b).

AE1 shares an overall amino acid sequence identity of 53% and 56% with AE2 and AE3, respectively (Romero et al., 2004). The membrane domains are more highly conserved with AE1 sharing 58 and 67% identity with AE2 and AE3, respectively (Romero et al., 2004). The AE1 sequence identity in mammals is very high (Table 2): at 45% overall and 70% in the membrane domain. Human AE1 is especially strongly conserved with mouse AE1 with an sequence identity of 81% overall and 89% in the membrane domain. Human and chicken display 59% and 74% identity for their full lengths and membrane domains, respectively. The yeast AE1 homologue is the most divergent, showing 23% identity with human AE1.

Human AE1 possesses a total of 11 Trp residues with 7 of them found in its membrane domain (Table 3). All 7 Trp residues are conserved in AE1, AE2 and AE3, except position 831, which is Leu in AE2. All Trp residues are conserved in mammalian and chicken AE1, with few Trp residues conserved in yeast AE1 (Table 4).

1.3.2 The structure and function
Human erythrocyte AE1 (eAE1) is 95 kDa in molecular mass as determined by SDS gel electrophoresis (Lepke and Passow, 1976; Falke et al., 1985b) and consists of 911 amino acid residues (Tanner et al., 1988), comprising two structurally and functionally distinct domains (Figure 7). The two domains can be physically separated at Tyr$^{359}$ and Lys$^{360}$ by mild chymotrypsin and trypsin digestion, respectively (Steck et al., 1976). The 43 kDa N-terminal cytosolic domain (cdAE1: Met$^{1}$-Lys$^{360}$) has a globular structure with 11 β-
Table 2. Sequence identity compared to human AE1 homologues

<table>
<thead>
<tr>
<th>Species</th>
<th>Protein</th>
<th>Overall</th>
<th>Membrane domain</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Paralogue</td>
</tr>
<tr>
<td>human</td>
<td>AE2</td>
<td>53</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>AE3</td>
<td>56</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Orthologue</td>
</tr>
<tr>
<td>mammals*</td>
<td>AE1</td>
<td>45</td>
<td>70</td>
</tr>
<tr>
<td>mouse</td>
<td>AE1</td>
<td>81</td>
<td>89</td>
</tr>
<tr>
<td>chicken</td>
<td>AE1</td>
<td>59</td>
<td>74</td>
</tr>
<tr>
<td>yeast</td>
<td>AE family protein</td>
<td>23</td>
<td></td>
</tr>
</tbody>
</table>

* Sequence identity among mammalian AE1 including human, mouse, house, and dog

Table 3. Location of Trp residues in the membrane domain of human AE1.

<table>
<thead>
<tr>
<th>Residues</th>
<th>Location</th>
<th>Mutations constructed</th>
</tr>
</thead>
<tbody>
<tr>
<td>492</td>
<td>TM4</td>
<td>W492A/F</td>
</tr>
<tr>
<td>496</td>
<td>TM4</td>
<td>W496A/F</td>
</tr>
<tr>
<td>648</td>
<td>EC loop 4</td>
<td>W648A/F</td>
</tr>
<tr>
<td>662</td>
<td>TM8</td>
<td>W662A/F</td>
</tr>
<tr>
<td>723</td>
<td>TM9</td>
<td>W723A/F</td>
</tr>
<tr>
<td>831</td>
<td>Not defined</td>
<td>W831A/F</td>
</tr>
<tr>
<td>848</td>
<td>TM11</td>
<td>W848A/F</td>
</tr>
</tbody>
</table>

Table 4. Conservation of Trp residues in the membrane domain of AE1 paralogues and orthologues.

<table>
<thead>
<tr>
<th>Species</th>
<th>Protein</th>
<th>Membrane domain homology</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Paralogues</td>
</tr>
<tr>
<td>human</td>
<td>AE1</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td>AE2</td>
<td>831L*</td>
</tr>
<tr>
<td></td>
<td>AE3</td>
<td>✓</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Orthologues</td>
</tr>
<tr>
<td>dog</td>
<td>AE1</td>
<td>✓</td>
</tr>
<tr>
<td>mouse</td>
<td>AE1</td>
<td>✓</td>
</tr>
<tr>
<td>horse</td>
<td>AE1</td>
<td>✓</td>
</tr>
<tr>
<td>chicken</td>
<td>AE1</td>
<td>✓</td>
</tr>
<tr>
<td>yeast</td>
<td>YNL275W</td>
<td>662A, 723A, 831I, 848F**</td>
</tr>
</tbody>
</table>

✓ = Trp residues are conserved in the membrane domain
* except position 831 in human AE2 is Leucine
** positions 662, 723, 831, and 848 in YNL275W are Ala, Ala, Ile, and Phe, respectively
Figure 7. Domain organization of erythrocyte AE1, kidney (k)AE1, and membrane (md)AE1. The entire human erythroid (e)AE1 consists of 911 amino acids with the N-terminal cytosolic domain encompassing residues 1-360. Kidney AE1 (kAE1) is a truncated version of eAE1 missing the first 65 amino acid residues from the N-terminal cytosolic domain. The 52 kDa C-terminal membrane domain (Asp$^{369}$-Val$^{911}$) has the transporter activity. The domain contains a single N-linked glycosylation site, Asn-X-Ser (Y), at Asn$^{642}$ in the fourth EC loop. An extracellular HA-tag inserted at the position Tyr$^{555}$ is indicated as a blue arrow. A C-terminal His$_6$-tag used for purification purposes is indicated as a green arrow.
strands and 10 α-helical segments (Zhang et al., 2000), providing interaction sites for a number of cytoskeletal and cytosolic proteins, and maintaining the biconcave disc shape of the erythrocyte (Low, 1986) (Figure 8). The 52 kDa C-terminal membrane domain (mdAE1: Gly$^{361}$-Val$^{911}$) is the functional unit of the transporter (Grinstein et al., 1978; Falke et al., 1985b) and is thought to share the same protein fold in the SLC4 bicarbonate transporter superfamily. The domain spans the membrane ~12 times and contains a single N-linked glycosylation site (Asn-X-Ser/Thr-) at Asn$^{642}$ in the fourth EC loop (Tanner et al., 1988). Erythrocyte AE1 is essential for the respiratory system (Figure 9) along with carbonic anhydrase (CA) (Vince et al., 2000). The anion exchange process enhances the carbon dioxide-carrying capacity of the blood by transporting bicarbonate into the blood in exchange for chloride. AE1 also plays important roles as a red cell antigen (Diego) and senescence. Mutations in AE1 thereby result in disease symptoms, such as hereditary spherocytosis (HS) and the Southeast Asian ovalocytosis (SAO).

Kidney AE1 (kAE1), whose structure of the N-terminal cytosolic domain is not available yet, is a truncated version of AE1 missing the first 65 amino acid residues from the N-terminal cytosolic domain. Missing a disordered acidic region (Met$^{1}$-Thr$^{54}$) and the first β-stand (His$^{55}$-Val$^{65}$) in the truncated cytosolic domain (Zhang et al., 2000; Pang et al., 2008), kAE1 is unable to interact with some of eAE1 binding partners like ankyrin (Ding et al., 1994) and glycolytic enzymes (Wang et al., 1995). The cytosolic domain of kAE1 has a less stable and more open structure than the erythroid version (Pang et al., 2008). The kidney AE1 (kAE1) facilitates bicarbonate reabsorption into the blood and acid secretion into the urine. An impairment of the kAE1 function is associated with distal renal tubular acidosis (dRTA), in which patients are unable to acidify their urine and linked to severe disease symptoms, such as metabolic acidosis, hypokalemia, nephrocalcinosis, kidney stones, and metabolic bone disease (Bruce and Tanner, 1999). These mutations typically result in folding defects and impaired trafficking of kAE1 to the plasma membrane (Quilty and Reithmeier, 2000; Kittanakom et al., 2004; Cheung et al., 2005a; Cordat et al., 2006).
Figure 8. Topology of human anion exchanger 1 (AE1).

Human erythrocyte AE1 is comprised of two structurally and functionally distinct domains. The 43 kDa N-terminal cytosolic domain (cdAE1: Met¹-Lys³⁶⁰) has a globular structure. The 52 kDa C-terminal membrane domain (mdAE1: Gly³⁶¹-Val⁹¹¹) is a functional unit of the transporter that presumably has the same protein fold as in the anion transport superfamily (SLC4). The domain spans the membrane ~12 times and contains a single N-linked glycosylation site (Asn-X-Ser/Thr-) at Asn⁶⁴² in the fourth EC loop. Kidney AE1 (kAE1) is a truncated version of eAE1 missing the first 65 amino acid residues, including a disordered acidic region (Met¹-Thr⁵⁴) and the first β-strand (His⁵⁵-Val⁶⁵), from the N-terminal cytosolic domain.
Figure 9. Anion exchanger 1 (AE1) in the erythrocyte and kidney.

(A) erythrocyte AE1 is essential for the respiratory system and enhances the carbon dioxide-carrying capacity of the blood by transporting bicarbonate into the blood in exchange for chloride. (CA: carbonic anhydrase, Hb: hemoglobin, aqp1: aquaporin 1).

(B) The kidney AE1 (kAE1) facilitates bicarbonate reabsorption into the blood and acid secretion into the urine. Mutations in kAE1 affect its functional expression and are associated with distal renal tubular acidosis (dRTA), and impaired urine acidification.
1.3.3 Tertiary and quaternary structural domains

A 2.6 Å crystal structure resolved from X-ray diffraction revealed that cdAE1 forms a tight symmetric butterfly-shaped dimer stabilized with interlocked dimerization arms (residue 304-357) (Zhang et al., 2000). The arms, consisting of largely helical structures with a single β-sheet, extend away from the globular peripheral protein binding domain and are interlocked by extensive interactions at the interface by eight backbone-to-backbone hydrogen bonds contributed by the intermonomeric antiparallel β-strands (Zhang et al., 2000). The crystal structure displays the native dimer structure of AE1 under physiological conditions. cdAE1 was expressed in *E. coli* and crystallized at pH 4.8, which may reflect the low pH structure of reversible pH-dependent conformational change that cdAE1 can undergo *in vitro* (Zhang et al., 2000; Zhou and Low, 2001). The crystal structure of cdAE1 tetramer is not available, yet the tetramer is the primary requirement of ankryrin binding in the erythrocyte.

The structure of mdAE1 reconstituted in lipids was determined at 20 Å resolution from two-dimensional crystals using negatively stained electron microscopic images (Wang et al., 1993; Wang et al., 1994) (Figure 11). The three-dimensional map was limited in providing information about the molecular details of mdAE1 structure; yet, consistent with biochemistry data, the structure was characterized as a dimer of two monomers related by two-fold symmetry and forms a central depression at the interface of the two domains (Wang et al., 1993; Wang et al., 1994). Three-dimensional crystals of deglycosylated mdAE1 covalently labeled with an AE1 inhibitor, DIDS, diffracted to a higher resolution (14 Å) (Lemieux et al., 2002). Further optimization is required to obtain crystals that diffract to atomic resolution. A recent cryo-electron crystallography study reported some progress in mdAE1 structure revealing the α-helical segments at 7.5 Å resolution in an outward-facing conformation (Figure 12) (Yamaguchi et al., 2010). The structure, interestingly, shares similar features and the angles of the α-helical segments with the ClC transporter (Dutzler et al., 2002; Dutzler et al., 2003). Several long and highly tilted helices were observed in the structure and assembled into two V-shaped peripheral densities, separated by a cleft (also see section 1.2.3).

Monomeric AE1 can only be observed in a denatured state in detergent; thus a functional monomer does not exist (Casey and Reithmeier, 1991). The dimer form is the
Figure 10. 2.6 Å crystal structure of the cytosolic domain of the human AE1 (cdAE1). cdAE1 forms a symmetric dimer stabilized with interlocked dimerization arms (residues 304-357, indicated in green ribbon diagram). Each monomer is represented in ribbon diagram with colour and gray. The arms have largely helical structures with a single β-strand and extend away from the globular peripheral protein binding domain (blue). (Trp residues: red and gray stick and ball). The figure is adapted from PDB (code 1HYN) (Zhang et al., 2000).
Figure 11. Three dimensional map of membrane domain of AE1 (mdAE1).

The AE1 structure was characterized as a dimer with a central depression at the interface of two subdomains. The size of the dimer is 110 Å x 60 Å x 80 Å in length, width, and thickness, respectively. The two planes indicate a lipid bilayer with 40 Å thickness. The structure was determined by electron microscopy of negatively stained two-dimensional crystals of mdAE1 reconstituted with phospholipids. The figure is adapted from (Wang et al., 1994).
Figure 12. A comparison of the AE1 membrane domain with the ClC chloride channel. Three-dimensional map of AE1 at 7.5 Å resolution projection in an outward-facing conformation was fitted with the ClC channel motifs (stereo view). AE1 and the ClC transporter are shown in mesh and ribbon, respectively (*Magenta and coral ribbon*: first and second motifs of a CIC monomer, *cyan and yellow ribbon*: first and second motifs of another CIC monomer). The AE1 structure shows similar features and the angles of α-helical segments with the CIC transporter. The figure is reprinted from Yamaguchi *et al.*, (2010) with permission of Elsevier.
fundamental unit for AE1 transporter activity, and the channel for anion transport likely exists, at least partially, at the dimer interface (Wang et al., 1993; Wang et al., 1994). AE1 presents as a mixture of dimers and tetramers in the erythrocyte membrane, yet isolated in detergent solution, mdAE1 is characterized as a dimer exclusively and is more stable than AE1 (Casey and Reithmeier, 1991; Vince et al., 1997). This implies that tetramer formation requires the cytosolic domain, and indeed, the intact AE1 has the tendency to undergo self-association in detergent, allowing formation of tetramers and higher oligomers (Vince et al., 1997). A higher proportion of tetramer and higher oligomeric forms of AE1 may be associated with the aging of red blood cells, in which AE1 aggregates creating a senescence antigen (Bruce et al., 1994).

1.3.4 Topology
Due to difficulty in obtaining a high resolutional structure of mdAE1, its topology has been extensively studied (Popov et al., 1997; Fujinaga et al., 1999; Popov et al., 1999; Kuma et al., 2002; Zhu et al., 2003; Zhu and Casey, 2007). These studies have provided information concerning the number and orientation of TM segments and the loops connecting them in AE1. Initial hydropathy analysis of the amino acid sequences of human AE1 revealed the presence of 10 hydrophobic peaks (Landolt-Marticorena et al., 1993), some of which correspond to sequences that are long enough to span the membrane twice (Reithmeier, 1996b). Later on, AE1 was proposed to have ~14 putative TM segments (Casey et al., 1995; Popov et al., 1997), though these models differ in detail especially in their C-terminal regions around TM 9 and TM 10. Several sites (also see section 1.3.8) in human AE1 sequences have been localized as a topology marker by biochemical experiments. Early proteolysis studies localized external chymotrypsin sites to the third EC loop at Tyr$^{553}$ and Leu$^{558}$ between TM 5 and 6 (Drickamer, 1978; Jennings et al., 1984). Two trypsin-sensitive sites were identified at Lys$^{360}$, cleaving the N-terminal cytosolic domain from the rest of AE1 (Steck et al., 1976), and at Lys$^{743}$, exposed at the cytoplasmic surface (Jennings et al., 1986; Kuma et al., 2002). Using epitope mapping, the C-terminus was identified to be on the cytoplasmic site in the erythrocyte (Lieberman and Reithmeier, 1988; Wainwright et al., 1989), and thus, AE1 spans the membrane an even number of times. The epitope for a monoclonal antibody,
BRIC 132 against residues Phe$^{813}$-Tyr$^{824}$, localized to the cytoplasmic side of the membrane (Wainwright et al., 1989). The Diego antigen, derived from a point mutation (P854L), and the Wright antigen (W$^s$) at E658K, providing a binding site for glycophorin A (GPA), were localized in an EC loop (Bruce et al., 1995). A palmitoylation site, Cys$^{843}$, is localized to the cytosolic site and modified by fatty acid (Okubo et al., 1991). An N-glycosylation site (Asn$^{642}$) was localized to the EC loop 4. The site must be exposed in the ER lumen to be N-glycosylated during biosynthesis and is then exposed extracellularly following trafficking to the plasma membrane. Thus, scanning N-glycosylation analysis was able to accurately map the extracellular ends of TM segments in AE1 and identified a re-entrant loop (T-loop) connecting TM9 and TM10 (Popov et al., 1997; Popov et al., 1999; Cheung et al., 2005b). The re-entrant loop does not fully span the lipid bilayer and is less hydrophobic than typical TM segments. It folds into the interior of the protein and often has a dynamic aspect related to transport function.

