STRUCTURAL BASIS FOR THE FUNCTION AND REGULATION OF THE EPITHELIAL SODIUM CHANNEL

Pradeep Kota

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biochemistry and Biophysics

Chapel Hill
2012

Approved by:

Nikolay V. Dokholyan, PhD

John R. Riordan, PhD

M. Jackson Stutts, PhD

Brian A. Kuhlman, PhD

Sharon L. Campbell, PhD

ABSTRACT

PRADEEP KOTA: Structural Basis for The Function and Regulation of the Epithelial Sodium Channel
(Under the direction of Dr. Nikolay V. Dokholyan)

Epithelial sodium channels (ENaC) mediate sodium transport across epithelia. Functional channels are assembled from three homologous α, β and γ subunits with ~30% similarity in amino acid sequence. Mutations in different subunits of this channel are responsible for diseases including Liddle's syndrome and type I pseudohypoaldosteronism. ENaC is synthesized on the ER membrane, aquires complex N-linked glycosylation in the Golgi and is trafficked to the plasma membrane where it is activated upon cleavage by numerous membrane-anchored and/or soluble serine proteases secreted into the extracellular milieu. Although it has been established that exogenous expression of all three subunits in oocytes is required for robust channel activity, the number and stoichiometry of subunits comprising one functional channel remains unclear. Different biophysical and electrophysiological studies have concluded that ENaC assembles as a trimer or a tetramer with possible larger molecular weight oligomers arising from higher order assembly of trimers or tetramers. Due to the lack of structural information on ENaC, the molecular aspects of channel activation and regulation of function remain less well understood. In the current study, using a battery of computational and experimental techniques, we address

specific questions concerning the structural aspects of regulation of channel activation and function by constructing a structural model of the channel. Significant advances through this study include determination of oligomerization state of ENaC using native gel electrophoresis and identification of allosteric communication within the channel and modulating channel activity by rational mutagenesis of the identified allosteric sites. In this study, we conclude that ENaC assembles as both trimers and tetramers in the same cell. The amount of tetramers correlates well with increase in function and more importantly, the gamma subunit plays a crucial role in the formation of tetramers in oocytes. We believe that the results presented here would be immensely helpful in the future for understanding the cellular aspects of channel regulation and function at the molecular level.

To my parents

ACKNOWLEDGEMENTS

More people than I can count, deserve credit for my successful completion of this dissertation. Due to limited space, I would like to acknowledge the contributions of the most important people in my life as well as my professional career.

Professional

I would like to thank Prof. Barry Lentz for giving me a position in the Molecular and Cellular Biophysics program and providing me with an opportunity to identify my interests in protein biochemistry and computational biophysics. I would like to thank my advisor Prof. Nikolay V. Dokholyan for inculcating passion and perseverance in me and motivating me at all times to succeed in completion of my dissertation. Dr. Brian Kuhlman and Dr. Sharon Campbell have provided valuable suggestions throughout my career as a graduate student, by serving crucial roles on my thesis committee. I would like to thank Dr. John Riordan for extending the honor of being the chair of my committee and for inspiring me to be a scientist. Dr. Jack Stutts has taken special interest in my professional development as a scientist. Working closely with him has given me a rich insight into the physiological aspects of epithelial sodium channel function. I would like to extend my heartfelt thanks to Dr. Martina Gentzsch for providing her support and motivation and aiding my development as an experimental biologist.

I would like to thank Dr. Douglas Cyr, Dr. Klaus Hahn and Dr. Aravind Asokan for very productive collaborations. I enjoyed working with Dr. Andrei Aleksandrov (Riordan lab), Dr. Lihua He (Riordan lab), Tim Jensen (Riordan lab), Dr. Andrei Karginov (Hahn lab), Dr. Dan Summers (Cyr lab) and Dr. Pulicherla Nagesh (Asokan lab).

I would like to extend special thanks to every member of the Dokholyan lab for making such a close-knit team and creating a friendly and motivating environment to be joyful and productive at all times.

Personal

I am indebted to my parents for providing me with the right opportunity and education at the right time. I am very grateful to them for being supportive of every step I took throughout my career so far. I would forever remember the momentous emotional support my wife, Dhivya Ramalingam, has provided me with, to stay sane and be productive in my graduate school. Finally, I would like to thank my paternal grand father for inculcating interest in science at a very early age.

PREFACE

Part of the work described in this dissertation was published as articles in Bioinformatics and Proteins: Structure, Function and Bioinformatics:

Pradeep Kota, Feng Ding, Srinivas Ramachandran and Nikolay V. Dokholyan, **Gaia: Automated quality assessment of protein structure models**, *Bioinformatics* (2011) 27:2209-2215

Srinivas Ramachandran, Pradeep Kota, Feng Ding and Nikolay V. Dokholyan, **Automated minimization of steric clashes in protein structures**, *Proteins: Structure, Function and Bioinformatics* (2011) 79:261-270

Required permission to reuse figures and text extracts from the article has been obtained from all the authors and Wiley-Blackwell and Oxford University Press (Journal Publishers).

Table of Contents

List of Figures

List of Abbreviations

ENaC	Epithelial sodium channel
DEG	Degenerin
BNaC	Brain sodium channel
FaNaCh	FMRFamide gated sodium channel
PHA-1	Type 1 pseudohypoaldosteronism
CAP	Channel activating protease
MTS	Methanethiosulfonate
MTSET	[2-(tri-methylammonium- *O*ethyl]methanethiosulfonate bromide
P_O	Open probability
CAP	Channel-activating protease
hNE	Human neutrophil elastase
MAP	Membrane associated protease
PIP2	Phosphatidylinositol 4,5-bisphosphate
PIP3	Phosphatidylinositol 3,4,5-trisphosphate
PS	Phosphatidylserine
CD	Circular dichroism spectroscopy
BN-PAGE	Blue native polyacrylamide gel electrophoresis
CN-PAGE	Clear native polyacrylamide gel electrophoresis
SASA	Solvent accessible surface area
MSA	Molecular surface area

MD	Molecular dynamics
DMD	Discrete molecular dynamics
PDB	Protein data bank
EEF1	Effective energy function 1
CG	Conjugate gradient
BLAST	Basic local alignment search tool
PMF	Potential of mean force

CHAPTER 1

Introduction

The epithelial sodium channel (ENaC) is a prototypic member of the DEG/ENaC superfamily of ion channels (Canessa et al., 1994b). The DEG/ENaC superfamily can be classified as, (i) amiloride-sensitive ENaCs involved in Na^+ reabsorption in epithelia including the distal colon, distal nephron and sweat glands (Duc et al., 1994; Renard et al., 1995), (ii) voltage-independent brain Na^+ channels (BNaC1 and BNaC2) (Garcia-Anoveros et al., 1997), (iii) degenerins (MEC-4, MEC-10 and DEG-1) that form part of a mechanotranduction complex for touch sensitivity in *Caenorhabditis* elegans (Driscoll and Tavernarakis, 1997; Garcia-Anoveros and Corey, 1997), and (iv) peptide neurotransmitter Phe-Met-Arg-Phe-NH_2 (FMRF) amide-gated sodium channels (FaNaCh) expressed in the ganglion of the snail *Helix* aspersa (Lingueglia et al., 1995). Other members of this superfamily include the acid-sensing ion channel (ASIC), dorsal root ganglia acid-sensing ion channel (DRASIC), and other mechanosensitive cation channels expressed in cochlear hair cells and oocytes (Benos et al., 1995; Corey and Garcia-Anoveros, 1996; Garty, 1994; Rossier et al., 1994). Due to its particularly crucial role in Na^+ transport across the aldosterone-sensitive distal nephron, regulation of expression and function of ENaC is critical in control of

blood pressure. Hormones such as aldosterone, vasopressin and insulin as well as PKA/cAMP, PKC, Ca2+ and G-proteins tightly regulate ENaC expression and function in kidneys (Benos et al., 1995; Garty and Palmer, 1997). The pathophysiological importance of ENaC has been evidenced by the identification of mutations in the channel responsible for diseases like Liddle's syndrome, an autosomal dominant variant of hypertension (Shimkets et al., 1994), and for type-1 pseudohypoaldosteronism (PHA-1), a salt-losing syndrome (Chang et al., 1996).

1.1 Molecular architecture of ENaC

ENaC is a heteromultimeric ion channel made of homologous (~30-40% sequence identity) α, β and γ glycoprotein subunits (75-90 kDa each) surrounding the channel pore (Canessa et al., 1993; Canessa et al., 1994b). The δ-subunit has functional similarities with the α-subunit, but its physiological role is less well-understood (Waldmann et al., 1995). All members of the DEG/ENaC superfamily share a common structural topology. Hydropathy analysis indicates that all ENaC subunits have two hydrophobic membrane-spanning regions separated by a large (~500 residues) hydrophilic loop (Canessa et al., 1994a). N- and C-termini are intracellular while the large hydrophilic loop is extracellular with highly conserved cysteine residues and multiple N-glycosylation sites (Canessa et al., 1994a; Snyder et al., 1994). α-subunits alone can assemble as channels at the plasma membrane, albeit of low conductance, while β and Y subunits cannot form functional channels when expressed alone (Harris et al., 2008). Moreover,

α-subunits are hypothesized to chaperone the assembly and trafficking of the heteromultimer to the plasma membrane (Harris et al., 2008).

1.2 Subunit stoichiometry of ENaC

Genes encoding different subunits of ENaC have been cloned ~20 years ago and yet the subunit composition of function epithelial sodium channels remains unsettled. Many groups have addressed this important issue regarding the subunit stoichiometry of ENaC and related channels. Although the recently solved crystal structures of related ion channels and receptors (Gonzales et al., 2009; Jasti et al., 2007; Kawate et al., 2009) suggest a trimeric organization of the channel, some groups propose a tetrameric structure (Anantharam and Palmer, 2007; Firsov et al., 1998; Kosari et al., 1998) while others suggest that functional ENaC channels are composed of three, six or nine subunits (Cheng et al., 1998; Snyder et al., 1998; Staruschenko et al., 2005; Staruschenko et al., 2004; Stewart et al., 2011). Quantitative analysis of cell surface expression of ENaC showed that assembly follows fixed stoichiometry with the α-ENaC as the most abundant subunit.

Firsov et al., provided several lines of evidence to argue that ENaC is a tetramer (Firsov et al., 1998). They developed a quantitative assay based on the binding of ^{125}I-labeled M_2 anti-FLAG monoclonal antibody directed against a FLAG reporter epitope introduced in the extracellular loop of different ENaC subunits (Firsov et al., 1996). Using this assay, they determined that channels have equal number of β and γ subunits and twice the number of α subunits, suggesting $\alpha_2\beta_1\gamma_1$ to be the most likely stoichiometry (Firsov et al., 1998; Firsov

et al., 1996). Kosari et al., independently concluded that ENaC is a tetramer by performing functional studies on *Xenopus* oocytes expressing mutant subunits (αS583C, βG525C, γG542C – based on the sequence of mouse ENaC subunits) with lower affinity to amiloride than the wildtype subunits (Kosari et al., 1998). Coscoy et al., performed sedimentation of the peptide-activated FaNaCh channels in sucrose gradients to determine subunit composition. They reported that FaNaCh was observed in the fraction corresponding to a molecular mass of ~350 kDa (Coscoy et al., 1998). Based on the molecular weights of individual subunits, which ranges between 75 and 90 kDa, they concluded that FaNaCh is a tetramer (Coscoy et al., 1998). Given that FaNaCh is a close member of ENaC in the DEG/ENaC superfamily, they hypothesized that ENaC could be a tetramer. The heterotetrameric assembly is particularly attractive because of its four-fold symmetry around a central conducting pore; a hallmark feature of several potassium (K^+) channels (Doyle et al., 1998).

In contrast to the tetrameric architecture proposed by Firsov et al., and Kosari et al., Snyder et al., proposed that ENaC is formed by nine subunits, with a stoichiometry of $\alpha_3\beta_3\gamma_3$ (Snyder et al., 1998). Methanethiosulfonate (MTS) derivatives have been traditionally used to study the gating properties, accessibility and structure of ion channel pore regions (Akabas et al., 1992). The chemical modification of an introduced cysteine by a charged MTS reagent may produce a measurable change in the function of the ion channel/transport protein, which can be measured by electrical recording or isotope flux. Such data give information concerning the time-course, state dependance and membrane

sidedness of the accessibility of the cysteine (Akabas et al., 1992; Stauffer and Karlin, 1994). Snyder et al., used the mutant γG537C (according to the sequence of rat ENaC), in which C537 can be modified by the positively charged MTS reagent MTSET ([2-(tri-methylammonium-*O*ethyl]methanethiosulfonate bromide). Modification by MTSET produces inhibition of the current of mutant channels but not of wild-type channels. They coexpressed several combinations of wildtype γ and the mutant γG537C subunits and measured the fraction of current blocked by MTSET. The reagent MTSET decreased ENaC currents in a much larger proportion than the fraction of injected mutant subunits. This result suggested that modification of γG537C produced a dominant effect and that there was more than one γ subunit per channel. The proportion of channels sensitive or resistant to MTSET modification can be correlated to the ratio of cRNA injected for the wildtype and mutant subunits using an approach originally applied for determination of stoichiometry of shaker K^+ channels (MacKinnon, 1991). Using this analysis approach, Snyder et al., determined that there are three γ subunits per functional ENaC channel. Similar experiments, using equivalent MTSET modifications on α583C and βG525C mutants, indicated the presence of three α and three β subunits per channel. Therefore, the model proposed by Snyder et al consists of nine subunits with the stoichiometry $\alpha_3\beta_3\gamma_3$. Cheng et al., performed sucrose gradient sedimentation experiments to conclude that ENaC is composed of nine subunits (Cheng et al., 1998). More recently, Staruschenko et al proposed higher order multimers with equal numbers of all three subunits (Staruschenko et al., 2005; Staruschenko et al., 2004). Stewart et

al., developed a method based on atomic force microscopy, using which, they visualize the complexes between isolated ENaC and antibodies/Fab fragments directed against specific epitope tags on different subunits. Their results indicate that pairs of antibodies decorate channels made of α, β and γ ENaC only at an angle of 120° with respect to each other. Using their method, they concluded that ENaC could form homo- and heterotrimers and higher order multimers with upto nine subunits (Stewart et al., 2011).

In summary, it is clear from the studies discussed above that a consensus regarding the subunit stoichiometry of ENaC has not yet been reached. Until the structure of ENaC is solved, further studies using new approaches are necessary to resolve the controversy on the stoichiometry of the subunits.

1.4 Activation of ENaC by proteases

Limited proteolysis is the last step in the attainment of a functional form of many proteins of biological significance and perhaps the first step in protein degradation (Neurath and Walsh, 1976). This important regulatory phenomenon is frequently observed in activation of many enzymes, hormones, receptors, and other biologically active proteins and is conserved through evolution (Neurath, 1984). ENaC is unique in its mode of regulation via cleavage by proteases, resulting in constitutive channel activation. Over the last decade it has become clear that an important determinant of ENaC activity is the extent of partial proteolysis of the channel subunits (Kleyman et al., 2009; Rossier and Stutts, 2008). Serine proteases, that constitute a large gene family encompassing ~2% of the identified genes, activate ENaC by cleaving at multiple sites in the

extracellular domain (Hedstrom, 2002; Rossier and Stutts, 2008). Proteolytic regulation of ENaC includes selective cleavage by furin-like convertases during biosynthetic maturation as well as cleavage at the cell surface by proteases that can be membrane-associated or soluble. The membrane-associated proteases that activate ENaC are termed channel-activating proteases (CAPs). Vallet et al., obtained first evidence of an epithelial membrane protease activating ENaC in an autocrine fashion (Vallet et al., 1997). The cloned channel-activating protease 1 (CAP-1), also called prostasin, from A6, a *Xenopus* kidney cell line, and established that it activaes ENaC when coexpressed in *Xenopus* oocytes (Vallet et al., 1997). Subsequently, two additional membrane-associated serine proteases, transmembrane protease serine 4 (TMPRSS4)/CAP2 and MT-SP1/Matriptase/Epithin/CAP3 were identified and found to increase the activity of ENaC when coexpressed in oocytes (Vuagniaux et al., 2002). The open probability (P_O) of near-silent ENaC channels can be increased upto 28-fold upon cleavage by trypsin or human neutrophil elastase (hNE) (Caldwell et al., 2004, 2005). Activation of ENaC by proteases results from combinatorial processing of α and γ subunits (Hughey et al., 2004; Hughey et al., 2003). Hughey et al., reported that the basic amino acid sequence following the consensus K/R-X-X-R is required for cleavage by furin-like convertases (Hughey et al., 2004). Using site-specific mutagenesis of the C-terminal arginine (P1 site) to block cleavage, they showed that ENaC cleavage correlates with channel activity (Hughey et al., 2004). Alternate approaches to study proteolytic processing of ENaC uncovered a 26 amino acid inhibitory tract in the α-subunit of ENaC, which when removed,

activated the channel (Carattino et al., 2008b). Until recently, the α-subunit was thought to play a dominant role in activation of ENaC. Recently, Carattino et al., achieved full activation of ENaC upon cleavage at the furin sites in the γ-subunit in a manner independent of cleavage of the equivalent sites in the α-subunit (Carattino et al., 2008a). This result shifted focus towards the γ-subunit and the current model for proteolytic regulation of ENaC is that the P_O of the heteromultimer is determined by cleavage events, so far confined to its α and γ subunits. In general, channels made of uncleaved subunits exhibit a very low P_O and a range of cleavage events increases P_O. These cleavage events appear to be more prominent in the γ-subunit than the α-subunit (Carattino et al., 2008a). Further evidence to the importance of the γ-subunit comes from the study conducted by Garcia-Caballero et al., with CAP2 where they characterized cleavage of all three subunits of ENaC by CAP2/TMPRSS4. They reported that cleavage of the γ-subunit led to most increase in sodium current (I_{Na}) compared to cleavage of the other two subunits (Garcia-Caballero et al., 2008).

1.5 Structural aspects of functional regulation of ENaC

Function of ENaC is regulated by many factors including hormones, phosphoinositides and post-translational modifications. The N- and C-termini of ENaC are critical for such regulation of ENaC activity. The N- and C-termini of ENaC are ~60-100 amino acid fragments. The N-terminus has several key functions including but not limited to subunit assembly, gating, and endocytic retrieval and degradation of ENaC (Adams et al., 1997; Chalfant et al., 1999; Grunder et al., 1997; Prince and Welsh, 1998). It is challenging to understand the

combinatorial role of the N-termini from different subunits in modulating the regulation of ENaC. Early evidence for the importance of the N-termini in regulation of channel function came from studies with deletion mutants lacking the entire N-terminal fragments. Deletion of the entire N-temrinus of α-, β-, or γ-ENaC completely eliminated amiloride-sensitive sodium currents in oocytes (Benos et al., 1995; Benos and Stanton, 1999). Sequence analysis suggests the possibility of myristoylation at two potential sites in the N-terminus. Recent studies using fatty acid exchange chemistry demonstrated that the N- and C-termini of the β and γ but not the α subunit are palmitoylated at more than one cysteine residue (Mueller et al., 2010). Site-specific mutagenesis of the palmitoylated cysteines and expression in *Xenopus* oocytes resulted in decreased amiloride-sensitive whole cell currents, enhanced sodium self inhibition and reduced single-channel P_O, with no apparent effect on surface expression and membrane trafficking (Mueller et al., 2010). It is widely accepted that the short lifetime of ENaC at the cell surface is related to modification by ubiquitination and consequent degradation. Several conserved lysine residues in the N-termini of α- and γ-ENaC are targets for ubiquitination and consequent internalization and recycling (Staub et al., 1997). Mutation of these conserved lysines to arginines resulted in elevated channel function when expressed in oocytes (Staub et al., 1997). Using the quantitative assay developed by Firsov et. al., Staub et al., demonstrated that increase in ENaC activity is a consequence of increase in number of channels at the surface (Firsov et al., 1996; Staub et al., 1997). Using yeast two-hybrid screen of the rat lung library of small peptide

fragments containing the PXXXY (PY-motif), frequently mutated or deleted in Liddle disease, Staub et al., identified Nedd4 as a ubiqutin ligase responsible for ubiquitination of ENaC (Hansson et al., 1995; Inoue et al., 1998; Shimkets et al., 1994; Staub et al., 1996). The N-termini of ENaC are also important for channel gating. Grunder et al., demonstrated that site-specific substitutions at a highly conserved glycine (αG95S, βG37S, γG40S – according to the amino acid sequence of rat ENaC) in the N-terminus of all three subunits drastically decreases amiloride-sensitive sodium currents by decreasing the P_O and altering the channel open and closed times (Chang et al., 1996; Grunder et al., 1997). The involvement of βG37S in the pathophysiology of PHA-1 highlights the importance of the N-termini of ENaC in channel gating.

Besides post-translational modification of the N- and C-termini of ENaC, interaction with phosphoinositides has been shown to regulate native and expressed sodium channels. Anionic phospholipids, normally located in the inner leaflet of the plasma membrane interact with ENaC to modulate its function. Among the anionic phospholipids, phosphotidylinositol 4,5-bisphosphate, commonly known as PIP2, is a minority phospholipid that modulates functional regulation of many ion channels and membrane receptors (Suh and Hille, 2008). PIP2 has been shown to stimulate ENaC activity in A6 cells (Yue et al., 2002). Ma et al., demonstrated the dependence of function of ENaC on presence of anionic phospholipids (Ma and Eaton, 2005; Ma et al., 2002). Using patch clamp techniques, Ma et al., showed that addition of anionic phospholipids such as phosphatidylinositol (4,5)-bisphosphate (PIP2), phosphatidylinositol (3,4,5)-

trisphosphate (PIP3) and phosphatidylserine (PS) to the cytoplasmic side of the patch increases sodium currents (Ma et al., 2002). Conversely, depletion of PIP2 using various endeavors like chelation with anti-PIP2 antibody, hydrolysis with exogenous phospholipase C (PLC) or activation of endogenous PLC, activated channel rundown and decreased I_{Na} (Ma et al., 2002). Using lipid-protein overlay studies, Zhang et al., determined that anionic phospholipids differentially regulate the function of ENaC by interacting with the α, β and γ subunits (Zhang et al., 2010). The basic amino acids immediately distal to the second transmembrane domain of β- and γ-ENaC have been identified as potential sites of interaction with PIP3, but not PIP2, resulting in channel activation (Booth et al., 2003; Pochynyuk et al., 2005; Pochynyuk et al., 2007). The highly conserved lysines in the N-terminus of β- and γ-subunits were hypothesized to interact with PIP2. Interestingly, these residues are common to two conflicting modes of regulation of ENaC; PIP2-mediated activation, ubiquitin-mediated degradation. Based on this observation, intracellular ubiquitination and extracellular cleavage were linked by a possible conformational change in the extracellular domain (Ruffieux-Daidie et al., 2008; Ruffieux-Daidie and Staub, 2011). Based on these findings, one can envision a complex series of events, probably guided by intramolecular allosteric signal propagation mediating regulation of ENaC activity at the cell

surface.

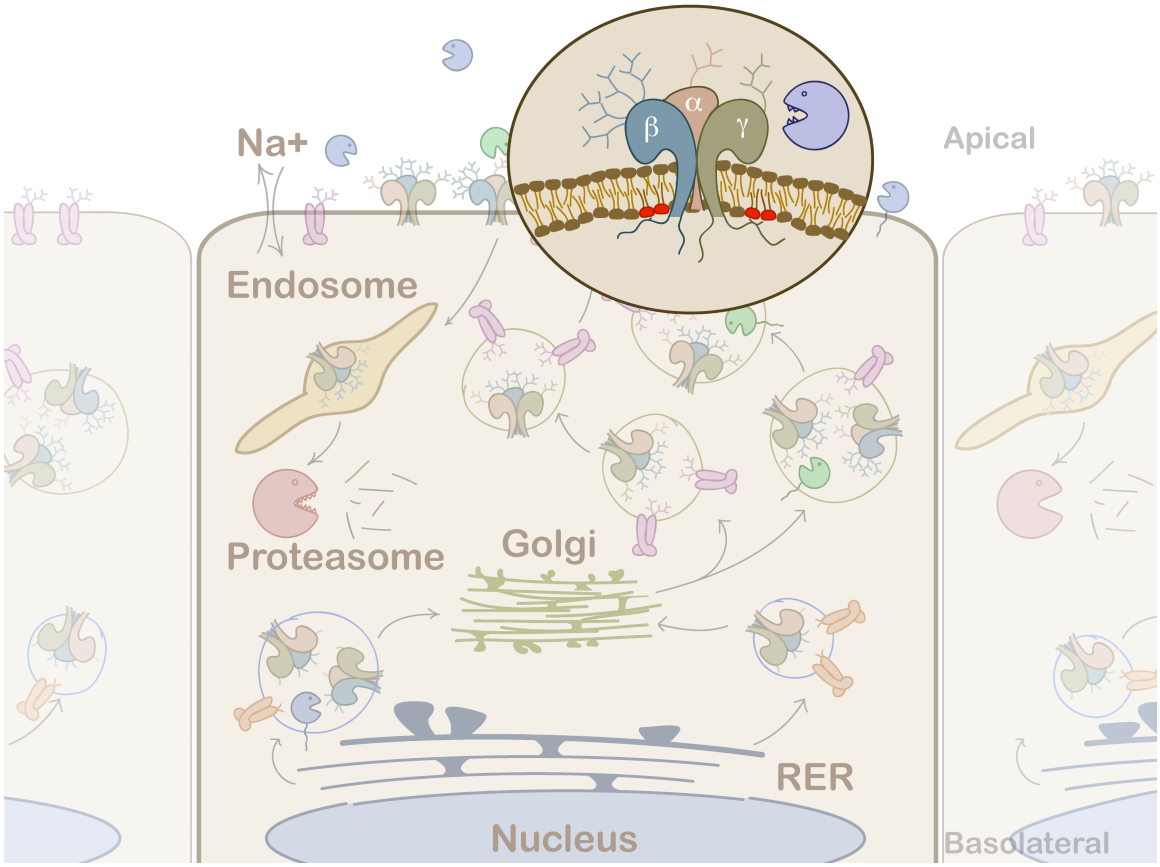


Figure 1.1 Activation and regulation of epithelial sodium channels.
Epithelial sodium channels are synthesized in the ER, trafficked to Golgi and to the apical plasma membrane, where they are activated by proteases either soluble or membrane-associated. Intracellular fragments interact with anionic phospholipids and mediate channel activation.

1.6 Motivation

In the current study, we addressed the structural basis for regulation of function of ENaC. Attempts to crystallize ENaC are averted by its transmembrane segments and dependence on expression levels of more than one subunit in the chosen expression system. Low-resolution structural studies are not as useful since understanding the functional aspects of a protein requires atomistic details of its structure. A promising alternative is homology based

model development, which has been used for successful drug screening and design in case of proteins that are difficult to crystallize, including the serotonin receptor (5-HT_{1A}) and CCR3 (an entry co-receptor for HIV1) (Becker et al., 2003; Flower, 1999; Hillisch et al., 2004; Klabunde and Hessler, 2002; Patny et al., 2006). Hence, we chose to build a homology model of ENaC based on the 3-dimensional co-ordinates of the recently crystallized chicken acid-sensing ion channel (cASIC) (Jasti et al., 2007)– a close homolog of ENaC. We set out to construct structural models of trimeric and tetrameric assemblies of ENaC. To experimentally address the controversy regarding the subunit composition of ENaC, we used blue native and clear native polyacrylamide gel electrophoresis (BN-, CN-PAGE) and biochemically characterized the assembly of ENaC in its native conditions.

