Mechanism of Calcium-dependent Chloride Channel Activation by the Secreted Regulator CLCA1

Zeynep Yurtsever

Washington University in St. Louis

Follow this and additional works at: https://openscholarship.wustl.edu/art_sci_etds

Part of the Biochemistry Commons, Biophysics Commons, and the Cell Biology Commons

Recommended Citation
Yurtsever, Zeynep, "Mechanism of Calcium-dependent Chloride Channel Activation by the Secreted Regulator CLCA1" (2016). Arts & Sciences Electronic Theses and Dissertations. 789.
https://openscholarship.wustl.edu/art_sci_etds/789

This Dissertation is brought to you for free and open access by the Arts & Sciences at Washington University Open Scholarship. It has been accepted for inclusion in Arts & Sciences Electronic Theses and Dissertations by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.
Mechanism of Calcium-dependent Chloride Channel Activation
by the Secreted Regulator CLCA1

by

Zeynep Yurtsever

A dissertation presented to the
Graduate School of Arts & Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2016
St. Louis, Missouri
# TABLE OF CONTENTS

List of Figures ........................................................................................................ v
List of Tables ........................................................................................................ vi
Acknowledgements ............................................................................................... vii
Abstract of the Dissertation .................................................................................... ix

## CHAPTER 1: CLCA in Airway Disease ................................................................. 1
  1.1 Mucus cell metaplasia in airway disease ......................................................... 2
  1.2 CLCA in mucus cell metaplasia .................................................................. 5
  1.3 CLCA in Cystic Fibrosis ............................................................................. 7
  1.4 References .................................................................................................. 10

## CHAPTER 2: Structural Organization of CLCA1 ................................................ 13
  2.1 Misannotation .............................................................................................. 14
  2.2 Structure predictions .................................................................................. 16
    2.2.1 Signal peptide ...................................................................................... 16
    2.2.2 N-linked glycosylations ..................................................................... 16
    2.2.3 Distinct domains ................................................................................. 18
    2.2.4 Novel CLCA1 metalloprotease is unlike MMPs ................................. 19
  2.3 Crystallization Trials ................................................................................. 21
    2.3.1 Rationale ............................................................................................ 21
    2.3.2 Experimental design ............................................................................ 22
    2.3.3 Results ................................................................................................ 29
  2.4 Summary ..................................................................................................... 34
  2.5 References .................................................................................................. 35

## CHAPTER 3: Proteolytic Processing ............................................................... 38
  3.1 Conserved Cleavage Sequence .................................................................... 39
    3.1.1 Determination of cleavage site by Edman degradation ..................... 39
    3.1.2 Results ................................................................................................ 39
  3.2 N-terminal zincin metalloprotease ................................................................ 41
    3.2.1 Point mutations in the active site ......................................................... 41
    3.2.2 In vitro digestion assay ....................................................................... 43
    3.2.3 Fluorogenic peptide assay .................................................................. 46
  3.3 Features of cleavage .................................................................................... 48
    3.3.1 Cross-cleavage ..................................................................................... 48
    3.3.2 Intermolecular Cleavage .................................................................... 50
    3.3.3 Regulation by GAGs .......................................................................... 52
  3.4 Summary ..................................................................................................... 57
  3.5 References .................................................................................................. 58

## CHAPTER 4: CLCA1 Function as CaCC/TMEM16A Regulator ......................... 59
  4.1 Cleavage is necessary for function ............................................................... 60
6.5 References ......................................................................................................................... 116
APPENDIX ............................................................................................................................... 118
CV ........................................................................................................................................ 206
List of Figures

FIGURE 1. SCHEME OF IL-13 STIMULATED MCM IN THE LUNG. ......................................................... 3
FIGURE 2. ION CHANNELS IN THE SMALL AIRWAYS. ................................................................. 9
FIGURE 3. DOMAIN ORGANIZATION OF THE HUMAN CLCA PROTEINS .......................................... 15
FIGURE 4. N-LINKED GLYCOSYLATION PREDICTIONS FOR FULL-LENGTH WT CLCA1 ...................... 17
FIGURE 6. CLCA1 METALLOPROTEASE DOES NOT HAVE A PRODOMAIN. ...................................... 21
FIGURE 7. REPRESENTATIVE SIZE EXCLUSION AND ION CHANGE (MONOQ) CHROMATOGRAMS ......... 30
FIGURE 8. PURIFICATION OF RECOMBINANT VWA (297-478) EXPRESSED IN 293F CELLS. ............... 31
FIGURE 9. HOMOLOGY MODEL OF VWA DEPICTING CLUSTERS FOR SURFACE ENTROPY REDUCTION. 33
FIGURE 10. CLCA FAMILY MEMBERS HAVE A CONSERVED CLEAVAGE SEQUENCE. ......................... 40
FIGURE 11. POINT MUTATIONS IN THE ACTIVE SITE BLOCK SELF-CLEAVAGE IN CELLS. .................. 42
FIGURE 12. CLEAVAGE CAN BE RECONSTITUTED IN VITRO ............................................................. 45
FIGURE 13. CLCA1 HAS A NOVEL METALLOPROTEASE FOLD. ...................................................... 46
FIGURE 14. FLUOROGENIC PEPTIDE ASSAY SHOWS CLCA1 ENZYME IS ZINC DEPENDENT ............. 47
FIGURE 15. CROSS-CLEAVAGE BETWEEN CLCA PROTEINS CAN OCCUR. ...................................... 50
FIGURE 16. INTERMOLECULAR CLCA1 CLEAVAGE CAN OCCUR ...................................................... 51
FIGURE 17. SEQUENCE ALIGNMENT OF CLCAS IN THE N-TERMINAL CATALYTIC REGION ............ 52
FIGURE 18. GAGS ARE NOT REQUIRED FOR CLCA1 SELF-CLEAVAGE OR CELL SURFACE BINDING. 54
FIGURE 19. SELF-CLEAVAGE OF CLCA1 IS REQUIRED FOR CACC ACTIVATION ............................ 62
FIGURE 20. CLCA1 CAN ACTIVATE CACC IN A PARACRINE FASHION ............................................ 65
FIGURE 21. CLCA1-CONDITIONED MEDIUM INCREASES CACC CURRENTS ...................................... 66
FIGURE 22. PURE N-TERMINAL CLCA1 PROTEIN INCREASES CACC CURRENTS ............................. 68
FIGURE 23. SCHEMATIC OF TMEM16A .......................................................................................... 70
FIGURE 24. CLCA1 INCREASES TMEM16A CURRENTS .................................................................... 73
FIGURE 25. CLCA1 INCREASES TMEM16A CELL SURFACE STAINING .......................................... 76
FIGURE 26. CLCA1 BINDS TMEM16A ON CELL SURFACE ............................................................. 81
FIGURE 27. CLCA1 INCREASES TMEM16A STAINING WITHIN MINUTES ........................................ 83
FIGURE 28. CLCA1 BLOCKS RECYCLING OF TMEM16A FROM CELL SURFACE ................................ 86
FIGURE 29. NOCODOZOLE AFFECTS TMEM16A CURRENTS ............................................................. 87
FIGURE 30. PROPOSED MODEL OF INTERACTION ......................................................................... 88
FIGURE 32. SEQUENCE ALIGNMENT OF CLCA1-CLCA2-CLCA4 WITHIN THE VWA DOMAIN. .... 95
FIGURE 33. MIDAS MOTIF IN THE CLCA1 VWA MODEL ............................................................... 96
FIGURE 34. THE DOUBLE MIDAS MUTANT DOES NOT COMPLETELY BLOCK CLCA1 FUNCTION ....... 97
FIGURE 35. TRIPLE MIDAS MUTANTS DO NOT BEHAVE LIKE WT VWA ........................................ 98
FIGURE 36. EXTRACELLULAR Divalent CATIONS ARE REQUIRED FOR CACC ACTIVATION .......... 99
FIGURE 37. S357N SNP DOES NOT AFFECT CLCA1 FUNCTION ON TMEM16A ............................... 101
FIGURE 38. CLCA1 VWA DOMAIN CAN BIND TMEM16A LOOP IN VITRO BY BLI ......................... 103
FIGURE 39. SMALL PEPTIDES DO NOT BIND VWA IN VITRO ...................................................... 105
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TABLE 1</td>
<td>TRANSFECTION REAGENT RATIOS FOR TRANSIENT TRANSFECTION</td>
<td>25</td>
</tr>
<tr>
<td>TABLE 2</td>
<td>ANTIBODIES USED FOR WESTERN BLOTTING</td>
<td>27</td>
</tr>
<tr>
<td>TABLE 3</td>
<td>METALLOPROTEASE CONSTRUCTS TESTED FOR STRUCTURAL STUDIES</td>
<td>30</td>
</tr>
<tr>
<td>TABLE 4</td>
<td>LIST OF POINT MUTATIONS IN THE PREDICTED ACTIVE SITE</td>
<td>41</td>
</tr>
<tr>
<td>TABLE 5</td>
<td>ANTIBODIES USED IN confocal microscopy</td>
<td>56</td>
</tr>
<tr>
<td>TABLE 6</td>
<td>BLOCKING ANTIBODIES USED IN FLOW CYTOMETRY ASSAY</td>
<td>80</td>
</tr>
<tr>
<td>TABLE 7</td>
<td>LIST OF SMALL PEPTIDE FRAGMENTS</td>
<td>104</td>
</tr>
</tbody>
</table>
Acknowledgements:

Firstly, I would like to thank Drs. Colin Nichols and Monica Sala-Rabanal. Without their contributions, this study would not have been possible.

Secondly, I would like to acknowledge American Heart Association, who awarded me with a pre-doctoral fellowship and funded the last two years of this research project.

Lastly, I would like to thank my thesis committee and especially my thesis advisor, Dr. Tom Brett, for believing in me and helping me become a better scientist every day.

Zeynep Yurtsever

Washington University in St. Louis

May 2016
To my family,

In honor of die jüngste Mitarbeiterin.
Abstract of the Dissertation

Mechanism of Calcium-dependent Chloride Channel Activation
by the Secreted Regulator CLCA1

by

Zeynep Yurtsever

Doctor of Philosophy in Biology and Biomedical Sciences

Biochemistry

Washington University in St. Louis, 2016

Dr. Tom J. Brett, Chair

The calcium-activated chloride channel regulator (CLCA) proteins are key signaling molecules, which are implicated in various diseases through their tissue-specific expression. Human CLCA1 protein, overexpressed in airway epithelia under pathophysiological conditions, is centrally involved in the manifestation of IL-13-driven mucus cell metaplasia (MCM), a hallmark feature of asthma and chronic obstructive pulmonary disease (COPD), for which there are currently no available therapeutics. Elucidating the poorly understood molecular basis of CLCA1 function is thus required to design specific inhibitors of CLCA1 activity to treat MCM in asthma and COPD.

Originally misannotated as ion channels, CLCA proteins are secreted soluble proteins that act as regulators of unknown calcium-activated chloride channels (CaCC) in the airways. Additionally, these proteins undergo proteolytic cleavage, adding yet another level of complexity to the mystery of how a secreted protein can regulate chloride currents.
This study focuses on the proteolytic cleavage of CLCA1 and provides mechanistic detail into how secreted CLCA1 protein modulates currents from a CaCC in mammalian cells, which is identified as TMEM16A. For this unique interaction to occur, secreted full-length CLCA1 undergoes a self-cleavage reaction using an N-terminal zincin metalloprotease with a novel fold, which releases the functionally active N-terminal fragment. This cleavage exposes the von Willebrand factor type A (vWA) domain to directly interact with membrane-associated TMEM16A at the C-terminal extracellular loop (loop 9-10). The direct interaction increases TMEM16A surface density by stabilizing the channel on the cell surface and blocking its rapid recycling, thus resulting in increased chloride currents.
CHAPTER 1: CLCA in Airway Disease
1. **CLCA in Airway Disease**

1.1 **Mucus cell metaplasia in airway disease**

Asthma and chronic obstructive pulmonary disease (COPD) are leading causes of death, with millions of people afflicted worldwide each year \((1, 2)\). These deadly pulmonary diseases are characterized by inflammation, airway hyperreactivity (AHR) and mucus cell metaplasia (MCM) \((3-6)\). Currently available treatments for these diseases fall under two categories: beta-2 adrenergic receptor (β2AR) agonists, used as bronchodilators, to treat AHR and corticosteroids to induce an anti-inflammatory response \((7, 8)\). However, 30-40% of asthmatics do not respond to steroids \((9, 10)\), the “gold-standard asthma treatment”, due to heterogeneity of disease on a molecular level \((11)\). For this particular reason, identification of new drug targets in the airways has become crucial.

Under physiological conditions, the airway epithelium maintains an absorptive phenotype and contains various cell types, including ciliated cells, secretory club cells (previously known as Clara cells) and a small amount of secretory goblet cells \((12, 13)\). Under inflammatory conditions, IL-13 and similar cytokines can convert the epithelium to a secretory phenotype \((4, 12)\). This process, when unregulated, can lead to abnormal airway remodeling, which is observed in hypersecretory diseases such as cystic fibrosis (CF), chronic bronchitis and asthma \((14-16)\). The hypersecretion is maintained mostly by the increased abundance of goblet cells in the surface airway epithelium \((16)\), which is mainly a result of the IL-13-driven cell transdifferentiation process and not of goblet cell division \((17, 18)\) (Figure 1a).
Figure 1. Scheme of IL-13 stimulated MCM in the lung. A. IL-13 signaling results in transdifferentiation of lung epithelial cells into goblet cells. The significant increase in goblet cell count results in mucus overproduction and subsequent hypersecretion into the small airways, which leads to a thick mucus layer. B. IL-13 binding to receptor activates STAT6 signaling and induces CLCA1 overexpression. After getting secreted and cleaved, CLCA1 acts on a target channel, CaCC, and also signals for the activation of MAPK13 pathway, which induces MUC5AC expression and drives mucus production. Circle encloses the pathways focused on in this thesis. Figure modified from Sala-Rabanal et al. (19).
Recognition of IL-13 by its receptor (IL-13Rα1+IL-4Rα) (7) activates signaling pathways, resulting in phosphorylation of signal transducer and activator of transcription-6 (STAT6) (20). The phosphorylation event is followed by STAT6 translocation to the nucleus, where it regulates expression of an array of genes by directly binding specific promoters. IL-13-induced STAT6 signaling was shown to be essential for the reprogramming of basal airway cells, resulting in MCM and AHR (21, 22)(Figure 1b).

Normal mucus in the airways is made up largely of water, with a small percentage of solids (23). The major solid component of mucus is mucins, which are large glycoproteins with water-absorbing properties (24). While there are at least eleven mucin genes that are expressed in the lung, MUC5AC and MUC5B are the predominant mucins in airway secretion (25). MUC5B is largely secreted from the submucosal glands and MUC5AC is found secreted from goblet cells that line the epithelial surface (5, 17). The latter is considered the main mucin overproduced in hypersecretory diseases, as it is the only one inducible by IL-13 (18, 23). Hypersecretion of MUC5AC increases the viscosity of the mucus significantly, because the mucin monomers polymerize and occlude water molecules through their glycan chains (23).