Non-penetrant reagents are also well established in their use for identifying extracellular or intracellular reactive sites. The H$_2$DIDS-reactive lysines, Lys$^{539}$ and Lys$^{851}$, for example, are localized at the extracellular side of AE1, as they were covalently modified when H$_2$DIDS was applied from the exterior but not interior side of the red cells (Okubo et al., 1994) (also see section 1.3.6). H$_2$DIDS reacts rapidly with Lys$^{539}$ at neutral pH due to the low pKa of this lysine and can be crosslinked to Lys$^{851}$ under more alkaline conditions. Cysteine-scanning accessibility mutagenesis was used to define the topology of AE1 (Tang et al., 1998; Fujinaga et al., 1999; Tang et al., 1999; Zhu et al., 2003; Zhu and Casey, 2007). Cysteine residues located in hydrophobic environments are poorly reactive allowing the identification of TM segments. Non-penetrating cysteine-directed reagents allow the determination of extracellular sites when applied to intact cells expressing AE1. The current topology model (Figure 13) displays AE1 with 12 TM segments, some of which are re-entrant and may not completely span the membrane (Tang et al., 1998; Fujinaga et al., 1999; Tang et al., 1999; Zhu et al., 2003; Zhu and Casey, 2007).
Figure 13. Predicted folding model of the C-terminal membrane domain (Asp$^{369}$-Val$^{911}$) of human AE1. The model shows 12 TM segments and a re-entrant T-loop. The N-and C-termini of the membrane domain face the cytoplasm. Positions of the seven tryptophans (blue), two H$_2$DIDS-reactive lysines (red), and reactive glutamate (green) residue are indicated. (C: chymotrypsin-sensitive site, Pa: papain-sensitive site, T: trypsin-sensitive sites, an N-linked oligosaccharide (magenta) are attached to Asn$^{642}$).
1.3.5 Ion translocation

AE1 catalyzes rapid Cl⁻/HCO₃⁻ exchange at a rapid turnover rate of ~5x10⁴ sec⁻¹ at 37°C (Jennings, 1989), just 10-fold slower than a typical ion channel. AE1 transport is not limited to Cl⁻/HCO₃⁻ as AE1 is able to mediate exchange of a wide range of anions including: sulfate, phosphate, pyridoxal 5-phosphate, and phosphoenolpyruvate (Nanri et al., 1983; Passow, 1986; Yamaguchi et al., 2010), yet with much lower turnover rates ~5 sec⁻¹ (Jennings, 1989). The obligatory exchange of anions by AE1 is accomplished by a ping-pong mechanism, in which movement of an ion in one direction is followed by that of another ion in the other direction (Falke and Chan, 1985; Jennings et al., 1998). This ping-pong model suggests the existence of two stable conformations of inward-facing and outward-facing, and this model is supported by many studies (Falke et al., 1984b; Falke and Chan, 1985; Jennings et al., 1998). The transition between the two states is facilitated by bound anions and does not occur at a measurable rate with an empty binding site. Both sites in the AE1 dimer can face inward or outward at the same time, suggesting that the two subunits can operate independently from one another.

Anion exchange activity can be carried out exclusively by the membrane-spanning region of mdAE1 (Lepke and Passow, 1976; Grinstein et al., 1978). Carbonic anhydrase II (CAII) binds to the C-terminal tail of AE1 at a specific site (886LDADD890) to link transport and metabolism, effectively converting CO₂ to HCO₃⁻, to maintain the rapid AE1 transport activity (Vince and Reithmeier, 2000; Reithmeier, 2001). AE1 is suggested to form a funnel with a large aqueous vestibule within the TM domain (Jennings, 1989) like in the ClC channel (Yamaguchi et al., 2010). According to the model, anions diffuse through an open aqueous pore-lining region in the channel until reaching the permeability barrier (Jennings et al., 1990). TM helices, 5, 8, 9, and 10 of AE1 are implicated in an anion translocation site by containing important polar residues, Lys, Arg, His, and Glu that face the ion translocation pathway (Muller-Berger et al., 1995a; Muller-Berger et al., 1995b; Karbach et al., 1998) (discussed in section 1.3.7.). Modification of certain Cys substitutions by sulfhydryl reagents in Ser⁶⁴³ to Ser⁶⁹⁰ in TM8 can inhibit anion transport (Tang et al., 1998), suggesting that this region forms part of the anion translocation pore. These reactive residues are spaced 3-4 residues apart consistent with a helical conformation.
1.3.6 Inhibitor binding sites

Stilbene disulfonate derivatives are known to inhibit anion exchanger activity by reversibly or irreversibly binding with AE1 (Rao et al., 1979). 4,4’ isothiocyano-2,2’-stilbene disulfonate (DIDS) and a reduced DIDS analogue, H$_2$DIDS, block the exchanger ($K_i = 0.04 \mu$M) by binding to an exofacial site in each AE1 monomer. These reagents can form a cross-link at high pH (H$_2$DIDS at pH9.5 and DIDS at pH13) between conserved Lys$^{539}$ and Lys$^{851}$ (Jennings and Passow, 1979; Kang et al., 1992; Schopfer and Salhany, 1995). The two residues are also covalently modified by pyridoxal-5-phosphate, which is transported by AE1, suggesting that these residues are located on the transport pathway (Wood et al., 1992). The site is characterized as a hydrophobic environment as revealed by enhanced fluorescence upon binding of inhibitors such as stilbene disulfonates, 4-benzamido-4’aminostilbene-2-2’-disulfonate: BADS ($K_i = 2 \mu$M) and 4,4’-benzamidostilbene-2-2’-disulfonate: DBDS ($K_i = \sim2 \mu$M) (Pimplikar and Reithmeier, 1986, 1988a). The bound inhibitors are inaccessible to anti-DIDS antibodies and are likely contained deep within a bundle of helices of AE1 (Landolt-Marticorena et al., 1995). Fluorescence energy transfer experiments showed that the inhibitors are close to each other in the dimer at a distance between 28-46 Å (Macara and Cantley, 1981) and can cooperatively interact with each other, with the first bound inhibitor increasing the affinity for the second. Inhibitor binding can stabilize the membrane domain of AE1 against thermal denaturation and proteolysis (Boodhoo and Reithmeier, 1984; Casey and Reithmeier, 1991). As shown by $^{35}$Cl NMR binding studies, 4,4’-dinitrostilbene-2-2’-disulfonate: DNDS ($K_i = 2 \mu$M), blocks the Cl$^-$ binding site by directly occupying the site (Falke et al., 1984a). These studies with competitive inhibitors of anion transport support the notion that the inhibitors occupy sites that encompass the anion binding site, which remains to be identified.

4-acetamido-4’isothiocyanostilbene-2-2’-disulfonate: SITS ($K_i = 10 \mu$M) is an AE1 inhibitor best-utilized as an immobilized ligand for characterizing AE1 structure and function. Mis-folded AE1 mutants like AE1 SAO fail to bind to SITS-Affi-Gel, while properly folded AE1 can bind and the binding can be blocked with free inhibitors. AE1 can bind reversibly to SITS-Affi-Gel at 4°C (Pimplikar and Reithmeier, 1988b). AE1 becomes more tightly bound in a time-dependent manner at higher temperatures.
(Pimplikar and Reithmeier, 1986), presumably through a change in conformation due to a partial translocation of the inhibitor into the binding site as derived by kinetic studies (Falke and Chan, 1985; Falke et al., 1985a). 4-benzamido-4’-isothiocyanostilbene-2-2’-disulfonate: BIDS ($K_i = \sim 2 \mu M$) and eosin-5-maleimide: EMA were also shown to induce a conformational change in AE1 upon binding, by quenching both intrinsic Trp fluorescence and fluorescence of the covalently bond inhibitors (Macara et al., 1983). 2,4-dinitrofluorobenzene: DNFB leaves the transport site and channel intact but slows translocation of bound Cl$^-$ in AE1, as shown by the slowing of other translocation inhibitors in $^{35}$Cl NMR studies (Falke and Chan, 1986a, b, c). EMA and DNFB, may therefore prevent anion translocation by locking AE1 into the outward-facing orientation without directly blocking the channel.

The inhibitor binding site is hydrophobic in nature as shown by the enhanced fluorescence of inhibitors like BADS upon binding to AE1 (Lieberman and Reithmeier, 1983). Being a hydrophobic environment, the inhibitor binding site may be lined by apolar residues such as Phe and Trp. In fact, energy transfer is observed between Trp residues and the bound fluorescent inhibitor, BADS at an estimated distance of $\sim 20 \text{Å}$ (Rao et al., 1979). The effect of mutating Trp residues on inhibitor binding by AE1 or its transport function has not been thoroughly studied and is the subject of this thesis.

1.3.7 Functionally important amino acid residues
Functionally important amino acid residues in anion binding and translocation of AE1 have been identified by chemical modification and site-directed mutagenesis studies (Jennings and Passow, 1979; Kang et al., 1992; Wood et al., 1992). The first characterized important residues were DIDS-reactive Lys$^{558}$ and Lys$^{869}$ of mouse AE1, homologous to Lys$^{539}$ and Lys$^{851}$ in human AE1 (also see 1.3.6.). Known as a binding site for pyridoxal-5-phosphate, either mutation of Lys$^{539}$ or Lys$^{869}$ can retain AE1 transport activity (Wood et al., 1992). While these Lys residues are important for inhibitor binding, they are not essential for the anion transport.

A glutamate residue Glu$^{681}$, located at the end of TM8 on the cytosolic side of the membrane, is involved in the AE1 anion transport process (Jennings and Anderson, 1987). This residue was first identified by showing altered anion exchange kinetics when
reacting with Woodward’s reagent K or reduced with sodium borohydride by Jennings and co-workers (Jennings and Anderson, 1987; Jennings and Smith, 1992). A mutagenesis study also confirmed its functionally important role (Muller-Berger et al., 1995a). Glu$^{681}$ directly participates in AE1 anion transport by protonating or deprotonating its carboxyl group, thereby providing the proton necessary for sulfate/chloride exchange (Jennings and Smith, 1992). As sulfate/proton co-transport takes place for both inward-facing and outward-facing conformations and is accessible from both the intracellular and extracellular sides of the membrane, Glu$^{681}$ may also function as a permeability barrier of AE1.

Anion exchange activity of AE1 is sensitive to a change of pH in the extracellular and intracellular environments (Jennings, 1989). The Cl$^-$ transport activity increases with pH until a plateau at 7.0 to 10.5 and further increases until pH 11.5 followed by a sharp decrease (Muller-Berger et al., 1995a). This pH-dependent anion transport behaviour suggests the presence of at least three dissociating groups with pKa values of approximately 6, 10.5, and 12, functioning in a rate limiting step (Muller-Berger et al., 1995a). Four conserved His residues (703, 735, 816, 834), presumably in TM helices of the access channels, were revealed to be essential players in pH dependent Cl$^-$ transport activity. They may form hydrogen bonds in the channel accounting for the pH dependence of transport (Muller-Berger et al., 1995b). Mutagenesis of the four mouse AE1 His residues (721, 752, 837, 852) corresponding to human His (703, 735, 816, 834), when expressed in Xenopus oocytes caused partial (H752S and H837Q/R) or complete (H723Q and H852Q) impairment of Cl$^-$ transport (Muller-Berger et al., 1995b).

Modification of His$^{752}$ (corresponding to His$^{735}$ in human) changes the pH-dependent anion transport behaviour with a change of the rate limiting step from pKa 5.8 to 6.8 (Muller-Berger et al., 1995a). The same effect was also observed by modification of Glu$^{699}$ (corresponding to Glu$^{681}$ in human), suggesting that residues His$^{735}$ and Glu$^{681}$ may form a hydrogen-bond with each other, functioning in the rate-limiting step in anion exchange (Muller-Berger et al., 1995a).
1.3.8 Tryptophan residues

AE1 contains four Trp residues in its cytosolic domain: Trp$^{75}$, Trp$^{81}$, Trp$^{94}$, and Trp$^{105}$. These Trp residues are located nearby each other in the structure. Their fluorescence is quenched in the folded structure and becomes de-quenched upon unfolding in denaturants before being fully quenched in the completely unfolded protein. Located in the interior of globular cdAE1, these Trp residues may provide hydrogen bonds with other amino acid residues. An inter-subunit hydrogen bond between Trp$^{105}$ and Asp$^{316}$ may function in a pH-dependent conformational change in cdAE1 by breaking the hydrogen bond resulting in opening a peripheral protein-binding domain in a higher pH conformation.

AE1 possesses seven Trp residues in its membrane domain, and all are conserved across species and in a family of anion exchanger proteins, except Trp$^{831}$, which is Leu in human AE2. The positions of these Trp residues in AE1 are relatively well defined from topology studies (Popov et al., 1997; Tang et al., 1998; Popov et al., 1999; Zhu et al., 2003; Zhu and Casey, 2007). Based on the folding model, Trp$^{492}$ and Trp$^{496}$, are located in the fourth TM segment, across from Lys$^{539}$ and Lys$^{542}$ in the fifth TM segment, and project towards the most hydrophilic surface of TM4. Trp$^{492}$ and Trp$^{496}$ may face into the transport channel along with DIDS-reactive lysine (Lys$^{539}$) and Lys$^{542}$ on TM 5, forming part of the ion translocation pore. A second possibility is that Trp$^{492}$ and Trp$^{496}$ may play a role in TM helix-helix interactions.

Trp$^{648}$ is located in the fourth EC loop, six residues distal from the N-glycosylation site (Asp$^{642}$). The region Ser$^{643}$-Leu$^{655}$ including Trp$^{648}$ is inaccessible to a water-soluble sulfhydryl-specific reagent, LYLA (649 Da), suggesting the region is buried within the folded structure of this region. Another possibility is that the carbohydrate structure attached at Asp$^{642}$ sterically blocks LYLA access by interacting directly with the surface of the protein (Tang et al., 1998). As mentioned previously, Trp residues are often involved in carbohydrate interactions. Trp$^{648}$ was identified as functionally important residues involved in ion translocation by reducing transport activity when mutated to cysteine (Tang et al., 1999).

Trp$^{662}$ is located at the extracellular TM interface at the beginning of TM 8 as defined by N-glycosylation (Popov et al., 1997; Popov et al., 1999) and cysteine scanning mutagenesis (Tang et al., 1998). Trp$^{662}$ was identified as functionally important
residues involved in ion translocation by reducing transport activity when mutated to cysteine (Tang et al., 1999). Trp$^{723}$ is located at the extracellular end of TM 9 and connected to a trypsin-sensitive loop (T-loop: Trp$^{723}$-Leu$^{764}$) (Jennings et al., 1986), which may have structurally and functionally important roles by forming part of the channel in AE1. This loop may be transiently exposed to the lumen of the ER during biosynthesis and folded towards the cytoplasmic side of AE1 after AE1 is completely folded (Popov et al., 1999). It is also possible that this loop has a dynamic aspect and may play a role in ion translocation.

Trp$^{831}$ exists in the region involved in malaria parasite infection (Asp$^{821}$-His$^{834}$) (Crandall et al., 1993) and is close to the extracellular antigen related to cell senescence (Leu$^{812}$-Arg$^{827}$) (Bruce et al., 1994). The region is not well defined in the folding model of AE1 as the region contains both hydrophobic and charged residues including important histidine and lysine residues, (His$^{816}$, His$^{834}$, Lys$^{814}$, and Lys$^{817}$) required for anion transport (Askin et al., 1998). The region may be flexible and become accessible from the outside under certain circumstances (Zhu et al., 2003). This was supported by an N-glycosylation study in which P820N was slightly glycosylated in vivo but not glycosylated in transfected HEK cells (Popov et al., 1999). The BRIC 132 antibody binds to the epitope Phe$^{813}$-Tyr$^{824}$ when AE1 is treated with the detergent and digested using trypsin and chymotrypsin under low ionic strength conditions (Wainwright et al., 1989). The BRIC 132 antibody can bind to leaky erythrocytes but not the intact cells (Wainwright et al., 1989). These observations suggest that the BRIC 132 epitope is buried under isotonic conditions.

Trp$^{848}$ was defined within TM11 in folding models of AE1 rather than in an aqueous loop. Region Phe$^{836}$-Lys$^{851}$ was not labeled with biotin maleimide in a cysteine mutagenesis study (Zhu et al., 2003) suggesting that this region is embedded in the membrane, and region Ile$^{840}$-Thr$^{853}$ was postulated to be in TM 11 in a N-glycosylation study (Popov et al., 1997). The TM, composed of 13 to 16 amino acid residues, may form a shorter helical structure than required to span the 30 Å membrane thickness. A helical wheel plot showed functionally important residues located in the hydrophilic helical face including Trp$^{848}$ (W848C showed 26% of transport activity) with the DIDS-reactive
Lys\textsuperscript{851} at the center of the helical face (Zhu \textit{et al.}, 2003). The hydrophilic helical side may face the anion translocation pore, and Trp\textsuperscript{848} may participate in ion translocation.