Two specific gaps in understanding proteolytic activation of ENaC involve the dominant role of γ-ENaC in proteolytic regulation of the channel. First, studies of at least five ENaC activating proteases identify essential sites in γ-ENaC, despite accepted importance of furin-like cleavage sites in α-ENaC. Second, the structural and/or energetic features of the region of γ-ENaC (residues 130-200), that renders it susceptible to cleavage by multiple proteases, have not been investigated. Previous studies on membrane associated proteases (MAPs) including prostasin/CAP1 and TMPRSS4/CAP2 revealed an important role for γ-ENaC in channel regulation via limited proteolysis (Adachi et al., 2001; Bruns et al., 2007; Garcia-Caballero et al., 2008). Additionally, the extracellular domain of γ-ENaC has been hypothesized to harbor an allosteric regulatory subdomain with

an important role in channel function (Winarski et al., 2010). Interestingly, proteolytic activity of CAP1 is not required for its stimulation of ENaC. This observation led to a suggestion that CAP1 may play a critical non-catalytic role in ENaC regulation, perhaps as part of a cascade of surface associated proteases (Vuagniaux et al., 2002). It was recently demonstrated that matriptase/CAP3 is a critical activator of CAP1 (List et al., 2007). Although CAP3 is believed to cleave and activate ENaC based on its homology to CAP1 and CAP2, the structural basis for such cleavage has not been studied. Here, we focus on the structural and energetic bases for activation of rat ENaC by CAP3-mediated cleavage of the γ subunit.

The heteromultimeric organization of ENaC gives rise to a channel that spans tens of angstroms in the extracellular environment. From structural homology, it is clear that the protease cleavage sites in ENaC are located in the extracellular domain, perhaps distant from the ion-selective pore. A fundamental question concerning activation of ENaC is the mode of signal propagation from the extracellular cleavage sites to the pore and the molecular details involved in such long-range communication. We hypothesized that ENaC undergoes significant conformational change in the extracellular domain that would eventually lead to channel activation. Here, we have developed a computational method to map the residues involved in allosteric signal propagation in proteins. Using our approach, we have identified critical sites in ENaC that mediate such long-range communication within the channel.

Overall, using a combination of computational and experimental techniques, we have attempted to understand the molecular details underlying the regulation of function of ENaC. We believe that the scientific advance achieved through this study will benefit a finer understanding of the molecular aspects of epithelial sodium channel structure, function and regulation.

CHAPTER 2

Materials and methods

2.1 Homology Model Building

Medusa is a comprehensive protein design toolkit developed in our laboratory (Ding and Dokholyan, 2006). Medusa uses Monte-Carlo based procedure to optimize the side-chains for a given backbone. The major strength of Medusa over other homology-modeling servers is its ability to efficiently sample the rotamer and sub-rotamer spaces of amino acid side-chains. Medusa has been successfully used in the past to model the structure of CFTR (Serohijos et al., 2008), dynein (a cytoseletal motor protein) (Serohijos et al., 2006) and the pore of the Ryanodine receptor (Ramachandran et al., 2009). The details regarding the protocol used for minimization and the force field parameters are described previously (Ding and Dokholyan, 2006; Ramachandran and Dokholyan, 2011).

2.2 Structure Refinement

Generation of surface dots. We use our implementation of the algorithm originally proposed by Le Grand and Merz (Le Grand and Merz, 1993) to compute solvent accessible and molecular surface area of proteins. The

algorithm represents each atom as a set of dots placed on the surface of the atom. For improved accuracy, we used 4096 dots to represent the surface of each atom compared to 256 in the original implementation. For scale-up, we represent the dots as pairs of spherical angles θ and Φ. In our convention, θ=[0,π] and Φ=[0,2π). We first generated dots on the surface of a unit sphere by randomly choosing θ and Φ values within their respective domains. We then performed Monte Carlo-based simulated annealing to minimize the following cost function using the Metropolis criterion,

$$W = \sum_{i<j}^{N} \frac{1}{d_{ij}^2} \quad (1)$$

where N is the number of dots on the surface and d_{ij} is the Euclidian distance between the dots i and j.

Solvent accessible surface area (SASA). We define the SASA of a protein as the area covered by the center of a solvent sphere, as it rolls over the protein surface. Considering the radius of the solvent sphere to be 1.4 Å (radius of one water molecule), we obtain SASA by calculating the surface area of the protein, when the radii of all its atoms are increased by 1.4 Å. We use our implementation of the algorithm proposed by LeGrand and Merz (Le Grand and Merz, 1993) for calculating SASA, where surface of each atom is represented by 4096 dots and boolean masks are used to delineate buried and exposed dots on each atom. The reported SASA of a protein therefore includes the surface area of the voids (if any) in the protein. We have modified the algorithm to ensure uniform

distribution of masks on the surface of a unit sphere. We define and use a metric h_{ij} given by,

$$h_{ij} = 1 - \cos\theta_{ij} \quad (2)$$

for generating masks instead of,

$$d_{ij} = \sqrt{2\left(1 - \cos\theta_{ij}\right)} \quad (3)$$

as proposed by LeGrand and Merz (Le Grand and Merz, 1993). We identify the dot closest to the point *D* on the line joining the centers of two atoms *i* and *j* and retrieve the appropriate mask to determine the fraction of surface of atom *i* not buried by atom *j*. We repeat this process for all atom pairs to determine the exposed surface of each atom. The surface area of the protein can then be computed by summing up the fractional surface areas contributed by individual atoms.

Molecular surface area (MSA). We define the MSA of a protein as the area covered by the edge of a solvent sphere, as it rolls over the protein surface. MSA is represented as a sum of three components - contact, toric and reentrant surfaces (Connolly, 1983).

Contact surface area: We compute the contact surface area using the same algorithm that we use for computing SASA, but without increasing the radii of atoms by the radius of the solvent. The contact surface area of the protein can be formally defined as:

$$A_c = \sum_{i=1}^{N} \frac{n_i}{D}(4\pi r_i^2) \quad (4)$$

where n_i and r_i are the number of exposed dots, and the radius of atom i, respectively, and D is the total number of dots on the atom, set to be 4096. A_c includes the contact surface area of voids (if any) in the protein core.

Toric surface area: We analytically calculate the toric surface area covered by the solvent probe on a pair of atoms i and j, using the following equation:

$$A_t = \sum_{i \neq j}^{N} A_{t,ij} = 2\pi \sum_{i \neq j}^{N} \tau_{ij} \left((r_i + r_w)\sin\theta_{ij} \left(\frac{\pi}{2} - \theta_{ij} \right) - r_w \cos\theta_{ij} \right) \quad (5)$$

where τ_{ij} is the fraction of the torus around the overlapping atoms i and j that is accessible to the solvent probe, r_i and r_w represent the radii of the atom i and the probe respectively, and θ_{ij} is the angle subtended by the atom j at the center of atom i. We compute τ_{ij} using edge masks as described by Bystroff (Bystroff, 2002). The solvent probe may roll over itself causing singularities in the toric surface. We treat such cases by computing the toric surface area by atoms i and j using the following equation when $(r_i+r_w)\sin\theta_{ij} < r_w$:

$$A_{t,ij} = 2\pi\tau_{ij} \begin{pmatrix} (r_i + r_w)\sin\theta_{ij} \left(\frac{\pi}{2} - \theta_{ij} - \arccos\left(\frac{(r_i + r_w)\sin\theta_{ij}}{r_w} \right) \right) \\ -r_w \cos\left(\theta_{ij} + \arccos\left(\frac{(r_i + r_w)\sin\theta_{ij}}{r_w} \right) \right) \end{pmatrix} \quad (6)$$

We use this algorithm to compute the toric surface area of all atom pairs including those forming voids (if any) in the protein core. Further details on the mathematical formulation are reported elsewhere (Bystroff, 2002).

Reentrant surface area: We apply the Gauss-Bonnet theorem to calculate the total reentrant curvature of the protein. Gauss-Bonnet theorem states that the total Gaussian curvature integrated over a closed manifold equals 2π times the Euler characteristic of the manifold. This theorem is applicable only if a normal can be generated unambiguously at every point on the surface of the manifold (orientable surface). Protein surfaces are orientable and hence the Gauss-Bonnet theorem can be used to calculate the Gaussian curvature of a protein. Since the Euler characteristic is geometrically invariant, the Gaussian curvature integral of a closed 3D surface, i.e. the Connolly molecular surface of a protein, is the same as that of a sphere and is equal to 4π. The total Gaussian curvature of the protein can be denoted as a sum of contact, toric and reentrant curvatures. However, proteins may contain voids, which are isolated continuous surfaces in the protein core. Each such void, if present, must be considered as an independent orientable manifold. Therefore, the reentrant curvature of the protein is given as

$$K_r = 4m\pi - K_c - K_t \quad (7)$$

where K_r is the total reentrant curvature integral and K_c, K_t represent the total contact and toric curvature integrals respectively. m represents the total number

of manifolds in the system including the solvent accessible surface and all the voids. The total contact curvature integral, K_c, can be obtained using,

$$K_c = 4\pi \sum_{i=1}^{N} \frac{n_i}{D} \qquad (8)$$

and total toric curvature integral K_t can be calculated by,

$$K_t = \sum_{i \neq j}^{N} k_{t,ij} = -2\pi \sum_{i \neq j}^{N} \tau_{ij} \cos\theta_{ij} \qquad (9)$$

Here, the curvature integral of toric surface is negative and we do not need to consider the overlapping of toric surfaces. K_r can be derived accordingly, which corresponds to the total reentrant curvature of the protein and voids since the contact and toric curvatures already take the corresponding curvatures from voids in the protein into account. We cluster all the exposed dots on the surface of all exposed atoms to obtain the number of independent manifolds within the protein. Our approach is different from that employed by MASKER (Bystroff, 2002) in that we derive the total re-entrant surface analytically using the global Gauss-Bonnet theorem instead of computing individual re-entrant surfaces, which have additional sources of errors. The reentrant surface area of the protein is then given by

$$A_r = K_r r_w^2 \qquad (10)$$

Void volume. We define voids as those internal cavities in the protein core that are inaccessible to the bulk solvent, but feature a volume greater than or equal to

at least one solvent molecule. We define void volume as the volume of such internal cavities inaccessible to the bulk solvent. To compute void volume, we first use our modified implementation of the algorithm proposed by LeGrand and Merz, to obtain all the dots on the surface of each atom that is not buried by other atoms. These exposed dots could either belong to the surface or internal voids in the protein. We identify voids by performing single-linkage clustering on these exposed dots using the distance between them as the clustering criterion. This process yields one large cluster corresponding to the solvent accessible surface and zero or more small clusters each corresponding to an internal void (Supplementary Figure S2). Since we increase the radius of each atom by the radius of a water molecule (1.4 Å) before void identification, the minimum volume of the identified voids is equal to the molecular volume of water. We delineate the volume of each identified void into i) solvent excluded volume – the region from the surface of the atoms to the surface traced out by the center of the solvent sphere as it rolls on the atoms lining the void and ii) solvent accessible volume – which is accessible to the solvent, should a solvent molecule be able to approach this space within the protein.

(i) Solvent excluded volume: The solvent excluded volume is composed of three components: the contact volume (fractional volume accessible to the solvent probe touching only one atom), the toric volume (fractional volume inaccessible to the probe touching two atoms at a time) and the reentrant volume (fractional volume inaccessible to the probe when it touches three atoms simultaneously).

The fractional volume accessible to the probe touching only one atom can be mathematically computed using,

$$V_c = \frac{4\pi}{3}\sum_{i}^{N}\frac{n_i}{D}\left((r_i + r_w)^3 - (r_i)^3\right) \quad (11)$$

v where V_c is the total contact volume, D represents the number of dots on the atom i, n_i is the number of exposed dots facing the void, N is the number of atoms lining the void, r_i is the radius of atom i and r_w is the radius of the probe. To calculate the toric volume of the void, we performed analytical integration to arrive at the following equation:

$$V_t = 2\pi\sum_{i\neq j}^{N} V_{t,ij} = 2\pi\sum_{i\neq j}^{N}\tau_{ij}\left[(r_i + r_w)\sin\theta_{ij}\left(\frac{\pi}{2} - \theta_{ij}\right)\frac{r_w^2}{2} - \cos\theta_{ij}\,\frac{r_w^3}{3}\right] \quad (12)$$

where the terms represent the same quantities as in A_t. To account for singularities, we used different lower limits for integration when $(r_i+r_w)\sin\theta_{ij} < r_w$, generating the following equation.

$$V_{t,ij} = \tau_{ij}\left[\begin{array}{l}(r_i + r_w)\sin\theta_{ij}\left(\frac{\pi}{2} - \theta_{ij} - \arccos\left(\frac{(r_i + r_w)\sin\theta_{ij}}{r_w}\right)\right)\frac{r_w^2}{2} \\ -\cos\left(\theta_{ij} + \arccos\left(\frac{(r_i + r_w)\sin\theta_{ij}}{r_w}\right)\right)\frac{r_w^3}{3}\end{array}\right] \quad (13)$$

We compute the total reentrant curvature for each identified void as described above. The reentrant volume can then be computed using

$$V_r = \frac{4\pi}{3} r_w^3 K_r \qquad (14)$$

where K_r is the total reentrant curvature for the void. Since the void is a single orientable manifold, we do not perform manifold correction in calculation of void volume, as we perform while calculating the molecular surface area.

(ii) Accessible void volume: We calculate the accessible void volume by numerical integration: we iteratively increment the radii of all the atoms (starting from atom radius plus solvent radius) forming the void by 0.01 Å and sum up the surface area of these voids times 0.01 at each increment till the area converges to zero. The total void volume is then obtained by summation of the independent components of solvent excluded volume and accessible void volume.

Unsatisfied hydrogen bond donor/acceptor. We define a polar nitrogen/oxygen atom as an unsatisfied hydrogen bond donor/acceptor if it is buried from the solvent and is not involved in a hydrogen bond. If a polar atom belongs to a residue whose total SASA is zero, it is marked as buried. On the other hand, if the polar atom itself is buried, but the residue it belongs to features a non-zero SASA, rotamer changes/side chain dynamics could expose the polar atom, and thus, the polar atom is classified as being in the shell: an intermediate layer between buried and solvent accessible regions of the protein. We first build all hydrogen bonds in a given protein structure using Medusa's directional hydrogen-bond potential (Ding and Dokholyan, 2006; Yin et al., 2007b), and then list all the buried/shell polar atoms that do not form hydrogen bonds.

Bond lengths, angles, torsions and side chain rotamers. To ensure the robustness of the covalent geometry of the input protein structure, we also calculate bond lengths, angles, backbone torsions and side chain rotamers to detect outliers. For side chain integrity, the nearest rotamer in the Dunbrack library (Dunbrack and Cohen, 1997) for a given side chain is determined, and then, the p-values of each of the applicable chi-angles of the given side chain with respect to the identified standard rotamer is calculated. A p-value less than 0.05 is reported as an outlier and presented in the output for a protein structure on the web server. The bond lengths for all standard bonds were calculated from our high resolution dataset, and the mean and standard deviation from the resulting distributions were used in determining p-values for bond lengths of the input structure. For bonds with standard deviation less than 2.5% of the mean (as calculated from the standard distribution), we reset the standard deviation as 2.5% of the mean. We reset the standard deviation because the force constants for the bonded term in the MD force fields allow between 2.5-4% deviations in bond lengths at 300K. Thus, to report realistic outliers in modeled structures, we require the standard deviation to be at least 2.5% of the mean. Similar analysis was performed for angles and the omega dihedral of the protein backbone. For the Φ-ψ dihedrals, a two-dimensional histogram with bin width of 2° was constructed combining all amino acid types excluding proline and glycine. A separate histogram was constructed for proline. In the input structure, residues whose Φ-ψ values belong to a lowly populated bin (roughly less than 2.5% of the population) are designated as outliers. The outliers in terms of Φ-ψ dihedrals in

an input structure are plotted on top of the heat map of the two-dimensional histogram.

Definition of steric clashes and the acceptable clash-score. We define a steric clash in a protein as any atomic overlap resulting in Van der Waals repulsion energy greater than 0.3 kcal/mol (0.5 k_BT), except i) when the atoms are bonded, ii) when the atoms form a disulfide bond or a hydrogen bond (i.e. the heavy atoms are involved in the hydrogen bond; we assign the Van der Waals radius of hydrogen to be zero), iii) when the atoms involved are backbone atoms and have separation of 2 residues (in order to accommodate the formation of tight turns). We calculate the Van der Waals repulsion energy using the non-bonded parameters from the CHARMM19 force field, which are identical to CNS parameters except for carboxyl oxygen atoms. Since clashes are local structural artifacts, we reduce the search time and space by restricting the search to the local environment of a given atom. We determine clashes using the above definition by constructing a grid around the protein with the dimension of each cell larger than the largest Van der Waals interaction distance between any two atom pairs (~4.5 Å) and walking along the chain to check if the overlap of the atom under consideration with the heavy atoms in the same or adjacent cells leads to a clash. We then define ‘contacts’ as the number of such overlaps tested. The clash-energy of a protein is the sum of Van der Waals repulsion energy of all the clashes in the protein’s structure. In order to arrive at a descriptor independent of protein size, we define the clash-score, which is the clash-energy divided by the number of contacts. Thus, clash-score describes the

clashes present in a protein-structure, but is independent of the size of the protein. To estimate the permissible Van der Waals repulsion in a given structure, we determine the clash-scores of high-resolution crystal structures (see below). The distribution of these clash-scores indicates the extent of clashes permissible in proteins as a consequence of tight packing. A clash-score that is deviant from the distribution for high-resolution structures would then point to clashes that are artifacts of model building rather than those inherent to the protein structure. Clash-score is acceptable if it is less than one standard deviation away from the mean on the higher side of the distribution of clash-score of high-resolution dataset of structures (which would include ~84% of the proteins in the dataset). From the distribution of clash scores of structures from the high-resolution dataset, we calculate the acceptable clash-score to be 0.02 kcal.mol^{-1}.contact^{-1}.

Protein datasets. In order to understand the extent of clashes in protein structures and arrive at an acceptable clash-score, we constructed datasets of protein structures of various resolutions. We obtained two sets of protein structures from the Protein Data Bank (PDB)(Berman et al., 2000) and one set of protein structures from Swiss-model repository. The sets obtained from PDB correspond to a high-resolution set (0-2.5 Å) and a low-resolution set (2.5-3.5 Å). The high-resolution set was used to arrive at the acceptable clash-score. Our high-resolution dataset comprises protein structures determined by X-ray crystallography with a reported resolution less than 2.5 Å. Other than protein chains, these structures did not contain any other biomolecules (i.e.

ligands/DNA/RNA). We then split these structures into individual peptide chains and clustered them based on sequence similarity. We used individual chains because we wanted clash statistics of globular proteins and not interfaces. We considered only one representative chain from each cluster of sequences that were at least 80% similar to each other, thus creating a dataset of 4495 unique chains. We further filtered the dataset based on radius of gyration to remove non-globular peptides/proteins from the dataset. The final dataset consisted of 4311 single chains at least 25 residues long. We used Medusa to accurately place any missing side-chain atoms in these structures.

We obtained a low-resolution dataset from PDB in order to explore if clash-score was worse in low-resolution structures compared to high-resolution structures (Results). The lower-resolution dataset contains 2942 unique protein structures determined using X-ray crystallography with a resolution between 2.5Å and 3.5Å. In addition to these two datasets, we obtained a set of 1000 homology models from the Swiss-model repository of random swiss-model entries (using the CGI-perl script provided by expasy: http://www.expasy.org/cgi-bin/get-random-entry.pl?S). We filtered these structural models based on radius of gyration resulting in a final dataset of 931 structural models.

2.3 Computational Methods

Minimization using DMD. The DMD simulation methodology is described in detail elsewhere (Ding, 2008; Dokholyan et al., 1998). DMD is a special type of molecular dynamics (MD) algorithm, which uses square-well potentials instead of continuous potentials. Thus, in DMD we seek to solve the ballistic equations of

motion instead of Newtonian equations of motion in a system of particles. We use CHARMM19 non-bonded potentials, EEF1 implicit solvation parameters (Lazaridis and Karplus, 1999) and geometry-based hydrogen bond potentials in DMD (Ding, 2008) to model various macromolecular interactions. The time unit of the all-atom DMD simulations is ~50 femtosecond (fs) and the temperature is maintained using Anderson's thermostat. The rate of velocity rescaling (for maintaining temperature) depends on the simulation we perform; in the present study we used either 200 ps^{-1} or 4 ps^{-1} as the rescaling rate.

Enzyme-peptide Docking. We represented γ-ENaC cleavage sites tentatively identified by mutagenesis by 8-mer peptides with the putative wild type P4-P1 cleavage sequence contained in the first four residues. We constructed three linear 8-mer peptides from rat γ-ENaC (Seq1: 135-RKRREAGS; Seq2: 178-RKRKISGK; Seq3: 132-KESRKRRE) such that the P1 site is between the fourth and fifth residues. We chose the initial configuration for the peptide by random placement at sites distant from the active site of either enzyme. We imposed two distance constraints – one between the epsilon nitrogen atom (NE2) of active site histidine and backbone amine of the P1' site on the peptide and the other between the gamma oxygen atom (OG) of active site serine and carbonyl oxygen of the P1 site – to draw the peptide close to the active pocket. In order not to bias the configuration of the peptide in the active site of the enzyme, we placed the peptide at ten different, randomly chosen starting positions and with different orientations with respect to one another. We performed replica exchange DMD simulations of each such initial configuration with eight replicas in a temperature

range of 0.35 to 0.75 reduced units at increments of 0.035 units (Ding, 2008). The total simulation time was 10^6 DMD time units. Each DMD time unit is approximately 50 fs in real time, accounting for a total simulation time of 50 ns per replica. The relationship between DMD time unit and real time is discussed elsewhere (Sharma et al., 2007). DMD uses Medusa, a CHARMM-based forcefield to treat interactions between atoms in the macromolecule. We used EEF1, an implicit solvation model to treat the solvation of the simulation system (Lazaridis and Karplus, 1999). Energy minimization of the crystal structure of furin and matriptase was performed using Chiron prior to replica exchange simulations (Ramachandran et al.). Proteases were maintained static during simulations allowing movement of only the loops surrounding the active site. We selected snapshots across simulation trajectories of all replicas that satisfied both the distance constraints and clustered them based on pair wise root mean square deviation (RMSD) of Cα atoms. We selected the representative structures from five such clusters and performed side-chain optimization using the fixed backbone custom design protocol from the Medusa suite (Ding and Dokholyan, 2006).

Peptide Disorder Prediction. We used Disopred2 (http://bioinf.cs.ucl.ac.uk/disopred) to analyze the peptide sequences of alpha, beta and gamma subunits of rat ENaC. We used the sequences L150-L290 from rat α-ENaC, K117-P240 from rat β-ENaC and K91-S222 from rat γ-ENaC.

2.4 Biochemical Characterization

Crude membrane preparation. MDCK or 3T3 Cells expressing ENaC were grown to ~95% confluency in 15 cm dishes, washed twice with ice cold PBS and scraped using a cell scraper in lysis buffer containing 2x protease inhibitor cocktail and incubated on ice for 10 min. Cells were then lysed using a Dounce homogenizer by applying 2-6 strokes depending on the cell type. Lysis was monitored using trypan blue stain and hemocytometer under a microscope. After 50-70% of lysis was achieved, equal volume of 0.5 M sucrose solution was added to the homogenizer. Lysates were centrifuged in an ultra centrifuge at 5000 rpm using an SW28 rotor for 15 min to remove the cell debris and organelles. The supernatant was centrifuged at 25000 rpm for 1 hr at 4 ^{0}C, the pellet was resuspended in ice cold buffer with protease inhibitors, aliquoted and stored at -80 ^{0}C until further use.

Detergent solubilization. Crude membrane fractions were thawed on ice and protein content was estimated using a Bradford assay. Membrane fractions were solubilized in solubilization buffer (150 mM NaCl, 50 mM Tris-HCl at pH 7.4) containing 2x protease inhibitors and 1% detergent of choice, incubated on an end-on-end rocker for 1 hr at 4 ^{0}C. The mixture was centrifuged in a tabletop ultracentrifuge using a TLA-100.2 rotor for 1 hr. Supernatants were collected and the pellets were resuspended in the solubilization buffer. Efficiency of solubilization was estimated by subjecting the supernatant and resuspended pellet to SDS-PAGE separation followed by western blot analysis.

Native Polyacrylamide Gel Electrophoresis (Native-PAGE)

Crude membrane fractions isolated from MDCK/3T3 cells or oocytes were solubilized in LPG prior to separation by native PAGE. ~5 – 10 ng of solubilized protein were mixed with 1X native buffer (Invitrogen) and 0.25% G-250 (Invitrogen) and was loaded onto Bis-Tris 4-16% gradient gels (Invitrogen) and subject to electrophoresis at 150 V for 150-180 minutes at 4 ^{0}C. Proteins were transferred onto a 0.22 micron PVDF membrane (Millipore Immobilon-P) for 150 minutes at 4 ^{0}C prior to recognition using subunit specific antibodies. Antibodies against alpha, beta and gamma subunits of ENaC were purchased from Stressmarq.

Plasmid Preparation. For biochemical analyses of ENaC subunit proteolysis, cDNAs encoding rat α,β and γ-ENaC with HA-N-terminal (HA-NT) and V5-C-terminal (V5-CT) epitope tags were generated. Wild type and mutant constructs (α-ENaC; R205A/R231A, β-ENaC, γ-ENaC; R135Q/K136Q/R137Q/R138Q, R135Q/K136Q/R137Q/R138; R135Q/K136Q/R137/R138Q; R135Q/K136/R137Q/R138Q; R135/K136Q/R137Q/R138Q, hepatocyte growth factor activator inhibitor 1 and 2 (HAI-1, HAI2) and CAP3 were generated by PCR and cloned into pCR-BluntII-TOPO (Invitrogen), linearized (HindIII) and in vitro transcribed using T7 RNA polymerase. A PolyA tail was added after transcription (Ambion). Mutations were done with the Quikchange multi site-directed mutagenesis kit (Stratagene). The WT ENaC plasmids were generously provided by Dr. Bernard Rossier. The sequence of all plasmids was verified at the University of North Carolina sequencing facility. Plasmid preparation was performed by Yan Dang in Dr. Stutts' laboratory.

Western Blot Analysis. Proteins were extracted from oocytes as described above. Biotinylated and total proteins were solubilized by boiling in Laemmli sample buffer for 10 min prior to loading onto 4-12% SDS-PAGE gels. Western blots were performed with anti-V5 (Invitrogen), anti-HA (Covance), anti-CAP3 (Bethyl, Laboratories, Inc.) and anti-actin (Chemicon International) antibodies.

Functional Studies of ENaC in Xenopus oocytes. V-VI stage healthy oocytes were harvested as described previously (Donaldson et al., 2002) and maintained in modified Barth's solution (MBS) at 18°C. Animals were maintained and studied under protocols approved by the University of North Carolina Institutional Animal Care and Use Committee. Oocytes expressing the desired combinations of ENaC subunits and CAP3 were obtained as before (Garcia-Caballero et al., 2008). Briefly, cRNAs encoding wild-type (WT) of both untagged and HA-NT/V5-CT epitope tagged subunits or mutant HA-NT/V5-CT tagged subunits of rat α, β and γ-ENaC (0.3 ng each) and CAP3 cRNA (typically 1 ng) were co-injected into oocytes. Twenty-four hr after injection, two-electrode voltage clamping was performed using a Genclamp amplifier (Axon Instruments) in a constant perfusion system. Currents were measured in the presence and absence of 10 μM amiloride, with membrane voltage clamped to -100 mV. Currents were digitized and recorded using a Digidata 1200 A/D converter (Axon Instruments) and Axoscope software. After basal amiloride sensitive current (I_{Na}) was recorded by washing out amiloride, oocytes were superfused with amiloride containing buffer and trypsin or hNE (2-20 μg/ml) for 5 m, followed by a second determination of I_{Na}. All results are expressed as the mean ± S.E. or as fold

stimulation by CAP3 or hNE. The means of two groups were tested for significant difference using an unpaired Student's t test, differences between three or more groups were evaluated using ANOVA analysis (GraphPad Prism software). Proteins extracted from control and injected oocytes were analyzed by Western blots to verify expression of ENaC and actin. Functional measurements were conducted by Yan Dang and Hong He in Prof. Jackson Stutts' laboratory.