Although this mucus can be cleared under normal conditions by the motion of ciliated cells, IL-13 signaling also reduces the number of ciliated cells and the ciliary beat frequency (CBF) of remaining cells (18, 26). Thus, MCM is characterized by the abundance of viscous, dehydrated mucus and impaired mucociliary clearance.
1.2 CLCA in mucus cell metaplasia

The first members of what we now call the CLCA family were initially characterized as lung endothelial cell adhesion molecules (27, 28) or ECAMs. A few years later, a new name was given to the family: “CLCA”, which stood for ChLoride channel Calcium-Activated, based on the predicted multiple transmembrane domains and the observed increase in calcium-activated chloride currents, $I_{CaCC}$, as a result of heterologous expression of various family members, such as hCLCA1, hCLCA2 and mCLCA4 (29-32).

hCLCA1 was described as the main family member expressed in the intestine under normal conditions and could not be detected in other human tissues (30). However, stimulation of primary epithelial cells with IL-4, IL-9 and IL-13 induced overexpression of this gene (18, 33, 34). Similarly, human tracheal epithelial cells (hTECs) treated with IL-13 significantly overexpressed CLCA1, whereas other family members were not induced by IL-13 (35). These hTECs present a good system to reconstitute “MCM in a dish”, since they produce MUC5AC in response to the IL-13-stimulated STAT6 signaling cascade (Figure 1b). In this system, CLCA1 mRNA levels correlated directly with MUC5AC levels and shRNA knockdown of CLCA1 suppressed mucin induction, suggesting CLCA1 signaling lies upstream of and is required for MUC5AC expression (35).
The homologue of hCLCA1 in mice, mCLCA3, has been the focus of researchers in developing a mouse model of asthma. Researchers were able to induce MCM in mice by intratrachaeal instillation of IL-13 (33, 36). The direct introduction of IL-13 resulted in the same disease attributes as in immunocompromised mice following ovalbumin challenge (37, 38) or viral infection in the airways (39). Following all three treatments, mCLCA3 levels were elevated in the bronchoalveolar lavage fluid (BALF) (37, 40). Similar to hCLCA1, no mCLCA3 expression could be detected in the lungs of naïve mice prior to induction with IL-4/IL-9/IL-13. This overexpression was also somewhat selective, since mCLCA1 was not induced under the same conditions (33).

Gene transfer of mCLCA3 alone was sufficient to induce MCM in mice (39), further strengthening the role of CLCAs in driving mucus overproduction in the lung. Based on the homology to hCLCA1, mCLCA3−/− was the first knockout mouse generated to study IL-13-stimulated MCM. Surprisingly, these knockout mice were also susceptible to MCM; disease progression and severity in these mice were comparable to wild-type (WT) mice. Further experiments showed an IL-13-dependent increase in mCLCA3, mCLCA5 and mCLCA7 (35), which had not been investigated by Zhou et al. (33) The mCLCA3−/− mice also had increased mCLCA5 levels after IL-13 stimulation (39). These observations suggested a functional redundancy between mouse CLCA proteins, where lack of a family member was compensated for by the activity of other family members, ultimately preventing a distinct lung disease phenotype with a single knockout mouse model.
Similar to mCLCA3, hCLCA1 was abundantly found in the BALF from asthmatic patients, whereas non-asthmatic samples were negative for CLCA1 (37). It is only expressed in the airways under pathophysiological conditions (41) and it is one the most highly induced genes in asthmatics (11), demonstrating the potential utility of CLCA1 as a biomarker of chronic airway diseases (42) and a viable drug target to treat MCM.

1.3 CLCA in Cystic Fibrosis

Mucociliary clearance in the airways is not dictated just by the mucin concentration, but also by hydration (18, 43). Since osmosis follows the salt gradient (mainly NaCl in this case), active ion transport becomes an important mechanism for airway surface liquid (ASL) maintenance. In lung epithelia, cystic fibrosis transmembrane regulator (CFTR) maintains ASL by secreting chloride/bicarbonate ions (44-47) and regulating sodium absorption through inhibition of epithelial sodium channel (ENaC) (48-50).

Cystic fibrosis (CF) is a hypersecretory disease characterized by MCM and decreased mucus clearance in the airways, both of which lead to inflammation and bacterial colonization (51, 52), similar to asthma and COPD. The driving force behind CF is inherited mutations in the CFTR gene, which result in decreased/diminished CFTR activity. As a result of the subsequent reduction in chloride secretion and the rapid increase in sodium absorption, ASL volume is significantly reduced (23, 49, 50).
CFTR currents are regulated by cyclic AMP (cAMP) and cannot be blocked by DIDS (4,4’-diisothiocyanato-2,2’-stilbenedisulfonic acid) (53). In the airways, a second type of chloride current was observed, which was sensitive to (local) calcium concentration, but insensitive to cAMP. Furthermore, these currents could be blocked by DIDS. Since the molecular identity was unknown at the time, the channel was named calcium-activated chloride channel (CaCC) (54). These types of currents were found upregulated in CFTR−/− mice, suggesting CaCC currents could be protective under CF conditions (48, 49, 54).

CLCAs were a good candidate for the unidentified CaCC, as their heterologous expression in HEK293 cells resulted in increased CaCC currents, similar to the ones observed in the airways. These currents had the kinetic properties of native CaCC activity and were blocked by non-specific chloride channel blockers, such as DIDS and niflumic acid (NFA) (29-32, 55). However, as will be discussed later, CLCAs are soluble, secreted proteins that can act as ion channel regulators, as opposed to being bonafide ion channels.

Although they are not channels themselves, CLCAs are still implicated in CF. While both CLCA proteins and CFTR have similar tissue expression patterns under normal conditions (29), CLCA proteins are found downregulated in CF (56). Aside from downregulation, there are also CLCA single nucleotide polymorphisms (SNPs) associated with worse CF disease in clinical investigations (55, 57, 58). Thus, if diminished CLCA activity accompanies worse CF symptoms, CLCAs might carry out a function beneficial under specific pathophysiological conditions. Indeed, heterologous expression of CLCAs has been shown to rescue chloride currents in CF patients (59) and ameliorate CF symptoms (57, 60). These results suggest that
CLCA can modulate chloride/bicarbonate currents in CF airways and help restore the ion gradient. The amelioration of CF symptoms is likely due to compensation of decreased/abolished CFTR activity, which is a result of the interaction between CLCA and its target CaCC, expressed alongside CFTR in the small airways (Figure 2).

Figure 2. Ion channels in the small airways. IL4/13 signaling induces a (hyper)secretory phenotype, which activates CFTR and CaCC activity. CFTR can, in turn, block ENaC activity. Ion concentrations due to channel activity dictate airway surface liquid (ASL) volume. CLCA1 as a secreted protein is found in the extracellular medium under pathophysiological conditions and interacts with CaCC. Figure modified from Sala-Rabanal et al. (19).
1.4 References

37. A. Gibson et al., hCLCA1 and mCLCA3 are secreted non-integral membrane proteins and therefore are not ion channels. *J Biol Chem* **280**, 27205-27212 (2005).


CHAPTER 2: Structural Organization of CLCA1
2. **Structural Organization of CLCA1**

2.1 Misannotation

The first insight into the structural organization/topology was made by hydropathy analysis of the primary sequence of CLCA proteins. The analysis was based on the Kyte-Doolittle scale, which calculates the hydropathy of a group of residues (n=19) to predict whether the segment is membrane-spanning (hydropathy score > 1.6) [1]. Data analysis of CLCA family members predicted multiple transmembrane domains. [2-5].

While most CLCAs are actually soluble proteins with no integral membrane domains [6-9], a few select members of the family are predicted to have different structural features, such as a single transmembrane domain in the C-terminus (like hCLCA2) or a glycoposphatidylinositol (GPI) anchor (hCLCA4) as predicted by PredGPI [10]. Although both hCLCA2 and hCLCA4 undergo self-cleavage [8], the C-terminal cleavage fragment might remain anchored to cells. The fate of the N-terminal cleavage fragment is uncertain. Recent experiments conducted in our lab suggest the N-terminal fragment of hCLCA4 remains associated with the cell membrane; while it was not freely detectable in the supernatant by Western blot, this cleavage fragment could be recovered in the supernatant of HEK293T cells following phospholipase C treatment, which releases GPI-anchored proteins from cell surface [11], such as the C-terminal fragment of hCLCA4.
Figure 3. Domain organization of the human CLCA proteins. CAT, catalytic domain, Cys, Cys-rich domain, vWA, von Willebrand domain type A, FnIII, fibronectin type III, TM, transmembrane domain, GPI, glycophasphatidylinositol anchor. Scissors denote experimentally determined cleavage site. Figure modified from Sala-Rabanal et al. (12).
2.2 Structure predictions

2.2.1 Signal peptide

The full-length human CLCA1 is a 120 kDa protein comprised of 914 residues (Figure 3). The first 21 residues constitute a signal peptide, as identified by Phobius (http://phobius.sbc.su.se/) (13) and SignalP4 (http://www.cbs.dtu.dk/services/SignalP/) (14). Signal peptides allow proteins to enter the endoplasmic reticulum (ER) for post-translational modifications (15, 16), which can regulate protein localization, structure and function.

2.2.2 N-linked glycosylations

N-linked glycosylation is an important component of quaternary protein structure, as these glycans can grossly affect protein function and interaction partners (17). Potential N-linked glycosylation sites in a given protein sequence can be predicted using NetNGlyc 1.0 server (http://www.cbs.dtu.dk/services/NetNGlyc/) (18), whose algorithm analyzes the sequence for N-X-S/T motifs (where X can be any residue except Proline) (19). For CLCA1, the server predicts a total of 9 N-linked glycosylation sites (Figure 4), however the first motif (around residue 270) is likely not glycosylated, due to the Proline residue in the second position, which creates conformational restraints that prevents glycosylation. Thus, there are 2 sites within the N-terminal fragment and 6 within the C-terminal fragment.
Figure 4. N-linked glycosylation predictions for full-length WT CLCA1. The signal sequence was omitted for the analysis by NetNGlyc 1.0 server (http://www.cbs.dtu.dk/services/NetNGlyc/). Green vertical lines indicate predicted N-linked glycosylation sites. Red horizontal line indicates confidence threshold for the prediction. Blue vertical line indicates cleavage site.
2.2.3 Distinct domains

Phyre2 is a protein structure prediction server (20), which can be used to analyze a new protein sequence for potential homology to known domains. For full-length CLCA1, the server identified three domains as homologues of previously determined structures. A 180-residue stretch in the middle was modeled as a von Willebrand factor type A domain (vWA), which is involved in mediating protein-protein interactions, especially in the extracellular matrix (ECM), where they play a role in cell adhesion (21). A 130-residue stretch at the C-terminus was identified as a fibronectin type III domain (FnIII), which mediates fibronectin-integrin interactions in the ECM (22). It is interesting to note that the vWA domain, homologous to the inserted (I) domain of integrins (23), could potentially interact with the FnIII domain. However, the CLCA vWA domain does not have the RGD peptide, which is the main recognition motif for fibronectin binding (22, 24). The third fold detected within CLCA1 was in the N-terminal region, which was predicted to assume a zincin metalloprotease-like fold (Figure 5). Although percent identity within the query (CLCA1) and template sequence in the region was low (<15%), the confidence of prediction was moderate to high (65-90%), as the HEXXH motif was identified in the CLCA1 N-terminal sequence, which is the zinc-binding region of metalloproteases (25). There were no predicted folds for the region spanning between the vWA and the FnIII domains, its secondary structure is predicted to mainly consist of beta sheets. Hence we coined the term beta-sheet-rich (BSR) for this region.
2.2.4 Novel CLCA1 metalloprotease is unlike MMPs

In order to determine how much CLCA1 enzyme resembled MMPs, we carried out sequence analysis, looking for certain features and conducted biochemical experiments, where applicable, to test the existence of said features. Based on the sequence analysis of the CLCA family, there is no conserved Met residue at the C-terminal end of the predicted catalytic site of the metalloprotease, which means this family of proteins lack the Met turn, a characteristic of metzincins (26).

The prodomain is also a characteristic of metalloproteases, as it is involved in the first step of regulating enzymatic activity. This is usually an 80-residue stretch in the N-terminus of the protein, which blocks the active site (27, 28) and has to be cleaved for the enzyme to be active. Phyre2 predicts the catalytic fold of CLCA1 would begin around residue 50. Almost half of the first 50 residues corresponds to the signal peptide, which would make the putative prodomain of CLCA1 only about 25 amino acids long. Although that would be a significantly shorter prodomain, which is unlikely, we wanted to experimentally determine whether it was intact following secretion and cleavage.
2.2.4.1. Experimental design

To analyze the potential prodomain, we inserted a FLAG-tag directly after the signal sequence and expressed the full-length WT protein in HEK293T cells via transient transfection (Figure 6). Proteins were detected in the supernatant by Western blotting with an anti-FLAG antibody, an anti-NCLCA1 antibody (8D3) and an anti-His6 antibody.

2.2.4.2. Results

The antibodies detected a total of three species, the full-length uncleaved protein around 120 kDa and the two cleavage fragments, N-terminal and C-terminal at 70 and 50 kDa, respectively. If the prodomain had been cleaved, the anti-FLAG antibody would not have recognized the N-terminal cleavage fragment and there would have been a protein band at very low molecular weight corresponding to the FLAG-tagged prodomain (8) (Figure 6). These results indicate that there is no prodomain preceding the catalytic fold of CLCA1.
Figure 6. CLCA1 metalloprotease does not have a prodomain. A, Full-length CLCA1 with an N-terminal FLAG tag and a C-terminal 6His tag. Epitope of the 8D3 antibody is between residues 500-695, after the vWA domain, as shown. B, Western blot against FLAG tag and N-terminal CLCA1 (8D3). C, Western blot against the C-terminal 6His-tag. Figure modified from Yurtsever et al. (8).

2.3 Crystallization Trials

2.3.1 Rationale

The lack of known metalloprotease regulatory elements, such as the Met turn and the prodomain, in CLCA structure suggests that the CLCA enzyme likely has a novel fold and that enzymatic activity is regulated by previously undefined structural features. Determination of such potential features requires structural information, especially on an atomic resolution, which can be obtained by protein X-ray crystallography.
2.3.2 Experimental design

2.3.2.1 Cloning

hCLCA1 constructs were cloned into pHLsec backbone (29) using Agel- KpnI restriction site cloning. There is a C-terminal 6-His tag in the backbone, which allows for detection of protein expression via Western blotting and purification from the cell supernatant.

Mutations in hCLCA1 vWA domain (297-478) were introduced using Quikchange Lightning Mutagenesis Kit (#210513 Agilent Genomics) according to manufacturer’s instructions. Primers were designed using neb website (http://nebasechanger.neb.com/) and ordered through IDTDNA (https://www.idtdna.com/site/order/oligoentry) (standard desalting). Amplification products were digested with Dpn I provided with the kit and then transformed onto XL10 Gold ultra-competent cells. After incubation of the cell culture in SOC media shaking for 1 hour at 37°C, the cells were plated onto an agar plate with carbenicillin (100 ug/ml) selection. The plate was incubated overnight at 37°C, then single colonies were cultured in 4 ml LB media supplemented with carbenicillin. The cells were pelleted using centrifugation at 2000 rpm and the DNA template was prepared using QIAprep Spin Miniprep kit (#27104 QIAGEN) according to manufacturer’s instructions. Briefly, cell pellets were lysed, the plasmid DNA was isolated and precipitated with ethanol, then eluted in 30 µl of elution buffer (10 mM tris-Cl pH 8.5) (Qiagen). The presence of target plasmid DNA was verified by analytical PCR reaction (10 µl 2x Phusion High Fidelity PCR Master Mix (m0531s NEB), 7 µl H2O, 1 µl mini-prep + (x2) 1 µl 10 µM primer). DNA bands were visualized on FlashGel DNA system (Lonza). The mutations were verified by sequencing (10 µl H2O, 4 µl 5x sequencing buffer (400 mM tris-HCl pH 9.0, 10
mM MgCl₂), 4 µl Big Dye3.1 (Applied Biosystems), 1 µl mini-prep, 1 µl 10 µM sequencing primer). Sequencing was done by PNACL at WUSM and sequences were analyzed by MacVector (Apex, NC).