Taken together, these Trp residues in the membrane domain may play an important role in the positioning of TM segments during biosynthesis, in the ion translocation pathway of AE1, and in forming the inhibitor binding sites. The purpose of the present study is to test roles of Trp residues in the folding, trafficking, and function of human AE1 when expressed in HEK-293 cells.

\subsection*{1.4 Functional expression of AE1}

\subsubsection*{1.4.1 Biosynthesis}
Similar to other membrane glycoproteins, the biosynthesis of AE1 initiates by translation of its mRNA by ribosomes attached to the ER membrane. The process is guided by a signal-recognition particle (SRP), which recognizes and binds both the signal sequence of the nascent polypeptide and the large ribosomal subunit, forming an SRP-ribosome complex (Walter, 1995; Halic \textit{et al.}, 2004). TM 1 is the first region in the sequence of AE1 to be recognized as an internal signal sequence, but distal TM segments can also interact with SRP and be inserted into the ER membrane (Braell and Lodish, 1982; Tam \textit{et al.}, 1994). Upon binding of the SRP-ribosome complex to the SRP receptor in the ER membrane, SRP and SRP receptor are released by guanosine triphosphate (GTP) hydrolysis, leaving the ribosome engaged to the translocation machinery and allowing the translocation of the nascent polypeptide across the ER membrane (Song \textit{et al.}, 2000). The translocon, called Sec61 in eukaryotes after the yeast orthologue, recognizes and binds the signal sequence of the translocating peptide chain to initiate translocation of the polypeptide segment that follows (Song \textit{et al.}, 2000; Halic \textit{et al.}, 2004; Van den Berg \textit{et al.}, 2004). For secreted proteins, the signal sequence is cleaved co-translationally on the luminal side of the ER membrane by membrane-bound signal peptidase allowing translocation of the complete polypeptide into the ER lumen. Membrane proteins contain “stop-transfer” sequence that engage the translocon and prevent translocation of the following polypeptide. The binding of the stop-transfer sequence triggers the opening of the translocon and the stop-transfer sequences moves laterally into the surrounding lipid bilayer and becomes a TM segment (Van den Berg \textit{et al.}, 2004).
AE1 synthesis begins on free ribosomes with biosynthesis of the N-terminal cytosolic domain and acylation of its N-terminal Met residue. Upon exposure of TM1, which acts as a stable signal sequence, the nascent chain engages SRP and the ER. Polypeptide synthesis continues with TM1 and TM2 inserting into the lipid bilayer likely as a hairpin loop. Further synthesis of AE1 continues with an alternating stable signal sequence and stop-transfer sequence stitching the polypeptide into the lipid bilayer, hairpin loop after hairpin loop. The mdAE1 then assumes its native fold in a post-translational step after the last TM segment is released into the lipid bilayer upon termination of translation. AE1 is not made in the mature red blood cell, which is devoid of organelles such as nuclei, ribosomes, ER and Golgi necessary for glycoprotein biosynthesis, but rather is made late during erythropoiesis in red cell precursors. AE1 biosynthesis initiates at the proerythroblast stage and continues until the terminal maturation of erythroblasts; but is not made in reticulocytes released into the blood stream (Braell and Lodish, 1981; Hanspal et al., 1993).

AE1 undergoes several posttranslational modifications, including, modification of its oligosaccharide during its production in red cell precursors. AE1 also undergoes fatty acylation, and phosphorylation in mature red cells. Palmitoylation in human AE1 occurs at residue Cys843 (Okubo et al., 1991). This modification can take place in the mature red cell by a membrane-bound palmitoyl transferase. Palmitoylation is not observed in AE1 functionally expressed in HEK-293 or COS-7 cells and mouse AE1 expressed in Xenopus oocytes, suggesting that this modification is not essential for AE1 trafficking and transport function (Cheung and Reithmeier, 2004).

Human AE1 is phosphorylated at residues, Tyr8, Tyr27, Thr42, Tyr359, and Tyr904, by interacting with p72syk or lyn protein tyrosine kinases. It is also phosphorylated at Ser and Thr residues by casein kinase, and serine/threonine kinases present in the mature red cell (Yannoukakos et al., 1991; Wang et al., 1997). Phosphorylation at Tyr8 results in inhibition of association of AE1 with the glycolytic enzymes including glyceraldehyde-3-phosphate dehydrogenase (GAPDH), aldolase, and phosphofructokinase (Low et al., 1987). As such, AE1 phosphorylation may modulate glycolysis in vivo. Tyrosine phosphorylation also recruits tyrosine phosphatases, such as SHP-2, and Src kinase inhibitor PP2 for dephosphoylation of AE1 (Pantaleo et al., 2011). Tyr359 and Tyr904 are
also suggested to be involved in AE1 trafficking by reversible phosphorylation to interact with the cellular targeting and/or internalization machinery (Williamson et al., 2008). AE1 phosphorylation may also regulate erythrocyte viscoelastic property by modifying binding with cytoskeletal proteins (Barbul et al., 1999; Pang et al., 2008; Pantaleo et al., 2011).

1.4.2 N-glycosylation

During biosynthesis, both secreted and membrane proteins are covalently modified by the addition of sugars in the ER and Golgi. The sugar modification in membrane proteins is classified into two types including O-linked, N-acetylgalactosamine (GalNAc) linked to Ser or Thr via their hydroxyl groups, and N-linked, N-acetylglucosamine (GlcNAc) linked to Asn (Alberts, 2008). O-linked oligosaccharides are typically shorter than N-linked oligosaccharides, and are formed in the Golgi. O-linked oligosaccharides are added a single sugar at a time, catalyzed by distinct glycosyltransferases (Alberts, 2008).

AE1 contains a single N-glycosylation site at Asn⁶⁴² and that is co-translationally N-glycosylated (Braell and Lodish, 1982). N-glycosylation occurs en bloc by attachment of a large preformed high mannose oligosaccharide from a dolichol pyrophosphate donor to acceptor sites (-Asn-X-Ser/Thr-) by a membrane-bound oligosaccharyl transferase (OST) complex (Parodi, 2000) (see Figure 14). The oligosaccharides are then processed to complex oligosaccharides by sequential modification in the ER and Golgi.

Glycoproteins in their native conformation are further trimmed by mannosidase I at the ER and cis Golgi and processed to complex oligosaccharide through the medial- and trans-Golgi en route to the plasma membrane by N-acetylglucosaminyltransferase (Gn) I/II, galactosyltransferase, and sialyltransferase as described in Figure 14 (Parodi, 2000; Cheung and Reithmeier, 2007).

Surveys of eukaryotic membrane proteins reveal that 90% of membrane proteins are glycoproteins, and even those few non-glycosylated proteins are typically associated with glycoproteins (Landolt-Marticorena and Reithmeier, 1994). In many cases, N-glycosylation is required for proper folding, trafficking, and stability of membrane proteins (Siffroi-Fernandez et al., 2001; de Souza and Simon, 2002; Vagin et al., 2003; Watanabe et al., 2004). N-glycosylation, however, is not essential for the functional
Figure 14. N-glycosylation processing pathway in mammalian cells.

N-glycosylation occurs as a co-translational process in the lumen of the endoplasmic reticulum (ER) by attachment of a high mannose oligosaccharide (Glc$_3$Man$_9$GlcNAc$_2$-) from a dolichol pyrophosphate donor (Dol-P-P) to acceptor sites (-Asn-X-Ser/Thr-) by a membrane-bound oligosaccharyl transferase (OST) complex (i). The oligosaccharides are then processed to complex oligosaccharides by sequential modification in the ER and Golgi. α-glucosidase I (ii) and α-glucosidase II (iii) deglucosylate the N-glycan, which is trimmed by α-mannosidase I (iv) and α-mannosidase II (v). The N-glycan is processed from the high mannose oligosaccharide to complex oligosaccharide forms by N-acetylglucosaminyltransferase I (vi), α-mannosidase II (vii), and N-acetylglucosaminyltransferase II (viii) in the medial Golgi. The N-glycan is further modified by galactosyltransferase (ix) and sialyltransferase (x). The figure is modified from (Cheung and Reithmeier, 2007).
expression of AE1 nor for its transport activity in red cells. Deglycosylated AE1 retains its transport ability in resealed erythrocyte ghosts (Casey et al., 1992). Mutation of the N-glycosylation site does not affect the functional expression of AE1 in HEK cells or when expressed in Xenopus oocytes (Groves and Tanner, 1994). Blocking oligosaccharide addition to Asn$^{642}$ using tunicamycin had no effect on AE1 trafficking rate to the cell surface (Sabban et al., 1981).

1.4.3 Folding and trafficking

Folding of AE1, like most membrane glycoproteins, is assisted by cytosolic chaperones (Hsc70 and Hsp70), ER resident chaperones, calnexin (CNX), calreticulin (CRT), and Bip/GRP78 (Popov and Reithmeier, 1999; Patterson et al., 2009). BiP/GRP78, a member of the Hsp70 family, binds its substrates through a promiscuously exposed hydrophobic sequence. GRP94, a Hsp90 family member, binds fully oxidized molecules after the action of BiP/GRP78 (Melnick et al., 1994). CNX and CRT are lectins that bind to monoglucosylated N-glycans after sequential removal of the outermost two glucose residues by glucosidase I (GI) and glucosidase II (GII) in the ER (Ruddock and Molinari, 2006). CNX/CRT are dissociated from the bound glycoprotein by further trimming of the final glucose by GII. Rebinding of CNX/CRT occurs through reglucosylating a non-native conformer by a glucosyltransferase (GT). The GT monitors the structure of the dissociated polypeptide through an interaction with exposed hydrophobic patches on the proteins and generates a monoglucosylated N-glycan on a non-native glycoprotein (Ruddock and Molinari, 2006). The interaction cycle of CNX/CRT and GT with the folding intermediate assures proper folding of glycoproteins and results in retention of improperly folded glycoproteins in the ER (Parodi, 2000).

Properly folded glycoproteins are packed into COPII-coated vesicles (Duden, 2003) and exit the ER in their mature conformation to the cis-Golgi (Ruetz et al., 1993). Membrane proteins then migrate to the trans-Golgi and are targeted to their final destination, such as the plasma membrane and lysosomes (Alberts, 2008). This sorting is accomplished by membrane fusion with the assistance of SNAREs (SNAP receptors) including v-SNAREs on the vesicle associating with s-SNAREs on the target membrane (Sollner, 2003).
Non-native conformers are trimmed by mannosidase I generating a Man$_8$ oligosaccharide, which was initially thought to be a signal for the ER-associated degradation (ERAD) pathway (Jakob et al., 1998; Cabral et al., 2001). More recently, the ER mannosidase I and EDEMs (ER degradation-enhancing α-mannosidase-like protein) produce a Man$_{7.5}$ oligosaccharide, which may function as an ERAD substrate, and prevent reglucosylation by a GT (Bagola et al., 2011). The ER mannosidase I and EDEMs may guide the ERAD substrate to the ERAD machinery, though the Man$_{7.5}$ structure is probably not the sole ERAD signal (Bagola et al., 2011). The ERAD machinery is still not deciphered in detail. Misfolded proteins detected by the ERAD machinery, including membrane proteins are retrotranslocated into the cytosol and degraded by the 26S cytosolic ubiquitin-dependent proteasome (McCracken and Brodsky, 2003). Retention of misfolded proteins can also occur at the ER-Golgi intermediate compartment (ERGIC) and cis-Golgi (Gilbert et al., 1998). An R760Q HS mutant of AE1, as an example, is retained in a pre-medial Golgi compartment (Quilty and Reithmeier, 2000).

1.4.4 Stability and degradation
AE1 in murine splenocytes traffics from the ER to Golgi with a half-time of 20 to 30 minutes and arrives at the cell surface in 30 to 40 minutes (Braell and Lodish, 1981), yet AE1 transfected in HEK-293 migrates more slowly from the ER to plasma membrane with a half-time of four hours (Li et al., 2000). This is likely due to the absence of red cell specific proteins like GPA and ankyrin in HEK cells. AE1 shows some intracellular retention with ~50% in the cell surface of the murine splenocyte (Braell and Lodish, 1981). In transfected HEK-293 cells about 30% of AE1 is found at the cell surface (Quilty and Reithmeier, 2000), although the majority of AE1 expressed in HEK-293 is in a high-mannose form (Li et al., 2000). In red cell development any ER- and Golgi-retained proteins would be destroyed during enucleation and elimination of the ER and Golgi before the reticulocyte stage (Quilty and Reithmeier, 2000). AE1 may increase its stability by interaction with GPA and cytoskeletal proteins such as ankyrin and protein 4.2 during biosynthesis in the ER and in the Golgi (Gomez and Morgans, 1993). AE1 present in the mature red cells is very stable and lasts the life-time of the cell (~120 days).
Mature AE1 transfected in HEK-293 and COS-7 cells is stable with a surface half-life of ~15 hours (Li et al., 2000; Quilty and Reithmeier, 2000). AE1 stability in transfected HEK-293 cells is not enhanced by introducing proteasome inhibitors, clasto-Lactacystin β-lactone (MG-262), suggesting that AE1 degradation may not occur via the proteasome (Quilty and Reithmeier, 2000). The turnover and stability of AE1 in erythroid precursors has not been studied in any great detail. However, a selective loss of chaperones, CNX, and ERp57 during late red cell development in cultured CD34+ human cells (Patterson et al., 2009) suggests that these cells lack an intact and robust quality control system during late stage erythroid development. This may allow massive synthesis of the red cell glycoproteins like AE1 that occurs late in differentiation.

1.4.5 Expression systems

Expression systems developed for AE1 include mammalian cells (such as human embryonic kidney cells (HEK)-293, COS-7 (African green-monkey kidney) and K562 (erythroleukemia cells), Xenopus oocytes, chicken embryo fibroblasts, QT6 quail cells, Spodoptera frugiperda (Sf9) insect cells, and yeast (Reithmeier, 1996b). HEK-293 cells are widely used systems for transiently expressing human and mouse anion exchangers by cDNA transfection. As HEK-293 cells show low endogenous anion exchange activity (likely due to the ubiquitously expressed AE2), they are thereby suited for testing the functions of AE1. Human AE1 can be functionally expressed at high levels (Fujinaga et al., 1999; Sterling and Casey, 1999) and processed to the cell surface (Li et al., 2000), though the expressed AE1 retains mainly a high mannose oligosaccharide form even at the cell surface (Li et al., 2000). This tendency was also observed in human AE1 expressed in COS-7 cells (Li et al., 2000) and mouse AE1 expressed in HEK-293 cells (Ruetz et al., 1993). Processing of the N-glycan to a complex form was restored in HEK cells when altering the glycosylation acceptor site from position Asn\(^{642}\) in EC loop 4 to Asn555 in EC loop 3 (Li et al., 2000), implying that the lack of oligosaccharide processing is not due to retention in the ER and that the Asn\(^{642}\) site may be relatively inaccessible for glycosylation processing enzymes. HEK cells expressing AE1 have a more rounded morphology and are easily removed from the surface of culture dishes. Stable AE1 expression in the HEK-293 system has not been attained, but an inducible
AE1 expression system is available (Timmer and Gunn, 1999). This inducible system can express functional AE1 in the plasma membrane using an inducible promotor element but has not been widely employed due to low expression levels.

Human or chicken AE1 is also functionally expressed to high levels in human K562 erythroleukemia cells using the pBabe reteoviral vector transfection system (Beckmann et al., 1998). K562 is thought to originate from erythrocyte precursor cells and lacks expression of endogenous AE1 and major erythroid surface antigens such as ABH, Wr, Fy3, and LW. This cell line contains endogenous glycophorin A, and blood group antigens, RhGP, and CD47, allowing for heterologous expression of AE1 and is suited for investigating the biosynthesis of AE1 (Beckmann et al., 1998). Human and mouse AE1 can be functionally expressed in *Xenopus* oocytes using cytoplasmic mRNA injection (Bartel et al., 1989; Garcia and Lodish, 1989; Groves and Tanner, 1992). An estimated 10% of AE1 expressed in the cells traffic to the cell surface, and co-expression with glycophorin A can improve the cell surface expression (Groves and Tanner, 1992). Chicken AE1 can be expressed in chicken embryo fibroblasts and QT6 quail cells using cDNA via an avian retrovirus vector SFCV-BIIIR (Fuerstenberg et al., 1990). Human AE1 can be functionally expressed in the plasma membrane of baculovirus infected Sf9 insect cells at a high level with ~ 0.5 x10^6 molecules/cell (Dale et al., 1996). However, the AE1 is not glycosylated in this expression system as its SDS-PAGE mobility corresponds to PNGaseF-treated AE1 (Dale et al., 1996).