Surface labeling. Xenopus oocytes were injected with desired combinations of WT or mutant double epitope (HA/V5) tagged α, β , and γ rat ENaC subunits (0.3 ng each) and with or without CAP3 cRNA (1 ng). After 24 hrs, 70 oocytes per experimental condition were pre-chilled on ice for 30 minutes and labeled with 0.7 mg/ml sulfo-NHS-biotin in MBS-Ca++ (mM), 85 NaCl, 1 KCl, 2.4 NaHCO3, 0.82 MgSO4, 0.41 CaCl, 0.33 Ca(NO3), 16.3 hepes titrated to pH 8.0 with NaOH, while tumbling gently for 20 min at 4°C. Oocytes were washed twice with chilled MBS-Ca++ buffer and incubated in MBS-Ca^{2+} buffer with 100 mM glycine for 10 min at 4°C to quench free biotin. Oocytes were washed again three times with chilled MBS-Ca++ buffer, then lysed with lysis buffer (in mM; 20 Tris, 50 NaCl, 50 NaF, 10 β-glycerophosphate, 5 Na4P2O7 pyrophosphate, 1 EDTA, pH 7.5 containing protease inhibitors (complete, Roche), aprotinin (Sigma)). Cell lysates were prepared by passing oocytes through a 27G1/2 needle twice and by centrifugation at 3,600 rpm for 10 minutes at 4°C. Supernatants were transferred to new tubes and samples were spun at 14,000 rpm for 20 minutes at 4°C. Supernatants were discarded and pellets were solubilized in solubilization buffer (in mM; 50 Tris, 100 NaCl, 1% triton X-100, 1%

NP-40, 0.2% SDS, 0.1% Na deoxycholate, 20 NaF, 10 $Na_4P_2O_7$ pyrophosphate, 10 EDTA + protease inhibitor cocktail, pH 7.5). Total inputs were taken from whole cell samples representing 4% of total protein. Solubilized proteins were incubated with 100 μl of neutravidin beads (Pierce) overnight while tumbling at 4°C. Samples were washed twice with (mM) 500 NaCl 50 Tris pH 7.5 buffer and once with 150 NaCl 50 Tris pH 7.5 buffer. Laemmli buffer was added and samples were loaded on a 4-12% gradient Tris-glycine gel after incubation for 10 minutes at 96°C. Samples were transferred to 0.45 μm polyvinylidene difluoride (PVDF) membranes (Millipore) and Western blot analysis was performed using an anti-V5 (Invitrogen), anti-HA (Covance) and anti-actin (Chemicon International) monoclonal antibodies. Surface ENaC biotinylated fragments were quantified using the metamorph imaging 4.5 program (Hooker Microscopy Facility, University of North Carolina). Densitometry of selected bands was performed, using uninjected oocyte samples as background signal. Surface labeling studies were performed by Dr. Martina Gentzsch.

CHAPTER 3

Structural modeling and biochemical characterization

3.1 Homology model building of ENaC

The three subunits of ENaC share ~30% sequence similarity among themselves and nearly 35% with ASIC. We generated a sequence alignment of ENaC with ASIC using ClustalW2 (Larkin et al., 2007). The generated alignment was not satisfactory because the program failed to align many cysteine residues highly conserved across the DEG/ENaC superfamily. We obtained amino acid sequences for different subunits of ENaC from various species to construct a more extensive sequence alignment. We obtained sequences from 15 species and aligned them using ClustalW2. The alignment was then manually adjusted in low complexity regions to ensure that insertions avoid any secondary structural elements in ASIC. An excerpt from the final alignment used for model building is shown in Figure 3.1.

The alignment thus generated was used for model building using Medusa. The three-dimensional atomic coordinates from the crystal structure of ASIC were used as a template to build the homology model of the α subunit of ENaC

(Figure 3.2). The stretches of amino acids present in ENaC but not in ASIC (insertions) were included as loops using Modeler (Fiser et al., 2000; Fiser and Sali, 2003). The amino acids present in ASIC but not in ENaC (deletions) were removed and the resulting ends were computationally connected using short DMD simulations. The side chains for all amino acids were repacked 1000 times using Medusa to fix the optimal rotamers for all the residues, and the model with least total energy was chosen. Similar models were generated for the beta and gamma subunit of ENaC.

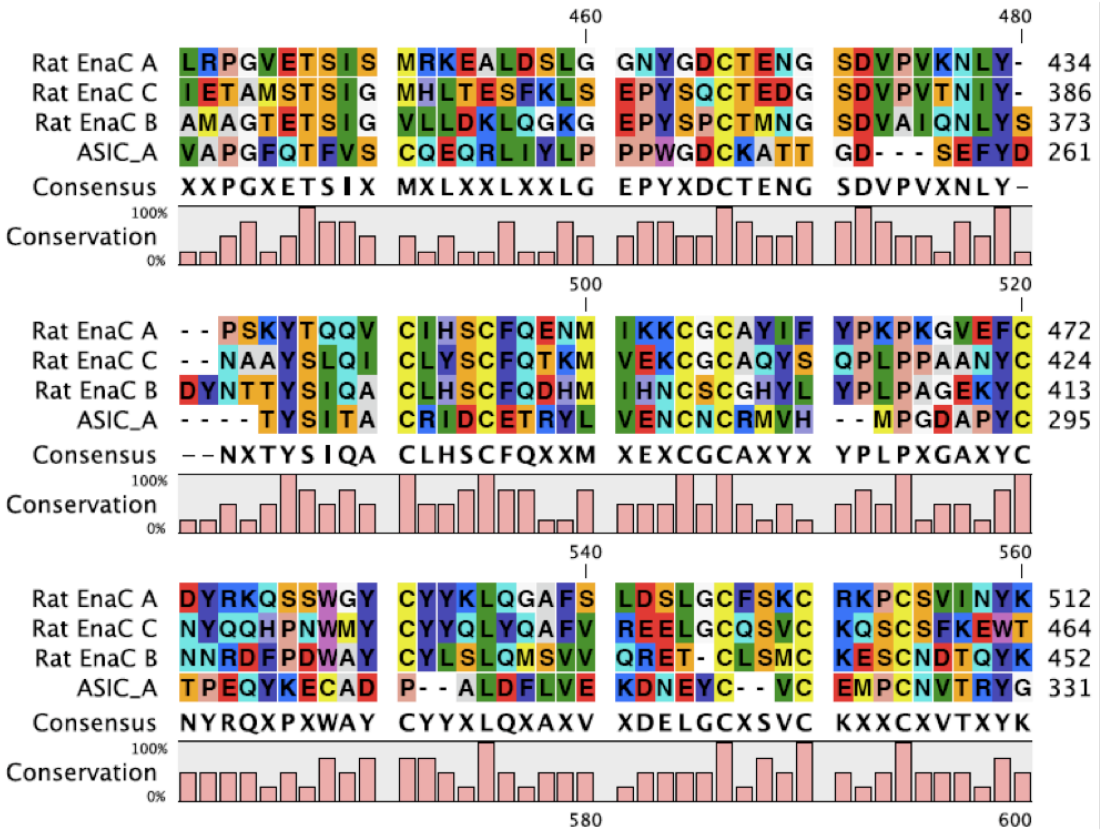


Figure 3.1: Sequence alignment of rat alpha, beta and gamma ENaC with chicken ASIC.
Sequence alignment was generated using ClustalW. The conserved cysteine residues are highlighted yellow.

The structure of the monomer is analogous to a hand holding a ball. As shown in figure 3.2A, the green helices form the thumb domain. Disulphide

bonds primarily contribute for the stability of this region of each subunit. The yellow beta strands form the palm domain and the red helices form the wrist domain. The helices colored purple and magenta form the fingers of the hand while the helices in teal form the knuckle domain. The orange beta twisted beta sheet in the middle of the protein is analogous to a ball and hence is called the beta ball domain. The three monomers were then assembled by aligning them structurally to their corresponding templates from cASIC.

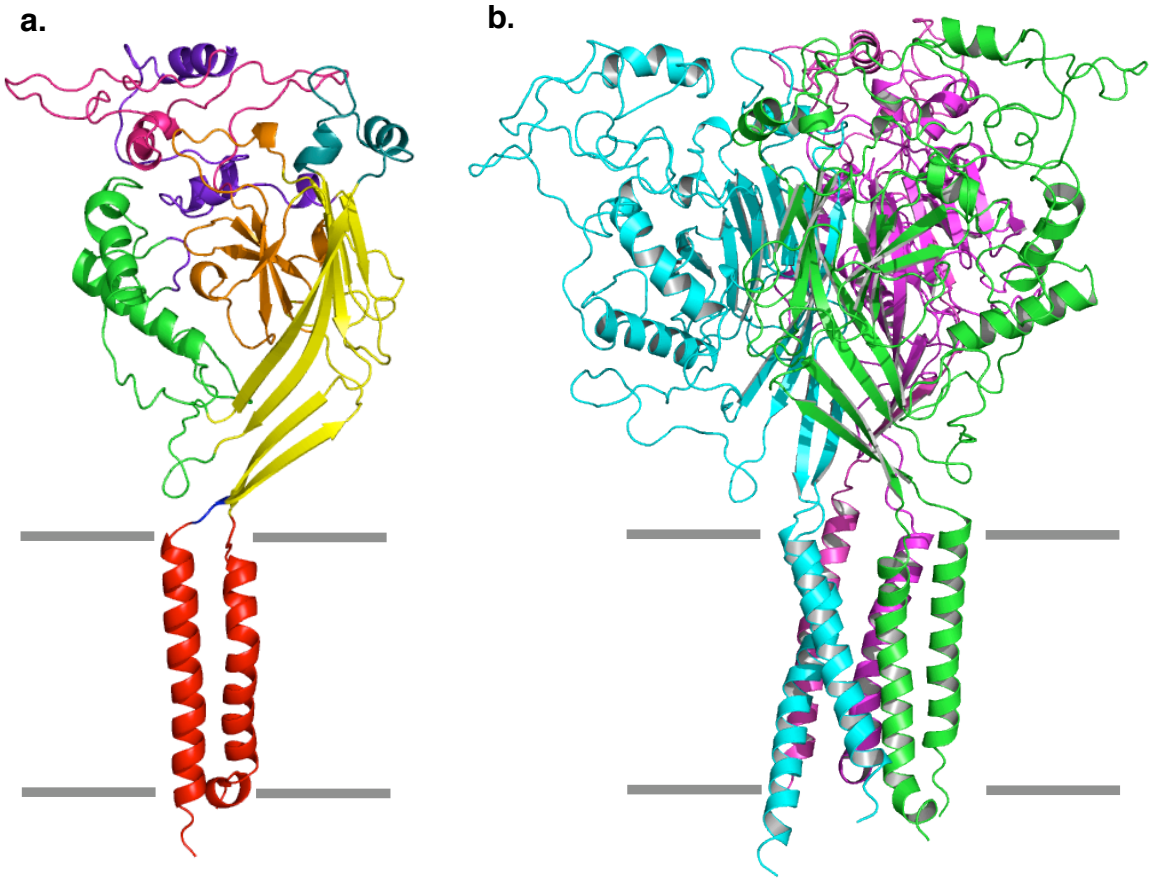


Figure. 3.2: Structural model of a-ENaC
a. Color scheme for domains is the same as used by Jasti *et. al*. for representing ASIC. TM helices are colored red. The palm domain is colored yellow; thumb colored green, finger colored purple, and the beta-ball is shown in orange. b. Structural model of the ENaC trimer. Individually modeled subunits were overlaid with the structure of ASIC to

generate the structure of the trimer of ENaC. Shown in green is α-ENaC, while β- and γ-ENaC are colored cyan and magenta respectively.

The three monomers in ASIC assemble using their palm domains to make asymmetric contacts with the thumb domain of the adjacent subunit. The structure of each subunit of ENaC differs from that of ASIC in that it forms an additional beta strand in the beta ball domain. Such strand formation in ENaC could impart additional stability to the structure of each subunit and also contribute to the stability of the trimer. Structural mapping of the precise cleavage sequence in the extracellular domains of different subunits suggests that the cleavage site in ENaC is present in the loop shown in magenta in figure 3.2 and is exposed to the solvent. In summary, we have built the structural model for the all three subunits of rat ENaC. Preliminary observations suggest that the structure of the alpha subunit recapitulates certain key features that have been previously observed experimentally.

3.2 Structual models of the N- and C-terminal segments

The N- and C-termini of each subunit of ENaC are indispensible for functional regulation of ENaC. Deletion of the termini has been shown to have adverse effects on channel activation and function. Recent studies have linked the N- and C-termini to the extracellular proteolytic processing of ENaC in an allosteric manner. Therefore, understanding the structural features of the N- and C-termini of ENaC is critical to understanding the molecular mechanisms involving the termini th activation of the channel. As described earlier, the N- and C-termini of all subunits of ENaC are of different lengths, intracellular and are relatively flexible compared to the rest of the protein. These fragments were

truncated in the constructs used by Jasti et al., for crystallographic studies of chicken ASIC (Jasti et al., 2007). In order to find a template structure to model the structures of these segments, we used the position-specific iterative basic local alignment search tool (PSI-BLAST) to search through the PDB database. We did not find any structures with a significant sequence similarity (at least 15%) to these segments. We modeled the structures of these segments *ab initio* using replica exchange DMD simulations (See methods). The outcomes of these simulations are reported in figure 3.3.

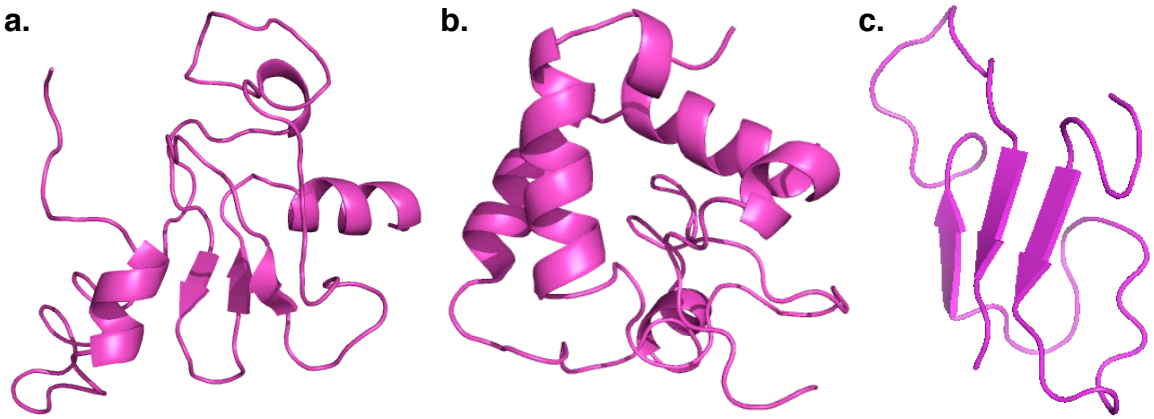


Figure 3.3. Structural models of N-terminal segments of ENaC
Shown here are the representative structures across eight replicas after ab initio folding simulations of N-terminal segments of alpha, beta and gamma ENaC. These structures are shown in panels a, b and c respectively.

The N-terminus of the beta subunit is mostly alpha helical while that of the gamma subunit is made of beta strands. The N-terminus of the alpha subunit features a mix of alpha helices and beta strands with a small three-strand beta sheet. In the reported model, the cysteines in the γ-subunit are surface-exposed and amenable for palmitoyl modification as observed by Mueller et al (Mueller et al., 2010). Furthermore, the cysteines are structurally close to the C-terminus of

the N-terminal fragment, such that the palmitoyl moiety can insert itself into the lipid bilayer in the fully assembled heteromultimer. Using circular dichroism spectroscopy and tryptophan fluorescence studies, Ismailov et al., showed that the C-terminal tails of β- and γ-ENaC have a propensity to form beta sheet structures (Ismailov et al., 1999). As predicted, we did not observe well-formed secondary structural elements in the C-terminal tails of ENaC, perhaps due to the presence of the proline-rich sequences forming the PY-motif.

3.3 Refinement of the structural model of ENaC

Steric clashes are one of the structural artifacts commonly seen in protein structural models. Current state-of-the-art tools for protein quality control identify clashes qualitatively, precluding an understanding of their possible energetic effects on protein structure. For instance, WHAT_CHECK (Hooft et al., 1996; Vriend and Sander, 1993) and Molprobity (Davis et al., 2007), commonly used in protein quality control, report a steric clash based on distances between two atoms with a distance cutoff for overlap set to 0.4 Å. However, the energetic penalty of such an overlap varies widely depending on the types of atoms involved in the clash (0-10 kcal/mol). We observe that low energy clashes are present even in high-resolution structures, however the number of severe clashes is very low. Thus, in order to correctly identify severe clashes, it is important to develop a quantitative measure to evaluate the effect of clashes present in a protein, and also it is necessary to benchmark the effectiveness of the measure by comparing against the extent of clashes seen in high-resolution crystal structures.

Several tools have emerged for resolution of such clashes upon identification. Steepest descent/Conjugate gradient minimization using all-atom Molecular Mechanics force fields is the most widely used method to resolve clashes in a protein structure before using the structure for further studies. However, minimization using Molecular Mechanics may not resolve severe clashes in some cases hampering subsequent Molecular Dynamics simulations. Molecular modeling tools like Rosetta are the alternate avenues for refining structures with severe clashes. These tools use knowledge-based potentials and small backbone moves to resolve clashes. However, these methods work best with smaller proteins (less than 250 residues in size) (Kaufmann et al.). Tools like MMTSB (Feig et al., 2004) and PULCHRA (Rotkiewicz and Skolnick, 2008) have emerged for structure refinement and for reconstruction of all-atom representation of proteins from Cα traces, which includes removal of clashes during refinement. We have developed an automated method for quantitative estimation and if required, resolution of clashes in a given protein structure. To accomplish the above, we developed a protocol using DMD simulations (Ding, 2008; Dokholyan et al., 1998). We also demonstrated that our protocol is more robust in comparison to other state-of-the-art tools widely used by the protein structural modeling community (Ramachandran et al., 2011). Using this method, we refined the structural model of ENaC and resolved steric clashes introduced during homology modeling building maintaining the backbone fixed (Figure 3.2b).

3.4 ENaC appears as both trimers and tetramers on the mammalian cell surface

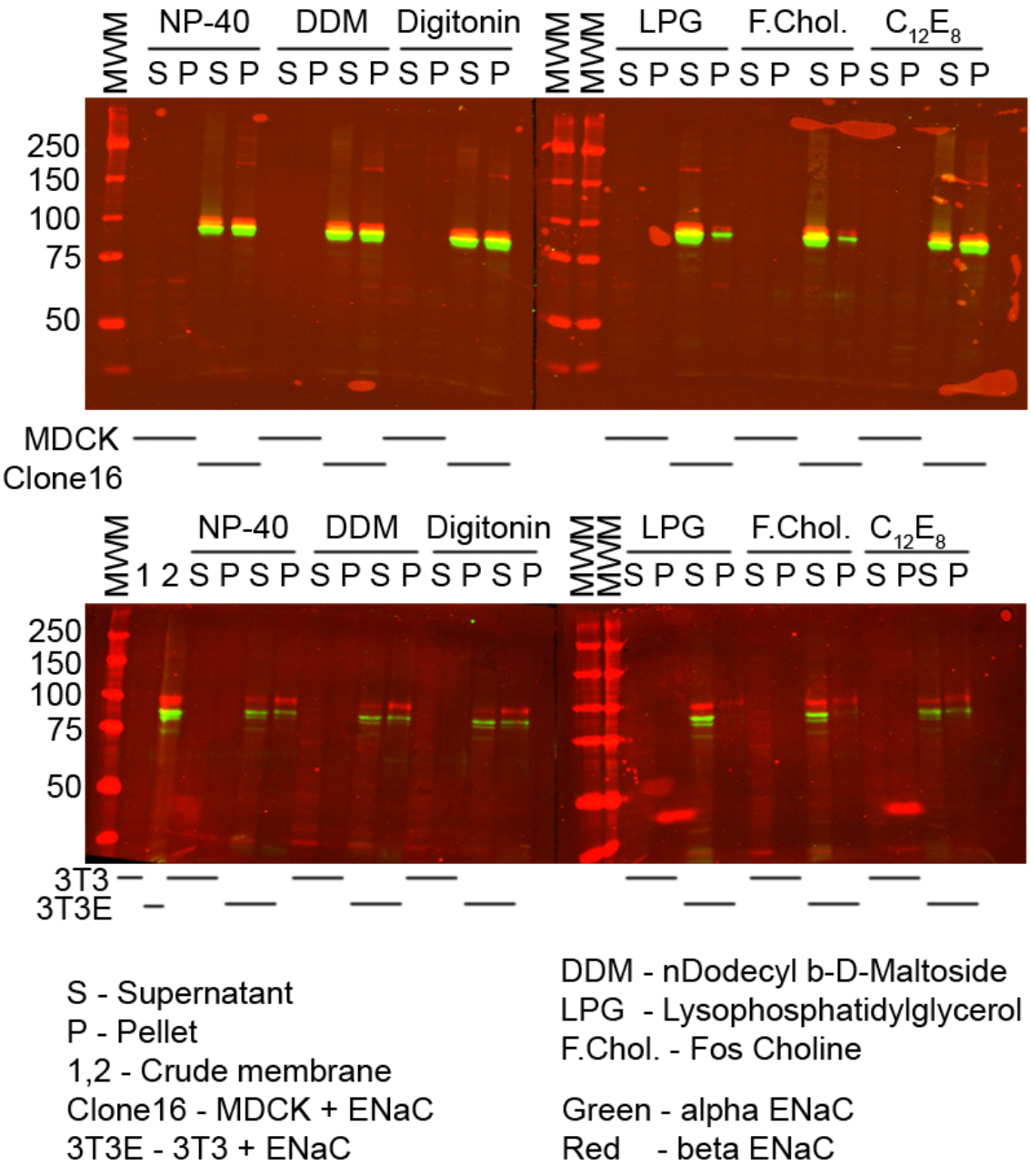


Figure 3.4 Detergent solubilization of ENaC
ENaC from MDCK or 3T3 cells was solubilized in six different detergents. The western blot was probed for alpha (green) and beta (red) subunits of ENaC. Samles from MDCK/3T3 cells without any observed ENaC expression were used as controls. Supernatant (S) and resuspended pellet (P) from all detergents were analyzed.

In order to determine the subunit composition of functional ENaC, we set out to study the protein in its native environment. As a first step, we intended to solubilize the heteromultimer in an appropriate detergent to maintain native interactions between subunits. Surprisingly, we did not find much evidence from the literature regarding the appropriate detergent for solubilization of ENaC. We studied the solubilization efficiency of ENaC in six different detergents and found that lysophosphatidylglycerol (LPG) and Fos-Choline solubilize ENaC more efficiently than the others (Figure 3.4). Unless otherwise specified, we solubilized ENaC in LPG for experimental characterization. Following solubilization, 5 ng of total solubilized protein was subject to native gel electrophoresis followed by western blotting. We studied the oligomerization states of native complexes by adding an increasing amount of SDS to the samples (Figure 3.5).

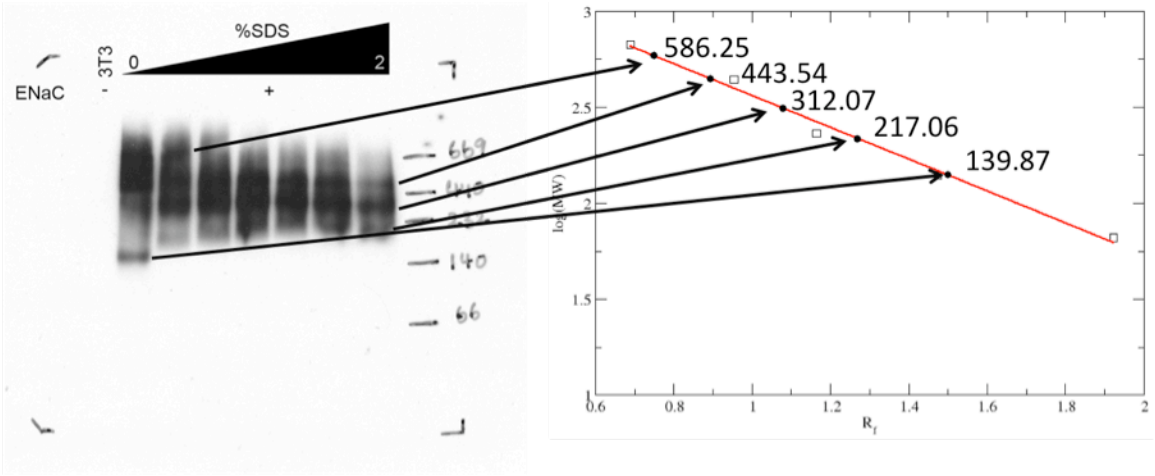


Figure 3.5 Oligomerization state of ENaC in 3T3 cells
(Left) Membrane fractions from 3T3 cells were isolated and solubilized in 1% LPG before loading on to a Bis-Tris 4-16% gradient native gel. Western blotting was performed following electrophoretic separation and the blot was probed using anti alpha ENaC antibody. (Right) Ferguson plot representing the relative mobility of the bands on the molecular weight marker on a log scale.

On a native gel, proteins are separated based on mass/charge ratio as well as the configuration in the species being separated. Therefore, the molecular weight of the protein does not change linearly as a function of the position on the gel.