The mutations K398A and K398A+K399A were introduced into the full-length WT CLCA1 protein in pHLsec vector. To clone the vWA domain separately, primers were made with 5’ AgeI and 3’ KpnI restriction sites starting at residues 297 and 478, respectively. vWA construct was amplified with these primers from the full-length CLCA1 template. The insert was PCR purified with QIAquick PCR Purification Kit (28104, Qiagen) according to manufacturer’s instructions and eluted in 30 µl EB. The insert and uncut pHLsec vector were double digested using AgeI-HF (R3552S, NEB) and KpnI-HF (R3142S, NEB) in 1x CutSmart Buffer (B7204S, NEB) at 37°C for 1 hour. 1 µl of alkaline phosphatase (M0290S, NEB) was added into the vector digestion reaction and incubated for 1 hour for removal of the 5’ phosphate groups. Digested samples were purified over a 1% agarose gel at 135 V for 40 minutes. DNA bands were cut out of the gel, the plasmids were extracted using QIAquick Gel Extraction Kit (28704, Qiagen) according to manufacturer’s instructions and eluted in 30 µl EB. The concentration of DNA was quantified on a NanoDrop 2000 instrument (Thermo Scientific). For the ligation, 50 ng of vector was used at a 1:3 mole ratio of vector-to-insert. The required volume of insert plasmid in the reaction was calculated based on the size of vector and insert using Ligation Calculator (http://www.insilico.uni-duesseldorff.de/Lig_Input.html). For the ligation reaction, the insert and the vector were mixed with 2x Quick Ligase buffer and 0.75 µl Quick Ligase (M2200S NEB). After incubating at RT for 15 minutes, 5 µl of the ligation reaction was transformed onto 20 µl of competent cells (either GC10, G2794 Sigma, or 10-beta, C3019I NEB) according to
manufacturer’s instructions. Following the 37°C incubation in 500 µl SOC, the reactions were plated onto carbenicillin (100 µg/ml) plates and placed in a 37°C incubator overnight. Colonies were picked and grown in 4 ml Luria broth (LB) + 1x carb overnight at 37°C. The plasmid was purified using QIAprep Spin Miniprep Kit (27104 Qiagen) and eluted in 30 µl EB. After the constructs were verified by sequencing as described above, colonies were streaked into 250 ml LB + 1x carb and the plasmid was purified using HiSpeed Plasmid Maxi Kit (12662 Qiagen) and eluted in 1ml EB.

For refolding, hCLCA1 vWA (297-478) K398A and K398A+K399A constructs were amplified with 5’ NheI and 3’ XhoI primers and ligated into pet23b vector. Inclusion bodies were prepared as described in Section 2.3.1.5. Solubilized inclusion bodies were refolded in arginine refolding buffer with 1 mM MgCl₂ or 1 mM MnCl₂ overnight at 4°C. The refoldings were concentrated to ~4 ml volume and purified over S75 in 1x sizing buffer. The protein refolded in MgCl₂ eluted off the column at a monomeric size.

2.3.2.2 HEK293T Cell culture and transient transfection

Adherent HEK293T cells were cultured in Dulbecco’s Modified Eagle Medium (DMEM), supplemented with 10% fetal bovine serum (FBS), 1% pen/strep (50 U/ml), 1% non-essential amino acids (11140050 Thermo Fisher) and 1% GlutaMAX (35050061 Thermo Fisher) at 37°C and 5% CO₂.
Maxi-prep DNA concentration was measured by absorbance at 260 nm using NanoDrop instrument. On the day of transfection, 1 µg DNA per 10⁶ cells was diluted in low-serum OPTI-MEM (31985070 Thermo Fisher). The transfection reagent (Table 1) was diluted in OPTI-MEM, separately from the DNA solution. After incubating for 5 minutes at room temperature (RT), diluted transfection reagent was added onto the DNA solution and incubated for 20 minutes at RT for complexation. Then, the solution was added onto the cells dropwise and the cells were returned to the incubator for 24-48 hours.

Small-scale expression tests were carried out with adherent HEK293T cells, usually on a 6-well plate (V_{total}= 2 ml) at 1x10⁶ cells/ml density, via transient transfection of the constructs. For large-scale expression, 293F cells grown in FreeStyle 293F Expression Medium (12338018 ThermoFisher) (V_{total}= 300 or 600 ml) were transiently transfected as we reported (30).

Table 1. Transfection reagent ratios for transient transfection

<table>
<thead>
<tr>
<th>Transfection Reagent</th>
<th>V_{reagent} / 1 µg DNA (µl)</th>
</tr>
</thead>
<tbody>
<tr>
<td>293Fectin</td>
<td>2</td>
</tr>
<tr>
<td>Lipofectamine 2000</td>
<td>2</td>
</tr>
<tr>
<td>FuGENE 6</td>
<td>3</td>
</tr>
<tr>
<td>Polyethelyneimine (PEI)</td>
<td>1.5</td>
</tr>
<tr>
<td>PEI-TMC-25 (31)</td>
<td>1.5</td>
</tr>
</tbody>
</table>
2.3.2.3 Recombinant Protein Purification

Proteins with a C-terminal 6His-tag were purified by nickel affinity chromatography as described by us (30). Briefly, the supernatant was collected from transfected 293F cells following centrifugation, pH-adjusted by adding 10% volume of 10x lysis buffer (500 mM K$_2$HPO$_4$ pH 8.0, 3 M NaCl, 100 mM imidazole) and poured over Ni-NTA superflow resin (Qiagen), which had been equilibrated in 1x lysis buffer. Once the flow-through was cleared, the column was washed with 2 column volumes (CV) of wash buffer containing 50 mM imidazole. The protein was eluted in 1 CV of elution buffer with 250 mM imidazole. The eluate fraction was concentrated to a final volume of 2 ml and buffer exchanged using a centrifugation filter. For structural studies, proteins were exchanged into 1x sizing buffer (20 mM HEPES pH 7.5 + 150 mM NaCl ± 10% glycerol) and purified over Superdex 75 (S75) or 200 (S200) columns on the ÄKTA FLPC instrument. If further purification was necessary, the protein peak was pooled and dialyzed into buffer A (20 mM tris pH 8.0, 10 mM NaCl and 10% glycerol) overnight at 4°C. The dialyzed protein sample was then injected onto the MonoQ column (GE Lifesciences) using buffer A and eluted off using a gradient from 0-100% of buffer B (20 mM tris pH 8.0, 1 M NaCl and 10% glycerol) over 40 CV. The protein peak was pooled and concentrated using a centrifugation filter.
2.3.2.4 Western blotting

HEK293T cells were lifted in phosphate buffered saline (PBS). Cells were lysed in 1% v/v triton x-100 in PBS with stirring at 4°C for 1 hour. Lysates were lightly spun and 10 µl of the clear supernatant was mixed with 10 µl 2x sodium dodecyl sulfate (SDS) loading buffer. The sample was boiled for a five minutes, loaded onto a 4-12% bis-tris polyacrylamide gel and run at 200 V for 35 minutes. The proteins were transferred onto a nitrocellulose membrane using gel transfer device (iBlot2). In the meantime, 50 mL of blocking buffer with 0.2 g milk powder dissolved in 0.1% Tween-20 in 10% PBS (PBST) was prepared. After blocking the membrane with 10 ml blocking buffer for 5 minutes, 10 ml of primary antibody in blocking buffer (Table 2) was incubated with the membrane for 15 minutes. Following three PBST washes, 10 ml of the appropriate HRP-conjugated secondary antibody in blocking buffer was placed on the membrane and incubated for 15 minutes. The membrane was then washed three times with PBST and proteins were detected using a western blotting substrate (Pierce ECL, Thermo Fisher 32106) at 4-ml/membrane volume.

Table 2. Antibodies used for Western blotting.

<table>
<thead>
<tr>
<th>Antibody</th>
<th>Species</th>
<th>Epitope/Target</th>
<th>Dilution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anti 6His-HRP</td>
<td>Rabbit</td>
<td>6His-tag</td>
<td>1:5000</td>
</tr>
<tr>
<td>8D3</td>
<td>Mouse</td>
<td>NTF of CLCA1</td>
<td>1:500</td>
</tr>
<tr>
<td>Anti-mouse IgG-HRP</td>
<td>Goat</td>
<td>Mouse IgG</td>
<td>1:5000</td>
</tr>
<tr>
<td>Anti-actin</td>
<td>mouse</td>
<td>Actin protein</td>
<td>1:2000</td>
</tr>
</tbody>
</table>
2.3.2.5 Inclusion body refolding and purification

hCLCA1 constructs were cloned into pet23b backbone using NheI-KpnI restriction site cloning. Constructs were transformed into Rosetta 2 (DE3) competent cells (Millipore) and expression was carried out under chloramphenicol and carbenicillin selection. A 100 ml starter culture was used to inoculate 6-8 L of Luria Broth (LB). After culture optical density reached 1.0, protein expression was induced using 2.0 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) at 25°C overnight (or 30°C for 4-5 hours). Cells were lysed in buffer (50 mM tris pH 8, 1% v/v triton x-100, 100 mM NaCl, 0.01% w/w NaN₃ and 10 mM dithiothreitol) and supplemented with lysozyme and DNAse I. Following rigorous washing with wash buffer containing 0.5% v/v triton x-100, a final wash without triton x-100 was performed and inclusion bodies were solubilized in 20 ml buffer (6 M guanidine hydrochloride, 10 mM tris pH 8 and 20 mM 2-mercaptoethanol). Inclusion bodies were diluted 1:4 in solubilization buffer and were added into oxidative arginine refolding buffer (100 mM tris, 400 mM arginine, 0.5 mM oxidized glutathione, 5 mM reduced glutathione, 0.1 mM ZnCl₂ and 10 mM CaCl₂) in small portions. Proteins were allowed to refold overnight at 4°C in 250-500 ml volume, then concentrated under pressure to a final volume of 5 ml and purified by gel filtration chromatography and ion exchange chromatography as described above.
2.3.3 Results

2.3.3.1 Metalloprotease domain

The first metalloprotease constructs I used for crystallization were 22-189, 22-199 and 22-283. These boundaries were determined using the structure predictions; the first two constructs contain only the metalloprotease domain, while the last construct also includes the Cys-rich domain (Figure 6). These constructs could be expressed in mammalian cells, but were not secreted (Table 3), which indicated problems with proper folding.

I then tried different truncations within the metalloprotease domain and among those, 22-293 was the construct that expressed and secreted the most robustly. Making the E157Q mutation, which has aided in the crystallization of other metalloproteases (32-35), did not affect expression or secretion of the protein. However, neither of these constructs could be purified to homogeneity via FPLC (Figure 7). With samples purified only by nickel affinity chromatography, crystal screens were set at 1-5 mg/ml using hanging-drop method with Mosquito. The commercial screens JCSG Core Suite I-IV, PEGS Suite I-II (Qiagen) were screened, but no crystal hits were obtained. With these samples, I also tried in situ proteolysis (36) with chymotrypsin to cut off flexible terminal residues, but the treatment did not improve the results.
Table 3. Metalloprotease constructs tested for structural studies.

<table>
<thead>
<tr>
<th>Construct</th>
<th>expressed</th>
<th>secreted</th>
<th>refolded</th>
<th>Purified</th>
<th>notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>22-189</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22-189 E157Q</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22-199</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22-199 E157Q</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22-283</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22-283 E157Q</td>
<td>x</td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22-293</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>heterogeneous</td>
</tr>
<tr>
<td>22-293 E157Q</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>heterogeneous</td>
</tr>
</tbody>
</table>

Figure 7. Representative size exclusion and ion change (monoQ) chromatograms. A. Refolded 22-283 E157Q purified over S75. Red star indicates the fraction collected for subsequent monoQ purification. B. Refolded 22-283 E157Q purified over monoQ after S75. C. Refolded 22-293 E157Q purified over S75. D. Refolded 22-293 purified over monoQ after S75.
2.3.3.2 vWA domain

CLCA1 vWA (297-478) construct was expressed and secreted robustly. During size exclusion chromatography, it eluted at a volume corresponding to its monomeric weight and from the pooled fractions the protein could be purified to 99%, as assessed by SDS-PAGE (Figure 8). Crystal screens were set at 10 mg/ml using hanging-drop method with Mosquito. The commercial screens JCSG Core Suite I-IV, PEGS Suite I-II (Qiagen) were screened, but no crystal hits were obtained.

Figure 8. Purification of recombinant vWA (297-478) expressed in 293F cells. Left, Size exclusion chromatogram (analytical S200) of the affinity purified protein. Protein elutes at a volume corresponding to monomeric protein (~20 kDa). Red star indicates the protein peak. Right, Imperial stain of the concentrated peak off of S200 column for crystallization trials. Protein is 99% pure Arrow indicates the major protein band at 20 kDa, corresponding to monomeric vWA. Note: There is only one Trp and two Tyr residues within the construct, so absorption at 280 nm is weak, despite the high concentration of protein.
2.3.3.3. Surface entropy reduction for crystallization

vWA domains are homologous to inserted (I) domains found in integrins (23). Both these domains adopt a Rossmann fold, also known as classic alpha/beta open, dinucleotide-binding or doubly wound fold (37). There is a closed conformation, which reverts to open conformation upon ligand binding (23, 38) and most structures determined so far have the vWA (or the I) domain in complex with a ligand (peptide) (38–40). Since the specific ligand of CLCA1 vWA is unknown, I tried other techniques to create stable and homogenous proteins amenable for crystallization.

To improve the probability of crystallization, the free energy of the reaction needs to be minimized and is determined by the following equation: \( \Delta G_{\text{crystallization}} = \Delta H_{\text{crystallization}} - T.\Delta S_{\text{crystallization}} = \Delta H_{\text{crystallization}} - T.(\Delta S_{\text{protein}} + \Delta S_{\text{solvent}}) \) (41). The enthalpy term of crystallization is minor, making the free energy directly correlated with the entropy term (42). Thus, the surface entropy of the protein becomes the major factor in determining the fate of crystallization. A high entropy term, in this case, translates as surface heterogeneity, which even at high sample purity, can be detrimental to protein crystallization. Residues with high side chain conformational entropy (Lys, Glu, Gln) introduce conformational flexibility on the surface and “interfere with the ability of the protein to form stable crystal contacts” (43).
There are two methods utilized that reduce surface entropy and create a conformationally homogeneous patch to help crystallization: lysine methylation (44) and surface mutagenesis (42). The latter method, in practice, requires a model of the protein to pick residues that would be surface exposed. Using the primary sequence of CLCA1 vWA, I built a high-confidence homology model using Phyre2 (Figure 9), which I used as my template. Surface entropy reduction prediction (SERp) server (http://services.mbi.ucla.edu/SER, UCLA, CA) (45) takes protein primary sequence as input to predict which charged residues would be clustered on the protein surface. For CLCA1 vWA domain, the server came up with three clusters, which are indeed surface exposed on the homology model built by Phyre2 (Figure 9). The first of the three clusters has the highest score, indicative of predicted success the mutagenesis would have in enabling crystallization.

Figure 9. Homology model of vWA depicting the predicted residue clusters for surface entropy reduction. Residues targeted for mutation are colored in the model and typed in bold. Numbers in the boxes denote SERp score, which is an indicator of entropy reduction upon mutation of these floppy charged residues to Ala. Model is built by Phyre2 and the figure is made with Pymol.
2.4 Summary

CLCAs are secreted soluble proteins, which are comprised of multiple distinct domains. Aside from the well-characterized vWA and FnIII domains, CLCAs also contain an N-terminal domain that is predicted to have an MMP-like fold. This metalloprotease domain is distinct from MMP and ADAM family members, because it lacks certain regulatory features such as the prodomain and the Met turn. For structural studies, I was able to determine domain boundaries, which would get expressed and secreted from cells, indicating proper folding. While I was able to get sufficient expression from mammalian cells, the metalloprotease constructs did not purify to homogeneity. The CLCA1 vWA domain, however, expresses robustly and can be purified to homogeneity. In order to improve probability of crystallization, I recently cloned a double surface Alanine mutant (K398A + K399A). These constructs express robustly and can be refolded for crystallization trials.
2.5 References

6. A. Gibson et al., hCLCA1 and mCLCA3 are secreted non-integral membrane proteins and therefore are not ion channels. J Biol Chem 280, 27205-27212 (2005).


CHAPTER 3: Proteolytic Processing
3 Proteolytic Processing

3.1 Conserved Cleavage Sequence

In order to determine how CLCA proteins were getting cleaved, it was important to identify the location of the cleavage site within the primary sequence. We also wanted to investigate if the cleavage site was conserved within the entire CLCA family or if it was occurring at a different site for each protein.