Human AE1 (amino acid residues 183-911) with a C-terminal hexahistidine purification tag can be expressed and purified from yeast, *Saccharomyces cerevisiae* (*S. cerevisiae*) (Sekler et al., 1995). About 0.7 mg AE1/liter of culture was obtained from the expression system; however AE1 was not glycosylated and the majority was retained subcellularly. Human mdAE1 (amino acid residues 361-911) expressed in a protease-deficient strain, FKY282, of *S. cerevisiae* under the control of the inducible GAL10-CYC1 hybrid promotor showed that 5 to 10% of expressed AE1 was functionally expressed and targeted to the plasma membrane (Groves et al., 1996). Currently, AE1MD-Rho, a recombinant AE1 (residues 388-911) with a C-terminal rhodopsin (Rho) epitope cloned into the YEpM vector was expressed at a high level and purified with 93% purity from a protease-deficient strain, BJ5457, of *S. cerevisiae* (Bonar and Casey, 2010).
The AE1 was structurally and functionally indistinguishable from the erythrocyte AE1, yet it was retained in the ER and other intracellular compartments.

kAE1 can be expressed in the basolateral membrane of polarized cells, like Madin-Darby canine kidney (MDCK) cells and porcine kidney epithelial cells (LLC-PK1 kidney) using a retroviral-derived expression system and cDNA transfection, respectively (Kittanakom et al., 2004; Cordat et al., 2006). Tyr-targeting motifs \( ^{359} \text{YKGL}^{362} \) and \( ^{904} \text{YDEV}^{907} \) identified at the N-terminus and C-terminus, respectively of kAE1 are required for localizing to the basolateral membrane of these polarized cells (Williamson et al., 2008). Polarized MDCK cells are a fine cellular model system for polarized \( \alpha \)-intercalated cells of the kidney (Cordat et al., 2006). kAE1 traffics exclusively to the basolateral membrane as a complex oligosaccharide form with a half-time \( \sim 2 \) hours (Kittanakom et al., 2004). Although these systems supply powerful tools to study the functional expression of normal and mutant forms of polarized kAE1, the over-expression of glycoproteins often overwhelms the secretory pathway resulting in the build-up of immature high-mannose forms of glycoproteins in the ER (Kittanakom et al., 2004; Cordat et al., 2006).

AE1, like most membrane proteins, can be synthesized using a cell-free translational system by supplementing rabbit reticulocyte lysates with microsomal membrane vesicles prepared from canine pancreas (Popov et al., 1999). The system was used for scanning N-glycosylation studies (Popov et al., 1999).

Soluble cdAE1 (1-379) was expressed and purified from \( \text{E. coli} \) (T7 expression system with either low or high levels of the T7 polymerase inhibitor, T7 lysozyme) as a stable dimer with the same Stokes radius as erythroid cdAE1 (Wang et al., 1992). cdAE1 expressed in this system also retained the capacity to bind ankyrin (also see section 1.3.3.). The \( \text{E. coli} \) expression system was used to produce sufficient quantities of cdAE1 and cdkAE1 for comparative structural studies including NMR (Pang et al., 2008).

### 1.5 Project focus and hypothesis

The seven Trp residues in mdAE1 are conserved across mammalian species and highly conserved in the SLC4 family of anion exchangers. Two Trp residues (Trp\(^{662}\) and Trp\(^{723}\)) are located at the membrane interface; Trp\(^{492}\) and Trp\(^{496}\) are on the same side of TM 4,
likely facing the inhibitor binding site; Trp<sup>648</sup> is located in an extracytosolic loop close to the N-glycosylation site (Asn<sup>642</sup>); Trp<sup>831</sup> is in a hydrophilic loop; and Trp<sup>848</sup> is in TM11. These Trp residues may play a crucial role in the positioning of TM segments during biosynthesis, in the ion translocation pathway, and in forming the inhibitor binding sites of AE1. We hypothesize that Trp residues play an essential role in the folding and functional expression of human AE1. This hypothesis was tested using transient expression of wild type (Wt) and Trp mutations of AE1 and mdAE1 in HEK-293 cells and were studied as follows:

1. **Expression**
   
   Expression levels of AE1 were measured using immunoblotting and quantified.

2. **N-glycosylation processing**
   
   N-glycosylation processing of mdAE1 was used to determine the extent of AE1 trafficking from the ER to the Golgi.

3. **Immunolocalization**
   
   AE1 was localized within cells by immunofluorescence and confocal microscopy.

4. **Cell surface expression**
   
   Cell surface expression was measured using the cell surface biotinylation assay.

5. **AE1 inhibitor binding**
   
   The ability of AE1 mutants to bind to a SITS inhibitor affinity resin was investigated.

6. **Anion transport**
   
   Cl<sup>-</sup>/HCO<sub>3</sub><sup>-</sup> transport was performed by collaborators (Dr. Joseph Casey laboratory in University of Alberta, Canada)
Chapter 2
Materials and methods

2.1. Materials
The following is a list of materials used and their suppliers: pcDNA3 vector, DH5α competent cells, Lipofectamine™ 2000 (Invitrogen, San Diego, CA, U.S.A.); QuickChange™ Site-directed Mutagenesis Kit, XL 1 blue competent cells (Stratagene, La Jolla, CA); DNA primers (ACGT Corp., Toronto, Canada); QIAGEN Plasmid Mini Kit, QIAGEN Plasmid Midi Kit, Ni-NTA agarose (Qiagen Inc, Mississauga, Ontario); Dulbecco’s modified Eagle medium (DMEM), calfserum, penicillin and streptomycin (Gibco BRL, Burlington, ON, Canada); chemiluminescence kit (Boehringer Mannheim, Laval, QC, Canada); EZ-Link™ sulfo-succinimidyl 2-(biotin-amido)ethyl-1,2-dithiopropionate (NHS-SS-biotin), ImmunoPure® immobilized streptavidin (Pierce Biotechnology, Inc., Rockford, IL, U.S.A.). SITS-Affi-Gel was prepared as described previously (Pimplikar and Reithmeier, 1986). C-terminal AE1 mouse monoclonal antibody was a generous gift from Dr. Michael Jennings. Mouse anti-GAPDH antibody (Chemicon International Inc., Temecula, CA, U.S.A.). Rabbit anti-CNX was a generous gift from Dr. David Williams Lab (University of Toronto, Canada). Cy3-conjugated anti-rabbit antibody (Jackson Immunoresearch Laboratories, West Grove, PA, U.S.A.); Alexa 488-conjugated anti-mouse antibody (Molecular Probes, Eugene, OR, U.S.A.). The extracellular AE1 mouse monoclonal antibody, BRIC 6, was purchased from the International Blood Group Reference Laboratory, University of Bristol, England. Octa (ethylene glycol) dodecyl ether (C₁₂E₈; Nikkol Chemical Co, Tokyo, Japan); goat peroxidase-conjugated anti-rabbit IgG, endoglycosidase-H, and N-glycanase F (New England BioLab). N-glycanase F was also a generous gift from Dr. James Rini Lab (University of Toronto, Canada).

2.2. Site-directed mutagenesis
cDNA encoding human erythrocyte AE1 (a gift from Drs. A.M. Garcia and H. Lodish, Whitehead Institute) was inserted into the Xhol (mutated from initial HindIII site) and BamHI sites of the pcDNA3 vector (Popov et al., 1999). PCR was used to construct a
plasmid encoding the membrane domain of AE1 (mdAE1) contains the sequence of erythrocyte AE1 encoding amino acids Asp$^{369}$-Val$^{911}$ with methionine start site, followed by subcloning into pcDNA3 (Bustos and Reithmeier, 2011). The C-termini of AE1 and mdAE1 in pcDNA3 were tagged with hexahistidine (His$_6$) by PCR (performed by Li, J). AE1 with extracellular hemagglutinin (HA)-tag was constructed by PCR, inserting the HA epitope into position 557 in the 3rd EC loop (performed by Li, J). The oligonucleotide primers (appendix), designed to mutate seven Trp residues in AE1 to Ala and Phe residues were purchased from ACGT Corp (Toronto, Canada). Trp mutants in full length AE1 and in the membrane domain with a C-terminal His$_6$-tag in expression vector pcDNA3 were constructed using the Quikchange Site-directed Mutagenesis Kit (Strategene). The plasmids were transformed into DH5$\alpha$ or XL 1 blue competent cells, and purified with DNA extraction kit (QIAGEN, Valencia). The constructs were sequenced by ACGT Corp (Toronto, Canada) to confirm the mutations using sequencing primers (appendix).

2.3. Transient transfection and expression in HEK-293 cells

Wt and mutant AE1 were transiently expressed in human embryonic kidney (HEK)-293 cells as described previously (Popov et al., 1999) using 1 µg plasmid DNA and 2.5 µl of Lipofectamine™ 2000 (Invitrogen) (Sells et al., 1995). The cells (0.7x10$^6$ cells/mL) were cultured in 6-well plates (Falcon, USA) with 10% (v/v) fetal bovine serum (FBS) supplemented Dulbecco’s Modified Eagle Medium (DMEM; Gibco Life Technologies, USA) at 37°C in 5% CO$_2$ for 24 to 48 hours and then harvested by scraping. Whole cell detergent extracts were prepared using 1% C$_{12}$E$_8$ in PBS solution containing protease inhibitors (1mM PMSF, 1% leupeptin, 1% pepstatin A, and 1% aprotinin). Samples were centrifuged at 14000 x g for 30 min to remove insoluble material, and 2 x SDS sample buffer was added at room temperature. Samples (50 µg of total protein per lane) were loaded on SDS-PAGE, and proteins were detected by Western blotting.

2.4. Electrophoresis and immunoblotting

Samples were resolved in 8% polyacrylamide gels (Laemmli, 1970) and transferred to nitrocellulose (Towbin et al., 1979). AE1 and GAPDH were detected using mouse monoclonal antibodies (anti-AE1, 1:2000 and anti-GAPDH, 1:20000 dilution,
respectively). The membranes were treated with 5% (w/v) skim milk powder and antibodies in TBS-T (137 mM NaCl, 20 mM Tris, pH 7.5 with 0.1% Tween-20) buffer. The blot was further incubated with HRP-conjugated goat anti-mouse IgG (1:5000 dilution) in TBS-T. The blot was developed with chemiluminescence reagent (Perkin Elmer) and by exposed to BioMax MR imaging film (Kodak) for various times. Relative protein expression levels were determined by densitometric analysis using ImageJ 1.42 software (National Institutes of Health). Total AE1 expression in the Western blot was quantified by calculating the densities of bands using image J, normalized to GAPDH for protein loading. The mean values ± S.D were calculated from three independent experiments.

2.5. Immunofluorescence and microscopy

The cell surface expression and intracellular localization of AE1 were detected using immunofluorescence and confocal microscopy. HEK-293 cells were cultured on poly-L-lysine-coated coverslips in 6-well plates. Transiently transfected cells were fixed with 3.7% paraformaldehyde for 10 minutes. The cells were blocked with 3% bovine serum albumin (BSA) and incubated with 1:500 diluted mouse monoclonal antibody, BRIC 6, which detects an extracellular epitope on AE1, in phosphate-buffered saline for 30 minutes at room temperature, followed by Alexa488 conjugated goat anti-mouse antibody (1:1000 dilution) (Molecular Probes, Eugene, OR). The fixed cells were washed with PBS buffer that contains 1 mM CaCl2 and 1mM MgCl2, and then permeabilized in 0.1% Triton X-100, blocking with 3% BSA, and incubated with BRIC 6 and rabbit anti-calnexin (CNX)(1:1000 dilution). The cells were further incubated with 1:1000 dilutions of Cy3-conjugated anti-mouse IgG and donkey anti-rabbit Cy5 (Jackson ImmunoResearch). The coverslips were mounted with DakoCytomation fluorescent mounting media before being observed using a laser scanning confocal Zeiss LSM 510 microscope.

2.6. Cell surface biotinylation

Cells grown on 6-well plates were washed with borate buffer (10mM boric acid, 154mM NaCl, 7.2 mM KCl, 1.8 mM CaCl2, pH9.0). The extracellular surface of cells was biotinylated twice with 1 ml of 0.5 mg/ml NHS-SS-biotin (Pierce) in borate buffer for 15
min at room temperature. The reagent was quenched by rinsing the cells with 0.192 M glycine. Total-cell lysates were prepared using RIPA buffer (1% deoxycholic acid, 1% Triton X-100, 0.1% SDS, 0.15 M NaCl, 1 mM EDTA, 10 mM Tris-HCl, pH7.5) with protease inhibitors (1 mM PMSF, 1% leupeptin, 1% pepstatin A, and 1% aprotinin). Mock-treatment was done on cells transfected with Wt AE1 with a C-terminal His₆-tag using the same conditions in borate buffer but without the biotinylation reagent. A fraction of the lysate was used for immuno-blots for total expression (T). Biotinylated proteins were captured with ImmunoPure (Pierce) streptavidin-agarose beads (100 µl) incubated for 2 hours at 4 °C. An aliquot of unbound supernatant (U) was saved for immuno-blots. The bound fraction (B) was eluted with 2 X SDS sample buffer containing 5% β-mecaptoethanol and incubated for 1 hour at room temperature to reduce disulfide bonds. Fractions were analyzed using 8% SDS-PAGE and followed by immuno-blotting with the mouse anti-AE1 monoclonal antibody. A mouse monoclonal anti-GAPDH was used for detecting intracellular GAPDH to confirm that biotinylation occurred only at the cell surface. The mean values ± S.D were calculated from three independent experiments.

2.7. Inhibitor (SITS) binding

The ability of AE1 mutants to bind to a SITS inhibitor affinity resin (SITS-Affi-Gel) was measured as described previously (Pimplikar and Reithmeier, 1986). Cells grown on 6-well plates were transfected with cDNA encoding erythrocyte AE1 and the Trp mutants. Cell extracts were lysed in 1% C₁₂E₈ PBS solution containing protease inhibitors (1 mM PMSF, 1% leupeptin, 1% pepstatin A, and 1% aprotinin) and centrifugated at 14000 x g for 30 min to remove insoluble material. Cell extracts were equally divided into 3 fractions for total AE1 expression (T), SITS binding (B), and non-specific binding in the presence of 1mM DIDS (D). A 50% slurry of SITS-beads was prepared in 228 mM sodium citrate buffer (pH 7.1). Lysates (100 µl) were incubated with 100 µl SITS-Affi-Gel with or without 1mM 4,4’-diisothiocyanostilbene-2,2’-disulfonic acid (H₂DIDS) in 228 mM sodium citrate buffer (pH 7.1) for 30 min at 4 °C. The unbound fraction was removed, and beads were washed with 228 mM sodium citrate buffer. The bound proteins on the SITS-Affi-Gel beads were eluted with 75 µL 2 × sample buffer. Samples were analyzed by 8% SDS-PAGE and immunoblotting. The intensities of bands were
quantified by densitometric analysis using ImageJ 1.42q software. As final elutants were twice concentrated compared to total AE1 expression in blots, percent binding was calculated as followed equation.

\[
\frac{[\text{Binding (B)} - \text{Non-specific binding (D)}]}{2 \times \text{Total AE1 (T)}} \times 100\%
\]

The mean values ± S.D of percent SITS-binding were calculated from three independent experiments.

2.8. Transport assay

Chloride and bicarbonate exchanger activity was tested as previously described (Fujinaga et al., 1999) in the laboratory of Dr. J.R. Casey (University of Alberta). Cells were grown on top of 7 x 11 mm glass cover slips in 60 mm tissue culture dishes at 37 °C in 5% CO₂ and transfected with 1.6 µg of calcium-precipitated DNA with 4.2 µg of pRBG4 carrier. Forty-eight hours post-transfection, coverslips were rinsed with serum-free DMEM and incubated with 4 ml DMEM and pH-sensitive dye, BCECF-AM (2µM), for 20-30 min at 37 °C. Coverslips were mounted in a custom-built quartz cuvette, with perfusion capabilities. Intracellular pH was measured by fluorescence of BCECF-AM using excitation wavelengths (440 and 502 nm) and emission wavelength (529nm) in a Photon Technologies International RPC spectrofluorometer. Cuvettes were alternatively perfused at 3.5 ml/min with Ringer’s buffer (5 mM glucose, 5mM potassium gluconate, 1mM MgSO₄, 2.5mM NaH₂SO₄, 2.5mM NaHCO₃, 10mM Hepes, pH7.4) containing 140 mM sodium chloride buffer, followed by chloride-free Ringer’s buffer (5 mM glucose, 5mM potassium gluconate, 1mM MgSO₄, 2.5mM NaH₂SO₄, 2.5mM NaHCO₃, 10mM Hepes, 140 mM sodium gluconate, pH7.4). Both buffers were bubbled with air containing 5% carbon dioxide. Intracellular pH was calibrated using the nigericin-potassium method (Thomas et al., 1979), with pH values from pH 6.5 to 7.5. Transport activities were determined using linear regression of initial linear rate of change of pH by Synergy Software, the Kaleidagraph program.