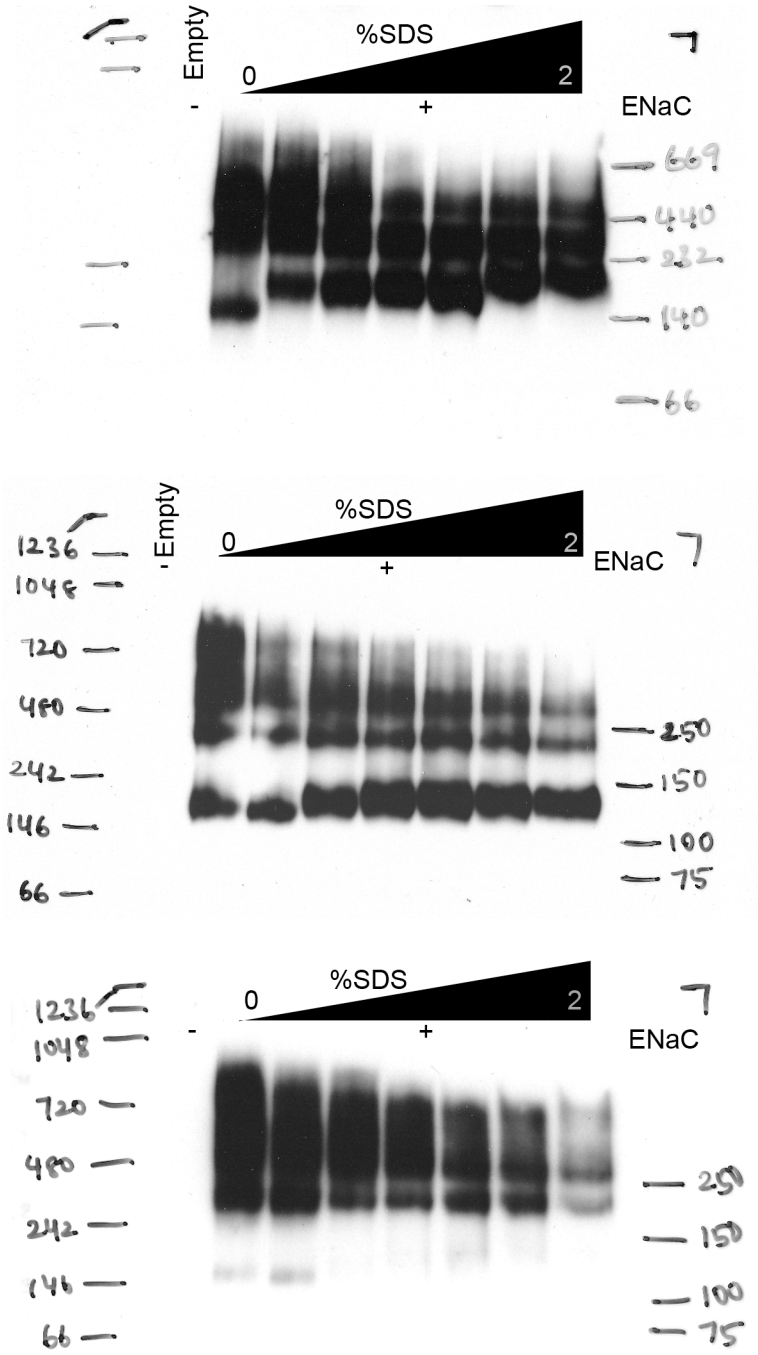


Figure 3.6 Oligomerization state of ENaC in *Xenopus* oocytes
Oocytes were injected with cRNA encoding different subunits of ENaC and solubilized membrane fractions were subject to native PAGE and western blot analysis. The blot was probed for alpha (top), beta (middle) and gamma (bottom) ENaC subunits using the respective antibodies.

To estimate molecular weights, we plotted the relative mobility of each band on the molecular weight marker as a function of its distance from one end of the gel (ferguson plot) (marked by the corners in Figure 3.5 – left panel). From the ferguson plot, we estimated the molecular weights of the two bands at 2% SDS to be 443.54 kDa and 312.07 kDa respectively.

Given that the molecular weights of the monomers are in the range of ~75-90 kDa, these molecular weights correspond to those to trimeric and tetrameric oligomeric states of ENaC. The additional constant (~40 kDa) in the molecular weight of the complex corresponds to that of the detergent micelle in which the complex is solubilized. At low SDS concentrations, we noticed higher order species probably corresponding to tertiary interactions between two or more trimers or tetramers. These interactions are weak and do not sustain treatment with higher concentrations of SDS.

In order to determine the functional oligomerization state of ENaC, we expressed different subunits in *Xenopus* oocytes. Before we measure function, we set out to determine whether ENaC appears in the same oligomeric states in oocytes as it does in mammalian cells (MDCK and 3T3). We injected equal amounts of cRNA for rat α, β and γ ENaC into oocytes and incubated the oocytes for 24 hrs or 48 hrs before measuring amiloride-sensitive whole cell sodium currents. Biochemical analysis of membrane samples (see Methods) from oocytes using native PAGE and western blotting revealed that ENaC exists as

both trimers and tetramers in oocytes (Figure 3.6). Interestingly, the total amount of gamma subunit is the least among all three subunits, suggesting that the expression of the gamma subunit is the limiting factor in assembly of ENaC, as reported previously.

3.5 Tetramers are more functional than the trimers

To evaluate the importance of each subunit on total ENaC-mediated whole cell currents, we injected cRNA encoding α-ENaC, α,β-ENaC, α,γ-ENaC, β,γ-ENaC or α,β,γ-ENaC into oocytes. We measured whole cell amiloride-sensitive currents 24h post injection and found that currents were maximal when all three subunits were co-expressed. This result is in agreement with previous reports (Harris et al., 2008) indicating the requirement for expression of all subunits for robust activation by proteases.

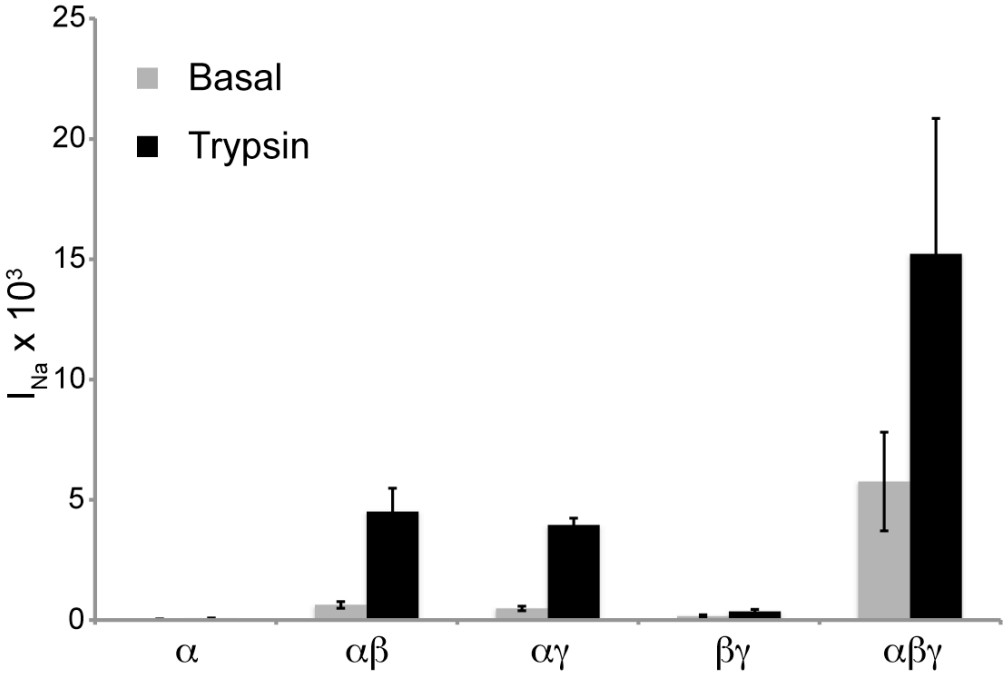


Figure 3.7 All three subunits are required for functional recapitulation
Amiloride-sensitive whole cell sodium currents were measured 24 h post injection of cRNAs encoding α, β, γ ENaC and the combinations thereof. Gray bars represent basal current, before treatment with trypsin. Black bars represent activated currents, post trypsin treatment.

To biochemically characterize the oligomeric states responsible for whole cell currents when all three subunits are expressed, we performed native PAGE separation of solubilized membrane fractions from occytes expression the afore-mentioned combinations of subunits. Interestingly, appearance of tetramers correlates with the presence of all three subunits (Figure 3.8). In order to quantify the amount of total tetramer to the total trimer for each case, we computed the total intensity of the gray pixels in the western blot using ImageJ. The ratio of tetramers to trimers increases from 0.2 for α,β alone to 1.39 for co-expression of all three subunits indicating an enrichment of tetramer population only in the presence of all subunits (Figure 3.8, 3.9). Comparing the biochemical data to the outcomes from whole cell current measurements, formation of tetramers is clearly dependent on the presence of all three subunits.

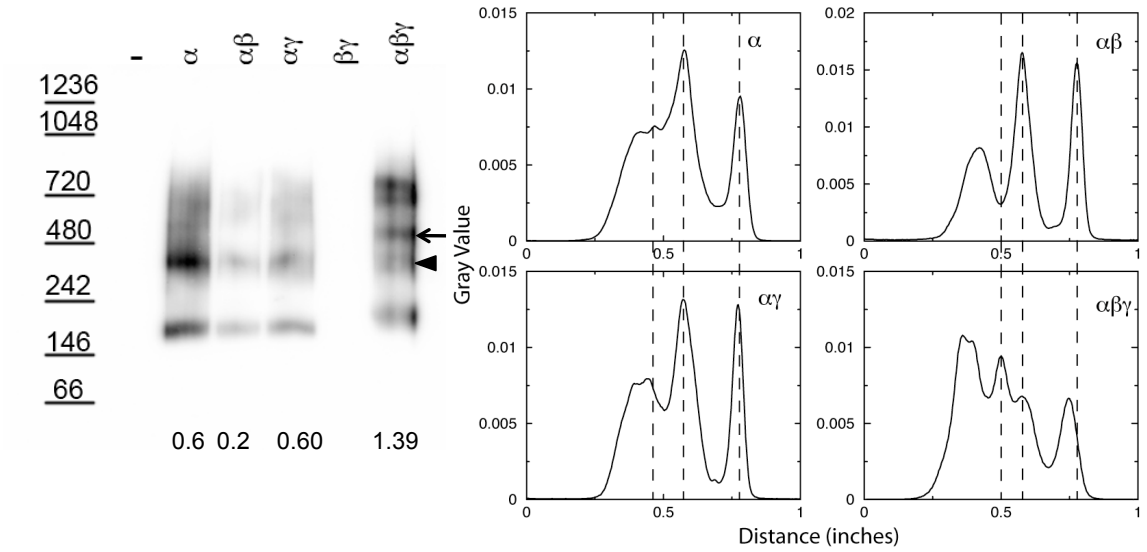


Figure 3.8 Correlation between tetramer formation and expression of subunits

(Left) Expression of all three subunits is required for tetramer formation (left – arrow). Trimers are formed even without one of the subunits (left – arrow head). The ratio of tetramer to trimer increases with expression of all three subunits. The western blot was probed for alpha-ENaC (Right) Profile for pixel intensity along the molecular weight axis. Peaks represent higher intensity. Right most peak represents monomer. The dotted line in the middle represents trimer and the one at a distance of 0.5 inches from the top of the blot represents the tetramer.

3.6 Expression of γ-ENaC is critical for formation of tetramers

Previous studies have clearly indicated that cleavage of γ-ENaC is critical for activation of the channel. We have demonstrated that the presence of all three subunits is critical for obtaining maximal activation of ENaC. Since cleavage and function of ENaC are tightly regulated, we hypothesized that the expression levels of the gamma subunit might directly translate to efficiency of assembly of tetramers. In order to understand the role of γ-ENaC in tetramer assembly, we co-injected different amounts of cRNA (0-0.6 ng) encoding γ-ENaC with equal amounts of cRNA (0.3 ng) encoding the α and β subunits in *Xenopus* oocytes.

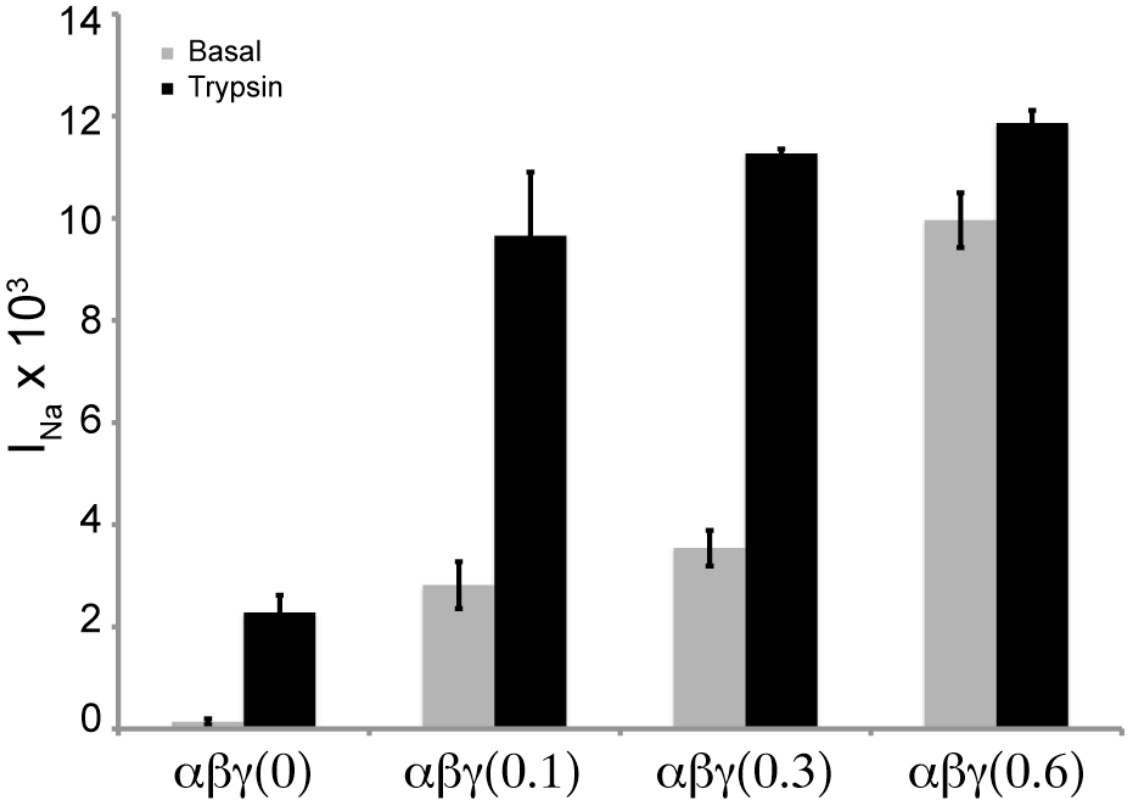


Figure 3.9 Modulation of γ-ENaC levels modulates whole cell currents
Amiloride-sensitive sodium current (I_{Na}) was measured using a two-electrode clamp from oocytes expressing increasing levels of rat gamma ENaC. The amount of cRNA (in ng) injected into oocytes is mentioned in parentheses. Gray bars represent whole cell current before activation by trypsin and the black bars represent activated currents after trypsin stimulation. Basal currents increase with increase in the gamma subunit.

As hypothesized, the basal current from oocytes injected with 0.6 ng of γ-ENaC is higher compared to the other batches of oocytes expressing equal or lower amounts of gamma ENaC, suggesting that expression level of the gamma subunit is the limiting factor for obtaining maximal activation of ENaC. However, interestingly, the activated current does not increase significantly with increase in the amount of co-expressed gamma subunit. It is tempting to speculate that the observed effect is due to increased access to the proteolytic cleavage sites in ENaC, upon formation of the tetramer, leading to increased basal current. To ensure that the gamma subunit is indeed expressed in excess upon injection of

increased amount of cRNA, we extracted crude membranes from the oocytes and solubilized them for biochemical analysis. Western blotting for all three subunits followed by SDS-PAGE separation revealed that the level of expression of gamma subunit increases steadily with increase in the amount of injected cRNA (Figure 3.10). Further, proteolytic processing of the gamma subunit is substantial at higher levels of expression (Figure 3.10 middle panel). The higher molecular weight bands, approximately at double the molecular weights of individual subunits correspond to dimers of the respective pairs. The fact that the intensity of this higher molecular weight band decreases with increase in the amount of co-expressed gamma ENaC suggests that the dimers assemble into higher order complexes (trimers and/or tetramers) and are processed. The doublet seen in case of β-ENaC corresponds to the mature, complex glycosylated form. In summary these results indicate that increase in gamma ENaC expression levels increases the amount of proteolytically processed, functional protein in oocytes.

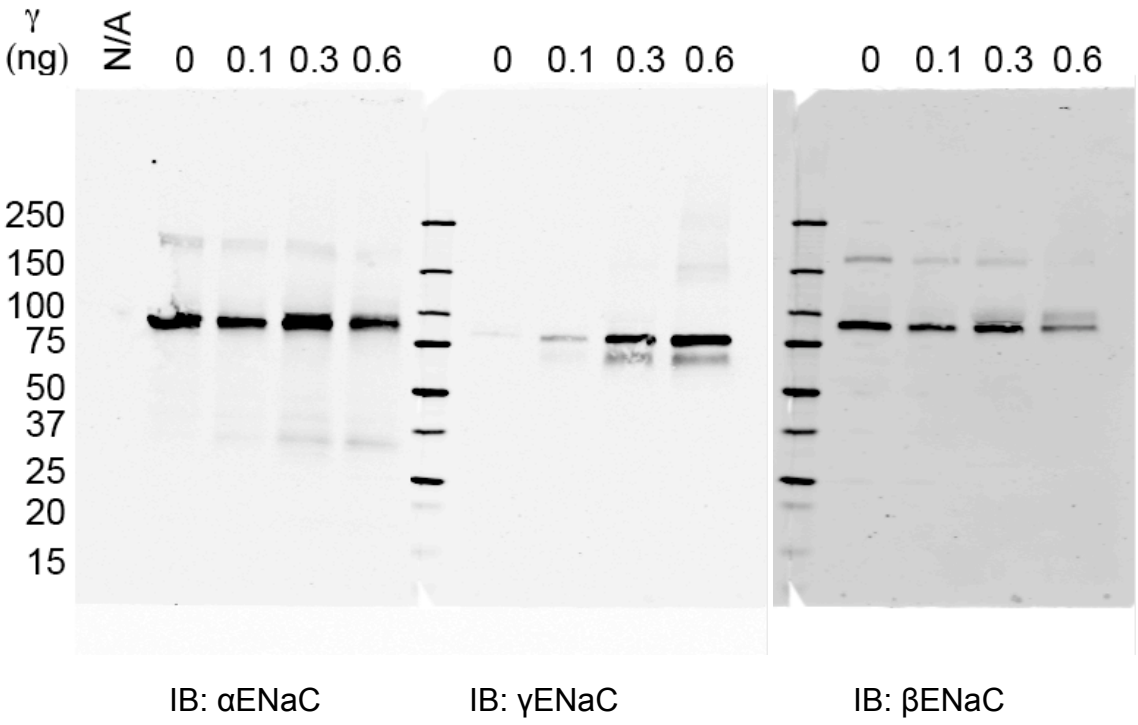
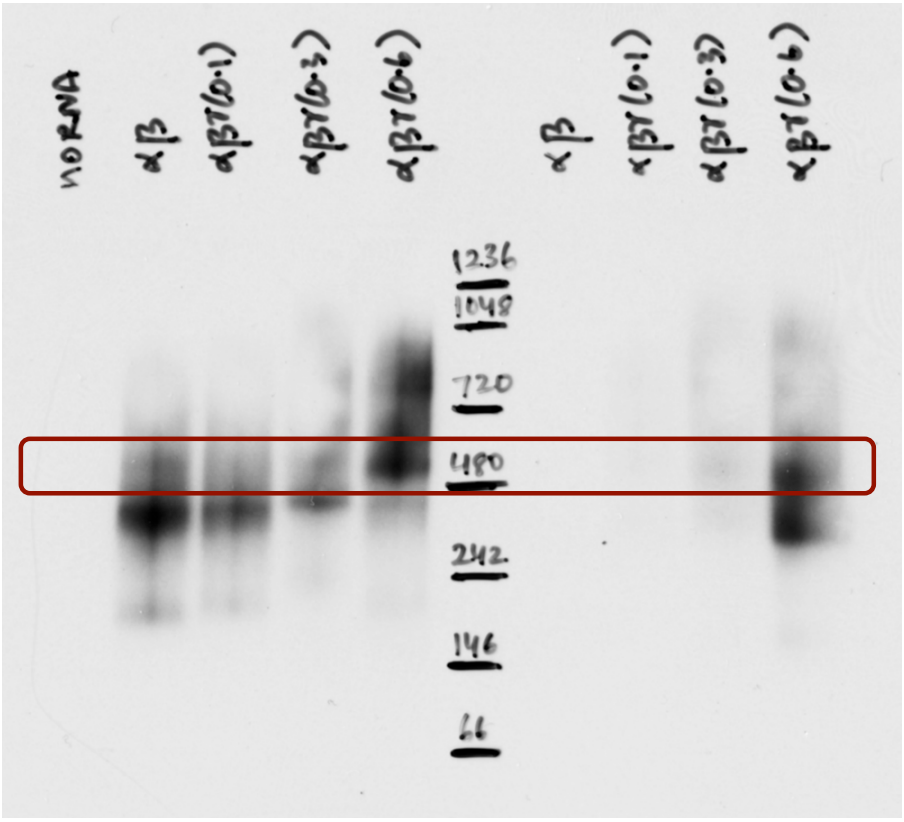


Figure 3.10 Gamma subunit expression levels regulate maturation
~10 ng of solubilized membranes from oocytes expressing different levels of gamma ENaC were subject to SDS-PAGE separation followed by western blot analysis. Full length, unprocessed ENaC migrates as a monomer between ~75 and 90 kDa. Lower molecular weight fragments represent proteolytically processed forms (see middle panel) and higher molecular weight species represent glycosylated forms (see right panel).

We performed native PAGE analysis to determine whether the increase in function upon increased expression of the gamma subunit correlates with increase in tetramer assembly. We observed that increase in expression level of gamma ENaC leads to increase in the amount of mature form of ENaC (Figure 3.10) and the amount of tetrameric complexes in oocytes (Figure 3.11).



IB: αENaC IB: γENaC

Figure 3.11 Increase in γENaC expression level promotes tetramer assembly
Crude membranes from oocytes expressing increasing levels of the gamma subunit of rat ENaC were solubilized in LPG and subject to native PAGE. Tetramer assembly is promoted by increased expression level of gamma subunit (red box).

In summary, expression of all three subunits of ENaC is required for robust channel activation and function. This result is in agreement with previous reports. ENac appears as trimers and tetramers on the cell surface when expressed in oocytes. The function of ENaC correlates with expression of tetramers, suggesting that the tetramers are more functional than trimers or other higher order multimers. Furthermore, expression level of the gamma subunit correlates with formation of tetramers in oocytes.

CHAPTER 4

Energetic and structural basis for activation of ENaC

To uncover the energetic basis for activation of ENaC upon cleavage, we performed computational analyses on peptides from ENaC that are susceptible to proteolysis by serine proteases. We performed discrete molecular dynamics (DMD) simulations to elucidate the structural and energetic bases for ENaC peptide recognition by CAP3. Our computational studies of various peptide-binding configurations to CAP3 establish the structural and energetic bases for CAP3 activity. Using potential of mean force (PMF) analyses, we determined the energetic basis for substrate recognition and cleavage by CAP3. We compared the results to those obtained via similar analyses for furin and elucidate the energetic basis for the lower sequence specificity of CAP3 compared to furin. We designed experimental studies to identify CAP3-mediated cleavage sites in ENaC that lead to channel activation. We conclude that CAP3, with less stringent sequence requirements than furin, robustly activates ENaC by cleaving at multiple basic residues in the extracellular domain. Using computational and experimental studies, we show that a site upstream of the traditional furin site in γ-ENaC is a potential substrate for CAP3.

4.1 CAP3 has less stringent sequence requirement for cleavage than furin

To assess the accessibility of the known protease cleavage tracts of different subunits of ENaC, we computed residue-wise disorder probability in the respective segments using Disopred2 (Figure 4.1) (Ward et al., 2004).

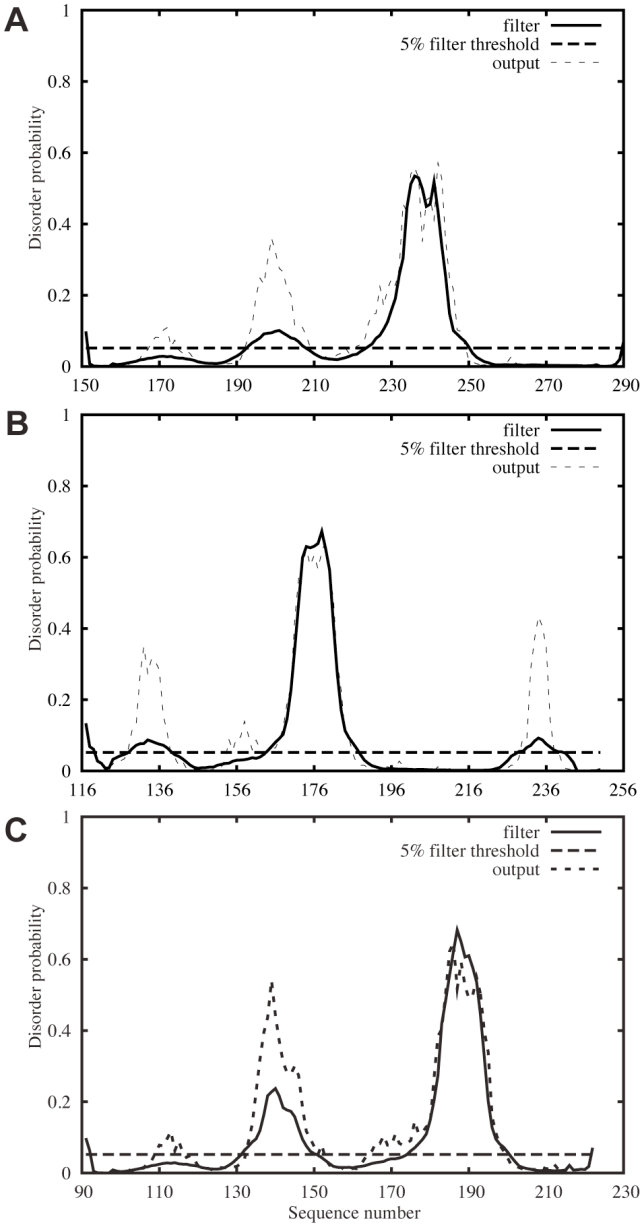


Figure 4.1 Disorder prediction for the hypervariable region in rat ENaC subunits
Residue-wise disorder was predicted using Disopred2 (http://bioinf.cs.ucl.ac.uk/disopred/). The 'filter' curve (continuous black line) represents the predicted probability of disorder for the corresponding amino acid. The horizontal dashed line represents the order/disorder threshold for the default false positive rate of 5%. The 'output' curve (dashed curve) represents the level of confidence in prediction of disorder for the corresponding amino acid. A, B, C) Disorder prediction for corresponding regions in α-, β- and γ-ENaC respectively.

To eliminate any bias in the prediction, we considered segments (α: V151-L290; β: K117-P240; γ: K91-S222) such that most of the cleavage sites are enclosed but are not near either end of the segment. Interestingly, the regions susceptible to cleavage by furin-like convertases in α- and γ-subunits (α: 202-RSSR, 228-RTAR; γ: 135-RKRR, 178-RKRK) of rat ENaC are intrinsically disordered (Figure 4.1A, C). The two peaks in the disorder plot for γ-ENaC correspond to the traditional furin site (135-RKRR-138) and the polybasic tract (178-RKRK-181) identified as cleavage sites for complete activation by furin and CAP1 (Bruns et al., 2007) (Figure 4.1C). This observation is in agreement with previous computational analyses reporting preference for intrinsic structural disorder in cleavage by serine proteases (Hubbard et al., 1991).

To elucidate the structural basis for substrate recognition and activation of ENaC by CAP3, we performed replica exchange DMD simulations (Ding, 2008; Dokholyan et al., 1998) of three peptide sequences (Seq1, 2, 3) from rat γ-ENaC (Materials and Methods) to study peptide binding to the active pocket of furin and CAP3 starting with the unbound state (Figure 4.2A, B). Using RMSD as the clustering criterion, we selected five final structures for each enzyme-substrate combination (Figure 4.2C). We found that the maximum RMSD between structures of peptide complexes with either furin or CAP3 is 1.5 Å. The final

substrate-bound configurations of both furin and CAP3 satisfy the distance constraints imposed during simulations (Materials and Methods, Figure 4.3A, C).