3.1.1 Determination of cleavage site by Edman degradation

Full-length CLCA proteins (hCLCA1, hCLCA2, hCLCA4, mCLCA3, mCLCA4), without their endogenous signal peptide, were cloned into pHLsec vector, which contains an optimized signal sequence and a C-terminal 6His-tag (I). hCLCA2, hCLCA4 and mCLCA4 were made into soluble constructs by the truncation of the C-terminal transmembrane domains or GPI-anchors. All proteins were expressed in 293F cells by transient transfection and the C-terminal fragments were captured by nickel affinity chromatography. SDS-PAGE was used to separate the C-terminal cleavage fragment for Edman degradation analysis.

3.1.2 Results

By Edman degradation, the first 5 N-terminal residues of the fragments were identified, which determined the proteolytic cleavage site of the protein. All proteins tested were cleaved in a heavily conserved region (Figure 10), which starts with an Ala in the P1’ site for most family members and includes an invariant Tyr in the P3’ position, followed by an invariant Gly in the P6’ position. There is a Pro in the P5’ for most family members and the P4’ site is occupied by small hydrophobic residues, either an Ile or a Val. These results indicate that CLCAs have a
conserved cleavage sequence within the family that transcends species, which is likely recognized by a specific protease (family).

Figure 10. CLCA family members have a conserved cleavage sequence. Red arrow indicates site of cleavage. Underlined green residues have been experimentally determined by Edman degradation/N-terminal sequencing. Figure modified from Yurtsever et al. (2).
3.2 N-terminal zincin metalloprotease

3.2.1 Point mutations in the active site

Based on the metalloprotease-like fold prediction, the active site of CLCA1 would consist of H156, E157, H160, D167 and E168. Residue E157 is the catalytically required Glu that initiates the nucleophilic attack on the substrate, i.e. the cleavage sequence of CLCA1. The other residues are required for coordinating the zinc ion (table 4). In order to test the hypothesis that there really was an N-terminal metalloprotease, I mutated each residue to Ala and tested the effect of the mutations on CLCA cleavage.

3.2.1.1 Experimental design

These residues were separately mutated to Ala using Phusion Site-directed mutagenesis kit (F541 Thermo Fisher), according to manufacturer’s instructions. Full-length CLCA1 constructs were expressed in HEK293T cells via transient transfection. Supernatants were collected after 72 hours and analyzed by Western blot for cleavage product formation.

Table 4. List of point mutations in the predicted active site.

<table>
<thead>
<tr>
<th>Constructs</th>
<th>Deficiency</th>
<th>Expressed</th>
<th>Cleaved</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild-type</td>
<td>none</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Q150A</td>
<td>none</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>H156A</td>
<td>Zinc binding</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>E157Q</td>
<td>Nucleophilic center</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>H160A</td>
<td>Zinc binding</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>D167A</td>
<td>Zinc binding</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>E168A</td>
<td>Zinc binding</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>
3.2.1.2 Results

The point mutations did not affect protein expression or secretion. However, each point mutation alone was sufficient to block cleavage product formation, as assayed by Western blotting for both the N-terminal cleavage fragment (8D3 mouse monoclonal antibody) and the C-terminal fragment (C-terminal 6His tag was recognized by anti-6His antibody). Residue Q150 was used as a control for this experiment, where the mutation to alanine had minimal effect on the amount of cleavage product formed. (Figure 11). These results suggest that CLCAs contain an N-terminal zincin metalloprotease and that they are capable of self-cleavage.

Figure 11. Point mutations in the predicted active site of the metalloprotease block self-cleavage in cells. Western blot monitoring cleavage product formation. Top panel, anti-6-His antibody that recognizes the full-length and the C-terminal cleavage fragment. Bottom panel, custom 8D3 antibody that recognizes the full-length and the N-terminal cleavage fragment. Figure modified from Yurtsever et al. (2).
3.2.2 *In vitro* digestion assay

3.2.2.1 Rationale

While the effect of mutations in the predicted active site of CLCA1, i.e. complete abrogation of cleavage product formation, indicated that CLCA1 metalloprotease was responsible for the cleavage reaction, the experiment was carried out in a cellular environment, where the mutations could possible affect other regulators of enzymatic activity. In order to definitely show that CLCA1 metalloprotease is the only enzyme required for CLCA1 cleavage, I wanted to recapitulate the cleavage reaction *in vitro* using recombinant proteins. Chelators were used to test the divalent cation dependency of the enzymatic reaction. MMP inhibitors were used to determine whether the CLCA1 active site could accommodate known MMP inhibitors and whether the interaction would be sufficient to block all enzymatic activity.

3.2.2.2 Experimental design

CLCA1 full-length protein was separated into two parts for the reconstitution of the cleavage reaction *in vitro*: The enzyme construct (22-478) contained the metalloprotease domain and the vWA domain, the substrate construct (297-914) was only missing the catalytic domain (Figure 12a). The vWA domain was included in the enzyme construct to confer stability to the protein.

The 50 µl reaction was set up in the digestion buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 10 µM ZnCl₂, 10 mM CaCl₂), with 0.5 µM substrate. A 10 µl aliquot was taken immediately upon enzyme addition (final concentration 2 µM) and quenched in 10 µl 2x SDS. The rest of the reaction was incubated in a water bath at 37°C. After 1 hour, a second 10 µl aliquot was taken and quenched. Both aliquots were stored at -20°C, while the remainder of the
reaction was incubated in the water bath overnight. The final aliquot was collected after ~16-18 hours and quenched similarly. The samples were boiled in a 90°C heating block and proteins were detected with Western blotting.

To test the hypothesis that CLCA1 enzymatic activity depended on divalent cations, ethylenediaminetetraacetic acid (EDTA) and 1,10-phenanthroline, as chelators, were added into the reaction at a final concentration of 15 mM. Halt protease inhibitor cocktail (78430 Thermo Fisher) was used at 1x concentration. Commercially available MMP inhibitors GM6001 (2983 Tocris), Marimastat (M2699 Sigma) and Batimastat (SML0041 Sigma) (Figure 13b) were added into the reaction mixture at a final concentration of 40 µM.

The MMP inhibitors mentioned above are part of the first-generation MMP inhibitors (3) and they all have a peptide-mimetic scaffold with a hydroxamate moiety to chelate the catalytically required zinc (4-7). For our custom inhibitor Zeynepstat, we used the same scaffold and functionalized the P1’ and P2’ residues to Ala and Leu, respectively, to mimic the CLCA1 cleavage sequence (Figure 13b). The synthesis was carried out by Dr. Artie Romero at Washington University in St. Louis, as described in Yurtsever et al. (2).

### 3.2.2.3 Results

Incubation of the enzyme and substrate together resulted in cleavage product formation over time, as assayed by Western blot (8D3), while the substrate alone was quite stable over the time period observed (~18 hours) (Figure 12b). Addition of EDTA or 1,10-phenanthroline blocked cleavage product formation, verifying that the protease activity was dependent on divalent cations (most likely zinc) (Figure 12b).
HALT protease inhibitor cocktail contains protease inhibitors targeted against serine, cysteine proteases, threonine proteases and aspartyl proteases, but not metalloproteases. Addition of this cocktail had no effect on CLCA1 activity (Figure 12b), further supporting the hypothesis that the CLCA1 enzyme responsible for the self-cleavage reaction is a metalloprotease.

The MMP inhibitors GM6001, Marimastat and Batimastat had varying efficacies of blocking CLCA1 activity, suggesting that the enzyme is indeed a zincin metalloprotease, but likely with a novel fold, which cannot be completely blocked by available MMP inhibitors (Figure 13a). The custom inhibitor, Zeynepstat, blocked approximately 50% of CLCA1 enzymatic activity (Figure 13a).

Figure 12. Cleavage can be reconstituted in vitro. A. CLCA1 enzyme and substrate constructs that were recombinantly expressed and purified. B. Western blot monitoring cleavage product formation over time under different conditions. Figure modified from Yurtsever et al. (2).
Figure 13. Clca1 has a novel metalloprotease fold. A. Western blot monitoring cleavage product formation over time in the presence of commercially available MMP inhibitors and our custom inhibitor. B. Chemical structures of the commercial MMP inhibitors and the custom inhibitor Zeynepstat. Figure modified from Yurtsever et al. (2).

3.2.3 Fluorogenic peptide assay

3.2.3.1 Experimental design

We designed a fluorescent substrate, by coupling DABCYL and EDANS FRET pair to the termini of the 10-residue CLCA1 cleavage sequence QQSG-ALYIPG (AnaSpec). When the substrate is intact, EDANS fluorescence is quenched due to the proximity of DABCYL. Thus, fluorescence only increases when the CLCA1 enzyme cleaves the substrate (Figure 14a).

For the reaction, 10 µM refolded CLCA1 enzyme (22-478) was mixed with 3.25 µM fluorogenic peptide in the digestion buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 10 µM ZnCl₂,10 mM CaCl₂) in a 96-well format. The experiments with inhibitors were carried out at 20 or 40 µM final inhibitor concentration. Using a BioTek plate reader, end point fluorescence at 37°C was read continuously for 1 hour at 1-minute intervals (exc=340 nm, emm= 490 nm).

3.2.3.2 Results

This assay was used as a more quantitative measurement of CLCA1 enzymatic activity, with results similar to those obtained with the in vitro digestion experiment. When mixed with the enzyme, increase of fluorescence from the reaction mixture could be observed over time as
the functional enzyme cleaved the substrate (Figure 14b). In the presence of a catalytically inactive enzyme construct (E157Q), there was no increase in fluorescence above background (data not shown), indicating that the increase is solely due to substrate cleavage by functional CLCA1 enzyme. Addition of EDTA into the reaction abolished the increase in fluorescence (Figure 14b), similar to the metal chelators (Figure 12b). The zinc dependence of the reaction was further demonstrated in this assay, where addition of excess Zn over EDTA rescued some of the enzymatic activity, leading to increase in fluorescence (Figure 14b).

Figure 14. Fluorogenic peptide assay shows CLCA1 enzyme is zinc dependent. A. Design of the fluorogenic peptide. Fluorescence increases if the substrate is cleaved. B. Graph depicting increase of fluorescence over time. Green line, CLCA1 enzyme and peptide. Red line, CLCA1 enzyme, peptide and EDTA. Gray line, CLCA1 enzyme, peptide, EDTA and excess Zn. C. Quantification of CLCA1 inhibition by MMP inhibitors. Green bar indicates 100% CLCA1 activity. Blue bars and red bars indicate 20 µM and 40 µM of inhibitors, respectively. Figure modified from Yurtsever et al. (2)
We were able to quantify CLCA1 activity inhibition with the fluorogenic peptide assay using the inhibitors from the digestion assay (Figure 13b). Using a refolded enzyme (instead of the purified recombinant protein) and a different substrate (fluorogenic peptide vs. purified recombinant CLCA1 substrate protein), we were able to recapitulate the findings qualitatively displayed in the Western blot (Figure 13a). While GM-6001, Marimastat and Zeynepstat reduced CLCA1 activity by 50%, Batimastat inhibited 80% of CLCA1 activity at the low concentration (20 µM) tested (Figure 14c). These quantitative results mirrored the data obtained using recombinant purified CLCA1 proteins in the presence of the same inhibitors (Figure 13a) and further showed that CLCA1 activity was not fully blocked by these molecules, at concentrations sufficient to completely block other MMPs.

3.3 Features of cleavage

3.3.1 Cross-cleavage

3.3.1.1 Experimental design

In order to determine if CLCA family members could recognize and cleave each other, we cloned the full-length CLCA1 constructs with either CLCA2 or CLCA4 cleavage sequence (Figure 15a), as well as CLCA2 and CLCA4 protease constructs (Figure 15c), similar to the CLCA1 construct in Figure 12a.

CLCA1 constructs with altered cleavage sequences were transiently transfected into HEK293T cells. Cells and supernatants were harvested after 72 hours and cleavage was detected via Western blotting against both the N-terminal and the C-terminal cleavage fragments.
For the *in vitro* digestion assay, CLCA2 and CLCA4 protease constructs were refolded and incubated with CLCA1 substrate construct. Aliquots were taken at 0, 1 and 18 hour-time points and assayed for cleavage product formation via Western blotting.

### 3.3.1.2 Results

Transient transfection of full-length CLCA1 with altered cleavage sequence successfully resulted in cleavage product formation in HEK293T cells (Figure 15b), indicating the CLCA1 enzyme could recognize and cleave slightly different sequences (Figure 15a). Similarly, cleavage product was formed over time *in vitro*, when the CLCA1 substrate (Figure 12a) was incubated with CLCA2 or CLCA4 proteases, although the amount of product was significantly less than the result of CLCA1 activity (Figure 15d). Thus, the CLCA2 and CLCA4 enzymes could also recognize the CLCA1 sequence. These results indicate that cross-cleavage can occur within family members.
Figure 15. Cross-cleavage between CLCA proteins can occur. A, Full-length CLCA1 protein with altered cleavage sequences. B, Western blot monitoring cleavage product formation in HEK293T cells. Right, anti-6His blot. Left, anti-N-terminal fragment blot. C, Construct design for CLCA2 and CLCA4 protease. D, Western blot monitoring cleavage product formation over time in vitro. Figure modified from Yurtsever et al. (2).

3.3.2 Intermolecular Cleavage

3.3.2.1 Experimental design

In order to determine if CLCA1 proteolysis could occur as an intermolecular reaction, I used two full-length CLCA1 mutants: E157Q, which is catalytically inactive, and contra, which does not have the CLCA1 cleavage sequence. I transiently transfected these constructs into HEK293T cells, both separately and together. 24-48 hours after transfection, I harvested cells and supernatant and detected proteins via Western blotting.
2.3.2.2 Results

As expected, no cleavage product was formed when either of these constructs were transiently transfected (Figure 16, lanes 1 and 2), since the entire population has the same cleavage deficiency. However, co-transfection of both constructs resulted in cleavage product formation (Figure 16) within the same time frame. This observation suggests that two CLCA1 molecules can form a functionally active complex and allow the cleavage reaction to take place intermolecularly.

Figure 16. Intermolecular CLCA1 cleavage can occur. Western blot displays cleavage product formation in the supernatant of HEK293T cells. Lane 1, cells transfected with full-length E157Q. Lane 2, cells transfected with full-length contra. Lane 3, cells co-transfected with both mutants. Black arrow denotes cleavage product formation.
3.3.3 Regulation by GAGs

3.3.3.1 Rationale

Glycosaminoglycan (GAG) binding motifs within proteins are described as a cluster of basic and hydrophilic residues, such as XBBXBX, XBBBXXBX, XBBXXBBBXXBBX, where B is a basic residue (mostly Lys, followed by Arg) and X is a hydrophilic residue (Ala, Val, Phe etc.) (8, 9). There is a predicted GAG binding motif BBXXXB within CLCA1 that is conserved within family members (Figure 17). Since there is precedence for binding and regulation of metalloprotease activity via GAG interaction (10-16), I wanted to verify if CLCA1 cleavage activity could be regulated by the same mechanism.

Figure 17. Sequence alignment of CLCA1-CLCA2-CLCA4 in the N-terminal catalytic region. Highlighted in green are all the identical residues. Red arrow at residue 157 denotes the catalytically required glutamate residue. Box at residues 70-76 highlights the conserved GAG binding motif BBXXXB, where basic residues are colored blue and the hydrophilic residues are colored red. Primary sequences were aligned with Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/) and annotated with ESPript 3.0 (http://escript.ibcp.fr/ESPript/ESPript/index.php).
3.3.3.2 Experimental design

As a cellular assay for GAG dependence of CLCA1 cleavage, I used Chinese hamster ovary (CHO) cells, because both WT/K1 and GAG-deficient cell lines were readily available. The GAG-deficient CHO cell line, CHO-745, is xylosyltransferase-deficient (17, 18). Adherent CHO K1 and 745 cells were cultured in RPMI containing 10% serum, 1% pen/strep and 1% GlutaMAX at 37°C and 5% CO₂.