2.9. Purification

To clearly visualize the N-glycosylation pattern of mdAE1 and Trp mutants, Ni-NTA purification was performed to enrich the protein concentration. Cells grown on 6-well
plates were transfected with cDNA encoding Wt mdAE1 and its Trp mutants with a C-terminal His6-tag. Transfected cells were grown for 48 hours and then harvested using 0.5 ml of PBS buffer containing 1% C12E8, 300 mM NaCl, and 10 mM imidazole in PBS with protease inhibitors (1mM PMSF, 1% leupeptin, 1% pepstatin A, and 1% aprotinin). Cell lysates were centrifuged at 14000 x g for 30 min at 4 °C to remove insoluble material. Cell extracts (500 µl) were incubated with 50 µl of a 50% slurry of Ni-NTA-agarose in binding buffer (1% C12E8, 300mM NaCl, and 10mM imidazole with protease inhibitors in PBS) for 2 hours at 4 °C. The unbound fraction was removed and the beads were washed with 1 ml of washing buffer (0.1% C12E8, 300mM NaCl, and 30mM imidazole with the protease inhibitors in PBS). The bound fraction was eluted with elution buffer (0.1% C12E8, 300mM NaCl, and 500 mM imidazole in PBS). Samples were solubilized with an equal volume of 2 X SDS sample buffer, and each fraction was analyzed by 8% SDS-PAGE and immuno-blotted as described above. Bands intensities were measured using ImageJ 1.42q software. The percent complex bands were determined as described below. The mean values ± S.D were calculated from three independent experiments.

2.10. N-glycosylation analysis

Cells grown on 6-well plates were transfected with cDNA encoding erythrocyte AE1, kAE1, and mdAE1 and incubated for 48 hours. Cell were solubilized in 1% C12E8 PBS solution containing protease inhibitors (1mM PMSF, 1% leupeptin, 1% pepstatin A, and 1% aprotinin) followed by centrifugation at 14000 x g for 30 minutes at 4 °C. Total cell-extracts (100 µl) were treated with 1000 units of endoglycosidase-H (endoH ) or 500 units of N-glycanase F (PNGaseF) at 37°C for 1 hour. Samples were analyzed by 8% SDS-PAGE and immuno-blotting as described above. Bands intensities were measured using ImageJ 1.42q software. Percent complex oligosaccharide form was quantified from endoH treated samples by dividing intensities of upper multiple bands by the intensity of all bands in the lane. The endoH treated samples were used to identify the high mannose component and to better resolve the complex bands from the high mannose band. The mean values ± S.D were calculated from three independent experiments.
3.1. Expression and stability of Trp mutants

Wt and Trp mutants of AE1 were expressed in HEK-293 cells in order to determine the effect of mutation of conserved Trp residues on the expression and trafficking of AE1. The seven endogenous Trp residues (492, 496, 648, 662, 723, 831, and 848) located in the membrane domain of AE1 were individually mutated to Ala or Phe in a full-length AE1 construct. A C-terminal His<sub>6</sub>-tag in these constructs was introduced to facilitate protein purification. Cells grown on 6-well plates were transfected in parallel with identical amounts of DNA in multiple transfection experiments using optimized procedures developed in the Reithmeier lab (Quilty and Reithmeier, 2000).

Although we did not routinely test the transfection efficiency during investigation of the expression of AE1 and Trp mutants, the mutations we made were point substitutions of Trp to either Ala or Phe and, therefore, it is not expected that the transfection efficiency of using the vectors would vary. The transfection methods have been optimized for maximal expression of AE1 in transfected HEK cells with typically greater than 50% of cells being transfected as determined by immunofluorescence staining (Quilty and Reithmeier, 2000). The intensity of staining does, however, vary from cell to cell. Protein expression mainly depends on the growth and final number of cells rather than on transfection efficiency, therefore we repeated the experiments multiple times in parallel transfections and normalized using GAPDH as an internal control for protein expression during these experiments.

Figure 15 shows immuno-blots of total-cell extracts prepared in 1% C<sub>12</sub>E<sub>8</sub> from cells transfected with Wt AE1 and the 14 Trp mutants. AE1 and mutant proteins were detected as a single band at approximate 95kDa in molecular mass. Previous results (Li et al., 2000) have shown this band contains high-mannose oligosaccharide, although AE1 can traffic to the cell surface in HEK cells. No AE1 band was observed in cells transfected with empty vector, pcDNA3 (Figure 15 A, B, lane 1), showing the specificity of the mouse monoclonal anti-AE1 antibody. The same blots were reprobed with anti-GAPDH antibody in order to account for differences in protein loading. As can be seen
Figure 15. Immunoblot analysis of C-terminal His6-tagged erythrocyte (e)AE1 and tryptophan mutants expressed in transfected HEK-293 cells. Cells grown in 6-well plates were transfected with 1 µg cDNA per well and were solubilized in 1% (w/v) C12E8 after 48 hours growth. Total cell proteins were resolved by SDS-PAGE and transferred to a
nitrocellulose membrane. A mouse monoclonal anti-AE1 antibody was used to detect AE1 (*upper panel*) and a mouse monoclonal anti-GAPDH antibody was used to detect endogenous GAPDH (*lower panel*), as a protein loading control. The positions of molecular weight markers are shown on the left. Panel (A) shows the expression levels of Ala substitutions and panel (B) shows Phe substitutions in eAE1. Lane 1 is control transfection with empty vector, pcDNA3. Panel (C) shows protein expression of the Trp mutants normalized to GAPDH expression (*dark bar*: Ala substitutions, *light gray bar*: Phe substitutions) in three independent transfection experiments. The percent expression of Wt AE1 was set to 100. The *error bars* show the S.D. with n = 3. Mutants W831A/F, W848A/F, W648F, W723F had Wt expression levels (p > 0.01), while all other mutants had lower expression levels (p < 0.005).
from these blots, the expression levels of the W831A/F, W848A/F, W648F, W723F were similar to Wt AE1, while the expression levels of the other mutants were diminished.

The percent expression levels of the Trp mutants relative to the Wt AE1, normalized to GAPDH, in three separate transfection experiments are shown in (Figure 15 C). A reduction of AE1 expression was observed with Ala substitutions at positions 492, 496, 648, 662, and 723 and Phe substitutions at 492, 496, with some reduction at position 662. No significant reduction of AE1 expression was observed with Ala substitutions at positions 831 and 848, and Phe substitutions at positions 648, 723, 831 and 848. Thus, mutations of Trp residues 831 or 848 had little effect on AE1 expression, while mutation of residues 492 or 496 reduced expression. Interestingly, mutations of residues 648 or 723 to Phe had little effect on expression, while mutation to the smaller amino acid, Ala, decreased expression.

3.2. Cell surface expression of Trp mutants

The localization of C-terminal His$_6$-tagged AE1 and Trp mutants in HEK-293 cells was detected by immunofluorescence and confocal microscopy. Previous studies have shown that Wt AE1 expressed in HEK-293 cells trafficked to the cell surface with about 1/3 of the protein at the plasma membrane and the remainder in the ER and secretory pathway (Gomez and Morgans, 1993; Li et al., 2000). The present study was performed with a C-terminal His$_6$-tagged version of Wt and Trp mutants of AE1. Cells cultured on cover slips in 6-well plates were transfected in parallel with identical amounts of DNA. The intact cells were stained with a mouse monoclonal antibody, BRIC 6, which detects an extracellular epitope on AE1 followed by detection with Alexa488-conjugated anti-mouse 2' antibody (green), indicating cell surface expression of AE1. These cells were washed and subjected to detergent permeabilization and incubation with the antibody, BRIC 6, followed by detection with Cy3-conjugated anti-mouse 2' antibody to show total AE1 expression (red). These images were merged to show AE1 cell surface distribution in yellow and intracellular AE1 in red. Fields of cells showing the strongest intensities were selected and visualized using the same confocal microscopy settings in order to determine the localization of poorly expressed mutants.
The images in Figure 16 show the localization of Wt AE1 and Trp mutants in transfected HEK-293 cells. Panel (Figure 16 A) shows Wt AE1 is expressed at the plasma membrane (yellow) but is also localized intracellularly as in previous studies (Gomez and Morgans, 1993; Li et al., 2000). Panel (Figure 16 B) shows that not all Trp mutants are expressed on the cell surface. All cells, however, showed at an intracellular pool of AE1 indicating the all mutants were expressed, some at low levels. Trp mutants, W831A/F, W848A/F, W648F, and W723F, were localized at the plasma membrane, while Trp mutants, W492A/F and W496A/F were not detected at the cell surface. W648A, W662A, and W723A were poorly expressed at the cell surface. The results indicate that mutations of Trp residues 831 or 848 to Phe or Ala had little effect on AE1 expression at the cell surface, while mutation of residues 492 or 496 diminished cell surface expression. Mutations of residues 648, 662, and 723 to Phe had little effect, while mutations to the smaller amino acid, Ala, decreased AE1 expression at the cell surface.

Cell surface biotinylation was conducted to quantify the amount of cell surface expression of the Trp mutants. C-terminal His<sub>6</sub>-tagged AE1 and Trp mutants were expressed in HEK-293 cells in parallel with identical amounts of DNA in multiple separate transfection experiments. Cells were biotinylated with NHS-SS-biotin in 6-well cell culture plates, and total-cell lysates were prepared using RIPA buffer. Transfected Wt AE1 was mock-treated under the same conditions in borate buffer without the biotinylation reagent. Biotinylated proteins were captured with streptavidin-agarose beads and AE1 was detected by immunoblotting.

Figure 17 shows the total-cell extract (lane T), the unbiotinylated fraction (lane U), and the biotinylated streptavidin bead bound fraction (lane B) of eAE1 and Trp mutants expressed in HEK-293 cells. GAPDH, an endogenous cytosolic protein, was used to confirm that biotinylation only occurred at the cell surface and that the reagent did not penetrate the cell under the biotinylation conditions used. The mock-treated cells showed that only biotinylated AE1 was able to bind to streptavidin beads (lane B), and that the biotinylated protein binding to streptavidin bead was specific (Figure 17 A). The biotinylation conditions caused AE1 to aggregate and run as monomeric (open square) and higher oligomeric (closed square) forms, even in the absence of biotinylation.
Figure 16. Localization of C-terminal His6-tagged erythrocyte (e)AE1 (A) and tryptophan mutants (B) expressed in HEK-293 cells by immunofluorescence using a LSM 510 confocal microscope. Intact cells expressing Wt (A) and tryptophan mutants (B, left: Ala substitution, right: Phe substitution) of eAE1, were stained by a mouse monoclonal antibody (BRIC 6), which detects an extracellular epitope on AE1. The cell surface expression was visualized as green by secondary Alex 488 conjugated goat anti-mouse IgG in non-permeabilized cells (column 1). Total AE1 expression was visualized as red by secondary Cy3-conjugated goat anti-mouse IgG after cells were permeabilized (column 2). AE1 on the cell surface was displayed in yellow by merging the cell surface and total expression of AE1 (column 3). The mutants W831A/F, W848A/F, W648F, W662F, and W723F showed detectable cell surface expression as indicated by yellow in
the merged images, while the other mutants did not (W662A shows week yellow in the merged image). Wt AE1 and all mutants showed intracellular localization (red).
(A) Control | Wt
---|---
MN (X1000) & T & U & B & T & U & B
170 & & & & & &
130 & & & & & &
90 & & & & & &
50 & & & & & &
35 & & & & & &
---
(eAE1) & & & & & &
(GAPDH) & & & & & &

(B)

<table>
<thead>
<tr>
<th>Wt</th>
<th>W492F</th>
<th>W496F</th>
<th>W631F</th>
<th>W648F</th>
<th>W492A</th>
<th>W723A</th>
</tr>
</thead>
</table>
MN (X1000) & T & U & B & T & U & B & T & U & B & T & U & B & T & U & B
170 & & & & & & & & & & & &
130 & & & & & & & & & & & &
90 & & & & & & & & & & & &
50 & & & & & & & & & & & &
35 & & & & & & & & & & & &
---
(eAE1) & & & & & & & & & & & &
(GAPDH) & & & & & & & & & & & &

<table>
<thead>
<tr>
<th>W648F</th>
<th>W662F</th>
<th>W723F</th>
<th>W496A</th>
<th>W648A</th>
<th>W662A</th>
</tr>
</thead>
</table>
MN (X1000) & T & U & B & T & U & B & T & U & B & T & U & B & T & U & B
170 & & & & & & & & & & & &
130 & & & & & & & & & & & &
90 & & & & & & & & & & & &
50 & & & & & & & & & & & &
35 & & & & & & & & & & & &
---
(eAE1) & & & & & & & & & & & &
(GAPDH) & & & & & & & & & & & &

(C)

% Biodiversity

<table>
<thead>
<tr>
<th>Wt</th>
<th>492</th>
<th>496</th>
<th>648</th>
<th>662</th>
<th>723</th>
<th>831</th>
<th>848</th>
</tr>
</thead>
</table>
Wt & & & & & & & &
492 & & & & & & & &
496 & & & & & & & &
648 & & & & & & & &
662 & & & & & & & &
723 & & & & & & & &
831 & & & & & & & &
848 & & & & & & & &

---
(A) Ala

---
Phe
Figure 17. Cell surface biotinylation of C-terminal His$_6$-tagged erythrocyte AE1 and tryptophan mutants expressed in HEK-293 cells. Cells cultured in 6-well plates were transfected with 1 μg cDNA per well and the cells were biotinylated with 0.5mg/ml NHS-SS-biotin reagent in borate buffer for 30 minutes. Whole-cell lysates solubilized in RIPA buffer were incubated with streptavidin-agarose beads to capture biotinylated proteins. Total AE1 (lane T), unbound to streptavidin beads (lane U), and bound AE1 at five times the amount loaded relative to total AE1 (lane B) were detected by immuno-blotting performed as described in Figure 15, using a mouse monoclonal anti-AE1 antibody. Endogenous GAPDH was detected using a mouse monoclonal anti-GAPDH to confirm that biotinylation only occurred at the cell surface. Control (A) represents mock-treated Wt AE1 cells without biotinylation reagent. Monomeric (open square) and oligomeric forms (closed square) of eAE1 were indicated (A, B). Panel (C) shows percent cell surface biotinylation of Wt AE1 and the Trp mutants (dark bar: Ala substitutions, light gray bar: Phe substitutions) in three independent transfection experiments. Monomeric and higher oligomeric bands were included in the quantification. The error bars show the S.D. with n = 3. The W831A/F, W848A/F, W648F, W662F, and W723F mutants showed biotinylation level similar to Wt AE1 (p > .05), while the other mutants have a lower biotinylation level (p < .05).
reagent. This aggregation effect was seen in previous biotinylation experiments (Li et al., 2000; Quilty and Reithmeier, 2000).

Biotinylated Wt AE1 and Trp mutants were capable of binding to streptavidin beads (Figure 17 B), although, they show different levels of cell surface biotinylation. Cell surface biotinylation of Wt AE1 and Trp mutants results in monomeric (open square) and higher oligomeric (closed square) forms in immunoblots. None of the samples showed GAPDH in the bound fraction (lane B). W831A/F, W848A/F, W648F, W662F were biotinylated at high levels, clearly showing a monomeric form, while the other mutants were poorly biotinylated and ran mainly as a higher oligomeric form. Wt AE1 was biotinylated to a level of approximately 30% (Figure 17 C), which was similar to previous experiments (Li et al., 2000; Quilty and Reithmeier, 2000). Trp mutants, W831A/F, W848A/F, W648F, W663F, and W723F were comparable to the Wt, with higher than 25% biotinylation. Both Ala and Phe substitutions at positions 492 and 496 showed little cell surface biotinylation. Ala substitutions at 648, 662, and 723 showed approximately half the level of biotinylation relative to Wt. The biotinylation results were consistent with cell surface expression detected using immunofluorescence and confocal microscopy.

3.3. Inhibitor binding of Trp mutants

The effect of mutating Trp residues in AE1 on inhibitor binding was determined. SITS and DIDS are inhibitors of anion transport and bind to an external site on AE1. AE1 solubilized in detergent can bind to immobilized SITS and this binding can be competed by other stilbene disulfonates such as DIDS. C-terminal His_{6}-tagged AE1 and its mutants were expressed in HEK-293 cells in parallel with identical amounts of DNA in multiple separate transfection experiments. The cells were grown on 6-well plates for 48 hours. Total-cell extracts in 1% C_{12}E_{8} were subjected to SITS-Affi-Gel binding that tests for properly folded AE1.