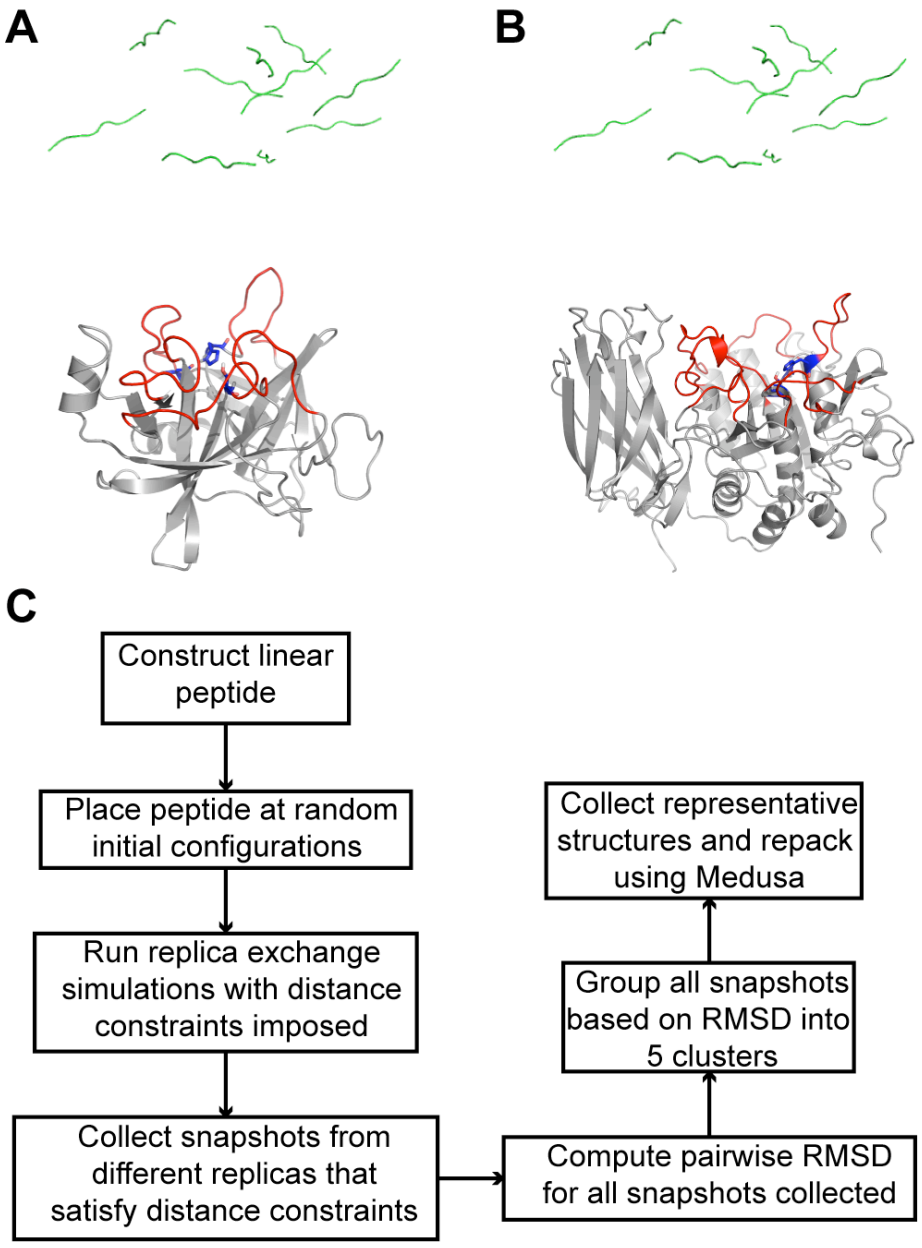


Figure 4.2. Simulation system and protocol
A, B) Starting conformations of CAP3 and furin respectively, with peptide Seq1 (colored green) placed at random positions at least 10Å away from the active site residues (colored blue and shown in stick representation). Portion of each protease shown in grey is maintained static during simulation while the loops shown in red are left flexible. Active site residues are static during the simulation. C) The steps outlined in the flowchart were

followed for modeling the enzyme-substrate complexes. Eight replicas were used for modeling with a starting temperature of 0.5 and a step of 0.035.

The peptide-bound configurations also portray the differences in the size, shape and charge distribution of the active sites of furin and CAP3 (Figure 4.3B, D).

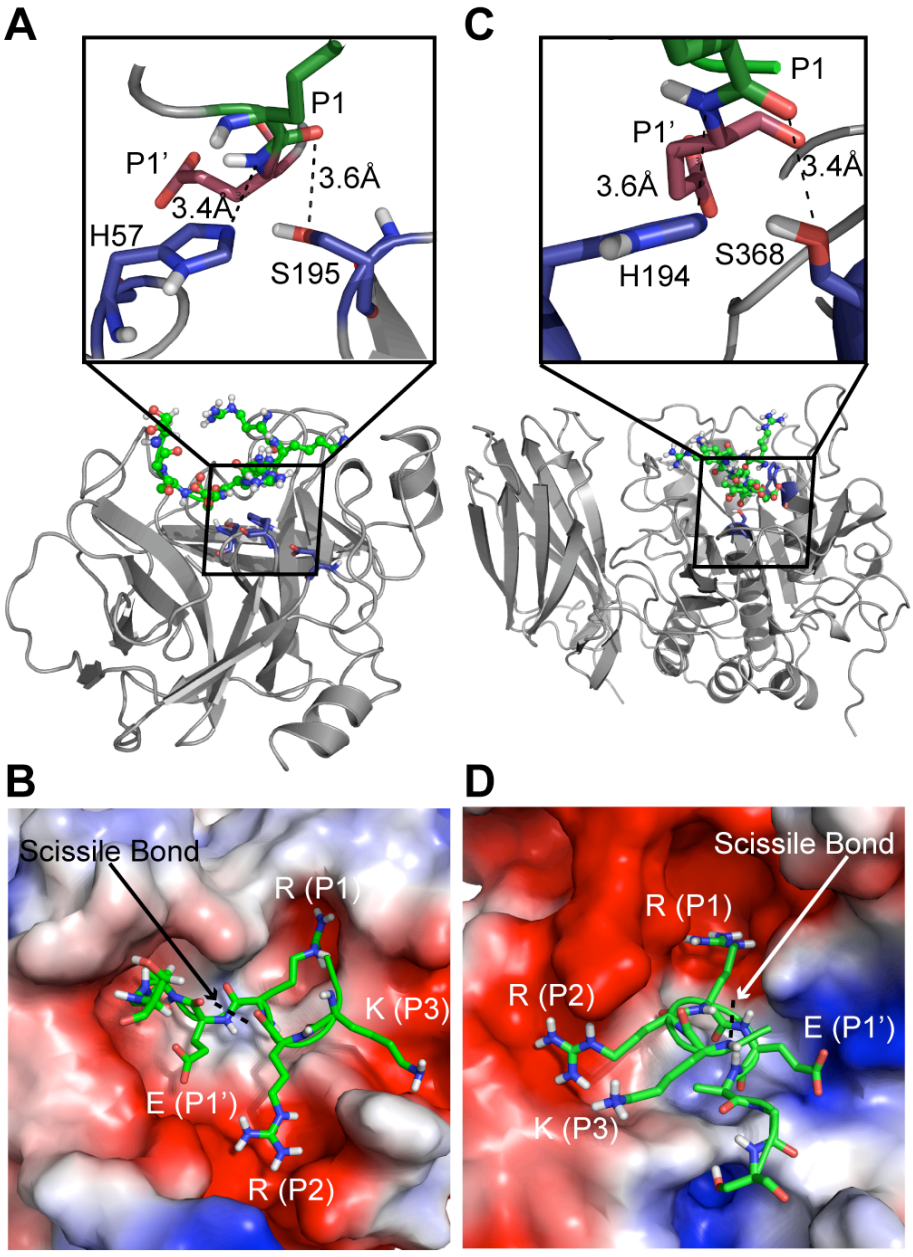


Figure 4.3. Structural models of peptides from γ-ENaC bound to furin and CAP3

A) Final docked configuration of peptide Seq1 in the active pocket of CAP3. Inset shows the distances between amine of the P1' site from the NE2 of active site histidine and that of the carbonyl oxygen of the P1 site from the OG of active site serine. B) Electrostatic surface representation of CAP3 with the side chains of residues in peptide Seq1 shown as sticks. The guanidium group of the arginine at the P1 site docks into a negatively charged groove in the enzyme. C) Final docked configuration of the peptide Seq1 in the active pocket of furin. Inset shows the distances between amine of the P1' site from the NE2 of the active site histidine and that of the carbonyl oxygen of the P1 site from the OG site of active site serine. D) Electrostatic surface representation of furin with the side chains of residues in peptide Seq1 shown as sticks. The guanidium group of arginine at P1 site docks into a negatively charged groove in the enzyme.

In our models, the residue at the P1 site of Seq1 is positioned in charge complementary pockets of both furin and CAP3 as observed in the corresponding crystal structures with bound inhibitors (Figure 4.3B, D) (Friedrich et al., 2002; Henrich et al., 2003). We conclude based on these observations that the peptides adopt similar final configurations dominated by electrostatic interactions between the enzyme and the peptide regardless of their initial configurations.

To establish the energetic basis for substrate recognition, we computed the two-dimensional potential of mean force (2D-PMF) with respect to *E* (normalized energy) and *d* (distance in Å between active site residues and the P1, P1' sites on the peptide). We observe from the 2D-PMF contours that Seq1 features low energy basins at a distance of ~3Å from the active site of both CAP3 and furin (Figure 4.4A, B). Similarly, the polybasic tract consisting of the alternate furin cleavage site (178-RKRK) presents a low energy basin at a distance of 3 Å or less from the active site of both furin and CAP3 suggesting that binding of Seq2 to either enzyme is energetically favorable (Figure 4.4C, D; black arrows). Takeuchi et al. proposed that the lysine residue (K132) upstream of the

traditional furin site could be the P4 site of a *bona fide* consensus cleavage motif (132-KESR) for CAP3 (Takeuchi et al., 2000).

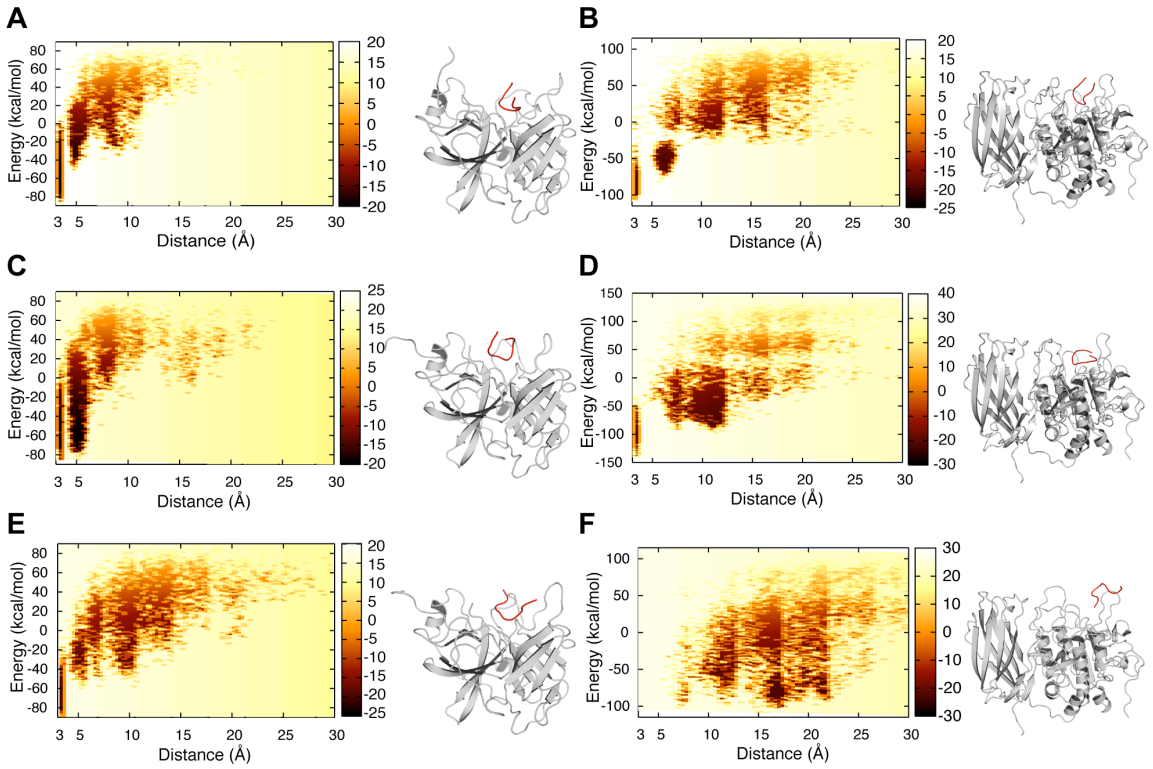


Figure 4.4. Energetic basis for peptide binding to furin and CAP3
2D-PMF as a function of normalized energy of enzyme-substrate complex and distance between the active site and respective peptides. Black arrows represent low energy basins in respective 2D-PMF plots. The peptide is colored red and the enzyme is colored grey in the accompanying models representing the minimum energy configurations for the corresponding enzyme-peptide combinations. A, B) 2D-PMF of Seq1 binding to CAP3 and furin respectively. C, D) 2D-PMF of Seq2 binding to CAP3 and furin respectively. Seq2 is trapped in an energy minimum at the surface loops of furin while it reaches the active site of CAP3. E, F) 2D-PMF of Seq3 binding to CAP3 and furin respectively.

To verify whether this sequence is susceptible to cleavage by CAP3, we performed simulations with Seq3 approaching the active sites of CAP3 and furin. We observe low energy configurations for Seq3 in the active site of CAP3 in simulations (Figure 4.4E). Surprisingly, none of the configurations from the simulation with Seq3 in presence of furin satisfied the distance constraints for

recognition, suggesting that binding of Seq3 to furin is energetically less favorable (Figure 4.4F). Interestingly, the plot for furin with Seq3 does not present a low energy basin at a distance of 3 Å from the active site while a low energy basin is observed for CAP3 (compare Figure 4.4E, F). These results indicate that CAP3 is more effective than furin at stimulating ENaC containing γ-ENaC 132KESR. Our computational results suggest low stringency sequence requirements for CAP3-mediated cleavage and presence of less ideal furin substrates in this region of γ-ENaC.

4.2 Catalytic activity of matriptase is required for activation of ENaC

To biochemically characterize the sequence requirements of CAP3, we first established that the catalytic activity of CAP3 is required for activation of ENaC. The motivation for this study arose from the fact that the catalytic activity of CAP1, a GPI-anchored membrane serine protease, is not required for its regulation of ENaC (Vuagniaux et al., 2002). We found that co-expression of CAP3 with ENaC for 24 h robustly increased basal amiloride-sensitive sodium current (I_{Na}), typically in the range of 3-5 fold (Figure 4.5A). Moreover, unlike the basal I_{Na} generated by ENaC alone, the larger basal I_{Na} with CAP3 co-expression was not further increased by application of trypsin (Figure 4.5A) or hNE. The decrement of I_{Na} sometimes seen following exposure of CAP3-expressing oocytes to exogenous protease (Figure 4.5A, ENaC+CAP3, black bar) reflects run down of the stimulated I_{Na}. We observed that CAP3 stimulation of ENaC is inhibited by co-expressed hepatocyte activator inhibitor-1 (HAI-1) (Figure 4.5B).

HAI-1 is a Kunitz-type serine protease inhibitor identified as the physiologic cognate inhibitor of CAP3 catalytic activity (Szabo et al., 2007). Western blots of CAP3 in lysates of the injected oocytes show that CAP3 was robustly expressed in its active form, as indicated by the expected fragmentation pattern of this self-activating protease (Figure 4.5C) (Benaud et al., 2001).

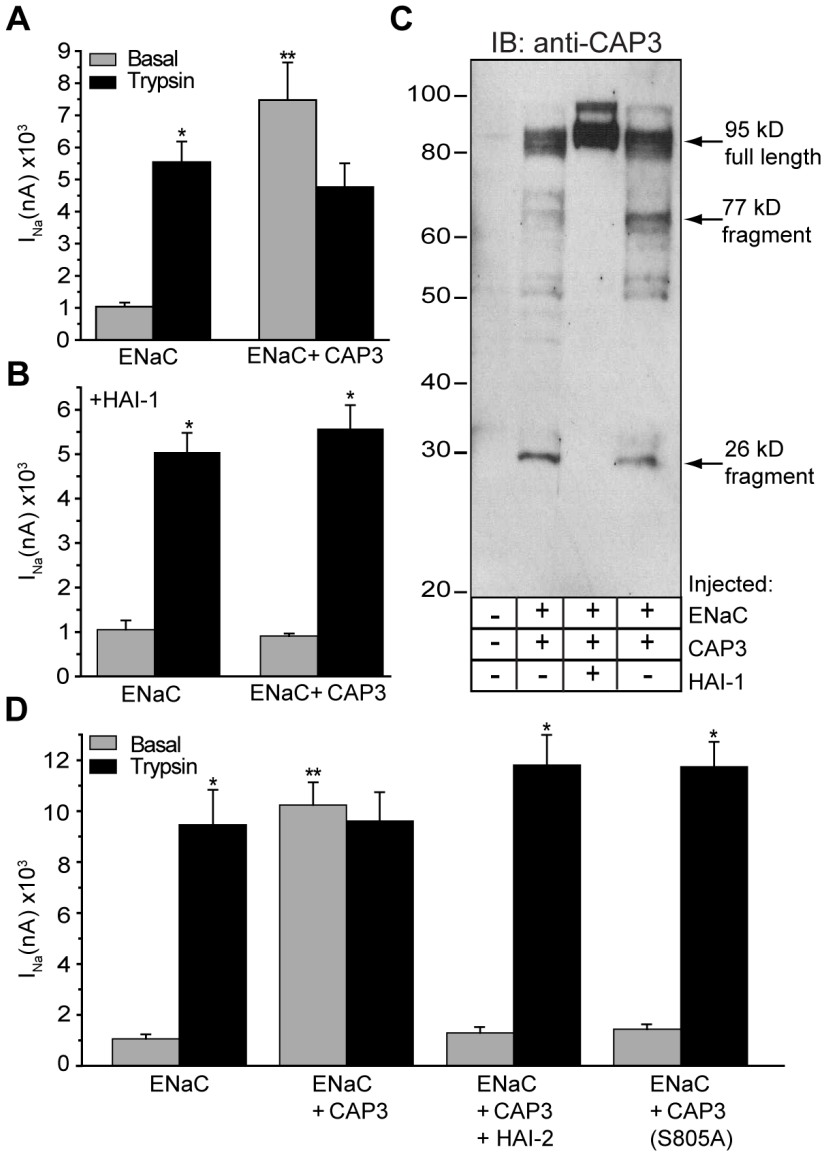


Figure 4.5. Catalytic activity of matriptase/CAP3 is required for stimulation of ENaC

A, B) Coexpression of the Kunitz domain containing inhibitor, HAI-1 blocks stimulation of ENaC. Oocytes were co-injected with 0.3 cRNA of α, β and γ-ENaC alone (Panel A), or

in combination with 1 ng cRNA of HAI-1 (Panel B). After 24 h incubation, amiloride-sensitive current (I_{Na}) was recorded before (gray bars) and following (black bars) 5 min exposure to 2 µg/ml trypsin. C) Expression and autocleavage of matriptase/CAP3, and effect of HAI-1. Oocyte lysates from (A) and (B). were studied by Western blotting using anti-CAP3 antibody. D) CAP3 stimulation of ENaC was blocked by co-expressed HAI-2 or by mutating serine 805 of the catalytic triad. Experiments were repeated on 2-4 batches of oocytes, with a total of 12-26 oocytes per condition. *Trypsin stimulated I_{Na} different from basal I_{Na}. **Basal I_{Na} different from ENaC alone. ANOVA, $p < 0.05$.

Furthermore, co-expression of HAI-1 with CAP3 prevented cleavage associated with CAP3 activation. Co-expression of HAI-2, a related Kunitz-type inhibitor, also completely prevented CAP3 stimulation of ENaC (Figure 4.5D) (Szabo et al., 2008). Finally, CAP3 inactivated by mutation of S805 of the catlaytic triad had no effect on I_{Na} of co-expressed ENaC (Figure 4.5D) (Miyake et al., 2009). Thus, ENaC co-expressed with CAP3 is fully activated through a mechanism that requires catalytic activity of CAP3.

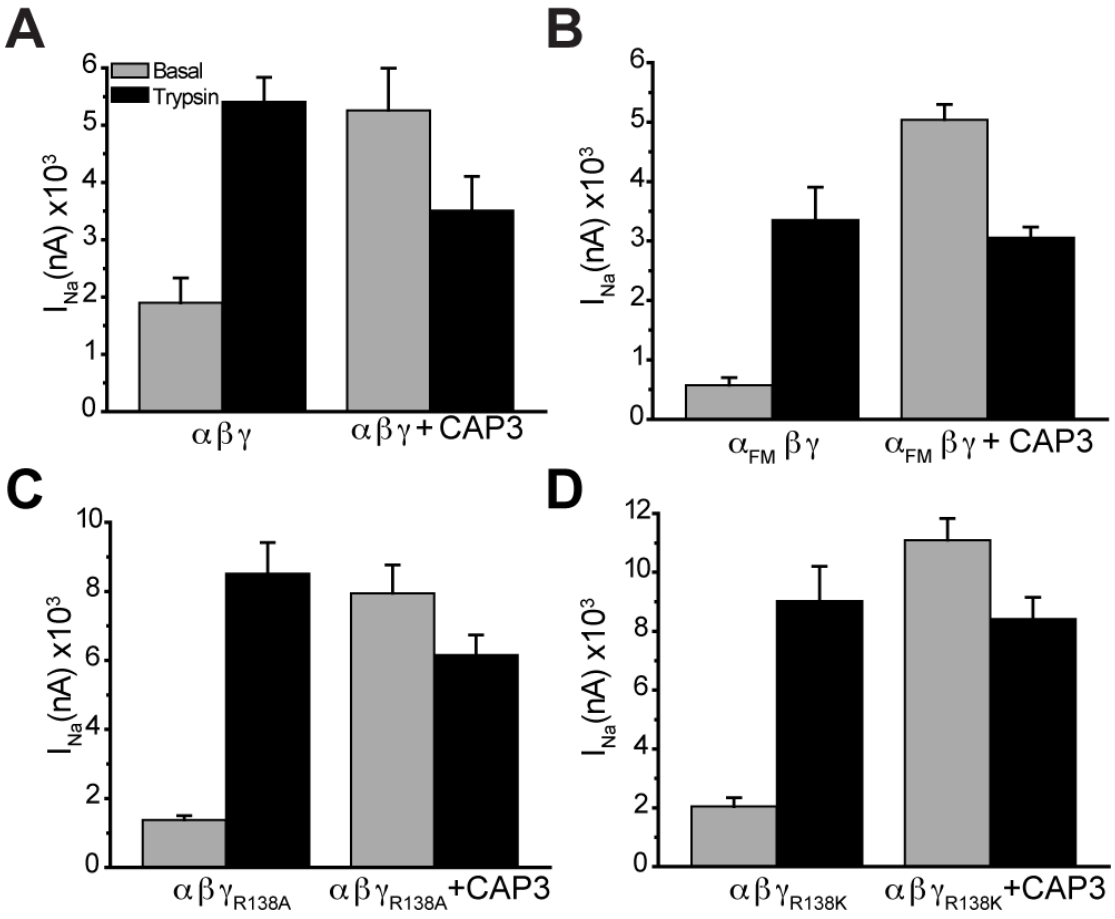


Figure 4.6. CAP3 coexpression stimulates ENaC containing mutant furin sites
A) Control experiment with wildtype ENaC and CAP3. B) Mutation of two α-ENaC furin sites (FM: $R_{205,231}A$) reduced basal I_{Na}, as expected, but did not prevent full proteolytic stimulation of I_{Na} by co-expressed CAP3. C, D) Two mutations of the furin site identified in γ-ENaC had no effect on CAP3 stimulation of ENaC. Methods, replications and analysis similar as described for Figure 4.5.

4.3 CAP3 cleaves γENaC at an alternate site N-terminal to the furin site

We studied the role of γ-ENaC in stimulation by CAP3, by co-expressing the furin site mutants of α-ENaC with WT β- and γ-ENaC in oocytes. CAP3 robustly activated ENaC without the furin sites in the α-subunit suggesting predominant cleavage of the γ-subunit (Figure 4.6A, B). In γ-ENaC, we specifically examined the importance of the basic P1 residue (R138) in the 135-

RKRR tract recognized as the furin site. Co-expressed CAP3 robustly stimulated ENaC containing γ-subunit furin resistant mutant R138K and the CAP2 insenstive furin mutant R138A (Figure 4.6C, D) (Garcia-Caballero et al., 2008) indicating that CAP3 activates ENaC by a mechanism involving cleavage at a site distinct from γ-R138. Although the consensus sequences for convertases (R/K-X-X-R) and CAPs (R/K-X-X-R/K) overlap, some studies have reported distinct preferences at individual residue positions (Bugge et al., 2009; Thomas, 2002). Therefore, we asked if extensive mutagenesis of the 135-RKRR tract in γ-ENaC to 135-QQQQ would affect the action of CAP3 toward ENaC. Co-expressed CAP3 did not stimulate this mutant ENaC, even though subsequent trypsin exposure lead to significant stimulation (Figure 4.7B). Further mutagenesis revealed that the presence of R135 is sufficient for CAP3-mediated cleavage of γ-ENaC (Figure 4.8).