To monitor CLCA1 cleavage product formation in these cells, I transfected them with the full-length WT CLCA1 plasmid in pHLsec backbone using either 293Fectin or Lipofectamine 2000. 24 and 48 hours after transfection, I collected the supernatants and the lysates for protein detection via Western blotting.

3.3.3.3 Results

Using the same amount of plasmid, transfection efficiency of 293Fectin was higher than Lipofectamine 2000 in both CHO cell lines (data not shown). The 8D3 antibody (against N-terminal CLCA1) detected both the full-length uncleaved protein and the N-terminal cleavage fragment in both cell lines and in both the supernatant and the cell lysate fractions. Band intensities were not significantly different between the two cell lines, indicating that GAGs are not required for CLCA1 cleavage activity (Figure 18a).
Figure 18. GAGs are not required for CLCA1 expression, secretion, self-cleavage or cell surface binding. A. Western blot of supernatants showing cleavage product formation in native vs. GAG-deficient CHO cells. Green box highlights the N-terminal cleavage product. B. Confocal microscopy image showing CLCA1 surface staining in native vs. GAG-deficient CHO cells.

3.3.3.4 Confocal

3.3.3.4.1 Rationale

There have been a number of reports observing cell surface staining of CLCA proteins (19, 20). In order to determine if surface exposed GAGs were mediating the cell surface binding of CLCA1, I used CHO WT/K1 and GAG-deficient 745 cells (17).
3.3.3.4.2. **Experimental design**

Both cell lines were transiently transfected with full-length WT CLCA1 with 293Fectin. Samples were fixed and stained using an anti-CLCA1 antibody 48 hours post-transfection.

At day 0, cells were plated on a 6-well-plate in 2 ml volume. On day 1, cells were transiently transfected with 1 µg CLCA1 pHlsec plasmid + 2 µl 293Fectin according to manufacturer’s instructions. On day 2, cells were trypsinized and replated on chamber slides (Lab-Tek ii chamber slide system / 4 well glass slide) in 400 µl volume and incubated overnight for adherence. On day 3, supernatant was aspirated and cells were immediately fixed with 200 µl 4% formaldehyde for 5 minutes at RT. Then, they were blocked in 200 µl 1% blocking buffer for 1 hour at RT and stained with primary antibodies (Table 5) in blocking buffer overnight at 4°C. Cells were washed twice with PBS and incubated with 200 µl 1:200 WGA-633 in PBS for 30 minutes at RT. Following two more PBS washes, 200 µl of the secondary antibodies in PBS were added for 2 hours at RT. Finally, cells were washed twice with PBS and the media was aspirated. Wells were removed from the glass slide. Samples were mounted with DAPI-containing medium (1 drop/well) and imaged. Confocal microscopy was carried out using a Zeiss LSM 880 Confocal Laser Scanning Microscope with Airyscan (Carl Zeiss Microscopy, Thornwood, NY). Images were acquired and analyzed using ZEN Blue software (Carl Zeiss Microscopy, Thornwood, NY).
Table 5. Antibodies used in confocal microscopy.

<table>
<thead>
<tr>
<th>Antibody type</th>
<th>Name</th>
<th>Target</th>
<th>Species</th>
<th>Dilution</th>
<th>Exc/Emm (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary</td>
<td>S20</td>
<td>TMEM16A</td>
<td>Goat</td>
<td>1:50</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1228</td>
<td>N-CLCA1</td>
<td>Rabbit</td>
<td>1:100</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2F4</td>
<td>vWA</td>
<td>Mouse</td>
<td>1:100</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6A4</td>
<td>vWA</td>
<td>Mouse</td>
<td>1:100</td>
<td></td>
</tr>
<tr>
<td>Secondary</td>
<td>WGA</td>
<td>Sialic acid (cell membrane)</td>
<td>1:200</td>
<td>633/647</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Anti-rabbit IgG</td>
<td>donkey</td>
<td>1:200</td>
<td>594/617</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Anti-goat IgG</td>
<td>donkey</td>
<td>1:250</td>
<td>488/525</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Anti-mouse IgG</td>
<td>Donkey</td>
<td>1:200</td>
<td>555/580</td>
</tr>
</tbody>
</table>

3.3.3.4.3 Results

Comparison of cell surface staining of CLCA1 in WT vs. GAG-deficient CHO cells revealed no observable difference, suggesting that cell surface binding of CLCA1 does not depend on GAG expression (Figure 18b).
3.4 Summary

CLCA proteins have a conserved cleavage sequence and proteolytic processing of family members occurs at this very specific site. The cleavage reaction is carried out by an internal N-terminal zincin metalloprotease. This unique self-cleavage activity likely takes place in the extracellular medium, as the full-length, uncleaved CLCA1 can be detected in the supernatant, alongside the two cleavage fragments.

Although the CLCA1 enzyme predicts to have an MMP-like fold, commercially available MMP inhibitors can only block up to 50% of CLCA1 activity, suggesting a novel conformation for the active site. Unlike MMPs, CLCAs also do not have a Met turn and do not require interaction with GAGs to be functionally active. The data obtained thus far indicate that the CLCA zincin metalloprotease has a novel fold, which requires further structural information to guide drug design.
3.5 References

3. L. Devel et al., Third generation of matrix metalloprotease inhibitors: Gain in selectivity by targeting the depth of the S1' cavity. Biochimie 92, 1501-1508 (2010).
CHAPTER 4: CLCA1 Function as CaCC/TMEM16A Regulator
4 CLCA function as CaCC Regulator

The bioinformatics and biochemical data shown so far demonstrate that CLCA1 is a fully soluble secreted protein, which undergoes a self-cleavage reaction. In elucidating the mechanism, by which CLCA1 regulates calcium-activated chloride currents, we first wanted to determine whether CLCA1 cleavage was important to its role as a CaCC regulator.

4.1 Cleavage is necessary for function

4.1.1 Whole-cell patch clamp assay

In collaboration with Drs. Colin Nichols and Monica Sala-Rabanal, we set up a whole-cell patch clamp assay as a functional readout of CLCA1 activity. For the experiments, adherent HEK293T cells were cultured as described before. On day zero, cells were plated on a 6-well plate ($V_{\text{total}} = 2 \text{ ml}$). On day one, cells were transfected with 1 µg DNA + 0.2 µg eGFP + 2 µl 293fectin. On day two, cells were replated onto UV-sterilized glass cover slips in fresh medium. On day three, cover slips were removed from the original well and placed into the patch clamp rig. Cells that fluoresced under UV light were specifically targeted for whole-cell measurements, as the likelihood of co-expressing CLCA1 would be higher in that population. Calcium-sensitive chloride currents were measured, which are conducted via a channel endogenously expressed in HEK293T cells.

The whole-cell patch clamp experiments were conducted as described in Yurtsever et al. (1). Briefly, 2 megaohm pipettes were filled with pipette solution (126 mM choline chloride, 10 mM HEPES, 10 mM EGTA) with 0 or 10 µM free Ca$^{2+}$. The bath solution contained 125 mM
NaCl, 10 mM HEPES, 30 mM sucrose, 2 mM CaCl₂, 2 mM MgCl₂. Following seal and break-in, cells were voltage clamped at -80 mV. A pulse protocol was applied in which membrane potential was held at -80 mV for 500 ms and stepped to a test value for 1000 ms before returning to the holding potential for an additional 500 ms. The test potential varied from -100 to +80 mV in 20 mV increments. Membrane capacitance was calculated from the integral of the current transient. Data were analyzed with ClampFit 10.1 (Molecular Devices).

4.1.2 Full-length mutants do not activate currents

As a negative control, cells were transfected with the empty vector backbone, which resulted in background currents (~20 pA/pF). Transfection with the full-length, enzymatically active protein resulted in a robust increase in slightly outward rectifying currents (~80 pA/pF) (Figure 19a). These currents could be blocked by gluconate and were sensitive to the calcium concentration in the pipette solution.

Enzymatically inactive full-length constructs, E157Q, H156A, D158A, did not increase currents over background levels. The full-length construct with a scrambled cleavage sequence (contra) behaved similarly as the enzymatically inactive constructs, and did not increase currents (Figure 19a). Since impairment of either of the catalytic sites was detrimental, CLCA1 enzymatic activity is required for the CaCC activation.

4.1.3 N-terminal fragment is necessary and sufficient for CaCC activation

In order to determine if the CaCC activation could be attributed to either of the cleavage fragments, cells were transfected with the N-terminal and the C-terminal cleavage fragment
separately. In the patch clamp assay, cells transfected with the N-terminal cleavage fragment displayed high calcium-sensitive chloride currents, at comparable levels to cells transfected with the full-length protein. The C-terminal cleavage fragment, on the other hand, did not cause an increase in the currents (Figure 19b).

Figure 19. Self-cleavage of CLCA1 is required for CaCC activation. A. Chloride current density measured from cells transfected with full-length CLCA1 constructs. B. Chloride current density measured from cells transfected with CLCA1 cleavage fragments. C. Model of proposed interaction for CLCA1 and CaCC; self-cleavage releases the C-terminal fragment from the N-terminal fragment, which is the functional portion of the protein. Figure modified from Yurtsever et al. (7).
4.1.4 Enzymatic activity is not required for CaCC activation

Some channel proteins need to be proteolytically cleaved to become active. For example, aerolysin, an *Aeromonas* toxin that forms pores on the cell membrane, is synthesized as a zymogen and is activated via proteolytic removal of the c-terminal peptide. Proteolysis, which makes the toxin insertion competent (2), can be carried out by trypsin (3) or furin (4).

There is another example for proteolytic channel activation, which occurs in the small airways where CLCA1 is expressed. The epithelial sodium channel ENaC is expressed in distal and proximal airways (5), where it controls ASL volume and mucociliary clearance (6). In both *in vitro* and *in vivo* lung models, ENaC activity has been shown to increase following cleavage by channel activating protease 1 (CAP1/prostasin) (7-9).

To determine if the metalloprotease within the N-terminal cleavage fragment was important for CaCC activation, the catalytically dead N-terminal fragment (E157Q) was used in the whole cell patch clamp assay, side-by-side with the WT N-terminal cleavage fragment. Both constructs increased currents to the same extent, thus indicating that CLCA1 enzymatic activity does not play a role in CaCC activation (Figure 19b).

Based on the functional data obtained from the whole-cell patch clamp experiment, self-cleavage of CLCA1 is necessary for CaCC activation to occur. As the N-terminal cleavage fragment is as active as the full-length protein and enzymatic activity is not required to increase current, self-cleavage is likely required to release the C-terminal cleavage fragment from the full-length compact structure of CLCA1. Once the C-terminal tail is released, the N-terminal fragment can obtain a conformation that is favorable for interacting with the CaCC (Figure 19c).
4.2 Paracrine effect

4.2.1 GFP co-culture experiment

4.2.1.1 Experimental design

As CLCA1 is a secreted protein, the question of where the interaction between CLCA1 and TMEM16A occurred was a relevant one. In order to distinguish between an autocrine, i.e. on the same cell, and a paracrine, i.e. on neighboring cells, effect, we designed a series of modifications on the whole-cell patch clamp experiment described previously.

For the first experiment, we tested if CLCA1 expressed and secreted from one cell could affect cells around it. I transfected one batch of cells with eGFP and another batch of cells with full-length WT CLCA1 (or pHLsec as vector control). On day 2, I lifted the cells with trypsin and replated both populations together for co-culture. For patching, we only targeted the green cells (Figure 20a), which would not be expressing CLCA1, and an increase in currents would indicate that CLCA1 from another cell could modulate currents from the green cells.

4.2.1.2 Results

As shown in Figure 20b, the currents measured from green cells significantly increased when the green cells were co-cultured with CLCA1-expressing cells, suggesting that CLCA1 could modulate currents of neighboring cells. As previously observed, these currents were Ca-sensitive and could not be activated with 0 µM free Ca$^{2+}$ in the pipette (Figure 20b). Cells that were co-cultured with vector-transfected cells did not display any increase in currents, regardless of the calcium concentration in the pipette.
Figure 20. CLCA1 can activate CaCC in a paracrine fashion. A. Schematic of the patch-clamp experiment. B. Chloride current density measured from GFP-expressing cells co-cultured with vector- or CLCA1-expressing cells. Pipette solution had 0 or 10 μM free Ca²⁺ to determine calcium dependency of the currents. Figure modified from Sala-Rabanal et al. (10).

4.2.2 Conditioned media

4.2.2.1 Experimental design

In order to verify that it is the secreted CLCA1 protein that activates I\textsubscript{CaCC}, we took the “GFP-coculture” experiment one step further: This time, one batch of cells was transfected with full-length WT CLCA1 (or pHLsec), while the other batch was left untreated. On the second day, I took the supernatant from CLCA1-transfected cells, “CLCA1-conditioned media”, and incubated the untreated cells in it for 24 hours prior to patching.
4.2.2.2 Results

Similar to the co-culture experiment (Figure 20), CaCC currents were significantly increased when cells were grown in CLCA1-conditioned media (Figure 21), indicating that CLCA1 could activate these currents in a paracrine fashion. The Ca sensitivity was verified in this setup, as well, where 1 μM free Ca$^{2+}$ in the pipette solution was still insufficient to activate currents.

Replenishing the media from transfected cells a few hours after the transfection to remove residual transfection reagent did not change the results of the patch clamp assay, indicating that the leftover transfection complexes in the supernatant, if there were any, were not responsible for activating currents.

![Figure 21. CLCA1-conditioned medium increases CaCC currents. A. Schematic of the patch-clamp experiment. B. Chloride current density measured from cells conditions with CLCA1. Pipette solution had 0, 1 or 10 μM free Ca$^{2+}$ to determine calcium dependency of the currents. Figure modified from Sala-Rabanal et al. (10).](image-url)
4.2.3 Purified protein

4.2.3.1 Rationale

While the conditioned media format was readily applicable, it did not definitively show that CLCA1 was the sole effector of CaCC activity, as the overexpression of a protein by transient transfection can result in the expression and secretion of other species. To address this concern, we decided to use purified CLCA1 protein.

4.2.3.2 Protein expression and purification

CLCA1 construct 22-694 was cloned into Avitag backbone (11), which contains a C-terminal 6His-tag, followed by a BirA ligase recognition sequence. The plasmid was transiently transfected into 293F cells, which grow in suspension in serum-free media. 72 hours after transfection, cell culture was spun down and the supernatant was collected. The protein was purified using nickel affinity chromatography, exchanged into biotinylation buffer (100 mM tris pH 7.5, 200 mM NaCl, 5 mM MgCl₂) and concentrated in a centrifugation filter to a final volume of 300 µl. Protein concentration was calculated using the Beer-Lambert law from sample absorbance at 280 nm: \[ A = \varepsilon \cdot b \cdot c. \]

The extinction coefficient \( \varepsilon \) was calculated from the primary sequence using ProtParam (http://web.expasy.org/protparam/) from ExPasy (12), which uses the number of tyrosines, tryptophans and cysteines in the sequence. The path length b is 1 cm for the cuvettes used.
50 µl of Biomix A (0.5 M bicine pH 8.3) and B (100 mM ATP, 100 mM MgOAc) were added. 100 µl 500 µM biotin was added. 20 µl recombinant BirA ligase was added. The solution was spun gently overnight at 4°C, then purified using a 0.5 ml Zeba desalting column (Thermo Fisher 89882).

Purified N-terminal CLCA1 (native or biotinylated) protein was added into the pipette solution at 10-50 µg/ml concentration. As a negative control, the exact buffer matches were used in the pipette.

4.2.3.3 Results

CaCC currents were activated by purified protein to the same level as observed with transfection and conditioned media formats, whereas the buffer controls did not induce currents above background (Figure 22). This experiment verified that CLCA1 protein is necessary and sufficient for CaCC activation.