Figure 18 shows total cell expression (lane T) and SITS-binding fraction without DIDS (lane B) and with DIDS (lane D). We note that the expression levels of the W831A/F, W848A/F, W648F, and W723F mutants in the total fractions (lane T) were
Figure 18. SITS-Affi-Gel binding of C-terminal His<sub>6</sub>-tagged erythrocyte AE1 and tryptophan mutants expressed in HEK-293 cells. Cells grown in 6-well plates were transfected with 1 µg cDNA per well and were solubilized in 1% (w/v) C<sub>12</sub>E<sub>8</sub> after 48 hours growth. The total cell lysates (lane T) were incubated with SITS-Affi-Gel resin in the absence (lane U) or presence (lane B) of 1mM H<sub>2</sub>DIDS, a competitive inhibitor of SITS binding. Immuno-blot analysis was performed as described in Figure 2, using a
mouse monoclonal anti-AE1 antibody. Vector (A) is a control transfection with empty vector, pcDNA3. The positions of molecular weight markers are shown on the left. Panel (C) shows the percent SITS binding 100 x (B-D/T) of Wt AE1 and the Trp mutants (*dark bar*: Ala substitutions, *light gray bar*: Phe substitutions) in three independent transfection experiments. The *error bars* show the S.D. with n = 3. The mutants, W831A/F, W848A/F, W648F, W662F, and W723F, showed SITS binding level similar to Wt AE1 (p > .05), while the other mutants showed diminished binding (p < .05).
similar to Wt, while other mutants had a lower expression level as shown previously in Figure 15. The Wt AE1 bound to SITS-Affi-Gel, which was effectively blocked by 1mM H$_2$DIDS, a competitive inhibitor of SITS binding. The percent AE1 binding to SITS-Affi-Gel was calculated as 100 x (B-D/T) that binding fraction (lane B) was corrected by subtracting fraction D as the binding in the presence of DIDS is considered non-specific.

Under the conditions employed, Wt AE1 (Figure 18 A) and the Trp mutants, W831A/F, W848A/F, W648F, and W723F (Figure 18 B), were effectively bound to SITS-Affi-Gel (lane B) with a small amount of non-specific binding detected in lane D. SITS-Affi-Gel binding of other mutants, W492A/F, W496A/F, W648A, W662A, and W723A, showed reduced binding with approximately the same amount in the absence (Figure 18 B) and presence (Figure 18 D) of DIDS. Approximately 50% of Wt AE1 bound to SITS-Affi-Gel (Figure 18 C). The Trp mutants, W831A/F, W848A/F, and W723F were comparable to Wt AE1, while W492A/F and W496A/F resulted in less than 10% binding. Trp mutations at positions 648 and 662 showed approximate half the SITS binding capacity for Ala substitutions, and 22 and 45% reduction for Phe substitutions relative to Wt AE1. The results indicated that Wt AE1 and Trp mutants, W831F/A, W848F/A, W648F, W662F, and W723F, had similar binding behaviour (p > .05), suggesting they may have a native conformation. The other mutants had impaired SITS-binding (p < .05) and may possess a misfolded conformation.

3.4. Transport activity of Trp mutants

Figure 19 shows the rates of intracellular pH change in Wt AE1, C-terminal His$_6$-tagged AE1, C-terminal His$_6$-tagged membrane domain (md)AE1, and C-terminal His$_6$-tagged Trp mutants of eAE1 expressed in HEK-293 cells. Intracellular pH changes were used to measure AE1-mediated bicarbonate transport activity. Cuvettes were alternatively perfused with Ringer’s buffer containing sodium chloride, followed by chloride-free Ringer’s buffer. Removal of extracellular chloride creates an outward-facing gradient that allows extracellular bicarbonate exchange for intracellular chloride. This results in alkanization of the cytosol, which is measured using the pH-sensitive dye (BCECF-AM). Wt AE1 and pcDNA3 vectors were used as positive and negative controls, respectively. We used His$_6$-tagged AE1 in our studies to allow purification of the proteins.
Figure 19. Preliminary results of anion exchange rates of Wt erythrocyte AE1, C-terminal His$_6$-tagged AE1, C-terminal His$_6$-tagged membrane domain (md)AE1, and C-terminal His$_6$-tagged Trp mutants of AE1 expressed in HEK-293 cells. Transport assay was conducted as described in Materials and Methods (section 2.8). Cells grown on cover slips in 60 mm tissue culture dishes were transfected with 1.6 µg cDNA per well and incubated for 48 hours. The pH-sensitive dye BCECF-AM was loaded and intracellular pH was measured by fluorescence at excitation wavelengths (440 and 502 nm) and emission wavelength (529 nm). The figure shows the rate of intracellular pH change per. pcDNA vectors and Wt AE1 were used as negative and positive controls, respectively.
As this assay is quite time consuming and was carried out by collaboration, we initially selected a subset of mutants to test. As this transport assay requires proteins to be expressed at the cell surface, we selected Phe substitutions of Trp\textsuperscript{648} and Trp\textsuperscript{662} since their Ala substitutions were poorly expressed and poorly processed to the cell surface. Ala and Phe substitution at Trp\textsuperscript{723} were both tested to compare the effects of the two different mutations where the Ala substitution had a moderate effect on expression while the Phe mutants showed normal expression levels. We chose Ala substitutions of Trp\textsuperscript{831} and Trp\textsuperscript{848} as these mutations had normal expression levels like the Phe mutations. Mutants W492A/F and W496A/F were not tested due to their low expression levels. Transport rates were normalized to total expression levels.

The transport assays showed that all selected mutants, and the C-terminal His\textsubscript{6}-tagged AE1 and mdAE1, are active as compared to control cells transfected with empty vectors. The His\textsubscript{6}-tagged AE1 constructs may have a lower transport rate than untagged constructs perhaps due to less efficient trafficking to the plasma membrane. Previous studies using cysteine scanning mutagenesis (Tang \textit{et al.}, 1999) showed that W648C and W662C had transport activities less than 10\% of Wt AE1. Our results indicate that W648F and W662F had higher transport activities than the cysteine mutants. Cysteine scanning mutagenesis was not tested on Trp\textsuperscript{723}, though nearby residues Gly\textsuperscript{714} and Ser\textsuperscript{725} were identified as structurally and catalytically important residues in anion transport activity (Fujinaga \textit{et al.}, 1999). As the S.D is large in some samples, especially in C-terminal His\textsubscript{6}-tagged Wt AE1, the transport experiments need to be repeated and also tested for other mutants including W831F and W848F. Transport rates are normalized relative to total expression levels but are also dependent upon the level of cell surface expression.

3.5. Expression and trafficking of AE1, kAE1, and mdAE1

We showed C-terminal His\textsubscript{6}-tagged Wt AE1 and Trp mutants transiently expressed in HEK-293 cells contain mainly the high mannose oligosaccharide (Figure 15). This is consistent with previous results as human erythrocyte AE1 expressed in transiently transfected HEK-293 cells exists mostly in a high-mannose form (Li \textit{et al.}, 2000). In contrast, the kidney isoform is processed to a more complex oligosaccharide form.
compared to AE1 (Cordat et al., 2006), and a truncated mutant that only contains the C-terminal membrane domain of AE1 (mdAE1), encompassing residues Asp$^{369}$-Val$^{911}$, displays a complex oligosaccharide form in the highest quantity (Bustos and Reithmeier, 2011). It is advantageous to use the mdAE1 instead of erythrocyte AE1 as it can be used as a trafficking marker to trace its exit from the ER, by quantifying N-linked complex oligosaccharide forms. We examined the differences in oligosaccharide processing by quantifying their complex oligosaccharide percentages in AE1, kAE1, and mdAE1 as described below.

AE1, kAE1, and mdAE1 subcloned in pcDNA3 were expressed independently in HEK-293 cells with identical amounts of DNA in multiple separate transfection experiments. Figure 20 shows an immuno-blot of the total-cell extracts comparing the expression and N-glycosylation patterns of AE1 (lane 1-3), kAE1 (lane 4-6), and mdAE1 (lane 7-9). AE1 shows a single band at a molecular mass of approximately 95kDa (lane 1), while kAE1 in lane 4 (90kDa) and mdAE1 (55kDa) in lane 7 show a lower molecular weight band and multiple higher bands. The oligosaccharide content of these bands was distinguished by treating detergent lysates with glycosidase enzymes, endoH (H) and PNGaseF (F). The shift of bands to a lower molecular weight after endoH treatment (lane 2, 5, 8), indicates the presence of N-linked high mannose oligosaccharide (Popov et al., 1999). Disappearance of endoH insensitive multiple bands after PNGaseF treatment (lane 6, 8), which cleaves all N-linked glycans, demonstrates the presence of complex oligosaccharides (Popov et al., 1999).

The percentage of molecules carrying complex oligosaccharides was quantified by the ratios of the density of upper bands to that of total bands after endoH treated using immuno-blotting (Table 3). The percent complex oligosaccharides increased significantly in the order of AE1, kAE1, and mAE1 (Table 5), with AE1 containing little complex oligosaccharide, kAE1 containing about 40%, and mdAE1 greater than 50%. The results from Figure 20 indicate that differences in the structure of the N-terminal cytosolic domain affect oligosaccharide processing. The tags (extracellular HA, C-terminal His, and combination of the extracellular HA and the C-terminal His) had little effect on the processing of AE1, kAE1, or mdAE1. The presence of complex oligosaccharide forms
Figure 20. Immuno-blot showing N-glycosylation analysis of erythrocyte AE1, kidney (kAE1), and membrane (mdAE1) expressed in transfected HEK-293 cells. AE1 (lane 1-3), kAE1 (lane 4-6), and mdAE1 (lane 7-9). The panel at right is a long exposure of mdAE1. Detergent-solubilized cell extracts were subjected to glycosidase digestion (C: control, H: endoH, F: PNGaseF). Immuno-blot analysis was performed as described in Figure 15. A mouse monoclonal anti-AE1 antibody was used to detect AE1. The positions of molecular weight markers are shown on the left. The AE1 ran as a single major band, kAE1 ran as a major lower band with higher molecular bands, and the mdAE1 ran as a lower band and a series of higher molecular weight bands. In all cases, the lower band (o) was sensitive to endoH digestion, while the upper bands (highlighted with a bracket:)] were resistant.

Table 5. Percentages of complex oligosaccharides in erythrocyte AE1, kidney (kAE1) and membrane domain (mdAE1) expressed in HEK-293 cells with no tag and tags (HA, His6, and HA and His6) from several independent transfection experiments, including data showing in Figure 20. The percentages were obtained from immunoblotting images of endoH digested bands (blots are not shown).

<table>
<thead>
<tr>
<th>Oligosaccharide</th>
<th>No tag (n=3)</th>
<th>HA tag (n=3)</th>
<th>His6 tag (n=3)</th>
<th>HA and His6 tag (n=3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>eAE1</td>
<td>6.2 ± 1.0</td>
<td>17.5 ± 0.4</td>
<td>7.4 ± 1.5</td>
<td>3.8 ± 2.2</td>
</tr>
<tr>
<td>kAE1</td>
<td>39.6 ± 3.9</td>
<td>49.2 ± 8.4</td>
<td>40.6 ± 2.4</td>
<td>46.4 ± 4.6</td>
</tr>
<tr>
<td>mdAE1</td>
<td>67.5 ± 3.9</td>
<td>75.2 ± 3.2</td>
<td>70.9 ± 10.0</td>
<td>61.6 ± 1.8</td>
</tr>
</tbody>
</table>
Figure 21. Localization of erythrocyte AE1 (upper row), kidney (kAE1, middle row), and membrane (mdAE1, lower row) AE1 expressed in HEK-293 cells by immunofluorescence using a LSM 510 confocal microscope. Fixed cells were stained by a mouse monoclonal antibody (BRIC 6), which detects an extracellular epitope on AE1. The cell surface expression (column 1) was visualized as green using a secondary Alex 488 conjugated goat anti-mouse IgG in non-permeabilized cells. Total AE1 expression (column 2) was visualized as red by secondary Cy3-conjugated goat anti-mouse IgG in detergent-permeabilized cells. Distribution of AE1 on the cell surface was displayed in yellow by merging the cell surface and total expression of AE1 (column 3). The merged images show that all three proteins were expressed at the cell surface, (yellow), with some intracellular localization (red).
indicates the trafficking of glycoproteins from ER to Golgi. An exclusively high mannose form may suggest that the glycoproteins are in the ER, or are simply not processed (Li et al., 2000). Immunofluorescence staining revealed all AE1 forms (AE1, kAE1, and mdAE1) can traffic to the cell surface Figure 21. Although AE1 is expressed in mostly high-mannose oligosaccharide form in transfected HEK cells, the protein is capable of trafficking to the plasma membrane, but somehow escapes oligosaccharide processing (Li et al., 2000). The complex forms of kAE1 and mdAE1 represent proteins that had trafficked to the Golgi and most likely are present at the plasma membrane.

3.6. Oligosaccharide processing of Trp mutants of mdAE1

Since the C-terminal His$_6$-tag did not influence oligosaccharide processing (Table 5), the present study used the mdAE1 with a C-terminal His$_6$-tag for a purification purpose, which may apply to further biochemical and biophysical studies. To analyze oligosaccharide processing of the Trp mutants, Ala or Phe substitutions of Trp residues (492, 496, 648, 662, 723, 831, and 848) were constructed in the context of mdAE1 with a C-terminal His$_6$-tag. Wt mdAE1 and Trp mutants were expressed in HEK-293 cells. Cells grown on 6-well plates were transfected in parallel with identical amounts of DNA in multiple separate transfection experiments.

Figure 22 shows immuno-blots of total-cell lysates of mdAE1 and Trp mutants (Figure 22 A, B). The blot shows that mdAE1 runs as a sharp lower-band at an approximate molecular mass of 55kDa and a series of closely spaced upper-bands. The lower band corresponds to the high-mannose form (open circle), while the upper bonds correspond to complex oligosaccharide forms (bracket). A non-specific immuno-reactive band was observed between molecular mass of 72 and 55 kDa and was seen in vector-transfected cells (Figure 22 lane 1) and serves as a useful loading control. The same blots were reprobed with anti-GAPDH antibody in order to account for differences in protein loading. As can be seen from these blots, the expression levels of the W831A/F, W848A/F, and W723F were similar to Wt AE1, while the expression levels of the other mutants were diminished. The expression levels were consistent with expression of Trp mutants in full-length AE1, except W648F whose expression level was comparable to Wt
Figure 22. Immunoblot analysis of N-glycosylation processing of C-terminal His$_6$-tagged membrane domain (mdAE1) and Trp mutants expressed in transfected HEK-293 cells. Immuno-blot analysis of total-cell lysates was performed as described in Figure 15. The positions of molecular weight markers are shown on the left. Arrows at the left between molecular weight of 72000 and 55000 in (A, B) denote a non-specific band as the band appears in the vector control. Wild-type mdAE1 ran as a lower molecular band and a series of higher molecular bands, as did W831A/F, W848A/F and W723F. The other mutants were expressed at a diminished level and ran as a lower band. (C) Immuno-blot analysis of N-glycosylation processing of C-terminal His$_6$-tagged membrane domain (mdAE1) and Trp mutants after purification by Ni-affinity chromatography. Note the absence of the non-specific band. Wild-type mdAE1 ran as a lower molecular weight band and a series of higher molecular bands, as did W831A/F, W848A/F and W723F. Mutant W648A/F and mutant W662F showed a major lower band with a lower amount of upper bands while the other mutants only showed the lower band. Panel (D) shows percent complex oligosaccharide of Trp mutants (dark bar: Ala substitutions, light gray bar: Phe substitutions) determined by scans of the immunoblots of proteins purified by Ni-affinity chromatography in three independent transfection experiments. The error bars show the S.D. with n = 3. W831A/F, W848A/F, W648F, W662F, and W723F were comparable to Wt mdAE1 (p > .01).
in AE1 but is reduced in mdAE1. The difference in expression level of mdAE1 and Trp mutants was reproducible in multiple independent (n = 3) transfection experiments.

Low expression levels of Trp mutants, W492A/F, W496A/F, W648A/F, W662A/F, and W723A provide little information regarding the N-glycosylation status of these mutants (Figure 22 A, B). Ni-affinity purification was performed to concentrate the proteins and remove the background, providing a clearer picture of the glycosylation pattern (Figure 22 C). Not all Trp mutants contain complex oligosaccharide forms. Complex oligosaccharide forms are present in Trp mutants, W831A/F, W848A/F, and W723F similar to Wt AE1; the other mutants are primarily in the high mannose forms. After Ni-affinity purification, the lower band could often be resolved into two closely spanned bands, the lowest band may represent mdAE1 that was not glycosylated.