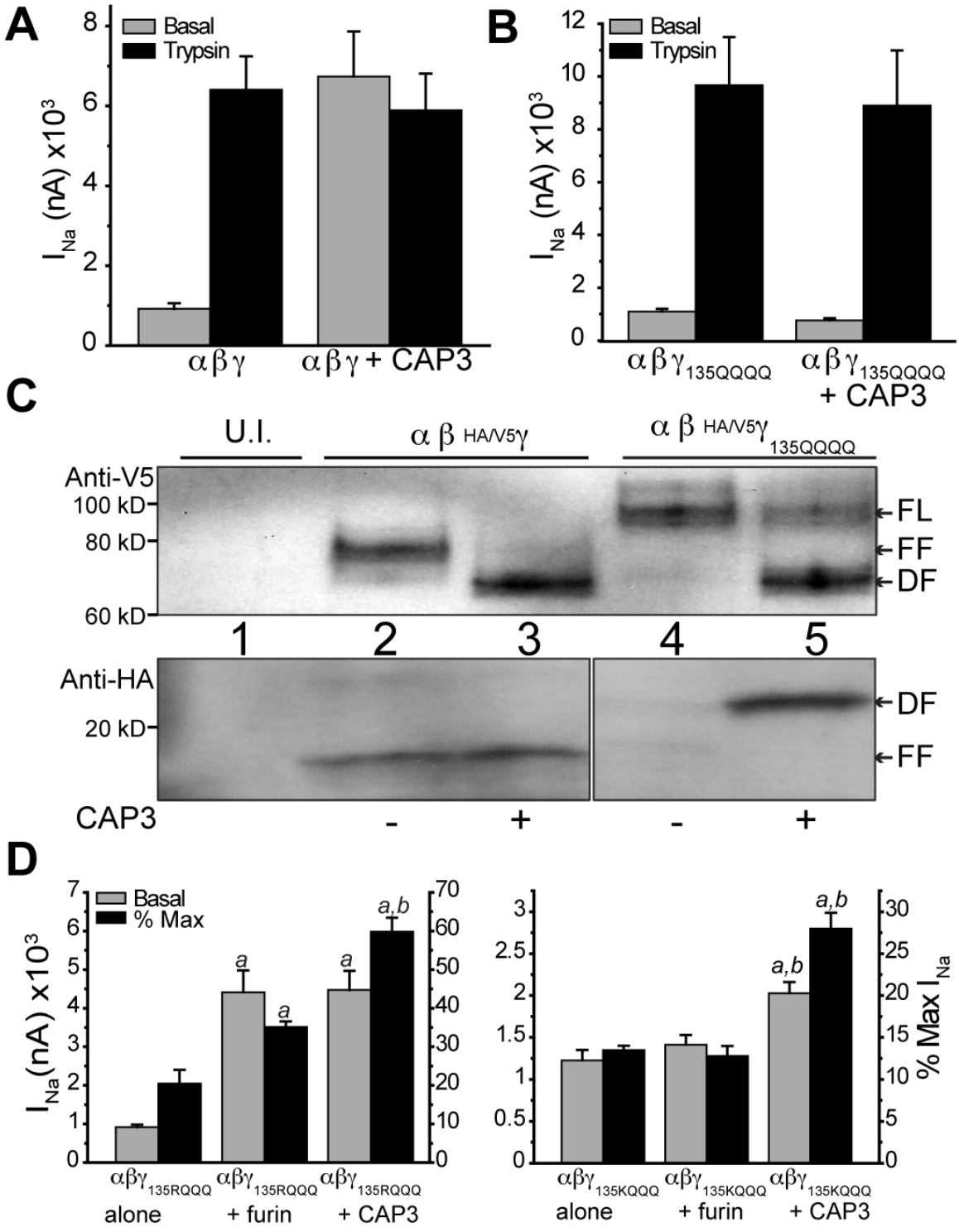


Figure 4.7. CAP3 mediates neither activation nor cleavage of γ-135_{QQQQ} ENaC
A) WT ENaC was expressed alone or with CAP3. I_{Na} was recorded 24 hr after cRNA injection, before (gray bars) and after (black bars) a 5 min trypsin exposure. B) ENaC containing the mutant γ135_{QQQQ} was expressed alone or with CAP3. I_{Na} was recorded as in A. C) ENaC made of WT subunits or containing the mutant γ-135_{QQQQ} were expressed

alone or with CAP3. After 24 h, uninjected (U.I.) oocytes and oocytes from each treatment group were surface biotinylated and the surface protein pool was captured on streptavidin beads. Full length and fragments of HA/V5 γ-ENaC were recognized by blotting for V5 (C-terminal epitope) or HA (N-terminal epitope). FL = full length; FF = furin fragment; DF =distal fragment. D) Oocytes were injected with α_{FM},β and either γ135RQQQ (left) or γ135KQQQ (right). Each of these ENaC combinations was expressed alone or co-expressed with furin or CAP3, and studied after 24 h. Basal I_{Na} was recorded before and following 5 min exposure to 2 μg/ml hNE. Basal I_{Na} as raw current (left ordinate) or as a percent of maximum I_{Na} following elastase (right ordinate) is shown. Mean values were compared by ANOVA and Tukey's test applied. a = different from mutant γ alone, b = different from mutant γ + furin, $p < 0.01$.

To biochemically characterize CAP3-mediated cleavage of γENaC, we performed Western blot analysis of biotinylated surface protein pool captured on streptavidin beads (Materials and Methods). We characterized the HA/V5 double epitope labeled γ-ENaC by the pattern of anti-V5 (C-terminal tag) and anti-HA (N-terminal tag) staining on Western blot (Figure 4.7C). In oocytes expressing WT ENaC subunits, V5-tagged γ-ENaC at the cell surface exists as a mixture of full length ("FL", ~93 kD band) and furin fragments ("FF", ~75 kD band) (Figure 4.7C, top panel). Under basal conditions, HA-label was found in a complementary ~18 kD band (Figure 4C, lower panel) and in a FL band (not shown). With co-expression of CAP3, a more rapidly migrating fragment of ~70 kD ("DF", distal fragment) replaces the FF band, a result now seen as characteristic of proteolysis of γ-ENaC at a site 20-40 residues downstream of the furin site (Bruns et al., 2007; Harris et al., 2007; Passero et al., 2008). The 18 kD anti-HA reactive (N-terminal) fragment indicates that cleavage at the traditional furin site is not affected. As predicted, mutant γ-ENaC containing the 135-QQQQ tract shows no FF, either V5- or HA-labeled, when co-expressed with α- and β-ENaC alone (Figure 4.7C, lane 4). However, when co-expressed with CAP3, the proportion of FL V5-labeled γ-135-QQQQ at the cell surface decreases, with a

significant increase in the intensity of the DF fragment (Figure 4.7C, lane 5). A complementary N-terminal HA-labeled mutant γ-ENaC fragment of ~23 kD appears at the surface of CAP3 co-expressing cells (Figure 4.7C, lower panel, lane 5). These results indicate that cleavage within the 135-138 furin site is blocked in γ-135-QQQQ while CAP3 induces cleavage C-terminal to the furin site. In addition, we conclude that cleavage within residues 135-138 is essential for CAP3 stimulation of ENaC.

Based on previous reports and our computational analyses, we hypothesized that K132 could be part of a *bona fide* consensus motif (132-KESR) targeted by transmembrane serine proteases (TSPs) or furin-like convertases (Kishi et al., 2001; Takeuchi et al., 2000). To test this hypothesis, we generated γ-ENaC with 132-HESRQQQ, which associates with WT α- and β-ENaC to produce reduced I_{Na} that responds briskly to hNE (2 μg/ml) or trypsin (20 μg/ml) (Figure 4.8, black bars). Although the presence of 132-KESR is evidently not optimal for endogenous convertases, we reasoned that over-expressed furin might recognize this site. Because the preference of furin at the P1 residue is strong for arginine over lysine (Matthews et al., 1994), while CAP3 is reported to tolerate lysine at P1 (Takeuchi et al., 2000), we compared the ability of co-expressed human furin and CAP3 to stimulate the γ-ENaC mutants 135-RQQQ and 135-KQQQ. To simplify interpretation, we expressed these mutant γ-subunits with WT β-ENaC and with α-ENaC furin site mutant (Figure 4.7D). Interestingly, co-expressed human furin partially activated ENaC containing γ-132-KESRQQQ, to 35% of maximum stimulated I_{Na}, albeit less

efficiently than CAP3 which led to 60% of maximum stimulated I_{Na} (Figure 4.7D, left panel). Co-expressed furin did not stimulate ENaC with a lysine at position 135 (Figure 4.7D, right panel) while CAP3 stimulated this mutant, albeit to a lesser extent than the mutant with R135 preserved. These results suggest that CAP3, due to its less stringent sequence requirements, can target basic residues in the 132-138 tract of γ-ENaC that are more resistant to endogenous convertases and over-expressed furin.

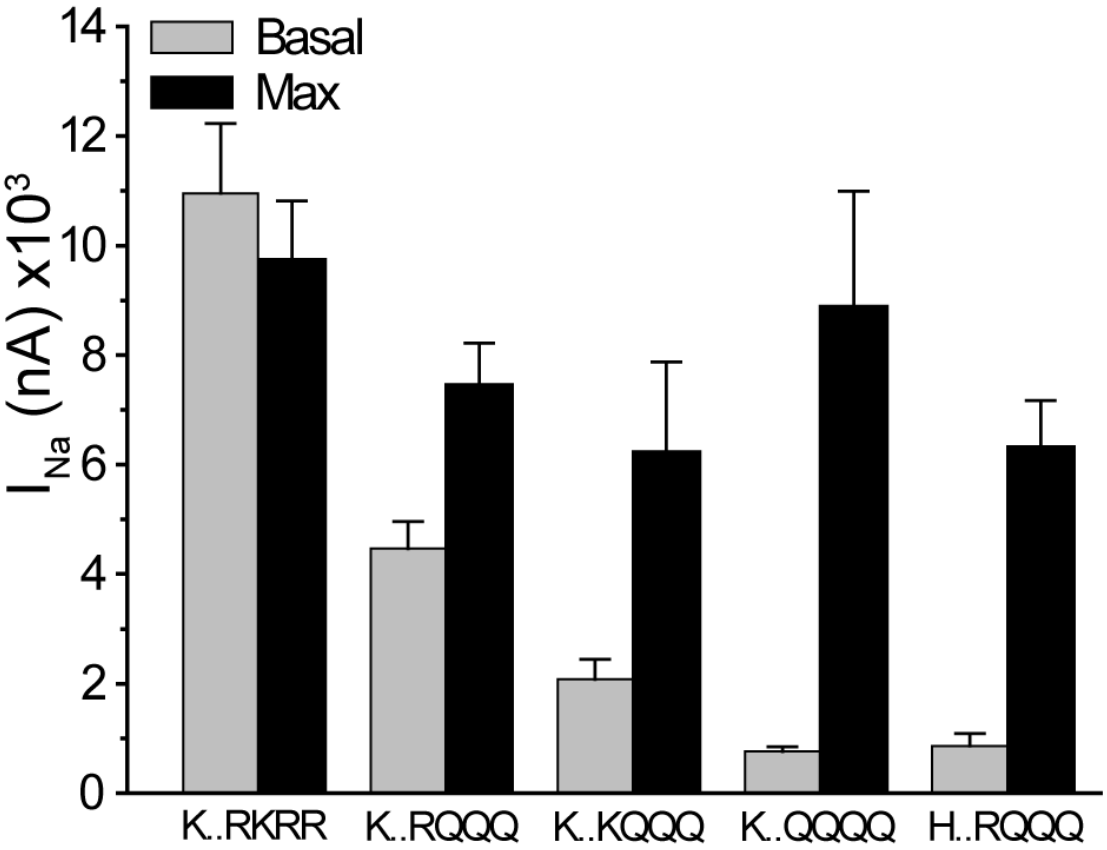


Figure 4.8. Residue 135R in γ-ENaC can form a CAP3-sensitive cleavage site with 132K

Oocytes were injected with 0.3 ng each of WT α- and α-subunits and γ-subunits bearing the 132..135-138 residues indicated above. CAP3 was co-expressed in each group. I_{Na} was recorded after 24 hr expression (grey bars). Following 5 min exposure to 20 μg/ml trypsin or 2 μg/ml hNE, I_{Na} was recorded again (black bars). The proteolytic stimulation of I_{Na} by co-expressed CAP3 was maximal with WT γ-ENaC. CAP3 partially stimulated

132KESRQQQ and 132KESKQQQ to a lesser degree. This stimulation was ablated by a non-basic residue at 135 or by Histidine at residue 132.

4.4 CAP3 cleaves ENaC at multiple sites C-terminal to the furin site

A candidate site for cleavage events responsible for the broadly staining DF (~70 kD) (Figure 4.7C) is the tract 178-RKRK, as this polybasic region is required for cleavage of γ-ENaC by CAP1 (Bruns et al., 2007), and shares the same minimal sequence requirements for cleavage with CAP3 (Shipway et al., 2004). DMD simulations revealed that CAP3/furin binding at this site is energetically favorable. Surprisingly, however, γ-ENaC with 178-QQQQ was stimulated by co-expressed CAP3 to about the same extent as ENaC containing WT γ-subunit (Figure 4.9A). The CAP3-stimulated basal current in either WT or mutant channel groups was not further increased by hNE, indicating that CAP3 attained full proteolytic stimulation of mutant ENaC at the surface. In addition, the patterns of C-terminal V5-labeled fragments of WT or mutant γ-ENaC contained in the cell surface pool of each expression group were affected similarly by co-expressed CAP3, each showing a shift from a mixture of FL or FF to a population dominated by DFs, consistent with CAP3-induced cleavage at sites downstream from the furin site (Figure 4.9A, lower panel). While these results do not refute the fact that 178-RKRK is a potential cleavage site, the data indicate that other potential cleavage sites exist in the vicinity of this basic tract.

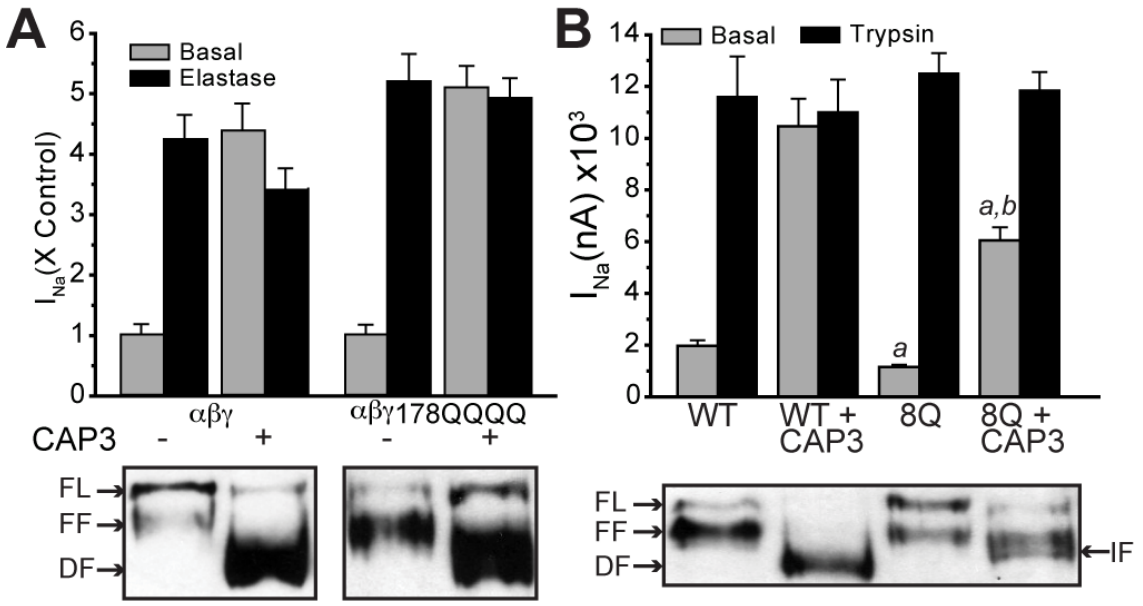


Figure 4.9. The basic tract 178-RKRK in γ-ENaC is not essential for CAP3 stimulation of I_{Na}

A) WT HA/V5-γ-ENaC or HA/V5-γ-RKRK(178-181)QQQQ cRNA was co-injected with WT α- and β-ENaC cRNA (0.3 ng/subunit). Half of each group was also injected with 1 ng cRNA for CAP3. I_{Na} was recorded after 24 h, before and following 5 min of exposure to 2 µg/ml hNE (upper panel, N = 10-12 oocytes per condition from two batches). 30-60 oocytes per condition were surface biotinylated, as described elsewhere. WT and mutant γ-ENaC fragments present in the cell surface pool were analyzed by Western blot (lower panel). B) WT α- and β-ENaC subunits were expressed for 24 h with WT or mutant (8Q) HA/V5 γ-ENaC, with or without CAP3 (see text for description of mutant 8Q). I_{Na} was recorded before and following 5 min exposure to 20 µg/ml trypsin (a = different from WT basal; b = different from 8Q basal, $p < 0.01$) Surface biotinylated proteins from the same injection groups were analyzed by anti-V5 Western blotting are shown in lower panel. FL = Full Length; FF + Furin Fragment; DF = Distal Fragment; IF = Intermediate Fragment.

As mutation of the polybasic tract 178-RKRK has no effect on CAP3-mediated stimulation or banding pattern of γ-ENaC fragments, we tested the contribution of flanking basic residues in the region from residues 172 to 202. We investigated the importance of 172R, 185K, 189K and 201K, 202K, individually, and in various combinations with 178-QQQQ. We observed no significant effects of mutating any single basic residue in this region on CAP3 cleavage and stimulation of ENaC. However, from a threshold of mutating 6-7 basic residues, up to 9 basic residues in this region replaced by glutamine, we observed

progressively decreased CAP3-mediated stimulation of I_{Na}, staining density of DF, and resting whole cell P_o. Particularly, ENaC containing mutant γ-subunits with eight glutamines substituted for basic residues within the 172-202 tract ("8Q") was only partially stimulated by co-expressed CAP3 (Figure 4.9B). Western blot analysis of the C-terminal V5-tagged fragments of 8Q γ-ENaC on the cell surface suggests that in this extensively mutated channel, CAP3 generates an intermediate fragment (IF) that migrates between FF and DF, characteristic of WT γ-ENaC (Figure 4.9B). Thus, CAP3 cleaves ENaC at multiple basic sites including those that do not conform to the furin consensus sequence requirements. These results are in agreement with our conclusion that CAP3 has less stringent sequence requirements for cleavage than furin. It is likely that the local structure of the protein is altered upon extensive mutagenesis, thereby hampering cleavage. Assuming that the structure of ENaC is intact, our results suggest that CAP3 cleaves γ-ENaC at multiple sites C-terminal to the furin site resulting in robust channel activation.

CHAPTER 5

Allosteric signal propagation within ENaC

Protein allostery is a ubiquitous mechanism central to the regulation of many cellular processes including enzyme catalysis and signal transduction (Changeux and Edelstein, 2005; Goodey and Benkovic, 2008). This phenomenon, often mediated by conformational rearrangement, refers to change at one site (allosteric site) affecting a distal site, resulting in functional modulation of corresponding proteins or protein complexes (Changeux and Edelstein, 2005). For instance, G-protein coupled receptors trigger downstream signaling cascades via a multitude of allosteric changes upon ligand binding at an extracellular site (May et al., 2007). One of the emerging views of protein allostery is that allosteric change is a redistribution of a protein's conformational ensemble upon ligand binding or mutation (Goodey and Benkovic, 2008; Gunasekaran et al., 2004). Therefore, it is feasible to convert a traditionally non-allosteric protein into an allosteric protein by shifting the conformational distribution toward or away from the functionally-relevant conformations with appropriate ligand binding and rational mutagenesis (Gulnik et al., 2000; Rose et al., 1998; Santamaria et al., 2002). Recent NMR studies of dynamic coupling between residues in proteins support the idea that allostery may be a common

intrinsic property of many proteins (Fuentes et al., 2006; Gunasekaran et al., 2004; Popovych et al., 2006). However, the molecular mechanism governing protein allostery remains one of the fundamental unanswered questions pertaining to protein biophysics.

Although conformational rearrangements involved in allostery can be detected, understanding how changes in structure translate to those in function remains a challenge. Specifically, determination of inter-residue interaction networks coupling distant sites and identification of “hot spot” sites or hubs with highest impact on the propagation of such coupling is essential to uncover the molecular origin of protein allostery (Hardy and Wells, 2004). Previous efforts have focused on identifying possible mechanisms governing protein allostery and searching for new allosteric sites (Hardy and Wells, 2004). X-ray crystallography studies of both the bound and unbound structures offered us an important structural insight into allosteric regulation. However, the analysis of static structures cannot provide a complete picture of the inter-residue interactions that result in allostery. NMR studies of protein dynamics have been pivotal in identifying “hidden” networks of residues with strong dynamic coupling (Fuentes et al., 2006; Tzeng and Kalodimos, 2009). Thermodynamic mutation cycles (Schreiber and Fersht, 1995), which measure the coupling between two mutation sites by their mutual contribution to protein stability, provide a direct method to systematically probe such relations between protein sites. However, due to experimental limitations and practical considerations, these methods make large-scale studies of proteins laborious and time consuming.

Various computational methods have also been proposed to probe the coupling of amino acids and to identify the networks of residues important for protein conformational changes. Sequence-based approaches have been applied to reveal the allosterically important residues in proteins, based on the argument that energetically/functionally coupled (allosteric) residues also co-evolve (Lockless and Ranganathan, 1999; Socolich et al., 2005; Suel et al., 2003; Zheng et al., 2006). However, the application of sequence-based approaches is limited by the availability of homologous sequences, and is complicated by the fact that evolutionary conservation is driven by factors other than function, such as stability and folding kinetics (Ding and Dokholyan, 2006). Motivated by the observation that a protein's dynamics to a large extent is often determined by its structure (Okazaki et al., 2006; Plaxco et al., 1998), structure-based local thermodynamic analyses (Freire, 2000; Pan et al., 2000), normal mode analysis (Changeux and Edelstein, 2001; Van Wynsberghe and Cui, 2006; Zheng et al., 2006) and Gaussian-network models (Tehver et al., 2009; Temiz and Bahar, 2002; Xu et al., 2003) have been used to study allostery in several proteins. These experimental and computational studies share the view that the communication between distal residues within the protein is mainly determined by complex interaction networks in the protein structure (Chang et al., 1993; Go, 1983; Onuchic and Wolynes, 2004).

Recently, we have demonstrated that function of the cystic fibrosis causing mutant chloride channel is modulated by allostery within one of its nucleotide binding domains (Aleksandrov et al., 2010; Aleksandrov et al., 2012)

[Kota et al., submitted 2012]. Here, we applied this method to understand allosteric signal transduction within ENaC.

5.1 Role of the N-terminus in activation of ENaC

Mounting evidence suggests that the positive charges in the N-terminus of different subunits of ENaC interact with anionic phospholipids in the inner leaflet of the plasma membrane. Putting the results from these reports in context of proteolytic activation of ENaC, we were motivated to understand how interaction of intracellular domains with PIP2 translates to proteolytic cleavage of the extracellular domain and result to channel activation. We therefore mutated the positive charges in the N-terminus of the β- and γ- subunits of rat ENaC to study their effect on channel function.

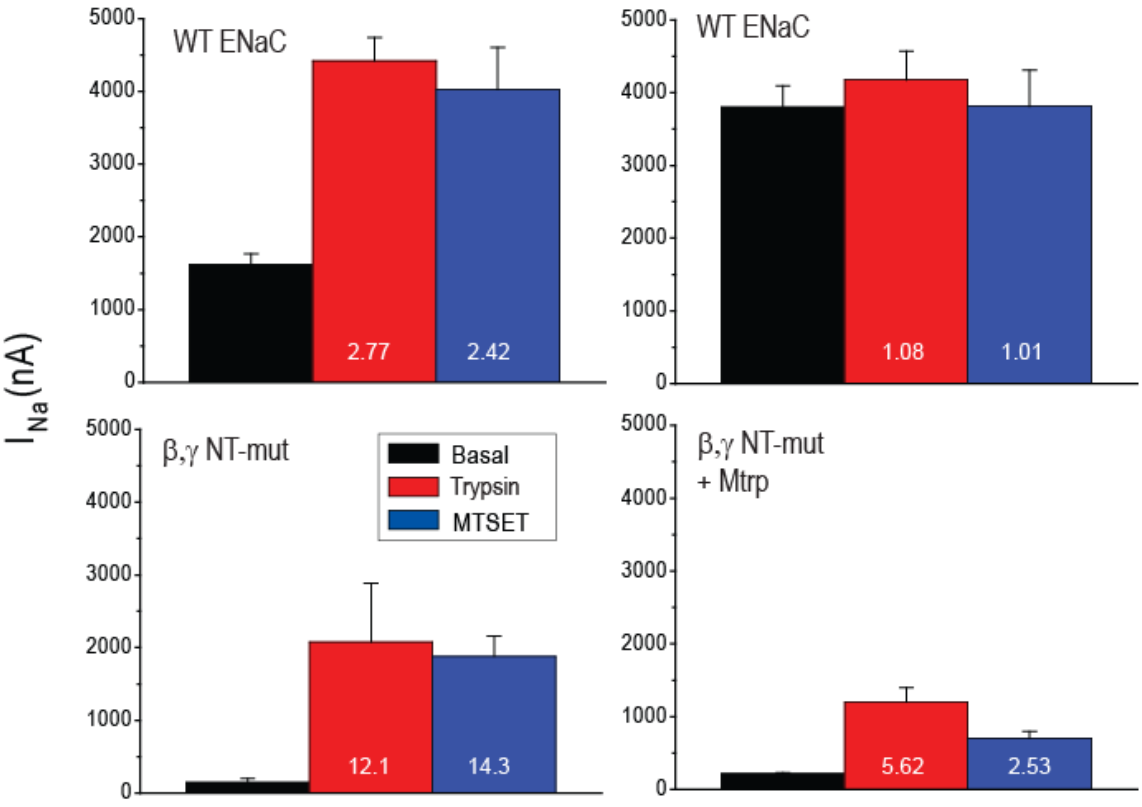


Figure 5.1 N-terminal basic stretch is critical for channel function
(Top left) Co-expression of WT ENaC subunits representing characteristic basal currents (black) that can be increased upon activation of the channel both by MTSET (blue) and trypsin (red). (Top right) Co-expression of matriptase increases basal currents with no apparent increase in trypsin-mediated or MTSET recovered currents. (Bottom left) Co-expression of WT αENaC with β- and γ-subunits lacking the N-terminal positive charges in oocytes indicates significant decrease in recovery by MTSET and trypsin. (Bottom right) Co-expression of matriptase does not improve basal currents, with only meager improvement in total stimulated currents using trypsin or MTSET.

The top left panel of Figure 5.1 shows a characteristic pattern of basal currents from oocytes expressing ENaC, which can be activated upon treatment with MTSET or trypsin. The top right panel of figure 5.1 shows a similar response to co-expressed matriptase where ENaC is proteolytically processed before reaching the plasma membrane, indicated by a marked increase in basal currents, with no significant improvement upon treatment with trypsin or MTSET. Interestingly, mutation of the N-terminal positive charges decreased basal currents and the maximal trypsin-mediated or MTSET-recovered currents (Figure 5.1 – bottom left). This result suggests that the positive charges in the N-terminus are important for interaction with PIP2. Strikingly, however, when WT αENaC was co-expressed with N-terminal mutants of β- and γ-ENaC in presence of matriptase, the basal currents could not be recovered by matriptase, suggesting that the accessibility of the protease cleavage sites is decreased upon mutation of the positive charges in the N-terminus. These results provide evidence for a conformational change in the extracellular domain upon interaction of the N-termini with PIP2, thereby revealing the protease cleavage sites involved in robust channel activation.

5.2 Long-range interaction networks within γENaC

Our experiments with mutant forms of the β and γ subunits provided evidence for allosteric communication between the N-termini and the proteolytic cleavage sites in ENaC. Since the γ subunit is the most cleaved subunit, we focused on the effect of mutation of the N-terminus in γENaC on proteolytic activation of the channel. Using the method we have proposed to study allosteric coupling networks within macromolecules, we have identified networks of interacting residues within the gamma subunit of ENaC.