Figure 22. Pure N-terminal CLCA1 protein increases CaCC currents. A, Representative traces from cells patched with native or biotinylated N-terminal CLCA1. B, Chloride current density measured from cells patched with native or biotinylated N-terminal CLCA1 or exact buffer matches (-). Figure modified from Sala-Rabanal et al. (10).
4.3 Identification of TMEM16A as Target CaCC

4.3.1 TMEM16A Structure and Function in the Airways

The following section is taken from the review “Novel Roles of Chloride Channels, Exchangers, and Regulators in Chronic Inflammatory Diseases” authored by Monica Sala-Rabanal, myself, Kayla N. Berry and Thomas J. Brett (13). Each author contributed one or more chapters within the review. While all authors were part of the editing and proofreading process, I reviewed the relevant literature and composed the section on TMEM16A.

While CaCC conductance was a long-observed phenomenon in the airways and could be separated from CFTR currents, the molecular identity of the channels responsible for these currents remained elusive until the late 2000s. The TMEM16/Anoctamin family was identified in 2008 as the first bonafide CaCCs (14-16). However, based on their electrical and pharmacological characterization, only two of the ten family members, TMEM16A and TMEM16B, displayed properties previously observed for CaCCs in the airways (17, 18), whereas most of the other members function as lipid scramblases. Of these two, TMEM16A expression has been verified in airway epithelium and airway smooth muscle cells (19).

The predicted topology for TMEM16 family members (Figure 23) is based on the recent landmark crystal structure of the fungal Nectria haematococca TMEM16 (nhTMEM16), which has 10 transmembrane domains instead of the previously predicted 8 (20). The purified and reconstituted protein, which was shown to be a homodimer (21), constitutes a channel on its own and does not require other proteins to be active (22). While the Ca$^{2+}$ sensitivity of the channel is well documented (20, 23, 24) and the protein directly binds Ca$^{2+}$, the possible involvement and mechanism of interaction with calmodulin as a calcium sensor and binding partner are still controversial (23, 25, 26).

Similar to CLCA1, expression of TMEM16A is upregulated by IL-4/IL-13 stimulation (Figure 2) (27-30). Upon upregulation, TMEM16A co-localizes to the apical plasma membrane of goblet cells (28, 31, 32), along with the mucin MUC5AC. It is expressed in airway smooth muscle cells, and has been shown to play a role in AHR (32). Additionally, inhibitors of TMEM16A have been shown to block mucus secretion (32, 33), whereas small molecule activators increase secretion (34).
Figure 23. Schematic of TMEM16A. Pink boxes at the N-terminus denote splicing isoforms. The extracellular loop in green is the proposed binding site for CLCA1. Figure modified from Sala-Rabanal et al. (13).

4.3.2 TMEM16A inhibitors

4.3.2.1 Experimental design

T16\textsubscript{inh}-A01 (27) and MONNA (35) were identified in high-throughput screens as TMEM16A-specific blockers. Neither of the compounds block CFTR currents (27, 35) and MONNA also does not block Bestrophin1 or CLC2 (35). T16\textsubscript{inh}-A01 is commercially available (4538 Tocris), but MONNA was synthesized in house by Dr. Arthur G. Romero. Stock solutions were made by dissolving powder in dimethyl sulfoxide (DMSO) and stored at -20°C.
The small molecule inhibitors were used in the whole-cell patch clamp assay to determine their effect on CLCA1-activated CaCC currents. For the assay, cells were transfected with full-length CLCA1 and replated on cover slips after 24 hours. The inhibitors were added onto the culture media at a final concentration of 10 µM and incubated overnight with the cells. Equal volume of DMSO was used as a vehicle control.

4.3.2.2 Results

CaCC currents were activated by CLCA1, as previously described. DMSO treated cells did not show any significant changes in the currents measured. However, incubation with either of the inhibitors significantly decreased CLCA1-activated currents to levels comparable to untransfected cells (Figure 24a). These results suggested that these small molecule inhibitors were decreasing CLCA1-activated CaCC currents.

4.3.3 TMEM16A siRNA

4.3.3.1 Experimental design

Commercially available TMEM16A siRNA (HSS123904, Life Tech) was dissolved in ultra-pure water to yield a 20 mM stock solution. As a transfection control, medium GC content RNA (12935300, Life Tech) was also prepared at 20 mM concentration. On the day of transfection, siRNA was diluted in OPTI-MEM. The transfection reagent, Lipofectamine 2000 at 20 pmol siRNA: 2 µl reagent ratio, was diluted separately in OPTI-MEM. The diluted transfection reagent was added onto the siRNA solution and incubated for 5 minutes at RT for complexation. Then, the solution was added onto the cells dropwise and the cells were returned
to the incubator. The final concentration of siRNA in cell culture volume should be 200 nM. The format that retains cell viability while maintaining siRNA transfection efficiency is plating 293T cells in a 48-well-plate in a volume of 200 µl 24 hours before transfection. The total volume of added siRNA-Lipofectamine 2000 complex should be 20 µl, increasing the total volume to 220 µl.

To replate cells 24 hours after transfection, I resuspended them in their own supernatant, instead of using trypsin, as the siRNA transfection made cells very fragile and they would not survive the trypsin treatment. siRNA- or control RNA-transfected cells were then replated in CLCA1-conditioned media and incubated for 24 hours at 37°C prior to patch clamp.

4.3.3.2 Results

Cells transfected with siRNA did not respond to CLCA1 treatment. In contrast, cells transfected with a control RNA responded to CLCA1 treatment and measured currents were significantly higher (Figure 24b). These results showed that TMEM16A siRNA could block the currents being activated by CLCA1, definitively showing that CLCA1 was activating TMEM16A.
Figure 24. CLCA1 increases TMEM16A currents without affecting TMEM16A expression. A. Chloride current density measured from cells conditioned in CLCA1-media and incubated with T16inh-A01 (A01 in the figure) or MONNA. Chemical structures of inhibitors were drawn with PubChem Sketcher V2.4. B. Chloride current density measured from cells transfected with TMEM16A siRNA (or control RNA) and conditioned with CLCA1. C. Western blot for TMEM16A expression in HEK293T cells that are transfected or conditioned with CLCA1. Figure modified from Sala-Rabanal et al. (10).
4.3.4 TMEM16A expression

4.3.4.1 TMEM16 mRNA levels after CLCA1 treatment

TMEM16A mRNA levels were measured by quantitative PCR (qPCR) using commercially available probes (PrimeTime Std qPCR assay, IDT: Hs.PT.58.2239278, Hs.PT.56a.3654734, Hs.PT.58.22490791, IDTDNA). The probes were reconstituted in 500 µl EB, which makes a 20x stock, and stored at -20°C.

RNA was isolated from 293T cells using TRIzol (15596026 Thermo Fisher) according to manufacturer’s instructions and diluted to 100 µg/ml concentration in ultra-pure water. cDNA library was constructed with High-capacity cDNA Reverse Transcript Kit (4374966 Thermo Fisher) (4 µl 10x RT buffer + 1.6 µl 25x dNTP + 4 µl 10x random primer + 2 µl multiscribe RT + 0.2 µl RNase inhibitor + 8.2 µl water + 20 µl 100 µg/ml RNA). For the qPCR reaction, 2 µl of the cDNA samples was mixed with 10 µl cocktail, which contained 4.3 µl water, 0.5 µl Std qPCR assay, 0.2 µl Rox low and 5 µl KAPA (KK4716 Kapabiosystems).

The reaction was run on an Applied Biosystems 7500 Fast instrument. These assays have a FAM reporter and NFQ-MGB quencher. Rox was selected as passive reference. Expression levels of TMEM16A were normalized to human GAPDH (4352934-1010033, Applied Biosystems) levels from the same sample and all samples were normalized to vector transfected HEK293T cells. TMEM16A mRNA levels were not significantly increased after the cells had been either transfected or conditioned with full-length WT CLCA1.
4.3.4.2 TMEM16A protein levels after CLCA1 treatment

To determine the effect of CLCA1 protein on TMEM16A protein levels, HEK293T cells were transfected or conditioned with full-length WT CLCA1. Alongside the treatments with CLCA1, some cells were also transfected with TMEM16A siRNA to verify reduced TMEM16A protein levels.

After 24 hours of incubation, cells were harvested and lysed. TMEM16A and actin (as loading control) proteins were detected via Western blotting. The developed film was scanned and the image was converted to an 8-bit tiff file for densitometry quantification. Any modifications to the brightness and/or contrast were applied to the entire image. Pixel intensities within boxes of equal size were quantified using Imagej 1.48 (http://imagej.nih.gov/ij/).

Compared to untransfected cells, TMEM16A protein levels were unchanged when the cells were transfected or conditioned with CLCA1, as expected from the mRNA data. Cells that were transfected with TMEM16A siRNA had significantly reduced protein levels, which did not increase by additional CLCA1-conditioned media treatment (Figure 24c).

4.3.5 Confocal

In order to visualize cellular localization of TMEM16A and CLCA1, HEK293T cells were transfected or conditioned with CLCA1 and then stained for both CLCA1 and endogenous TMEM16A.
While these HEK293T cells do express TMEM16A, there was no detectable surface staining for it in untreated cells (Figure 25- first and third columns), which could be explained by the low basal levels or rapid trafficking/shuttling of the channel to and from the cell surface.

Surprisingly, transfecting the cells with CLCA1 increased the surface staining of TMEM16A (Figure 25- second column). Results were similar when conditioned medium was used instead of transient transfection (Figure 25- fourth column). In both cases, CLCA1 and TMEM16A were co-localized on the cell surface.

Figure 25. CLCA1 increases TMEM16A cell surface staining. Confocal microscopy images showing cells co-stained for CLCA1 and TMEM16A following CLCA1 treatment by either transfection or conditioned media. Figure taken from Sala-Rabanal et al. (10).
4.3.6 Pull-down

4.3.6.1 Rationale

Since the images acquired with the confocal microscope suggested a co-localization for CLCA1 and TMEM16A on the cell surface, I hypothesized that there would be a direct interaction between the two proteins and that I could use a co-immunoprecipitation assay to pull-down TMEM16A using CLCA1.

4.3.6.2 Experimental design

*In vitro* biotinylated CLCA1 Nterm (22-694) was prepared as described before. On the day of the experiment, adherent HEK293T cells were lifted in PBS, resuspended in lysis buffer and lysed shaking at 4°C for 30 mins.

**Lysis buffer recipes:**

- 150 mM NaCl, 50 mM HEPES pH 7.4, 1% NP40
- 1% NP40 in PBS
- 1.5 mM KH$_2$PO$_4$, 4.3 mM Na$_2$HPO$_4$, 2.7 mM KCl, 137 mM NaCl, 1% Triton-100
- 140 mM NaCl, 20 mM tris-HCl, 5% glycerol, 1% Triton-100

The cell lysate was incubated with biotinylated CLCA1 protein, shaking at 4°C for 30 mins. When using a crosslinker, fresh dithiobis(succinimidyl propionate) (DSP) was added at a final concentration of 1 mM, shaking at RT for 1 hour. The reaction was quenched with tris pH 8 at a final concentration of 100 mM. Then, streptavidin-coated magnetic beads (DynaBeads MyOne C1, Thermo Fisher 65001 / Spherotech SVFM-40-5) were added and incubated with the protein mixture at 4°C for 1 hour. The beads were isolated from the solution using a magnet stand and washed beads with buffer.
Wash buffer recipes:

- 150 mM NaCl, 50 mM HEPES pH 7.4
- 1.5 mM KH$_2$PO$_4$, 4.3 mM Na$_2$HPO$_4$, 2.7 mM KCl, 137 mM NaCl, 1% triton-100

Finally, aliquots were mixed with 2x SDS and proteins were detected by Western blotting with antibodies against CLCA1 and TMEM16A.

4.3.6.3 Alterations on the protocol

- For *in vivo* biotinylation, 293T cells can be transfected with Nterm Avitag and mammalian BirA constructs (1 µg/10⁶ cells) using 293Fectin. After the transfection, add biotin into the cell culture at a final concentration of 5 µM. Incubate cells overnight at 37°C, 5% CO$_2$.
- The order of lysis, CLCA1 incubation and crosslinking can be changed; i.e. intact cells can be incubated with protein, cross-linked and then lysed in buffer.
- Instead of specifically biotinylating CLCA1, antibodies can be non-specifically biotinylated *in vitro* using EZ-Link NHS-PEG4-Biotin (Thermo Fisher 21330). For the protocol, biotinylated antibody addition would be after the step with CLCA1 incubation.
- Instead of directly boiling beads, the proteins can be eluted off the beads in an elution buffer, like 10 mM EDTA pH 8.2.
- Instead of adding biotinylated CLCA1 onto cells and then adding beads into solution, biotinylated CLCA1 can be incubated with the beads and then added onto intact or lysed cells.
4.3.6.4 Results

TMEM16A is a 10-transmembrane-domain protein and the co-immunoprecipitation of the endogenous channel, especially one with low expression, is a challenging experiment. Successful pull-down experiments have been conducted only with heterologous expression of tagged TMEM16A proteins (21, 36-39).

Using the methods described above, I was not successful in co-immunoprecipitating endogenous TMEM16A using biotinylated CLCA1 protein in a reproducible manner. TMEM16A stuck to the DynaBeads non-specifically, which was not observed with the Spherotech beads. Although freshly biotinylated CLCA1 could be loaded onto the Spherotech beads successfully, there was no detectable enrichment of the TMEM16A band on the Western blot.

4.3.7 Flow cytometry

4.3.7.1 Experimental design

As an alternative to the pull-down experiment, we designed a flow cytometry experiment to show a physical interaction between CLCA1 and TMEM16A on HEK293T cell surface. For the binding assay, biotinylated recombinant CLCA1 protein was used as a staining reagent.
Biotinylated N-terminal CLCA1 was prepared as described above. HEK293T cells were plated on a Petri dish. On the day of the experiment, cells were lifted and resuspended in ice-cold FACS buffer (1% bovine serum albumin (BSA) in PBS) at 2x10^6/ml density. Cells were sterile filtered through 50 µm and aliquoted into 200 µl samples. 2 µl/sample human Fc receptor (FcR) block (Miltenyi Biotech 130-059-901) was added and incubated for 10 minutes on ice. In the meantime, biotinylated CLCA1 protein was incubated with streptavidin-coated phycoerythrin (SA/PE) (Thermo Fisher S-866) for complexation for 15 minutes at RT at 4:1 mole ratio (Figure 26A). Blocking antibodies (Table 6) were added onto the cells at 1:10 dilution and incubated for 20-40 minutes on ice. Then CLCA-SA/PE complexes were added onto the cells at 1:50 dilution and incubated for 1-2 hours on ice in the dark. The same volume of SA/PE reagent was added to a sample for background staining assessment. Following the incubation, cells were lightly spun down, washed with FACS buffer three times and finally resuspended in 400-800 µl of FACS buffer. Samples were analyzed by FACScan instrument and data were analyzed with FlowJo 9.

Table 6. Blocking antibodies used in flow cytometry assay

<table>
<thead>
<tr>
<th>Name</th>
<th>Catalogue number</th>
<th>Epitope</th>
</tr>
</thead>
<tbody>
<tr>
<td>S20</td>
<td>69343 Santa Cruz</td>
<td>TMEM16A loop9-10</td>
</tr>
<tr>
<td>C5</td>
<td>377115 Santa Cruz</td>
<td>TMEM16A cytosolic N-term</td>
</tr>
<tr>
<td>G19</td>
<td>9890 Santa Cruz</td>
<td>Aquaporin5</td>
</tr>
</tbody>
</table>
### 4.3.7.2 Results

Staining the cells with CLCA1-SA/PE increased the mean fluorescence intensity (mfi) by a 10-100 fold over SA/PE background staining (Figure 26b, green line), which indicated the binding of CLCA1 to a target on the surface of HEK293T cells. In the follow-up experiment, cells were pre-incubated with an anti-TMEM16A antibody (S20, epitope within the loop 9-10). When the CLCA1-SA/PE reagent was added onto these cells, the mfi was higher than background SA/PE staining, but significantly lower than CLCA1 staining without antibody in a reproducible manner (Figure 26b). Neither of the control antibodies (Table 6) blocked CLCA1 binding to cell surface (Figure 26c) indicating the blocking effect of the loop antibody was specific. This finding suggests CLCA1 might be interacting with the 9-10 loop of TMEM16A on cell surface.