The percent complex oligosaccharide was quantified from immuno-blots of three independent transfection experiments after Ni-purification to eliminate non-specific bands (Figure 22 D). Trp mutants, W831F/A, W848F/A, and W723F, contain high percentage of complex oligosaccharides compared to Wt AE1. W492A/F and W496A/F contained little complex oligosaccharide. Percent complex oligosaccharides were approximately half of Wt mdAE1 in the Phe mutants (residues 648 and 662). The results indicate that Trp mutants, W831F/A, W848F/A, and W723F, were effectively processed to complex oligosaccharide forms, while the other mutations had an effect on the processing of mdAE1 into complex oligosaccharide forms, consistent with impaired exit from the ER.

3.7. Cell surface expression of mdAE1

Figure 23 shows the localization of Wt mdAE1 and Trp mutants in transfected HEK-293 cells using immunofluorescence and confocal microscopy as described in Figure 16. Panel (Figure 23 A) depicts Wt mdAE1 expressed at the cell surface (green) and intracellularly (red). Yellow in the merged image provides the cell surface distribution of AE1. Panel (Figure 23 B) shows that all Trp mutants could be detected at the cell surface, except substitutions at residues 492 and 496. All cells show intracellular localization of AE1. The intensity of staining detected by confocal microscopy was saturated and was not able to show the differences among mutants. The results indicate that mutation of Trp
Figure 23. Localization of C-terminal His$_6$-tagged membrane domain AE1 (A) and mdAE1 tryptophan mutants (B) expressed in HEK-293 cells by immunofluorescence.
using LSM 510 confocal microscope. Wt (A) and tryptophan mutants (B, left: Ala substitution, right: Phe substitution) of mdAE1, were stained by a mouse monoclonal antibody (BRIC 6), which detects an extracellular epitope on AE1. The cell surface expression was visualized as green by secondary Alex 488 conjugated goat anti-mouse IgG in non-permeabilized cells (column 1). Total AE1 expression was visualized as red by secondary Cy3-conjugated goat anti-mouse IgG in permeabilized cells (column 2). Distribution of AE1 on the cell surface was displayed in yellow by merging the cell surface and total expression of AE1 (column 3). The mdAE1 mutants W648A/F, 662A/F, 723A/F, 831A/F, 848A/F showed evidence of cell surface expression (yellow), while W492A/F and W496A/F did not. Wild-type and all mutants showed intracellular expression (red).
residues 492 or 496, but not the other Trp residues, profoundly influenced trafficking to the cell surface. The results are consistent with cell surface expression of AE1 shown by immunofluorescence and confocal microscopy except Ala substitutions at 648 and 662. W648A and W662A affects cell surface expression less in mdAE compared to AE1.
Chapter 4
Discussion

My project examined the role of individual Trp residues in the functional expression of AE1 in transfected HEK cells. Trp residues are commonly found between the hydrophobic core and surface in soluble proteins (Miller et al., 1987). It is also a common residue in the binding sites of enzymes (Katz, 1995; Ordentlich et al., 1995; Zeltins and Schrepf, 1997; Yengo et al., 1998; Adamian and Liang, 2001). In membrane proteins, the amphipathic nature of the Trp side chain is important in stabilizing TM segments to position them in the bilayer at the interface of the lipid bilayer during biosynthesis (Yau et al., 1998; von Heijne, 1999). As Trp residues are conserved across species in the SLC4 family of anion exchangers, this suggests that they may be important in the functional expression of these membrane transport proteins. Based on the folding models of AE1 (Popov et al., 1997; Fujinaga et al., 1999; Popov et al., 1999; Zhu et al., 2003; Zhu and Casey, 2007), two Trp residues (Trp\textsuperscript{662} and Trp\textsuperscript{723}) out of the seven Trp residues (Trp\textsuperscript{492}, Trp\textsuperscript{496}, Trp\textsuperscript{648}, Trp\textsuperscript{662}, Trp\textsuperscript{723}, Trp\textsuperscript{831}, and Trp\textsuperscript{848}) in the membrane domain are located at the membrane interface; Trp\textsuperscript{648} is close to the external N-glycosylation site (Asn\textsuperscript{642}); Trp\textsuperscript{492} and Trp\textsuperscript{496} reside on the same side of TM4 adjacent to the DIDS-reactive Lys\textsuperscript{539} in TM5 likely facing the inhibitor binding site; Trp\textsuperscript{831} is located in a hydrophilic loop, and Trp\textsuperscript{848} is in TM11 close to the second DIDS reactive Lys\textsuperscript{851}.

In this study, we mutated all seven Trp residues individually into Ala and Phe to investigate the roles of Trp residues in AE1 functional expression in transfected HEK cells. Total protein expression levels, cell surface expression, inhibitor binding, and anion transport activities were measured. His\textsubscript{6}-tagged constructs were used to allow the purification and analysis of the mutants, including poorly expressed constructs. Our investigation showed that the Trp residues could be sorted into three groups based on the outcomes of total expression, immuno-localization, cell surface biotinylation, inhibitor (SITS) binding, and oligosaccharide processing. We categorized the seven Trp residues into three classes: Class 1, AE1 expression not affected by mutation of Trp to either Ala or Phe (Trp\textsuperscript{831} and Trp\textsuperscript{848}); Class 2, AE1 expression reduced by mutation of Trp to either Ala or Phe (Trp\textsuperscript{492} and Trp\textsuperscript{496}); Class 3, AE1 expression decreased by Ala substitution...
and slightly or not influenced by Phe substitution (Trp<sup>648</sup>, Trp<sup>662</sup>, and Trp<sup>723</sup>) (The three classes are summarized in Table 6). Our observations indicate that Trp<sup>831</sup> and Trp<sup>848</sup> in class 1 may be exposed to an aqueous environment and not be involved in a critical protein interaction. Trp<sup>492</sup> and Trp<sup>496</sup>, in class 2 are essential in protein folding during biosynthesis as mutation of these residues leads to ER retention of AE1 and poor expression. These mutants are likely misfolded and subject to rapid protein degradation. Trp<sup>648</sup>, Trp<sup>662</sup>, and Trp<sup>723</sup>, in class 3 may have a structural role facilitated by aromatic amino acid residues since substitution with Phe had little effect on protein expression, while a substitution with the small residue Ala leads to reduced expression, again likely due to protein misfolding and degradation.

We used different experimental approaches to examine the functional expression of the Trp mutants in transfected HEK cells. Firstly, we showed, using immunoblots, differences in total expression levels among Trp mutants of AE1 expressed in transfected HEK-293. Mutations of Trp at position 831 or 848 resulted in little effect on total AE1 expression, while mutations of the residues 492 or 496 showed dramatically decreased expression with less than about 25% relative to Wt AE1. Mutations of Trp residues, 648, 662, and 723, showed reduced expression levels when substituted by Ala (~25% relative to Wt AE1) but not as much when substituted by Phe. The same tendency was observed through examining cell surface expression by immuno-localization and cell surface biotinylation.

Secondly, we used immunofluorescence and biotinylation assays to determine cell surface expression of the Trp mutants in transfected cells. Cell surface expression of Trp mutants, W831A/F and W848A/F, was comparable to Wt AE1, while W492A/F and W496A/F affected cell surface expression profoundly. Substitutions of residues 648, 662, and 723 to Phe did not influence cell surface expression, while Ala substitutions of these residues reduced cell surface expression as measured by cell surface biotinylation.

Thirdly, we investigated the effect of mutating Trp residues to Ala and Phe on inhibitor binding and transport function of AE1. We showed differences in SITS-Affi-Gel binding capacity of Trp mutants and the Wt AE1. W831A/F, W848A/F, W648F, W662F, and W723A/F have similar SITS binding capacity to Wt AE1 (p > .05) but W492A/F, W496A/F, W648A, W662A, and W723A are significantly reduced (p < .05).
Table 6. Summary of expression studies. Relative levels of total expression, percent complex oligosaccharide, cell surface immuno-fluorescence and biotinylation, and SITS-Affi-Gel binding of Trp mutants relative to Wt AE1 and mdAE1 set at +++ are shown.

<table>
<thead>
<tr>
<th></th>
<th>AE1</th>
<th>mdAE1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Samples</td>
<td>Expression</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
</tr>
<tr>
<td>Wt</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>W492A</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>W496A</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>W648A</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>W662A</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>W723A</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>W831A</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>W848A</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>W492F</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>W496F</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>W648F</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>W662F</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>W723F</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>W831F</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>W848F</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>-</td>
<td>&lt;10</td>
<td>*</td>
</tr>
<tr>
<td>+</td>
<td>10-50</td>
<td>**</td>
</tr>
<tr>
<td>++</td>
<td>50-75</td>
<td></td>
</tr>
<tr>
<td>+++</td>
<td>&gt;75%</td>
<td></td>
</tr>
</tbody>
</table>

* Intracellular, ** Cell surface and intracellular localization
This tendency is also consistent with the protein and cell surface expression levels. That is poorly expressed proteins had impaired inhibitor binding capacity consistent with them being misfolded proteins. We also show preliminary results of the Cl⁻/HCO₃⁻ transport assay on Trp mutants that are expressed at Wt levels, carried out in collaboration with Dr. Casey’s laboratory at the University of Alberta. These results showed all Trp mutants tested are active with transport activity above vector control, with some mutants having activities close to Wt AE1. We did note a reduction in the transport activity of the His₆-tagged constructs relative to untagged AE1. Previous work from our lab has shown that the integrity of the C-terminal tail of AE1 is important for trafficking of AE1 from the ER to the plasma membrane (Cordat et al., 2003). The transport assays are being repeated and normalized to protein and cell surface expression levels to produce statistically meaningful results. Transport assays will also be performed on mutants such as W831F and W848F. Mutants that are expressed at low levels and retained in the ER will not be tested, as these mutants will show no transport activity in a whole cell assay.

We tested expression and trafficking of erythrocyte AE1, kidney AE1, and the membrane domain of AE1 and showed that differences in the structure of the N-terminal cytosolic domain resulted in differences in oligosaccharide processing. The percent complex oligosaccharide increased significantly in the order of AE1, kAE1, and mdAE1. The tags (extracellular HA, C-terminal His, and combination of extracellular HA and C-terminal His) had little effect on the processing of oligosaccharides in AE1, kAE1, and mdAE1. It is puzzling how changes to the cytosolic domain of AE1 can affect N-glycosylation processing on the opposite side of the membrane. The simplest explanation is that AE1 is retained in the ER; kAE1 has impaired exit for the ER and mdAE1 exits efficiently and moves through the Golgi where its oligosaccharide is processed. All AE1 forms, however, traffic to the cell surface with similar efficiencies. Even AE1, expressed in mostly the high-mannose oligosaccharide form in transfected HEK cells, is capable of trafficking to the plasma membrane as shown previously (Li et al., 2000).

Oligosaccharide processing is dependent upon accessibility of the polypeptide chain to processing enzymes. It may be that the cytosolic domain interacts with intracellular loops on AE1 that affects the disposition of TM segments. Alternatively, it may result from oligomeric states of AE1 as AE1 can form tetramers, while the mdAE1 is strictly
dimeric. Perhaps, subunit interactions affect N-glycosylation processing. The N-terminal cytosolic domain is required for tetramer formation and the association of cytoskeletal proteins, such as ankyrin (Gomez and Morgans, 1993). In studies of the biosynthesis of chicken AE1, tetramer formation is acquired in the Golgi during the internalization of dimeric AE1 from the plasma membrane and assists AE1 to target to the cell surface by the association of cytoskeletal proteins (Ghosh et al., 1999).

Since mdAE1 was processed the most effectively, we mutated the Trp residues to follow N-glycosylation processing in Trp mutants. We demonstrated that mutants, W831A/F, W848A/F, and W723F, were effectively processed to complex oligosaccharide forms, while W492A/F, W496A/F, W648A, W662A, and W723A mutants were predominantly in the high mannose form. The W648F and W662F mutants were less effectively processed into complex oligosaccharide forms. We also showed cell surface expression of Trp mutants in mdAE1 using immunofluorescence and confocal microscopy. Cell surface expression was observed in all Trp mutants except W492A/F and W496A/F; all mutants also showed intracellular localization of AE1. The results showed that all Trp mutants were expressed and could traffic to the cell surface except W492A/F and W496A/F, which were retained intracellularly. The processing of Ala substitutions at W648, W662, and W723, showed a more significant defect in trafficking from the ER than Phe substitutions.

**Class 1: Trp\textsuperscript{831} and Trp\textsuperscript{848}**

We observed that mutations of Trp residues, 831 and 848 to Ala or Phe had no effect on expression nor on trafficking from the ER to the cell surface. Scanning cysteine mutagenesis studies showed that biotin maleimide could partially label W831C indicating a reactive cysteine exposed to aqueous media (Zhu et al., 2003). W831 is in a region involved in malaria parasite infection (Asp\textsuperscript{821}-His\textsuperscript{834}) and close to the extracellular antigen related to cell senescence (Leu\textsuperscript{812}-Arg\textsuperscript{827}). The region is not well defined in the folding model of AE1. The region contains both hydrophobic and charged residues including important histidine and lysine residues, (His\textsuperscript{816}, His\textsuperscript{834}, Lys\textsuperscript{814}, and Lys\textsuperscript{817}) required for anion transport, and a proline-rich portion (Pro\textsuperscript{815}-Pro\textsuperscript{825}) (Askin et al., 1998). Popov et al. showed Leu\textsuperscript{785}-Ile\textsuperscript{841} corresponds to the cytoplasmic loop between segment 10 and 11 during biosynthesis. The region Pro\textsuperscript{815}-Lys\textsuperscript{829} was however efficiently
labeled by sulfhydryl-specific reagents in cysteine scanning mutagenesis and postulated to be on the extracellular portion of the membrane with an extended structure of Val\textsuperscript{828}-Leu\textsuperscript{835}.

The region may be flexible and induced to be accessible from the outside under certain conditions (Zhu et al., 2003). This was supported by an N-glycosylation study in which P820N was slightly glycosylated in cell-free translation experiments but not glycosylated in transfected cells (Popov et al., 1999). The BRIC 132 antibody binds to the epitopes Phe\textsuperscript{813}-Tyr\textsuperscript{824} when treated by detergent and can only immuno-precipitate AE1 from open and leaky erythrocyte (Wainwright et al., 1990). The BRIC epitope may be buried under isotonic conditions and conformation changes at low ionic strength may allow the region to be accessible (Wainwright et al., 1990). An NMR study (Askin et al., 1998) of a synthetic peptide of Gly\textsuperscript{796}-Ilu\textsuperscript{841} showed the region can be fitted into the extramembrane protrusion of the cytoplasmic region in the three dimensional map of Wang et al. but the region possesses an unusual structure with two distorted helices that (Ile\textsuperscript{803}-Leu\textsuperscript{810} and Tyr\textsuperscript{824}-Phe\textsuperscript{836}) are connected by the proline-rich loop (Askin et al., 1998). The authors postulated the structure of this region was stabilized by inter-helical interactions of conserved hydrophobic residues, Leu\textsuperscript{800}, Val\textsuperscript{828}, and Trp\textsuperscript{831}. Our results indicate that Trp\textsuperscript{831} is not essential for functional expression of AE1.

Although mutation of Trp\textsuperscript{848} had little impact on expression of AE1 in the present study, Trp\textsuperscript{848} has been defined within TM11 in folding models of AE1 rather than in an aqueous loop. The entire region Phe\textsuperscript{836}-Lys\textsuperscript{851} was not labeled with biotin maleimide in a cysteine mutagenesis study (Fujinaga et al., 1999) suggesting that this region is in a hydrophobic environment embedded in the membrane. The region Ile\textsuperscript{840}-Thr\textsuperscript{853} was also postulated to be in TM11 in a N-glycosylation study (Popov et al., 1997). The segment composed of only 13 to 16 amino acid residues, may form a helical structure too short to fully span the 30 Å membrane thickness, as ~20 residues are required. A helical wheel plot (Reithmeier, 1996b) showed functionally important residues located in a hydrophilic helical face including Trp\textsuperscript{848} (W848C showed 26% of transport activity) and DIDS-reactive Lys\textsuperscript{851} at the center of the helical face. The hydrophilic helical side defined by Lys\textsuperscript{851} may face the anion translocation pore, and Trp\textsuperscript{848} may also face the same side and participate in ion translocation by lining the pore. However, W848A had less effect on
anion transport than W848C by showing a similar Cl⁻/HCO₃⁻ transport activity as Wt. Presumably mutation of W848C to hydrophilic residues like Cys causes greater impairment in the region than mutation to a small hydrophobic residue. The Cys residue may destabilize the interaction with a substrate or between other amino acid residues in the region.