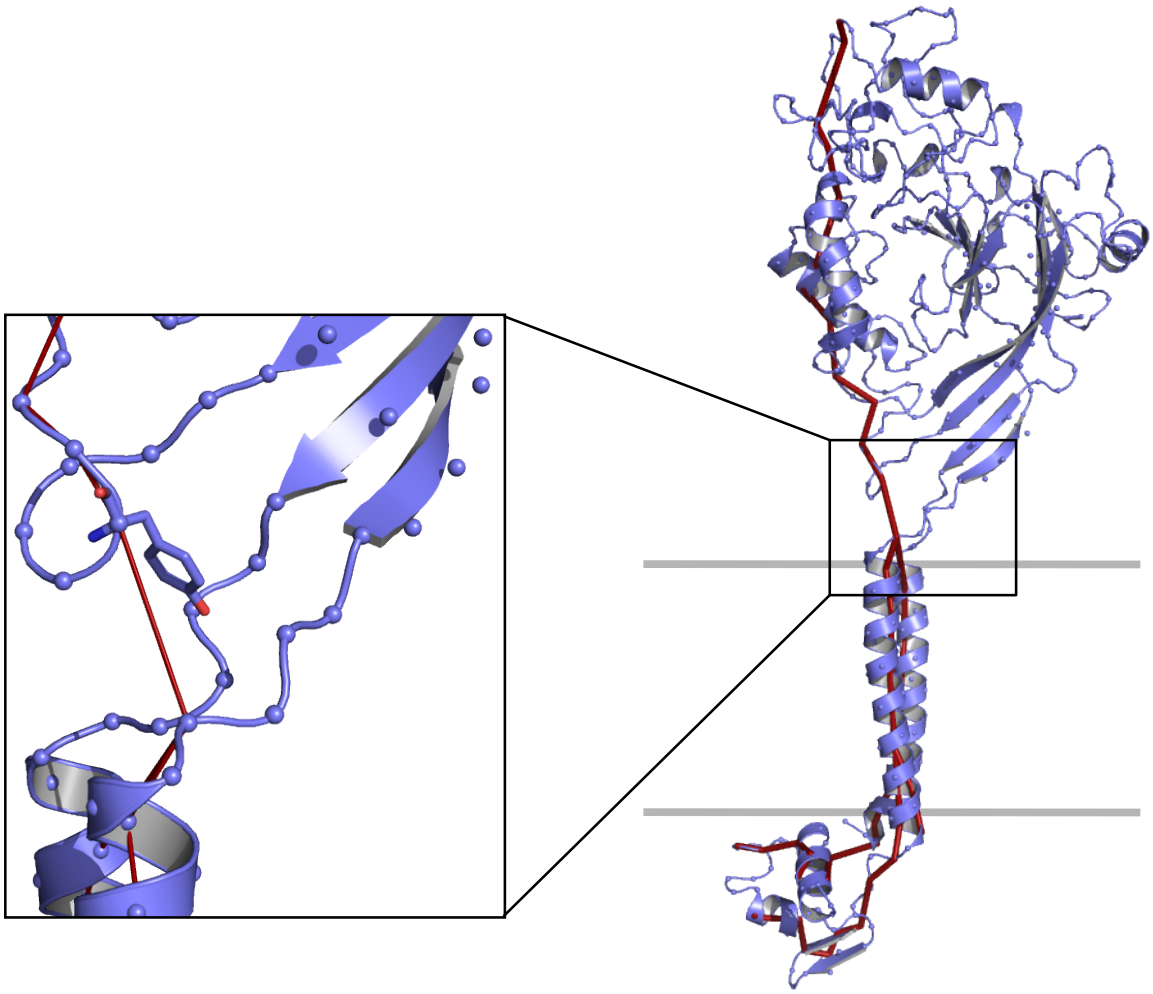


Figure 5.2 Allosteric networks in gamma ENaC
Structural model of gamma ENaC is shown in cartoon representation with the alpha carbon atoms shown as spheres. The network of interactions originating from the intracellular side to the protease cleavage site is colored red. The residue in the hinge between the transmembrane helices and the extracellular domain is show in stick representation (inset).

Based on our structural model, we hypothesized that the hinge between the transmembrane and the extracellular domains is critical for signal propagation across the protein. We therefore rationally mutated the hinge residue (Y370 – rat sequence numbering), to other possible amino acids using Eris (Yin et al., 2007a). We hypothesized that stabilization of the hinge residue would uncouple the allosteric sites involved in PIP2 binding and those involved in proteolytic cleavage. We identified two potential substitutions using Eris; Y370K and Y370N. We then performed site-specific mutagenesis of the chosen position to K or N and expressed the mutant ENaC subunit with WT ENaC subunits. We tested the effect of these mutations in context of the well-known degenerin mutant (β518C) to understand the mechanism of action of the chosen mutations. The chosen mutants in the gamma subunit decreased trypsin-mediated currents while MTSET mediated currents were unaffected. Y370K was more effective in decreasing the trypsin response compared to Y370N. These results suggested that the mechanism of channel opening by MTSET is independent of the mechanism of action of trypsin in increasing ENaC activity (Figure 5.3 – Top). To understand whether MTSET recovery can be intercepted by this manouver, we made a site-specific substitution in beta ENaC at a position equivalent to Y370 in the gamma subunit. We co-injected oocytes with cRNA encoding the degenerin mutant and wildtype alpha, gamma subunits. The equivalent substitution in the

beta subunit dramatically decreased both MTSET and trypsin recovery of whole cell currents in oocytes (Figure 5.3 – bottom).

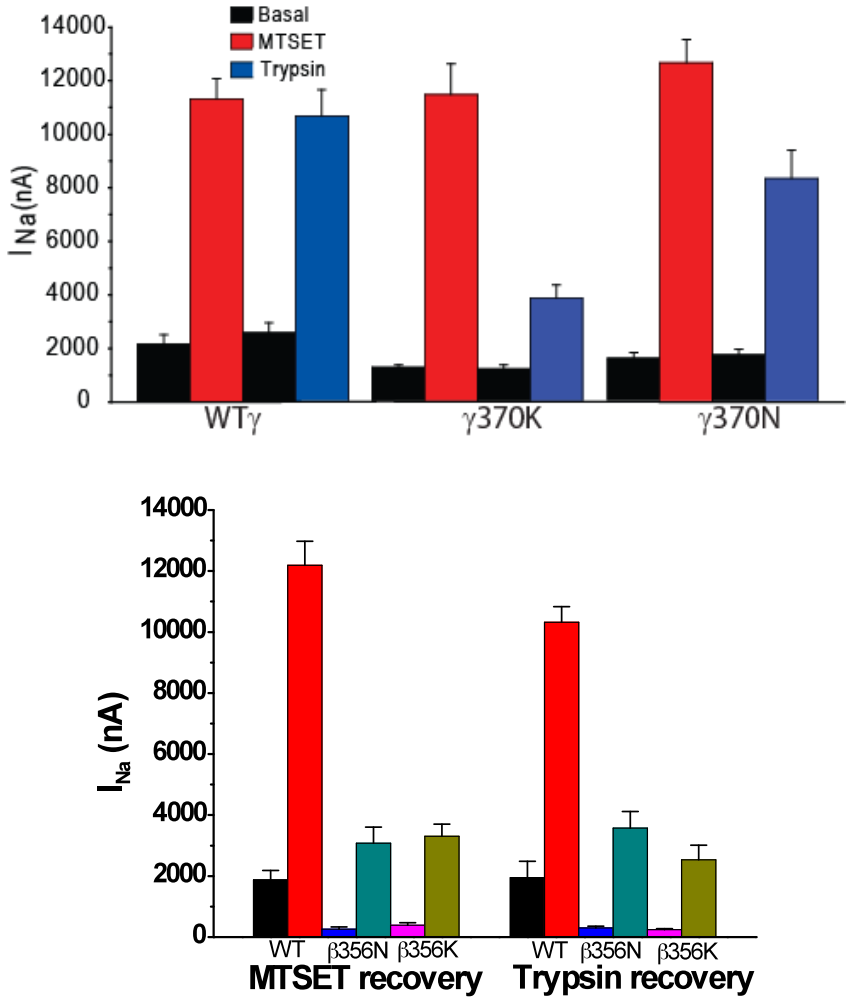


Figure 5.3 Y370 is a critical residue mediating allosteric propagation within ENaC
(Top) Oocytes were injected with cRNA encoding α,βS518C,γ ENaC, α,βS518C,γY370KENaC, α,βS518C,γY370NENaC and amiloride-sensitive sodium currents were measured 24 h post-injection. Black bars represent basal currents from oocytes. Red bars represent MTSET activated currents and the blue bars represent trypsin mediated currents. (Bottom) Oocytes were injected with cRNA encoding α,βS518C,γ ENaC, α,βS518C/Y370K,γENaC, α,βS518C/Y370N,γENaC and amiloride-sensitive sodium currents were measured 24 h post-injection. Left half of the plot corresponds to MTSET recovery of whole cell currents, while the right half represents trypsin recovery.

5.3 Interaction of the N-terminus with PIP2

As described earlier, the N-terminus of beta and gamma subunits of ENaC interact with PIP2 in the inner leaflet of the plasma membrane. In order to understand the structural aspects of the N-terminus in presence of PIP2, we performed DMD simulations of the N-terminal fragment in the presence of PIP2.

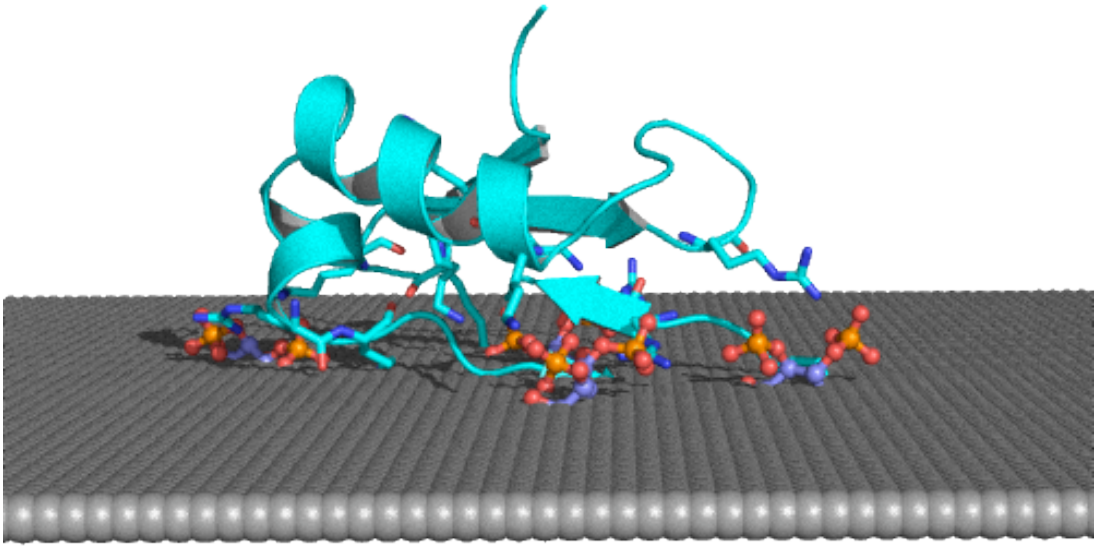


Figure 5.4 N-terminus of γENaC forms an α/β fold in presence of PIP2
Ab initio folding simulations of the N-terminus of gamma ENaC in presence of PIP2 impart stable secondary structures to the fragment. The protein is color cyan and PIP2 is represented as balls connected by sticks. The gray surface represents a hypothetical membrane surface to which the protein and PIP2 are constrained during the simulation.

We also imposed electrostatic constraints between the peptide and PIP2, to enable coulombic attraction/repulsion between the charged groups. The C-terminal amino acid of the peptide fragment was constrained to a plane, along with the head groups of PIP2, mimicking the native interaction at the membrane-cytosol interface. The positions of PIP2 on the plane were minimized before performing folding simulations. The peptide interacts with PIP2 such that the lysine's and arginines participate in electrostatic interactions with the head group

of the lipid molecule. The positive charges do not form any structural motifs to interact with PIP2. In order to verify experimentally, that the peptide interacts with PIP2 directly, we performed tryptophan fluorescence studies.

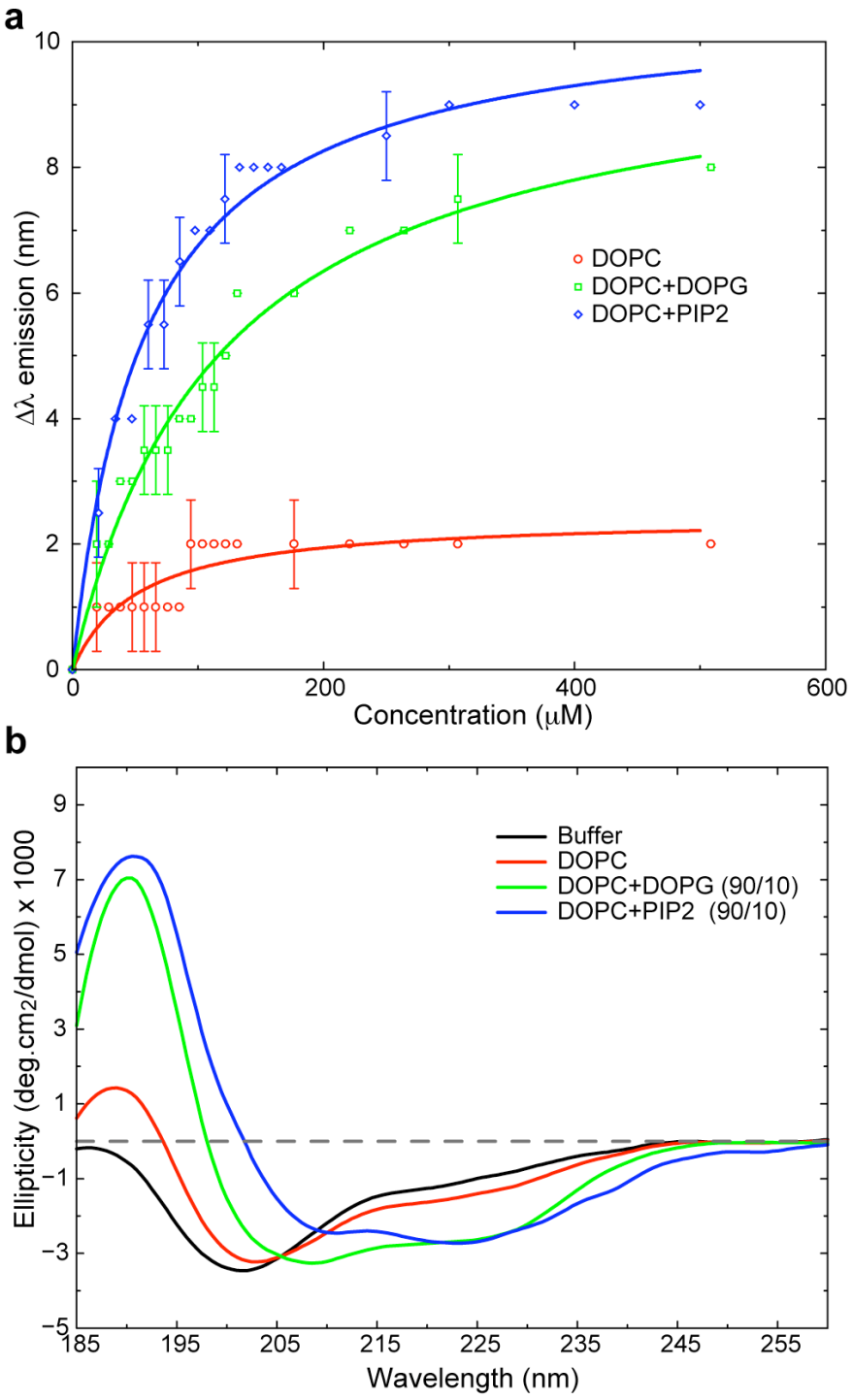


Figure 5.5. N-terminus changes secondary structural content upon binding PIP2
(a) Change in emission upon interaction with PIP2 is plotted as a function of concentration of lipid. Solid lines indicate fitted values for the experimental data. Higher the difference in emission, more shift towards the blue end of the visible spectrum. (b) Circular dichroism spectroscopy of the peptide in presence and absence of PIP2. Peaks at 220 and 208 nm indicate alpha helical content, while peaks below 200 nm indicate predominance of loops and unstructured regions. The peptide gains alpha helical content upon interaction with PIP2.

Pure peptide was synthesized and provided by Dr. Jan Kubelka and Dr. Ginka Buchner. Tryptophan fluorescence assay directly reflects changes in the environment of the target tryptophan in the protein under investigation. In our studies, we show that the tryptophan in the N-terminus of gamma ENaC undergoes more blue-shift in presence of PIP2 than in its absence, suggesting that the environment of the tryptophan becomes more hydrophobic in the presence of PIP2 (Figure 5.5A). This result can be interpreted in two ways; the tryptophan might directly be interacting with the hydrophobic tail of PIP2 or the peptide might change configuration and adopt a folded form in which the tryptophan is buried in the core of the folded domain. To determine whether the peptide undergoes change in secondary structural content upon interaction with PIP2, we determined the mean ellipticity of each amino acid in the peptide using circular dichroism (CD) spectroscopy (Figure 5.5B). From CD studies, we conclude that the N-terminus of the gamma subunit of rat ENaC gains alpha helical content upon interaction with PIP2.

CHAPTER 6

Conclusions and future directions

The study presented here focuses on the structural, biochemical and biophysical aspects of expression, activation and regulation of function of epithelial sodium channels. ENaC is a multimeric ion channel protein with a debatable subunit stoichiometry. Based on our results, we conclude that ENaC can form both trimers and tetramers, with the tetramer species being the more dominant and functional when all three subunits are expressed together. Interestingly, oligomerization state of ENaC is influenced by the expression level of the gamma subunit. Using biochemical and computational techniques, we believe that we marked the beginning of the end of a 20-year old controversy converning the subunit makeup of the channel. Further analysis regarding the biosynthetic processing and trafficking of ENaC would advance the field in understanding critical assembly steps involved in making of the tetramers.

ENaC is activated by a multitude of proteases both inside the cell and at the plasma membrane. Although the roles of alpha and gamma subunits of ENaC have been shown to be dominant in mediating channel activation via proteolytic cleavage, the structural and energetic aspects of activation remain

unclear. In this study, we focused on this aspect of channel activation and our results shed light on the structural aspects of channel activation via cleavage of the extracellular domain. Further analysis of the accessibility of the cleavage sites using substituted-cysteine accessibility method (SCAM) will provide insights into the conformational changes involved in proteolytic cleavage and channel activation.

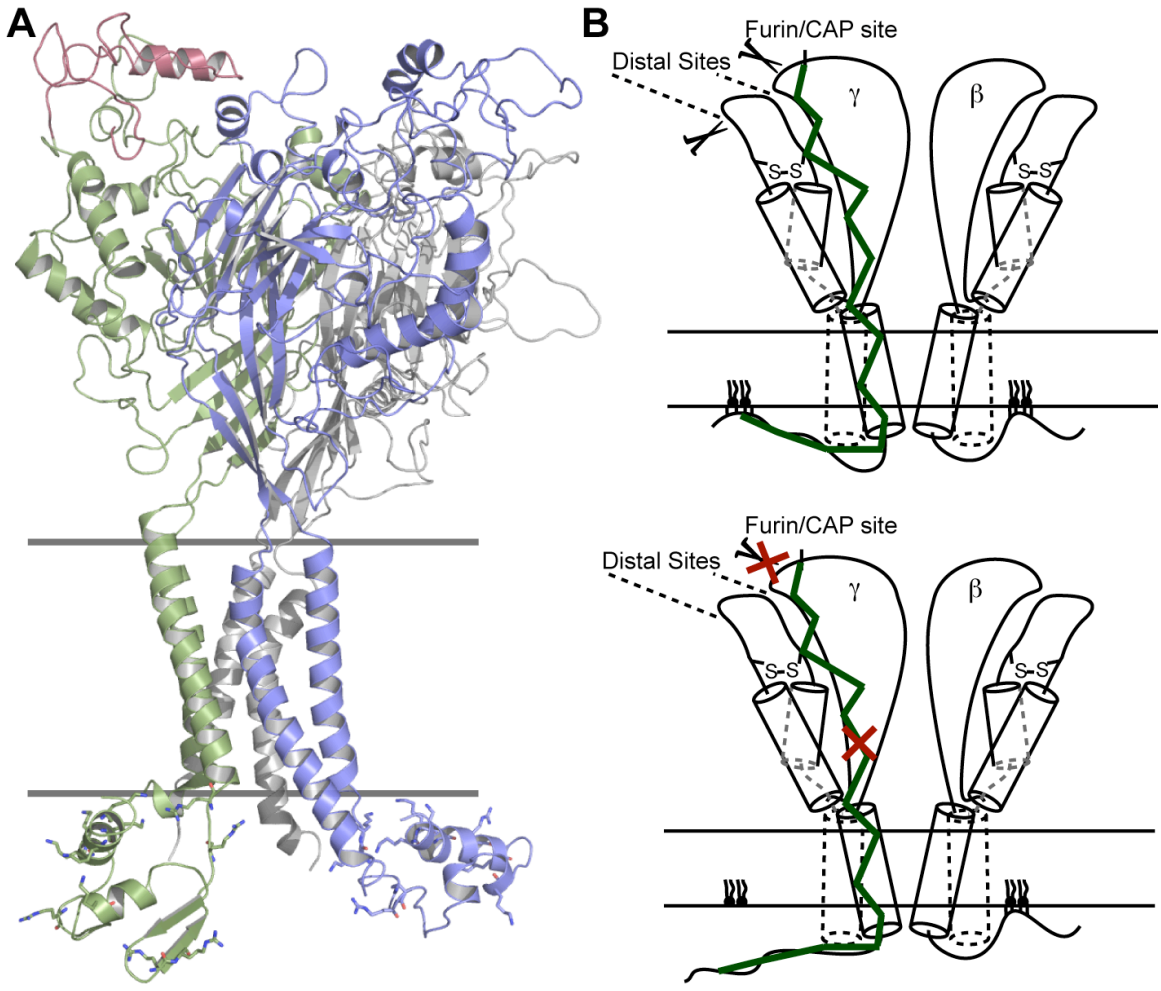


Figure 6.1 Model for regulation of ENaC activation
We propose an allosteric model for activation of ENaC upon interaction with PIP2. The trimeric assembly of ENaC with the structural models of the N-termini of beta and gamma subunits is shown in panel a. The N-terminus of alpha ENaC is hidden for clarity. We hypothesize that interaction of the N-terminal domains with PIP2 in the inner leaflet

leads to conformational change in the extracellular domain resulting in cleavage and channel activation (panel b- top), while mutations that abolish PIP2 binding hamper channel activation due to inaccessibility of the extracellular cleavage sites.

Using extensive computational simulations, we discovered long-range communication within ENaC that likely explains the mechanistic details of channel activation and regulation via interaction with PIP2. Based on our analysis of the N-terminal lysine residues and the hinge residue in gamma ENaC, we provide a putative model for activation of ENaC upon interaction of the N-terminus with PIP2 (Figure 6.1). Further work is needed to validate our hypothesis depicted in Figure 6.1. We believe that this study provide the essential framework to understand the allosteric aspects of ion channel activation in general and that of ENaC in particular. Furthermore, we believe that the results presented here are coherent with existing literature and aid in resolving existing controversies and providing new directions for understanding the molecular details of epithelial sodium channel structure, function and regulation.

Bibliography

Adachi, M., Kitamura, K., Miyoshi, T., Narikiyo, T., Iwashita, K., Shiraishi, N., Nonoguchi, H., and Tomita, K. (2001). Activation of epithelial sodium channels by prostasin in Xenopus oocytes. J Am Soc Nephrol *12*, 1114-1121.

Adams, C.M., Snyder, P.M., and Welsh, M.J. (1997). Interactions between subunits of the human epithelial sodium channel. J Biol Chem *272*, 27295-27300.

Akabas, M.H., Stauffer, D.A., Xu, M., and Karlin, A. (1992). Acetylcholine receptor channel structure probed in cysteine-substitution mutants. Science *258*, 307-310.

Aleksandrov, A.A., Kota, P., Aleksandrov, L.A., He, L., Jensen, T., Cui, L., Gentzsch, M., Dokholyan, N.V., and Riordan, J.R. (2010). Regulatory insertion removal restores maturation, stability and function of DeltaF508 CFTR. J Mol Biol *401*, 194-210.

Aleksandrov, A.A., Kota, P., Cui, L., Jensen, T., Alekseev, A.E., Reyes, S., He, L., Gentzsch, M., Aleksandrov, L.A., Dokholyan, N.V.*, et al.* (2012). Allosteric Modulation Balances Thermodynamic Stability and Restores Function of DeltaF508 CFTR. J Mol Biol.

Anantharam, A., and Palmer, L.G. (2007). Determination of epithelial Na+ channel subunit stoichiometry from single-channel conductances. J Gen Physiol *130*, 55-70.

Becker, O.M., Shacham, S., Marantz, Y., and Noiman, S. (2003). Modeling the 3D structure of GPCRs: advances and application to drug discovery. Curr Opin Drug Discov Devel *6*, 353-361.

Benaud, C., Dickson, R.B., and Lin, C.Y. (2001). Regulation of the activity of matriptase on epithelial cell surfaces by a blood-derived factor. Eur J Biochem *268*, 1439-1447.

Benos, D.J., Awayda, M.S., Ismailov, II, and Johnson, J.P. (1995). Structure and function of amiloride-sensitive Na+ channels. J Membr Biol *143*, 1-18.

Benos, D.J., and Stanton, B.A. (1999). Functional domains within the degenerin/epithelial sodium channel (Deg/ENaC) superfamily of ion channels. J Physiol *520 Pt 3*, 631-644.

Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., and Bourne, P.E. (2000). The Protein Data Bank. Nucleic Acids Res *28*, 235-242.

Booth, R.E., Tong, Q., Medina, J., Snyder, P.M., Patel, P., and Stockand, J.D. (2003). A region directly following the second transmembrane domain in gamma ENaC is required for normal channel gating. J Biol Chem *278*, 41367-41379.

Bruns, J.B., Carattino, M.D., Sheng, S., Maarouf, A.B., Weisz, O.A., Pilewski, J.M., Hughey, R.P., and Kleyman, T.R. (2007). Epithelial Na+ channels are fully activated by furin- and prostasin-dependent release of an inhibitory peptide from the gamma-subunit. J Biol Chem *282*, 6153-6160.

Bugge, T.H., Antalis, T.M., and Wu, Q. (2009). Type II Transmembrane Serine Proteases. Journal of Biological Chemistry *284*, 23177-23181.

Bystroff, C. (2002). MASKER: improved solvent-excluded molecular surface area estimations using Boolean masks. Protein Eng *15*, 959-965.

Caldwell, R.A., Boucher, R.C., and Stutts, M.J. (2004). Serine protease activation of near-silent epithelial Na+ channels. Am J Physiol Cell Physiol *286*, C190-194.

Caldwell, R.A., Boucher, R.C., and Stutts, M.J. (2005). Neutrophil elastase activates near-silent epithelial Na+ channels and increases airway epithelial Na+ transport. Am J Physiol Lung Cell Mol Physiol *288*, L813-819.

Canessa, C.M., Horisberger, J.-D., and Rossier, B.C. (1993). Epithelial sodium channel related to proteins involved in neurodegeneration. Nature *361*, 467-470.

Canessa, C.M., Merillat, A.M., and Rossier, B.C. (1994a). Membrane topology of the epithelial sodium channel in intact cells. Am J Physiol *267*, C1682-1690.

Canessa, C.M., Schild, L., Buell, G., Thorens, B., Gautschi, I., Horisberger, J.-D., and Rossier, B.C. (1994b). Amiloride-sensitive epithelial Na+ channel is made of three homologous subunits. Nature *367*.

Carattino, M.D., Hughey, R.P., and Kleyman, T.R. (2008a). Proteolytic processing of the epithelial sodium channel gamma subunit has a dominant role in channel activation. J Biol Chem *283*, 25290-25295.

Carattino, M.D., Passero, C.J., Steren, C.A., Maarouf, A.B., Pilewski, J.M., Myerburg, M.M., Hughey, R.P., and Kleyman, T.R. (2008b). Defining an inhibitory domain in the alpha-subunit of the epithelial sodium channel. Am J Physiol Renal Physiol *294*, F47-52.

Chalfant, M.L., Denton, J.S., Langloh, A.L., Karlson, K.H., Loffing, J., Benos, D.J., and Stanton, B.A. (1999). The NH(2) terminus of the epithelial sodium channel contains an endocytic motif. J Biol Chem *274*, 32889-32896.

Chang, S.S., Grunder, S., Hanukoglu, A., Rosler, A., Mathew, P.M., Hanukoglu, I., Schild, L., Lu, Y., Shimkets, R.A., Nelson-Williams, C*., et al.* (1996). Mutations in subunits of the epithelial sodium channel cause salt wasting with hyperkalaemic acidosis, pseudohypoaldosteronism type 1. Nat Genet *12*, 248-253.