![Figure 26. CLCA1 binds TMEM16A on cell surface. A. model of tetrameric CLCA1 staining reagent coupled to SA-PE. B. Histograms showing SA-PE background staining (black line), CLCA1 binding (green line) and blocking effect of the antibody (red line). C. Histograms showing SA-PE background staining (black line) and negative control antibody with CLCA1 (orange and blue lines). Figure modified from Sala-Rabanal et al. (10).](image)

81
4.4 Time course of interaction

The two functional assays that we established, the whole-cell patch clamp and the confocal imaging, included an overnight incubation step, which means the cells were exposed to CLCA1 for approximately 18 hours to affect TMEM16A. However, if the interaction between CLCA1 and TMEM16A is direct, the actual time course of activation, in terms of surface staining and currents, should be significantly shorter than 18 hours.

4.4.1 Fixed cell imaging

To determine the minimal amount of incubation time required for robust TMEM16A cell surface staining, I incubated naïve HEK293T cells with CLCA1-conditioned media and immediately fixed cells at distinct time points (0, 5, 10, 20, 30, 60 and 120 minutes). All the samples were fixed and blocked for the same amount of time and the timelines matched up for the overnight incubation step with the primary antibody. The samples were co-stained for TMEM16A and CLCA1.

In the first round of experiments, 30, 60 and 120-minute time points were assayed. The staining results indicate that TMEM16A surface staining is robust within 30 minutes of exposure to CLCA1-conditioned media. The difference in fluorescence intensity between 30, 60 and 120 minutes of incubation was negligible, indicating that a half an hour was enough for CLCA1 to act on TMEM16A (Figure 27a).
Figure 27. CLCA1 increases TMEM16A staining within minutes. A, Images from cells exposed to CLCA1 for 0-120 minutes. B, Images from cells exposed to CLCA1 for 0-30 minutes. Note: panels A and B are not on the same scale.
In the second round of experiments, I assayed a shorter time window and fixed cells at 5, 10 and 20-minute time points to compare with the 30-minute time point. Increase in TMEM16A surface staining was detectable within 5 minutes of exposure to CLCA1 and comparable in intensity to the 30-minute time point (Figure 27b), indicating that the interaction between the two molecules and the subsequent increase in surface staining are occurring within a few minutes.

4.4.2 Live cell imaging

The results from fixed cell immunohistochemistry indicated that exposure to CLCA1 for 1-5 minutes was long enough to increase TMEM16A surface staining to robust levels. Based on these results, combined with the observation that TMEM16A mRNA and protein levels remained the same on a whole cell level, I hypothesized CLCA1 was stabilizing TMEM16A on cell surface and blocking the recycling of the membrane protein from cell surface. In order to monitor TMEM16A recycling, I used the live cell imaging option of the confocal microscope and monitored the cellular trafficking of GFP tagged TMEM16A following CLCA1 treatment. I used nocodazole in these experiments as a positive control for blocked trafficking.

4.4.2.1 Experimental design

Adherent HEK293T cells were cultured as described previously. Cells were plated on a 6-well plate and transiently transfected with 1 µg mTMEM16A-GFP or 1 µg CLCA1 (for conditioned medium) with 293Fectin according to manufacturer’s instructions. Nocodazole (1404, Sigma Aldrich) stock was prepared by dissolving the white powder in DMSO for a stock concentration of 2 mg/ml. 24 hours after transfection, mTMEM16A-GFP transfected cells were trypsinized and replated on imaging dishes (MatTek P35G-0-14-C) and grown overnight for
adherence. 1 hour before the experiment, either the medium was replaced with CLCA1-conditioned media or 1:100 nocodazole stock (or DMSO as vehicle) was added into the media and returned to the incubator. Cells were imaged without any further disturbances with 488 nm laser for 25 cycles with 5 s intervals at 63x magnification (+2x digital magnification).

4.4.2.2 Results

Untreated cells displayed great mobility of bright green dots corresponding to TMEM16A channels both inside the cell and on the cell surface (Figure 28a). The movement is not as distinct on the graph for the puncta on cell surface, as the background fluorescence for the surface is significantly higher than for the cytoplasm. Cells treated with DMSO vehicle looked the same as untreated cells (data not shown).

Following application of nocodazole into the supernatant, the green dots were stationary and there was no increase or decrease in the fluorescence intensity (Figure 28b). This indicated that the transport of GFP-tagged TMEM16A was stopped in both directions, trafficking to and from the membrane, as the trafficking machinery was impaired.

Incubating the cells with CLCA1-conditioned medium did not slow down the intracellular trafficking of the green dots, as they retained mobility. However the green TMEM16A dots on the cell surface remained stationary throughout the experiment (Figure 28c), suggesting that CLCA1 was affecting recycling from the cell surface, but not intracellular shuttling to the surface. Results were similar between cells conditioned overnight (data not shown) and cells conditioned for 1-2 hours.
Figure 28. CLCA1 blocks recycling of TMEM16A from cell surface. Live cell images from $t=0$ and graphs showing change of fluorescence within the colored boxes over the course of the experiment (120 seconds). A, Untreated cells. B, Cells incubated with nocodazole for 1 hour. C, Cells conditioned with CLCA1 for 1 hour.
4.4.3 Effect of nocodazole

In order to ensure that nocodazole was not disrupting the integrity of the cells, the application conditions were mimicked for the whole cell patch clamp assay. Cells treated with nocodazole displayed increased currents, somewhat higher than cells conditioned with CLCA1 (Figure 29). Since the small molecule blocks microtubule polymerization/dynein activity (40), which is a trafficking mechanism common to most channels, we hypothesized that the treatment would increase surface density of multiple channels and result in significantly high currents. Surprisingly, the effect of nocodazole on these currents could be completely blocked by transfecting the cells with TMEM16A siRNA (Figure 29), indicating that the increase in currents is solely due to increased TMEM16A activity.

Figure 29. Nocodazole affects TMEM16A currents. Chloride current density measured from cells treated with nocodazole ± TMEM16A siRNA or control RNA.
4.5 Summary

Based on the functional data obtained, we have come up with a model of interaction between CLCA1 and TMEM16A (Figure 30): CLCA1 gets expressed as a full-length protein, gets secreted and the self-cleavage reaction releases the N-terminal fragment as the functionally active portion of the protein. This fragment in the extracellular medium interacts with cells and binds the cell surface through TMEM16A. The direct interaction between the two molecules has rapid consequences of stabilizing TMEM16A on cell surface within minutes and blocking its recycling from the surface, which lead to higher channel surface density and thus higher chloride currents.

Figure 30. Proposed model of interaction. Secreted full-length CLCA1 undergoes self-cleavage. The released N-terminal cleavage fragment interacts with the CaCC TMEM16A on the cell surface. The direct interaction stabilizes the channel on cell surface and blocks recycling. Figure modified from Sala-Rabanal et al (10).
4.6 References


Chapter 5: Minimal Domains of Interaction
Minimal Domains of Interaction

The N-terminal cleavage fragment of CLCA1 is the fragment that is necessary and sufficient to increase TMEM16A currents. This 70 kDa protein is comprised of three distinct domains (the metalloprotease domain, the vWA domain and the BSR domain, see Figure 5) and the entire N-term as a recombinant protein construct (22-694) is not stable enough to facilitate more detailed structural studies required to gain more insight into the interaction mechanism. In addition, TMEM16A is also a large protein (~110 kDa), which presents 5 loops to the extracellular space (Figure 23) and any of them could be involved in the interaction with CLCA1. In order to identify the minimal domain of CLCA1, I cloned each of the domains separately and we tested their effect on TMEM16A by whole-cell patch clamp and immunohistochemistry. For the minimal domain of TMEM16A, I took a targeted approach and focused on the 9-10 loops based on the results from flow cytometry (section 4.3.7).

5.1 Minimal domain of CLCA1

5.1.1 Cloning

CLCA1 N-terminal cleavage fragment is an 80 kDa multidomain protein. In order to determine the minimal domain necessary, I cloned the three domain constructs (22-293 for metalloprotease, 297-478 for vWA, 483-694 for BSR) separately into pHLsec vector using AgeI- KpnI restriction sites.
5.1.2 Patch clamp

HEK293T cells were transiently transfected with the three short constructs. 24 hours later, cells were trypsinized and replated on UV-sterilized cover slips. Following overnight incubation, currents were measured using the whole-cell patch clamp assay. Only the vWA domain significantly increased endogenous calcium-activated chloride currents; transfection with either the catalytic domain or the BSR domain induced currents comparable to vector-transfected cells (Figure 31).

Figure 31. CLCA1 vWA is necessary and sufficient to increase TMEM16A surface staining and currents. Top panel, Chloride current density from cells transfected with short N-terminal CLCA1 constructs. Bottom panel, confocal microscopy images from cells transfected with truncated CLCA1 constructs, co-stained for TMEM16A (green) and the membrane (blue).
5.1.3 Fixed cell confocal

HEK293T cells were transfected with either the full-length CLCA1 or truncated N-terminal fragment constructs (Nterm 22-694, NV 22-478). 24 hours later, cells were trypsinized and replated on chamber slides in 400 µl volume. Following overnight incubation, media was aspirated from the slides and cells were immediately fixed. Samples were stained for TMEM16A.

Surface staining was detected with constructs that contained the vWA domain, including the full-length, N-term and NV construct. The NV construct contains only the catalytic domain and the vWA domain; since TMEM16A staining increased following transfection with NV, I concluded that the BSR domain was not necessary for this function, as supported by the patch clamp data (Figure 31). Hence, the two short constructs of interest were the catalytic domain and the vWA domain. Transfection with the catalytic domain did not increase TMEM16A staining above vector control levels (Figure 31), confirming that the vWA domain is necessary and sufficient to increase TMEM16A staining.

The time course experiment was repeated with the vWA domain alone (data not shown), where supernatant was taken from vWA-transfected cells and incubated with untreated cells for 5, 10, 20 and 30 minutes. Results were comparable with the previous time course observations with full-length CLCA1 and TMEM16A surface staining increased within 5 minutes of application.
5.1.4 Structural motifs of vWA

Since the vWA domain of CLCA1 was determined to be necessary and sufficient to directly interact with TMEM16A, I wanted to test structural motifs within the domain and to see whether they play a role in stabilizing CLCA1-TMEM16A interaction on cell surface.

5.1.4.1 MIDAS motif

The metal ion dependent adhesion site (MIDAS) motif is highly conserved among vWA domains, it coordinates a divalent cation (usually Mg\(^{2+}\)) and it is usually engaged in substrate/ligand binding (1-6). The ligand also donates an acidic residue, Asp (1, 7) or Glu (2), to coordinate the cation in the interface. The vWA domain of CLCA1 has a perfect MIDAS motif, which is significant, as it is not a conserved feature of the CLCA family (Figure 32).

![Sequence alignment of CLCA1-CLCA2-CLCA4 within the vWA domain. Identical residues are highlighted in magenta and similar residues are highlighted in yellow. Black arrows indicate the MIDAS residues. Bold residues within the CLCA2 sequence indicate variations from the perfect MIDAS motif of CLCA1.](image-url)
Based on previous literature, even a single mutation within this site is sufficient to abrogate ligand binding (1, 8-10). In order to determine if this motif was crucial in mediating CLCA1-TMEM16A interactions, the cation coordinating residues Asp312 and Ser314 were mutated to Ala (Figure 33).

![Figure 33. MIDAS motif in the CLCA1 vWA model. Residues highlighted in red in numbered order are the MIDAS residues D312, S314, S316, T383 and D412. Figure was drawn in PyMOL.](image)

The mutant vWA was used in the two functional assays and compared against the WT vWA protein. In the confocal imaging assay, which is a qualitative assay, the mutant behaved exactly the same as WT protein. However in the patch clamp assay, the amplitude of currents with mutant vWA is decreased from WT protein, but still higher than background currents (Figure 34). These results suggest, that maybe the ligand binding site, i.e. the MIDAS, is not in the ideal conformation, but it is not fully inactivated by the two mutations either.
Figure 34. The double MIDAS mutant does not completely block CLCA1 function on TMEM16A. Left, confocal microscopy images from cells transfected with vWA constructs. Right, chloride current density from cells transfected with vWA constructs.

There are examples when multiple mutations within the site were necessary to block ligand binding (11). Looking at the geometry of divalent cation coordination based on the homology model built by Phyre2 (Figure 33), the two mutations I have introduced into CLCA1 vWA may not be sufficient to completely abrogate cation coordination. So, I have also introduced the T383A and the D412A mutations onto the D312A + S314A backbone in the vWA. After the constructs were verified with sequencing, I transfected HEK293T cells with the triple MIDAS mutants and collected both supernatant and cell lysate 48 hours after transfection. I did an anti-6His Western blot to detect protein.
Figure 35. Triple MIDAS mutants do not behave like WT vWA in heterologous expression system. Left panel, 15 second exposure. Right panel, 15 minutes exposure. Lane 1, supernatant from cells transfected with vWA D312A+S314A+T383A. Lane 2, lysate from cells transfected with vWA D312A+S314A+T383A. Lane 3, supernatant from cells transfected with vWA D312A+S314A+D412A. Lane 4, lysate from cells transfected with vWA D312A+S314A+D412A. Green boxes enclose distinct species in the lysate fraction. Red box encloses the protein bands detected in the supernatant fraction.

Both constructs got expressed, as shown by the bands in the cell lysate. It should be noted that there are three distinct bands, one at the corresponding monomeric weight of the protein and two higher order sizes. Although the bands in the lysate could be readily detected with 15 seconds of exposure, the bands in the supernatant only became visible after exposing the films for 15 minutes. In this case, only a single band around 60 kDa was detected in the supernatant from both types of mutants; this size would roughly correspond to a trimeric complex (Figure 35). In the absence of a monomeric protein band in the supernatant, it is reasonable to be concerned about protein stability and secretion, both of which are required to use these constructs in the functional assays (whole-cell patch clamp and confocal microscopy).
5.1.4.2 Divalent cation-free patch clamp

As a complementary experiment to test the requirement of divalent cations in the extracellular medium (including the ones coordinated by CLCA1 MIDAS), we modified the whole cell patch clamp conditions where the bath solution contained 1 mM EDTA and 1 mM EGTA, instead of 1 mM MgCl$_2$ and 1 mM CaCl$_2$.

Native cells were used as negative control. Treated cells were conditioned with either full-length WT CLCA1 or the vWA domain. While currents were activated as expected in the normal bath solution containing Mg$^{2+}$ and Ca$^{2+}$, in the divalent cation-free buffer the activation was not significant over background (Figure 36). Overall cell appearance in this experiment was unhealthy; the cells started to shrink and round up a while after EDTA/EGTA application, as the Mg$^{2+}$ and Ca$^{2+}$, which are necessary for cell adhesion and survival, were being depleted.

![Figure 36. Extracellular divalent cations are required for CLCA1-activated currents. Chloride current density measured from cells conditioned with either full-length or vWA of CLCA1. Standard: 1 mM CaCl$_2$ + 1 mM MgCl$_2$. Divalent-free: 1 mM EDTA + 1 mM EGTA.](image-url)
5.1.4.3 Divalent-free confocal

When I used the conditions of the patch clamp experiment explained above in the context of confocal microscopy, I observed that cells exposed to media containing EDTA/EGTA even briefly (~15 mins) were starting to round up and showed signs of different morphology. Under these conditions, the experiment could not be completed.

5.1.4.4 S357N mutant

The S357N variant of CLCA1 was identified as a SNP in a clinical CF study. When the patients were ranked according to severity of CF, the frequency of this single nucleotide polymorphism (SNP) was found directly correlated with severity (12).

This point mutation falls within the vWA region of CLCA1. Since we knew that the vWA domain was the functional part in terms of CaCC activation, I hypothesized that this SNP might have a role in blocking the interactions between CLCA1 and TMEM16A.