**Class 2: Trp⁴⁹² and Trp⁴⁹⁶**

We determined that Trp⁴⁹² and Trp⁴⁹⁶ in class 2 are essential residues. Mutation to either Ala or Phe had profound effects on protein expression and trafficking. During biosynthesis, mutations of these residues likely lead to misfolding and, as a consequence, ER retention and lack of trafficking to the cell surface. The observations that both Ala and Phe substitutions impacted on AE1 expression suggest Trp⁴⁹² and Trp⁴⁹⁶ may be engaged in a crucial role in TM interactions and folding, and that the indole ring may be required in these interactions. Topology and helical models proposed that Trp⁴⁹² and Trp⁴⁹⁶ are located at the same surface of TM 4, opposite the most hydrophilic surface (Reithmeier, 1996b). Trp⁴⁹² and Trp⁴⁹⁶ may face into a channel with DIDS-reactive lysine (Lys⁵³⁸ and Lys⁵⁴²) on TM 5, forming a part of the ion translocation pore and/or be part of an essential interaction with adjacent TM helices. Since important roles for Trp are also found in enzyme active sites, there is a possibility that Trp⁴⁹² and Trp⁴⁹⁶ are involved in substrate binding or ion translocation in AE1. It is likely that these Trp residues are located close to the AE1 inhibitor binding sites and are the ones involved in energy transfer (Reithmeier, 1996b). Unfortunately, the low level of expression and loss of inhibitor binding of these mutants did not allow us to investigate the role of these residues in energy transfer to bound fluorescent inhibitors.

**Class 3: Trp⁶⁴⁸, Trp⁶⁶², and Trp⁷²³**

Mutation of residues, 648, 662, and 723 by Ala substitutions in class 3 reduced protein expression levels, cell surface expression, and SITS inhibitor binding, while Phe substitutions did not. The outcomes suggest that Trp⁶⁴⁸, Trp⁶⁶², and Trp⁷²³ have a structural role by forming protein-protein interaction sites provided by large aromatic amino acid residues. Trp⁶⁴⁸ is located at the fourth EC loop, six residues distal from the N-glycosylation site (Asn⁶⁴²). The region Ser⁶⁴³-Leu⁶⁵⁵ is inaccessible to a water-soluble sulfhydryl-specific reagent, LYLA (649 Da), suggesting the region is buried within the
folded structure of region Arg\textsuperscript{656}-Met\textsuperscript{663} or perhaps protected by the carbohydrate structure at Asn\textsuperscript{642}, sterically blocking LYLA access (Tang et al., 1998). Recall, however, that the oligosaccharide on AE1 in transfected cells retains a high-mannose structure. Thus, any protection afforded by the oligosaccharide to cysteine-reactive reagents is unlikely due to specific interactions beyond the core structure (GlcNAc\textsubscript{2}, Man\textsubscript{3}). We suggest that Trp\textsuperscript{648} is involved in specific interactions that stabilize this EC loop.

Scanning of N-glycosylation (Popov et al., 1999) and cysteine mutagenesis (Fujinaga et al., 1999) have defined TM8 as containing 21 residues and encompassing Met\textsuperscript{663}-Gln\textsuperscript{683}, placing Trp\textsuperscript{662} at the extracellular transmembrane interface at the beginning of TM8. TM 8 is a part of the anion translocation pore and contains three hydrophobic Phe residues facing the same direction, likely towards the lipid bilayer. The opposite side of these Phe residues are proposed to be lining the ion translocation pore along with Glu\textsuperscript{681}, which is the best-characterized residue involved in anion translocation (Tang et al., 1999). Based on a helical plot, Trp662 and other hydrophobic residues, such as Leu, Ile, and Ala, are located to the same side as Glu681, suggesting their participation in ion translocation. Trp\textsuperscript{662} was indeed identified as a functionally important residue involved in ion translocation by cysteine mutagenesis, as mutation to Cys greatly affect transport activities (Tang et al., 1999). Furthermore, we suggest that Trp\textsuperscript{662} also serves a structural role in AE1 packing, probably involved in a hydrogen bond interaction from its large side chain, since mutation to Phe also moderately influenced the functional expression of AE1. Thus both the size and indole ring of Trp\textsuperscript{662} may be required to optimize stable interactions at the position.

Trp\textsuperscript{723} is located at the extracellular end of TM9 and is connected to a trypsin-sensitive loop (T-loop: Trp\textsuperscript{723}-Leu\textsuperscript{764}), which may have structurally and functionally important roles by forming part of the channel in AE1. An N-glycosylation analysis indicates the loop may be transiently exposed to the lumen of the ER during biosynthesis and folded into the protein following protein biosynthesis. As shown by G714C, S725C, and S731C in the loop region having low transport activity and labeling by external LYLA, this region may be accessible to the extracellular media in the folded protein structure consistent with it forming part of the external facing channel. Our results showed even though W723A severely affected the expression of AE1, the mutation did
not impact the Cl-/HCO₃⁻ transport activity of AE1. We postulate that Trp⁷²³ participates in the tight packing of AE1 and in forming a vestibule of AE1. Mutating to a small residue, such as Ala or Cys, may have a negative effect on folding by hindering the close packing in AE1, though the small residue may generate a space at the vestibule region, enhancing accessibility by substrates.

The results indicate that Trp residues play different roles in the functional expression of AE1 depending on their location in the protein. Mutation of some Trp residues (Trp⁴⁹² and Trp⁴⁹⁶) profoundly affects expression suggesting that they are essential in attaining a proper folded state, while others (Trp⁸³¹ and Trp⁸⁴⁸) can be mutated without serious effect on the functional expression of AE1. Mutation of Trp⁶⁴⁸ at the extracellular loop and interfacial residues (Trp⁶⁶² and Trp⁷²³) to Phe did not severely affect expression while mutation to Ala decreased expression suggesting that aromatic residues at the extracellular loop and interface plays a role in stabilizing AE1. The effect of mutation of Trp residues in the third class was intermediate and depended on the nature of the mutation suggesting the nature of the residue is important in the protein structure.
Chapter 5
Future Directions

The present study has shown that Trp residues participate in differential roles in the functional expression of AE1. According to their distinct roles in AE1 functional expression, we were able to categorize them into three classes: Class 1 (Trp$^{831}$ and Trp$^{848}$) Trp residues that can be substituted to Ala or Phe; Class 2 (Trp$^{492}$ and Trp$^{496}$) Trp residues that cannot be mutated to either Ala or Phe; and Class 3 (Trp$^{648}$, Trp$^{662}$, and Trp$^{723}$) Trp residues that can be substituted to Phe but not to Ala. Future studies are required to expand on these results and further specify the roles of these residues in AE1 folding and function.

We did not routinely test the transfection efficiency during investigation of the expression of AE1 and Trp mutants, as the point mutations from Trp to either Ala or Phe were not expected to dramatically change the transfection efficiency in the transfected cells. Transfection efficiency may be measured by using a dual promoter expression system or pIRES Bicistronic expression system, in which cDNA of an expression marker protein, such as GFP which is in the same vector as cDNA encoding the proteins in the study. GFP expression measured by immunoblotting or fluorescence can be used as an indication of transfection efficiency. Testing mRNA amounts after transfection may also provide a control for transfection efficiency.

A pulse-chase experiment may provide important information concerning the stability of expressed proteins. By using the pulse-chase experiment, we may be able to indicate whether the differences in expression levels of the Trp mutants are due to high rates of degradation with poorly expressed proteins showing a more rapid turnover. It was previously showed that AE1 expressed in HEK-293 migrates from the ER to plasma membrane with a half-time of about four hours (Li et al., 2000) and that AE1 is stable with a half-life of ~15 hours (Li et al., 2000; Quilty and Reithmeier, 2000). We can compare the half-time of Trp mutants with Wt AE1 and these previous data. We can predict that those mutants with an exclusively high mannose form, low expression levels, impaired SITS-Affi-Gel binding capacity, and no cell surface expression (W492A/F, W496A/F, W648A, and W662A) would have a faster half-time of degradation due to
ERAD. The other Trp mutants (W831A/F, W848A/F, W648F, W662F, and W723A/F) may have similar biosynthesis and trafficking rates and also degradation half-time as Wt AE1. Pulse-chase experiments with mdAE1 would be even more valuable as it would be possible to measure the rate of conversion of the high-mannose to complex oligosaccharide and provide a measure of the rate of ER exit.

We were also able to show cell surface and intracellular localization among Wt and Trp mutants of AE1 using immunofluorescence and confocal microscopy. We showed cell surface and intracellular localization in the Wt and all Trp mutants except for W492A/F and W496A/F, which were not localized to the cell surface. Using cell surface biotinylation, we showed cell surface expression of W831A/F, W848A/F, W648F, W662F, and W723F were similar to AE1 but not so for the other mutants (W492A/F, W496A/F, W648A, W662A, and W723A). This tendency is consistent with the protein expression levels as analyzed by immunoblots. In future studies, it may be also worth confirming the amount of the cell surface and intracellular expression from images of immunofluorescence using Colocalizer express 1.2 (003) software to quantify the relative levels of expression of the mutants.

The effect of mutating Trp residues on AE1 transport function or its inhibitor binding has not been thoroughly studied yet. We demonstrated differences in SITS-Affi-Gel binding capacity of Trp mutants. W831A/F, W848A/F, W648F, W662F, and W723A/F have the similar SITS binding capacity to Wt AE1, while W492A/F, W496A/F, W648A, W662A, and W723A do not. This tendency is also consistent with the low protein and cell surface expression level shown by these mutants. We showed preliminary results of the Cl⁻/HCO₃⁻ transport assay, which have been addressed in collaboration with Dr. Casey’s laboratory. As this assay is time consuming and was performed by collaboration, we selected a subset of mutants to test. The preliminary results showed all Trp mutants tested are active, having transport activities above vector controls, some with similar activity to Wt C-terminal His₆-tagged AE1. Unfortunately, we cannot determine which Trp residues tested are essential for AE1 transport activity due to the large standard deviation observed in Wt C-terminal His₆-tagged AE1. The transport assay for these mutants requires repeating in a future study. Other mutants,
especially W831F and W848F also need to be tested. The transport activity of mutants with low expression will not be tested using a whole cell assay.

We predicted that certain Trp residues may form part of the inhibitor binding pocket, as observed by enhanced fluorescence upon fluorescent inhibitor binding (Dix et al., 1979). Energy transfer is observed between Trp residues and the bound fluorescent inhibitor, BADS at an estimated distance of ~20 Å (Rao et al., 1979). One goal of this project was to identify the Trp residues involved in energy transfer to bound inhibitor. These Trp residues may be positioned close to the inhibitor binding site. We designed His₆-tagged constructs to facilitate purification of the Trp mutants with the goal of performing the energy transfer experiment in detergent solution (Lieberman and Reithmeier, 1983). The fluorescence resonance energy transfer (FRET) requires greater than ~100 µg of pure protein, which may be possible using HEK cells as a source.

Based on the folding model of AE1, we predicted that Trp⁴⁹² and Trp⁴⁹⁶ at TM4 adjacent to DIDS-reactive Lys⁵³⁹ in TM5, are facing the channel of the DIDS binding site. However, mutations to either Ala or Phe resulted in impaired AE1 functional expression by showing dramatically reduced protein expression, no trafficking to the cell surface, and no SITS binding capacity. We conclude W492A/F and W496A/F were essential for proper folding and expression of AE1; furthermore, they are likely misfolded and retained in the ER. The low expression of W492A/F and W496A/F mutants make their purification in sufficient quantities for the FRET assays difficult. Mutation of Trp⁸⁴⁸, close to the second DIDS-reactive Lys⁸⁵¹ when mutated to Ala or Phe was expressed similar to the Wt. W848A/F mutants also retained SITS binding capacity and would be suitable for a FRET experiment. The study measures the energy emission of a fluorescent AE1 inhibitor, BADS, at 450 nm. The energy transfer occurs from Trp residues (λ excite = 295 nm) to the bound fluorescent inhibitor BADS (λ excite = 350 nm), which then fluoresces at 450 nm (Lieberman and Reithmeier, 1983). Briefly, purified His₆-tagged AE1 from HEK-cells grown in multiple plates or roller-bottles is titrated with BADS. The solution is excited at 295 nm, and the emission is measured at 450 nm. Mutation of Trp residues at the inhibitor binding site will be result in loss of the energy transfer.
Structures of anion-binding proteins determined by X-ray crystallography studies have shown that bound anion substrates are stabilized by hydrogen bonds with the backbone amide nitrogens near the amino-termini of helices, instead of being stabilized by charged residues such as Arg and Lys. This feature is also observed in the CIC channel, in which partial positive charges derived from the N-terminus helix dipole, main chain nitrogen atoms, and side chain oxygen atoms support stabilizing Cl⁻ ions (Dutzler et al., 2002). Pro and Gly are often found in the N-termini of a helix dipole; therefore, mutation of Pro or Gly residues at the anion binding site may result in destabilizing the helix and affect Cl⁻ ion binding. This hypothesis may be tested for AE1 in future studies by testing the effect of mutations on Pro and/or Gly of AE1.

In the present study, we were able to identify important Trp residues involved in the functional expression of AE1 and suggest their positional roles. Some of them may be involved in folding during biosynthesis, helix-helix interactions, formation of the anion translocation pore, and inhibitor binding. Nevertheless, we were not able to specify their precise action in the structure and function of AE1. The role of Trp residues in AE1 could be further illustrated by a crystal structure of this membrane protein, which remains a challenging goal for future study. A high resolution structure of AE1 will be fundamental of understanding the protein structure and function in a molecular level, and the structure may be applied to understand the transport mechanisms of related transporters and the effect of mutations at Trp and other residues.
References


second site mutation of Lys 558, the locus of covalent H2DIDS binding. Biochemistry 34, 9315-9324.


Appendix 1

List of Trp mutant primers

<table>
<thead>
<tr>
<th>Residues</th>
<th>Substitutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>492</td>
<td>forward: CGTGGGCCGCGGTGCGATCGGGCTTCTTGGC</td>
</tr>
<tr>
<td>reverse: GCCAGAAGCCGATCGCCACGCGGGCCACG</td>
<td></td>
</tr>
<tr>
<td>496</td>
<td>forward: GTGTGATCGGCCTCCTGCTATCTGTGCTG</td>
</tr>
<tr>
<td>reverse: CACCAGCAGGATGAGCGCGGCCCACG</td>
<td></td>
</tr>
<tr>
<td>648</td>
<td>forward: GTCAGCCCGGGCCTGCTGAG</td>
</tr>
<tr>
<td>reverse: CACCAGCAGGATGAGCGCGGCCCACG</td>
<td></td>
</tr>
<tr>
<td>723</td>
<td>forward: CAAGCAGGATCGCACTCCCATGAGTGTG</td>
</tr>
<tr>
<td>reverse: GTGAATAAATCGATGAGTGCATGCACT</td>
<td></td>
</tr>
<tr>
<td>831</td>
<td>forward: GGTGGACTTCACCAGCAGG</td>
</tr>
<tr>
<td>reverse: CAAGCAGGATCGCACTCCCATGAGT</td>
<td></td>
</tr>
<tr>
<td>848</td>
<td>forward: CGTGGACTTCACCAGCAGG</td>
</tr>
<tr>
<td>reverse: GTGAATAAATCGATGAGTGCATGCACT</td>
<td></td>
</tr>
</tbody>
</table>

Phe

<table>
<thead>
<tr>
<th>Residues</th>
<th>Substitutions</th>
</tr>
</thead>
<tbody>
<tr>
<td>492</td>
<td>forward: CGTGGGCCGCGGTGCGATCGGGCTTCTTGGC</td>
</tr>
<tr>
<td>reverse: GCCAGAAGCCGATCGCCACGCGGGCCACG</td>
<td></td>
</tr>
<tr>
<td>496</td>
<td>forward: GTGTGATCGGCCTCCTGCTATCTGTGCTG</td>
</tr>
<tr>
<td>reverse: CACCAGCAGGATGAGCGCGGCCCACG</td>
<td></td>
</tr>
<tr>
<td>648</td>
<td>forward: GTCAGCCCGGGCCTGCTGAG</td>
</tr>
<tr>
<td>reverse: CACCAGCAGGATGAGCGCGGCCCACG</td>
<td></td>
</tr>
<tr>
<td>723</td>
<td>forward: CAAGCAGGATCGCACTCCCATGAGTGTG</td>
</tr>
<tr>
<td>reverse: GTGAATAAATCGATGAGTGCATGCACT</td>
<td></td>
</tr>
<tr>
<td>831</td>
<td>forward: GGTGGACTTCACCAGCAGG</td>
</tr>
<tr>
<td>reverse: CAAGCAGGATCGCACTCCCATGAGT</td>
<td></td>
</tr>
<tr>
<td>848</td>
<td>forward: CGTGGACTTCACCAGCAGG</td>
</tr>
<tr>
<td>reverse: GTGAATAAATCGATGAGTGCATGCACT</td>
<td></td>
</tr>
</tbody>
</table>