Chang, X.B., Tabcharani, J.A., Hou, Y.X., Jensen, T.J., Kartner, N., Alon, N., Hanrahan, J.W., and Riordan, J.R. (1993). Protein kinase A (PKA) still activates CFTR chloride channel after mutagenesis of all 10 PKA consensus phosphorylation sites. J Biol Chem *268*, 11304-11311.

Changeux, J., and Edelstein, S.J. (2001). Allosteric mechanisms in normal and pathological nicotinic acetylcholine receptors. Curr Opin Neurobiol *11*, 369-377.

Changeux, J.P., and Edelstein, S.J. (2005). Allosteric mechanisms of signal transduction. Science *308*, 1424-1428.

Cheng, C., Prince, L.S., Snyder, P.M., and Welsh, M.J. (1998). Assembly of the epithelial Na+ channel evaluated using sucrose gradient sedimentation analysis. J Biol Chem *273*, 22693-22700.

Connolly, M.L. (1983). Solvent-accessible surfaces of proteins and nucleic acids. Science *221*, 709-713.

Corey, D.P., and Garcia-Anoveros, J. (1996). Mechanosensation and the DEG/ENaC ion channels. Science *273*, 323-324.

Coscoy, S., Lingueglia, E., Lazdunski, M., and Barbry, P. (1998). The Phe-Met-Arg-Phe-amide-activated sodium channel is a tetramer. J Biol Chem *273*, 8317-8322.

Davis, I.W., Leaver-Fay, A., Chen, V.B., Block, J.N., Kapral, G.J., Wang, X., Murray, L.W., Arendall, W.B., 3rd, Snoeyink, J., Richardson, J.S*., et al.* (2007). MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res *35*, W375-383.

Ding, F., and Dokholyan, N.V. (2006). Emergence of protein fold families through rational design. PLoS computational biology *2*, e85.

Ding, F., Tsao, D., Nie, H., and Dokholyan, N. V. (2008). Ab initio folding of proteins with all-atom discrete molecular dynamics. Structure *16*, 1010-1018.

Dokholyan, N.V., Buldyrev, S.V., Stanley, H.E., and Shakhnovich, E.I. (1998). Discrete molecular dynamics studies of the folding of a protein-like model. Fold Des *3*, 577-587.

Donaldson, S.H., Hirsh, A., Li, D.C., Holloway, G., Chao, J., Boucher, R.C., and Gabriel, S.E. (2002). Regulation of the Epithelial Sodium Channel by Serine Proteases in Human Airways. J Biol Chem *277*, 8338-8345.

Doyle, D.A., Morais Cabral, J., Pfuetzner, R.A., Kuo, A., Gulbis, J.M., Cohen, S.L., Chait, B.T., and MacKinnon, R. (1998). The structure of the potassium channel: molecular basis of K+ conduction and selectivity. Science *280*, 69-77.

Driscoll, M., and Tavernarakis, N. (1997). Molecules that mediate touch transduction in the nematode Caenorhabditis elegans. Gravit Space Biol Bull *10*, 33-42.

Duc, C., Farman, N., Canessa, C.M., Bonvalet, J.-P., and Rossier, B.C. (1994). Cell specific expression of epithelial sodium channel α, β and γ subunits in

aldosterone responsive epithelia from the rat: localization by in situ hybridization and immunocytochemistry. J Cell Biol *127*, 1907-1921.

Dunbrack, R.L., Jr., and Cohen, F.E. (1997). Bayesian statistical analysis of protein side-chain rotamer preferences. Protein Sci *6*, 1661-1681.

Feig, M., Karanicolas, J., and Brooks, C.L., 3rd (2004). MMTSB Tool Set: enhanced sampling and multiscale modeling methods for applications in structural biology. J Mol Graph Model *22*, 377-395.

Firsov, D., Gautschi, I., Merillat, A.M., Rossier, B.C., and Schild, L. (1998). The heterotetrameric architecture of the epithelial sodium channel (ENaC). EMBO J *17*, 344-352.

Firsov, D., Schild, L., Gautschi, I., Merillat, A.M., Schneeberger, E., and Rossier, B.C. (1996). Cell surface expression of the epithelial Na channel and a mutant causing Liddle syndrome: a quantitative approach. Proc Natl Acad Sci U S A *93*, 15370-15375.

Fiser, A., Do, R.K., and Sali, A. (2000). Modeling of loops in protein structures. Protein Sci *9*, 1753-1773.

Fiser, A., and Sali, A. (2003). ModLoop: automated modeling of loops in protein structures. Bioinformatics *19*, 2500-2501.

Flower, D.R. (1999). Modelling G-protein-coupled receptors for drug design. Biochim Biophys Acta *1422*, 207-234.

Freire, E. (2000). Can allosteric regulation be predicted from structure? Proc Natl Acad Sci U S A *97*, 11680-11682.

Friedrich, R., Fuentes-Prior, P., Ong, E., Coombs, G., Hunter, M., Oehler, R., Pierson, D., Gonzalez, R., Huber, R., Bode, W.*, et al.* (2002). Catalytic Domain Structures of MT-SP1/Matriptase, a Matrix-degrading Transmembrane Serine Proteinase. J Biol Chem *277*, 2160-2168.

Fuentes, E.J., Gilmore, S.A., Mauldin, R.V., and Lee, A.L. (2006). Evaluation of energetic and dynamic coupling networks in a PDZ domain protein. J Mol Biol *364*, 337-351.

Garcia-Anoveros, J., and Corey, D.P. (1997). The molecules of mechanosensation. Annu Rev Neurosci *20*, 567-594.

Garcia-Anoveros, J., Derfler, B., Neville-Golden, J., Hyman, B.T., and Corey, D.P. (1997). BNaC1 and BNaC2 constitute a new family of human neuronal sodium channels related to degenerins and epithelial sodium channels. Proc Natl Acad Sci U S A *94*, 1459-1464.

Garcia-Caballero, A., Dang, Y., He, H., and Stutts, M.J. (2008). ENaC proteolytic regulation by channel-activating protease 2. J Gen Physiol *132*, 521-535.

Garty, H. (1994). Molecular properties of epithelial, amiloride-blockable Na+ channels. FASEB J *8*, 522-528.

Garty, H., and Palmer, L.G. (1997). Epithelial sodium channels: function, structure, and regulation. Physiol Rev *77*, 359-396.

Go, N. (1983). Theoretical studies of protein folding. Annu Rev Biophys Bioeng *12*, 183-210.

Gonzales, E.B., Kawate, T., and Gouaux, E. (2009). Pore architecture and ion sites in acid-sensing ion channels and P2X receptors. Nature *460*, 599-604.

Goodey, N.M., and Benkovic, S.J. (2008). Allosteric regulation and catalysis emerge via a common route. Nature chemical biology *4*, 474-482.

Grunder, S., Firsov, D., Chang, S.S., Jaeger, N.F., Gautschi, I., Schild, L., Lifton, R.P., and Rossier, B.C. (1997). A mutation causing pseudohypoaldosteronism type 1 identifies a conserved glycine that is involved in the gating of the epithelial sodium channel. EMBO J *16*, 899-907.

Gulnik, S., Erickson, J.W., and Xie, D. (2000). HIV protease: enzyme function and drug resistance. Vitam Horm *58*, 213-256.

Gunasekaran, K., Ma, B., and Nussinov, R. (2004). Is allostery an intrinsic property of all dynamic proteins? Proteins *57*, 433-443.

Hansson, J.H., Nelson-Williams, C., Suzuki, H., Schild, L., Shimkets, R., Lu, Y., Canessa, C., Iwasaki, T., Rossier, B., and Lifton, R.P. (1995). Hypertension caused by a truncated epithelial sodium channel gamma subunit: genetic heterogeneity of Liddle syndrome. Nat Genet *11*, 76-82.

Hardy, J.A., and Wells, J.A. (2004). Searching for new allosteric sites in enzymes. Curr Opin Struct Biol *14*, 706-715.

Harris, M., Firsov, D., Vuagniaux, G., Stutts, M.J., and Rossier, B.C. (2007). A Novel Neutrophil Elastase Inhibitor Prevents Elastase Activation and Surface Cleavage of the Epithelial Sodium Channel Expressed in Xenopus laevis Oocytes. J Biol Chem *282*, 58-64.

Harris, M., Garcia-Caballero, A., Stutts, M.J., Firsov, D., and Rossier, B.C. (2008). Preferential assembly of epithelial sodium channel (ENaC) subunits in Xenopus oocytes: role of furin-mediated endogenous proteolysis. J Biol Chem *283*, 7455-7463.

Hedstrom, L. (2002). Serine protease mechanism and specificity. Chem Rev *102*, 4501-4524.

Henrich, S., Cameron, A., Bourenkov, G.P., Kiefersauer, R., Huber, R., Lindberg, I., Bode, W., and Than, M.E. (2003). The crystal structure of the proprotein processing proteinase furin explains its stringent specificity. Nat Struct Biol *10*, 520-526.

Hillisch, A., Pineda, L.F., and Hilgenfeld, R. (2004). Utility of homology models in the drug discovery process. Drug Discov Today *9*, 659-669.

Hooft, R.W., Vriend, G., Sander, C., and Abola, E.E. (1996). Errors in protein structures. Nature *381*, 272.

Hubbard, S.J., Campbell, S.F., and Thornton, J.M. (1991). Molecular recognition. Conformational analysis of limited proteolytic sites and serine proteinase protein inhibitors. J Mol Biol *220*, 507-530.

Hughey, R.P., Bruns, J.B., Kinlough, C.L., Harkleroad, K.L., Tong, Q., Carattino, M.D., Johnson, J.P., Stockand, J.D., and Kleyman, T.R. (2004). Epithelial sodium channels are activated by furin-dependent proteolysis. J Biol Chem *279*, 18111-18114.

Hughey, R.P., Mueller, G.M., Bruns, J.B., Kinlough, C.L., Poland, P.A., Harkleroad, K.L., Carattino, M.D., and Kleyman, T.R. (2003). Maturation of the epithelial Na+ channel involves proteolytic processing of the alpha- and gamma-subunits. J Biol Chem *278*, 37073-37082.

Inoue, J., Iwaoka, T., Tokunaga, H., Takamune, K., Naomi, S., Araki, M., Takahama, K., Yamaguchi, K., and Tomita, K. (1998). A family with Liddle's syndrome caused by a new missense mutation in the beta subunit of the epithelial sodium channel. J Clin Endocrinol Metab *83*, 2210-2213.

Ismailov, II, Shlyonsky, V.G., Serpersu, E.H., Fuller, C.M., Cheung, H.C., Muccio, D., Berdiev, B.K., and Benos, D.J. (1999). Peptide inhibition of ENaC. Biochemistry *38*, 354-363.

Jasti, J., Furukawa, H., Gonzales, E.B., and Gouaux, E. (2007). Structure of acid-sensing ion channel 1 at 1.9 A resolution and low pH. Nature *449*, 316-323.

Kaufmann, K.W., Lemmon, G.H., Deluca, S.L., Sheehan, J.H., and Meiler, J. (2010). Practically useful: what the Rosetta protein modeling suite can do for you. Biochemistry *49*, 2987-2998.

Kawate, T., Michel, J.C., Birdsong, W.T., and Gouaux, E. (2009). Crystal structure of the ATP-gated P2X(4) ion channel in the closed state. Nature *460*, 592-598.

Kishi, K., Yamazaki, K., Yasuda, I., Yahagi, N., Ichinose, M., Tsuchiya, Y., Athauda, S.B.P., Inoue, H., and Takahashi, K. (2001). Characterization of a Membrane-Bound Arginine-Specific Serine Protease from Rat Intestinal Mucosa. J Biochem *130*, 425-430.

Klabunde, T., and Hessler, G. (2002). Drug design strategies for targeting G-protein-coupled receptors. Chembiochem *3*, 928-944.

Kleyman, T.R., Carattino, M.D., and Hughey, R.P. (2009). ENaC at the cutting edge: regulation of epithelial sodium channels by proteases. J Biol Chem *284*, 20447-20451.

Kosari, F., Sheng, S., Li, J., Mak, D.-O.D., Foskett, J.K., and Kleyman, T.R. (1998). Subunit stoichiometry of the epithelial sodium channel. J Biol Chem *273*, 13469-13474.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R.*, et al.* (2007). Clustal W and Clustal X version 2.0. Bioinformatics *23*, 2947-2948.

Lazaridis, T., and Karplus, M. (1999). Effective energy function for proteins in solution. Proteins *35*, 133-152.

Le Grand, S.M., and Merz, K.M.J. (1993). Rapid approximation to molecular surface area via the use of boolean logic and look-up tables. J Comput Chem *14*, 349-352.

Lingueglia, E., Champigny, G., Lazdunski, M., and Barbry, P. (1995). Cloning of the amiloride-sensitive FMRFamide peptide-gated sodium channel. Nature *378*, 730-733.

List, K., Hobson, J.P., Molinolo, A., and Bugge, T.H. (2007). Co-localization of the channel activating protease prostasin/(CAP1/PRSS8) with its candidate activator, matriptase. J Cell Physiol *213*, 237-245.

Lockless, S.W., and Ranganathan, R. (1999). Evolutionarily conserved pathways of energetic connectivity in protein families. Science *286*, 295-299.

Ma, H.P., and Eaton, D.C. (2005). Acute regulation of epithelial sodium channel by anionic phospholipids. J Am Soc Nephrol *16*, 3182-3187.

Ma, H.P., Saxena, S., and Warnock, D.G. (2002). Anionic phospholipids regulate native and expressed epithelial sodium channel (ENaC). J Biol Chem *277*, 7641-7644.

MacKinnon, R. (1991). Determination of the subunit stoichiometry of a voltage-activated potassium channel. Nature *350*, 232-235.

Matthews, D.J., Goodman, L.J., Gorman, C.M., and Wells, J.A. (1994). A survey of furin substrate specificity using substrate phage display. Protein Sci *3*, 1197-1205.

May, L.T., Leach, K., Sexton, P.M., and Christopoulos, A. (2007). Allosteric modulation of G protein-coupled receptors. Annu Rev Pharmacol Toxicol *47*, 1-51.

Miyake, Y., Yasumoto, M., Tsuzuki, S., Fushiki, T., and Inouye, K. (2009). Activation of a Membrane-Bound Serine Protease Matriptase on the Cell Surface. J Biochem *146*, 273-282.

Mueller, G.M., Maarouf, A.B., Kinlough, C.L., Sheng, N., Kashlan, O.B., Okumura, S., Luthy, S., Kleyman, T.R., and Hughey, R.P. (2010). Cys palmitoylation of the beta subunit modulates gating of the epithelial sodium channel. J Biol Chem *285*, 30453-30462.

Neurath, H. (1984). Evolution of proteolytic enzymes. Science *224*, 350-357.

Neurath, H., and Walsh, K.A. (1976). Role of proteolytic enzymes in biological regulation (a review). Proc Natl Acad Sci U S A *73*, 3825-3832.

Okazaki, K., Koga, N., Takada, S., Onuchic, J.N., and Wolynes, P.G. (2006). Multiple-basin energy landscapes for large-amplitude conformational motions of proteins: Structure-based molecular dynamics simulations. Proc Natl Acad Sci U S A *103*, 11844-11849.

Onuchic, J.N., and Wolynes, P.G. (2004). Theory of protein folding. Curr Opin Struct Biol *14*, 70-75.

Pan, H., Lee, J.C., and Hilser, V.J. (2000). Binding sites in Escherichia coli dihydrofolate reductase communicate by modulating the conformational ensemble. Proc Natl Acad Sci U S A *97*, 12020-12025.

Passero, C.J., Mueller, G.M., Rondon-Berrios, H., Tofovic, S.P., Hughey, R.P., and Kleyman, T.R. (2008). Plasmin Activates Epithelial Na+ Channels by Cleaving the {gamma} Subunit. J Biol Chem *283*, 36586-36591.

Patny, A., Desai, P.V., and Avery, M.A. (2006). Homology modeling of G-protein-coupled receptors and implications in drug design. Curr Med Chem *13*, 1667-1691.

Plaxco, K.W., Simons, K.T., and Baker, D. (1998). Contact order, transition state placement and the refolding rates of single domain proteins. J Mol Biol *277*, 985-994.

Pochynyuk, O., Staruschenko, A., Tong, Q., Medina, J., and Stockand, J.D. (2005). Identification of a functional phosphatidylinositol 3,4,5-trisphosphate binding site in the epithelial Na+ channel. J Biol Chem *280*, 37565-37571.

Pochynyuk, O., Tong, Q., Medina, J., Vandewalle, A., Staruschenko, A., Bugaj, V., and Stockand, J.D. (2007). Molecular determinants of PI(4,5)P2 and PI(3,4,5)P3 regulation of the epithelial Na+ channel. J Gen Physiol *130*, 399-413.

Popovych, N., Sun, S., Ebright, R.H., and Kalodimos, C.G. (2006). Dynamically driven protein allostery. Nature structural & molecular biology *13*, 831-838.

Prince, L.S., and Welsh, M.J. (1998). Cell surface expression and biosynthesis of epithelial Na+ channels. Biochem J *336 (Pt 3)*, 705-710.

Ramachandran, S., and Dokholyan, N.V. (2011). Homology modeling: Generating structural models to understand protein function and mechanism (Springer).

Ramachandran, S., Kota, P., Ding, F., and Dokholyan, N.V. (2011). Automated minimization of steric clashes in protein structures. Proteins *79*, 261-270.

Ramachandran, S., Serohijos, A.W., Xu, L., Meissner, G., and Dokholyan, N.V. (2009). A structural model of the pore-forming region of the skeletal muscle ryanodine receptor (RyR1). PLoS computational biology *5*, e1000367.

Renard, S., Voilley, N., Bassilana, F., Lazdunski, M., and Barbry, P. (1995). Localization and regulation by steroids of the alpha, beta and gamma subunits of the amiloride-sensitive Na^+ channel in colon, lung and kidney. Pflü gers Arch *430*, 299-307.

Rose, R.B., Craik, C.S., and Stroud, R.M. (1998). Domain flexibility in retroviral proteases: structural implications for drug resistant mutations. Biochemistry *37*, 2607-2621.

Rossier, B.C., Canessa, C.M., Schild, L., and Horisberger, J.D. (1994). Epithelial sodium channels. Curr Opin Nephrol Hypertens *3*, 487-496.

Rossier, B.M., and Stutts, J.M. (2008). Activation of the epithelial sodium channel (ENaC) by serine proteases. Annu Rev Physiol *71*, 16.11-16.19.

Rotkiewicz, P., and Skolnick, J. (2008). Fast procedure for reconstruction of full-atom protein models from reduced representations. J Comput Chem *29*, 1460-1465.

Ruffieux-Daidie, D., Poirot, O., Boulkroun, S., Verrey, F., Kellenberger, S., and Staub, O. (2008). Deubiquitylation regulates activation and proteolytic cleavage of ENaC. J Am Soc Nephrol *19*, 2170-2180.

Ruffieux-Daidie, D., and Staub, O. (2011). Intracellular ubiquitylation of the epithelial Na+ channel controls extracellular proteolytic channel activation via conformational change. J Biol Chem *286*, 2416-2424.

Santamaria, B., Estevez, A.M., Martinez-Costa, O.H., and Aragon, J.J. (2002). Creation of an allosteric phosphofructokinase starting with a nonallosteric enzyme. The case of dictyostelium discoideum phosphofructokinase. J Biol Chem *277*, 1210-1216.

Schreiber, G., and Fersht, A.R. (1995). Energetics of protein-protein interactions: analysis of the barnase-barstar interface by single mutations and double mutant cycles. J Mol Biol *248*, 478-486.

Serohijos, A.W., Chen, Y., Ding, F., Elston, T.C., and Dokholyan, N.V. (2006). A structural model reveals energy transduction in dynein. Proc Natl Acad Sci U S A *103*, 18540-18545.

Serohijos, A.W., Hegedus, T., Aleksandrov, A.A., He, L., Cui, L., Dokholyan, N.V., and Riordan, J.R. (2008). Phenylalanine-508 mediates a cytoplasmic-membrane domain contact in the CFTR 3D structure crucial to assembly and channel function. Proc Natl Acad Sci U S A *105*, 3256-3261.

Sharma, S., Ding, F., and Dokholyan, N.V. (2007). Multiscale modeling of nucleosome dynamics. Biophys J *92*, 1457-1470.

Shimkets, R.A., Warnock, D.G., Bositis, C.M., Nelson-Williams, C., Hansson, J.H., Schambelan, M., Gill, J.R., Jr., Ulick, S., Milora, R.V., Findling, J.W.*, et al.* (1994). Liddle's syndrome: heritable human hypertension caused by mutations in the beta subunit of the epithelial sodium channel. Cell *79*, 407-414.

Shipway, A., Danahay, H., Williams, J.A., Tully, D.C., Backes, B.J., and Harris, J.L. (2004). Biochemical characterization of prostasin, a channel activating protease. Biochemical and Biophysical Research Communications *324*, 953-963.

Snyder, P.M., Cheng, C., Prince, L.S., Rogers, J.C., and Welsh, M.J. (1998). Electrophysiological and biochemical evidence that DEG/ENaC cation channels are composed of nine subunits. J Biol Chem *273*, 681-684.

Snyder, P.M., McDonald, F.J., Stokes, J.B., and Welsh, M.J. (1994). Membrane topology of the amiloride-sensitive epithelial sodium channel. J Biol Chem *269*, 24379-24383.

Socolich, M., Lockless, S.W., Russ, W.P., Lee, H., Gardner, K.H., and Ranganathan, R. (2005). Evolutionary information for specifying a protein fold. Nature *437*, 512-518.

Staruschenko, A., Adams, E., Booth, R.E., and Stockand, J.D. (2005). Epithelial Na^+ channel subunit stoichiometry. Biophys J *88*, 3966-3975.

Staruschenko, A., Medina, J.L., Patel, P., Shapiro, M.S., Booth, R.E., and Stockand, J.D. (2004). Fluorescence resonance energy transfer analysis of subunit stoichiometry of the epithelial Na^+ channel. J Biol Chem *279*, 27729-27734.

Staub, O., Dho, S., Henry, P., Correa, J., Ishikawa, T., McGlade, J., and Rotin, D. (1996). WW domains of Nedd4 bind to the proline-rich PY motifs in the epithelial Na+ channel deleted in Liddle's syndrome. EMBO J *15*, 2371-2380.

Staub, O., Gautschi, I., Ishikawa, T., Breitschopf, K., Ciechanover, A., Schild, L., and Rotin, D. (1997). Regulation of stability and function of the epithelial Na+ channel (ENaC) by ubiquitination. EMBO J *16*, 6325-6336.

Stauffer, D.A., and Karlin, A. (1994). Electrostatic potential of the acetylcholine binding sites in the nicotinic receptor probed by reactions of binding-site cysteines with charged methanethiosulfonates. Biochemistry *33*, 6840-6849.

Stewart, A.P., Haerteis, S., Diakov, A., Korbmacher, C., and Edwardson, J.M. (2011). Atomic force microscopy reveals the architecture of the epithelial sodium channel (ENaC). J Biol Chem *286*, 31944-31952.

Suel, G.M., Lockless, S.W., Wall, M.A., and Ranganathan, R. (2003). Evolutionarily conserved networks of residues mediate allosteric communication in proteins. Nat Struct Biol *10*, 59-69.

Suh, B.C., and Hille, B. (2008). PIP2 is a necessary cofactor for ion channel function: how and why? Annu Rev Biophys *37*, 175-195.

Szabo, R., Hobson, J.P., List, K., Molinolo, A., Lin, C.Y., and Bugge, T.H. (2008). Potent inhibition and global co-localization implicate the transmembrane Kunitz-type serine protease inhibitor hepatocyte growth factor activator inhibitor-2 in the regulation of epithelial matriptase activity. J Biol Chem *283*, 29495-29504.

Szabo, R., Molinolo, A., List, K., and Bugge, T.H. (2007). Matriptase inhibition by hepatocyte growth factor activator inhibitor-1 is essential for placental development. Oncogene *26*, 1546-1556.

Takeuchi, T., Harris, J.L., Huang, W., Yan, K.W., Coughlin, S.R., and Craik, C.S. (2000). Cellular localization of membrane-type serine protease 1 and identification of protease-activated receptor-2 and single-chain urokinase-type plasminogen activator as substrates. J Biol Chem *275*, 26333-26342.

Tehver, R., Chen, J., and Thirumalai, D. (2009). Allostery wiring diagrams in the transitions that drive the GroEL reaction cycle. J Mol Biol *387*, 390-406.

Temiz, N.A., and Bahar, I. (2002). Inhibitor binding alters the directions of domain motions in HIV-1 reverse transcriptase. Proteins *49*, 61-70.

Thomas, G. (2002). Furin at the cutting edge: from protein traffic to embryogenesis and disease. Nat Rev Mol Cell Biol *3*, 753-766.

Tzeng, S.R., and Kalodimos, C.G. (2009). Dynamic activation of an allosteric regulatory protein. Nature *462*, 368-372.

Vallet, V., Chraibi, A., Gaeggeler, H.P., Horisberger, J.D., and Rossier, B.C. (1997). An epithelial serine protease activates the amiloride-sensitive sodium channel. Nature *389*, 607-610.

Van Wynsberghe, A.W., and Cui, Q. (2006). Interpreting correlated motions using normal mode analysis. Structure *14*, 1647-1653.

Vriend, G., and Sander, C. (1993). Quality control of protein models: Directional atomic contact analysis. J Appl Cryst *26*, 47-60.

Vuagniaux, G., Vallet, V., Jaeger, N.F., Hummler, E., and Rossier, B.C. (2002). Synergistic activation of ENaC by three membrane-bound channel-activating serine proteases (mCAP1, mCAP2, and mCAP3) and serum- and glucocorticoid-regulated kinase (Sgk1) in Xenopus Oocytes. J Gen Physiol *120*, 191-201.

Waldmann, R., Champigny, G., Bassilana, F., Voilley, N., and Lazdunski, M. (1995). Molecular cloning and functional expression of a novel amiloride-sensitive Na channel. J Biol Chem *270*, 27411-27414.

Ward, J.J., Sodhi, J.S., McGuffin, L.J., Buxton, B.F., and Jones, D.T. (2004). Prediction and functional analysis of native disorder in proteins from the three kingdoms of life. J Mol Biol *337*, 635-645.

Winarski, K.L., Sheng, N., Chen, J., Kleyman, T.R., and Sheng, S. (2010). Extracellular allosteric regulatory subdomain within the gamma subunit of the epithelial Na+ channel. J Biol Chem *285*, 26088-26096.

Xu, C., Tobi, D., and Bahar, I. (2003). Allosteric changes in protein structure computed by a simple mechanical model: hemoglobin T<-->R2 transition. J Mol Biol *333*, 153-168.

Yin, S., Ding, F., and Dokholyan, N.V. (2007a). Eris: an automated estimator of protein stability. Nat Methods *4*, 466-467.

Yin, S., Ding, F., and Dokholyan, N.V. (2007b). Modeling backbone flexibility improves protein stability estimation. Structure *15*, 1567-1576.

Yue, G., Malik, B., and Eaton, D.C. (2002). Phosphatidylinositol 4,5-bisphosphate (PIP2) stimulates epithelial sodium channel activity in A6 cells. J Biol Chem *277*, 11965-11969.

Zhang, Z.R., Chou, C.F., Wang, J., Liang, Y.Y., and Ma, H.P. (2010). Anionic phospholipids differentially regulate the epithelial sodium channel (ENaC) by interacting with alpha, beta, and gamma ENaC subunits. Pflugers Arch *459*, 377-387.

Zheng, W., Brooks, B.R., and Thirumalai, D. (2006). Low-frequency normal modes that describe allosteric transitions in biological nanomachines are robust to sequence variations. Proc Natl Acad Sci U S A *103*, 7664-7669.