The mutation was introduced using Quikchange Lightning Mutagenesis kit according to manufacturer’s instructions. S357N mutation within the vWA did not change the effect of the recombinant protein on TMEM16A activity as assayed by whole cell patch clamping and confocal microscopy (Figure 37).
Figure 37. S357N SNP does not affect CLCA1 function on TMEM16A. Left, confocal microscopy images from cells transfected with vWA constructs. Right, chloride current density from cells transfected with vWA constructs ± TMEM16A siRNA.

5.2 Minimal domain of TMEM16A

The last extracellular loop of TMEM16A (loop 9-10) is the most likely region of extracellular interaction with CLCA1, as it is the largest surface exposed to the extracellular matrix by the channel – all the other loops are much shorter, which reduces the likelihood of a functional surface. Also, the epitope of the blocking antibody used previously in the flow cytometry assay is within this loop (Figure 26b), further strengthening the hypothesis that the interaction is localized to this loop.
5.2.1 Biolayer interferometry

Proteins were expressed and purified from mammalian cells as described. *In vitro* biotinylation was carried out with recombinant BirA ligase. Running buffer for the biolayer interferometry (BLI) experiment contained 20 mM HEPES pH 7.4, 150 mM NaCl, 0.05% TWEEN-20 and 1 mg/ml BSA. The detergent was added to wash off substrates more easily and BSA was added to reduce non-specific substrate binding to the pins. For a single kinetic assay (~1.5 hours), final volume in 96-well plate (655209, Greiner Bio-One) was 250 µl; for longer assays (up to three hours), the volume was increased to 300 µl. Both the loading proteins and the substrates were diluted in running buffer at the desired concentrations.

Streptavidin (SA) pins (18-5020 ForteBio) were equilibrated in running buffer prior to the kinetic assays using SideKick (ForteBio). Biotinylated protein was loaded (loading signal should be around 1 nm) and pins were equilibrated in running buffer. Then, pins were dipped into substrate containing wells, starting at the lowest concentration. Pins were washed and equilibrated between each substrate dilution. To account for non-specific binding to the pin, a pin-only control experiment was performed without loading any biotinylated protein. To account for buffer-specific effects, a buffer control experiment was performed with the flow-through of the substrate. The experiment was carried out on an Octet instrument using Data Acquisition software. Data are analyzed with Data Analysis 8.2 software using the double reference mode.
5.2.2. Loop binding experiments

To determine whether the interaction between vWA and the loop can be recapitulated in vitro, I designed a biolayer interferometry experiment, where the biotinylated vWA protein is loaded onto streptavidin (SA) pins and is incubated in the well containing the purified loop peptide as the substrate (at 2 µM, 4 µM, and 10 µM concentrations). While there was some observed non-specific binding to the pins, there was also a dose-dependent response, indicating binding to the vWA protein on the pin. After the double reference subtraction, there was a net binding signal (Figure 38).

Figure 38. CLCA1 vWA domain can bind TMEM16A loop in vitro by BLI. Top panel; raw binding data with buffer only and pin only controls. Green line, pin only in buffer. Cyan line, pin only in substrate. Red, vWA on the pin in buffer. Blue, vWA on the pin in substrate. Middle panel; net binding data after double-reference subtraction. Bottom panel; association curves fit to a global model. Substrate concentration: blue 2 µM, red 4 µM, and cyan 10 µM.
5.2.3 Small peptide fragmentation

To further narrow down the region of interaction within the 9-10 loop, we ordered small segments (20 aa) of the loop as custom peptides (GenScript) for use in BLI (Table 7). The peptides were dissolved in water (+5% DMSO) and used as substrates. Binding of vWA to peptides was assayed for all 5 samples, 10-100 µM. Peptides 1-4 interacted with the pin non-specifically and the double subtracted data indicated no binding was occurring to the vWA protein (Figure 39a). Peptide 5 did not interact with the pin or the vWA protein (Figure 39b).

TABLE 7. List of small peptide fragments synthesized based on the 9-10 loop sequence of hTMEM16A.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>Sequence</th>
<th>MW (g/mol)</th>
<th>pI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TSDFIPRLVLYLMSKNGTM</td>
<td>2399.8</td>
<td>8.10</td>
</tr>
<tr>
<td>2</td>
<td>KNGTMHGFVNHTLSFNVSD</td>
<td>2192.3</td>
<td>6.92</td>
</tr>
<tr>
<td>3</td>
<td>FNVSDFQNGTAPNPDLDGY</td>
<td>2184.3</td>
<td>3.42</td>
</tr>
<tr>
<td>4</td>
<td>LDLGYEVICRYKDYREPPW</td>
<td>2543.8</td>
<td>4.78</td>
</tr>
<tr>
<td>5</td>
<td>REPWSNKYDIKDFWA</td>
<td>2268.4</td>
<td>4.78</td>
</tr>
</tbody>
</table>

As shown in Figure 39, no net binding signal to the vWA protein could be obtained with the peptides in the wells, largely due to non-specific binding to the pins themselves. Another way of conducting this experiment would be to reverse the initial experiment setup, where biotinylated loop peptide would be on the pin and the vWA domain would be in the wells and treated as a substrate. In this context, the small peptides could be spiked into the substrate wells, where they could compete for vWA binding to the pin.
Figure 39. Small peptides do not bind vWA in vitro. Representative raw kinetic data from vWA binding to peptides. Top panel, peptide 1 as representative non-specific binder. Bottom panel, peptide 5 as representative non-binder. Left column, raw binding data. Middle column, pin-only controls are subtracted. Right column, pin-only and buffer-only controls are subtracted. Green line, pin only control in buffer. Red line, pin only control in substrate. Cyan line, vWA on the pin in buffer. Blue line, vWA on the pin in substrate.
5.3 Summary

The vWA domain of CLCA1 is the necessary and sufficient unit that directly interacts with TMEM16A on cell surface. The interaction is likely localized to the last extracellular loop of TMEM16A (9-10 loop), which is the largest surface the channel is presenting to the extracellular space. The MIDAS motif within the vWA might play a role in mediating the protein-protein complex, as mutations within the site decrease TMEM16A currents significantly.
5.4 References

11. C. Canti et al., The metal-ion-dependent adhesion site in the Von Willebrand factor-A domain of alpha2delta subunits is key to trafficking voltage-gated Ca2+ channels. Proc Natl Acad Sci U S A 102, 11230-11235 (2005).
Chapter 6: Implications for Airway Disease
6 Implications for airway disease

6.1 History of CLCA Proteins

The involvement of CLCAs in airway diseases was first reported by Zhou et al., where mCLCA3 and its homologue hCLCA1, which cannot be detected in the lungs under normal conditions, were selectively induced by Th2 cytokines (IL-4, IL-9, IL-13) in mouse and human cellular models of asthma, respectively (1). Manipulation of gene expression showed that siRNA knockdown decreased MCM and AHR (2), while overexpression induced MUC5AC expression and MCM (2-5), implying a critical role for CLCA function in inflammatory mucus production. Consistent with these observations, DNA vaccines and antibodies targeting CLCA (6-8) have displayed some effectiveness in reducing airway inflammation and MUC5AC levels in mouse asthma models (9).

CLCA proteins have been ascribed various functions, starting with lung-specific cell adhesion molecules in the early 1990’s (10, 11), followed by calcium-activated chloride channels in the late 1990’s (12, 13). While it has been repeatedly confirmed that heterologous expression of CLCA results in calcium-activated chloride currents, previous hydropathy algorithm predictions of multiple transmembrane domains (TMD) within these proteins (10, 12-15) were clearly incorrect. More accurate bioinformatics approaches developed in the 2000’s predicted that CLCAs were shown to be not integral membrane proteins, but soluble secreted proteins, and these predictions were further supported with biochemical methods (16, 17). The nomenclature was revised and CLCAs are now referred to as calcium-activated chloride channel regulator proteins. Thus started the mystery of how CLCA as a secreted soluble protein could influence calcium-activated chloride currents.
An additional mystery surrounding the CLCA family arose from the observation that these secreted proteins were recovered from the extracellular medium in two cleavage fragments. Analysis of the bronchoalveolar lavage fluid (BALF) from asthmatic patients and ovalbumin-challenged mice revealed that hCLCA1 and mCLCA3, which were not detected in control samples, were abundant in the supernatant at a molecular weight corresponding to a truncated protein fragment (~70 kDa), as opposed to the full-length protein (~110 kDa) detected in cell lysates (16). The nature of the proteolytic cleavage, the identity of the enzyme responsible and how this cleavage affected CLCA function were unknown.

When CLCAs were first implicated in airway disease, they were thought to carry out the function of CaCC in the airways, but as soluble ion channel regulators, it was not clear how CLCA signaling could contribute to MCM. It is further possible to hypothesize that CLCAs have two distinct functions as signaling molecules that drive MCM and as CaCC regulators.

### 6.2 A Brief History of TMEM16A

The molecular identity of CaCC in the airways remained elusive for a long time, as previous candidate genes CLCs, Bestrophins and CLCAs were disproven to be CaCC one after the other (summarized in Caputo et al. (18)). The bonafide CaCC TMEM16A was finally identified in 2008 by three separate research groups; Schroeder et al. identified xTMEM16A as the Xenopus oocyte CaCC using Axolotl oocytes as an expression system (19). Yang et al. isolated mTMEM16A cDNA from mouse eye and showed that heterologous expression of this gene in HEK293T cells recapitulated previously observed characteristics of the endogenous CaCC (20). Caputo et al. identified hTMEM16A as the membrane protein that was differentially expressed in bronchial epithelial cells stimulated with IL-4 (18), a Th2 cytokine like IL-13 (21).
The name of the family, TMEM16, denotes the family as a group of transmembrane proteins with unknown function 16. Based on the prediction that these proteins contained 8 TMDs (18), together with the observation that family members generated anion channels, the family was given the second name of Anoctamins. This name is now controversial, as the CaCC activity has only been definitively shown for TMEM16A and TMEM16B, while other members function as lipid scramblases and/or non-specific cation channels (22). The predicted 8-TMD topology of the channels has also been revised, since the crystal structure of fungal nhTMEM16 displayed 10 TMDs and this topology appears to be conserved in mammalian orthologues by sequence analysis (23).

Expression of TMEM16A in the airways is regulated by Th2 cytokines IL-4 and IL-13 (24). This membrane protein is expressed in airway smooth muscle cells and goblet cells in the airway epithelium, where it is implicated in AHR (25) and MCM (26). While the purified protein can function on its own as a homodimeric channel (27, 28), identification of TMEM16A regulators under pathophysiological conditions is an area of active investigation. Various proteins and small molecules (Mw < 900 Da) including calmodulin, ezrin, radixin, moezin and PIP$_2$ have been proposed to interact with TMEM16A (29-31), but these interactions are localized to the cytoplasm and there had been no reports of an extracellular protein regulator of TMEM16A activity until this study.
6.3 Contributions of this Study

The work carried out in this dissertation represents a detailed study of CLCA1 structure and function, using biochemistry, biophysics, cell biology, electrophysiology and structural biology as tools. The experiments answer questions pertaining to the nature of CLCA proteolytic cleavage and the link between this enzymatic reaction and CLCA function as a CaCC regulator. In addition, this study directly identifies TMEM16A as the channel regulated by CLCA1, as well as giving mechanistic insight into how CLCA1 modulates currents through the channel.

The first chapter provides an introduction into MCM and how CLCA is involved in chronic inflammatory airway diseases as a signaling molecule, while the importance of ion currents in mucociliary clearance is highlighted. The second chapter focuses on hCLCA1 structural organization as predicted by fold recognition servers and describes the protein expression and purification system designed for structural studies of CLCA1. These expression experiments are significant to the study, as the distinct secreted domains disprove the previous hypothesis of multiple TMDs and indicate the presence of a zincin metalloprotease in the unique N-terminal domain of CLCA1.

The third chapter investigates the hypothesis that there is an N-terminal metalloprotease using an array of in vivo and in vitro assays, which confirm the presence of a zinc-dependent metalloprotease that is responsible for the proteolytic cleavage of CLCA proteins. This conserved feature of the family enables cross-reactivity between the human CLCA proteins and enables two proteins to interact for the intermolecular cleavage reaction to occur.
The fourth chapter elucidates the role of CLCA1 as a CaCC regulator and identifies the target CaCC as TMEM16A. Self-cleavage of CLCA1 is required for channel activation, as the full-length uncleaved mutant CLCA1 proteins, either with a disabled catalytic site or with an unrecognizable cleavage sequence, were unable to increase currents over background levels. Following secretion and self-cleavage, the N-terminal cleavage fragment in the extracellular medium can directly interact with TMEM16A expressed on neighboring cells. The direct interaction between the two molecules stabilizes plasma membrane-associated TMEM16A, which otherwise is recycled across the membrane quickly enough that surface levels of TMEM16A are too low to be detected by immunohistochemistry. As a result of this stabilization, surface density of TMEM16A is increased, which leads to the higher currents induced by CLCA1. Two manuscripts resulting from this work detail CLCA1 self-cleavage and the mechanism of TMEM16A activation (32, 33). These results highlight a novel mechanism for ion channel regulation, where secreted protein CLCA1 is the first direct protein modifier of TMEM16A activity.

The fifth chapter investigates the minimal interacting domains of CLCA1 and TMEM16A. The distinct domains within the N-terminal cleavage fragment were separately used in the functional assays, which revealed that the vWA domain of CLCA1 is necessary and sufficient to interact with TMEM16A to increase currents. Preliminary direct binding assays using biolayer interferometry, supported by the blocking antibody in the flow cytometry experiment, suggest that vWA binds the C-terminal loop of TMEM16A (9-10 loop), which is roughly 70 residues long. The MIDAS motif within the vWA domain is potentially important for mediating CLCA1-TMEM16A interactions. The potential structural features within the loop that contribute to the interaction need to be investigated further.
6.4 Future Directions

This is the first study that has linked CLCA and TMEM16 family of proteins, not just by physical association, but also by functional association. While further studies will be needed to determine whether CLCA1-TMEM16A interaction is involved in airway physiology, or whether defective interactions might play any role in disease phenotypes such as like MCM, potential functional interactions between other family members should also be investigated. Both CLCA and TMEM16 family members display great variability in their tissue expression patterns in mammalian organisms (22, 34), which could indicate tissue- and disease-specific CLCA-TMEM16 pairings, similar to CLCA1-TMEM16A in chronic inflammatory airway diseases.

Based on the data obtained in this study, CLCA1 seems to carry out two distinct functions in the airways: on one hand, CLCA1 is required for MUC5AC expression, and contributes to MCM. From this perspective, secreted CLCA1 can be considered an accessible drug target for asthma and COPD. Unpublished experiments conducted outside of this thesis suggest that self-cleavage of CLCA1 is required to induce MUC5AC in cells (Dr. Tom Brett, unpublished observations). These results implicate the CLCA1 metalloprotease in the signaling function. Further structural studies are needed to determine the geometry of the enzyme and to guide structure-based drug design towards MCM therapeutics. On the other hand, CLCA1 is a regulator of TMEM16A, which is a chloride/bicarbonate channel, similar in its ionic permeability function to CFTR. The direct interaction between CLCA1 and TMEM16A results in increased chloride currents, which could be beneficial in a CF setting by rescuing the diminished ion gradient to improve mucus hydration and mucociliary clearance. It is specifically the vWA domain of CLCA1 that is necessary and sufficient to carry out this particular function. This minimal domain can be manipulated separately from the full-length CLCA1 to alleviate CF
symptoms in the airways. Judging from these complex functions, the concentration of CLCA1 in the airways is likely a major determinant for the fate of airway mucus and pathophysiology, as the self-cleaved protein is secreted and can affect airway epithelia in a paracrine fashion once disseminated within BALF.
6.5 References

16. A. Gibson et al., hCLCA1 and mCLCA3 are secreted non-integral membrane proteins and therefore are not ion channels. *J Biol Chem* **280**, 27205-27212 (2005).
APPENDIX
Self-cleavage of Human CLCA1 Protein by a Novel Internal Metalloprotease Domain Controls Calcium-activated Chloride Channel Activation

Received for publication, August 13, 2012, and in revised form, October 29, 2012. Published, JBC Papers in Press, October 30, 2012, DOI 10.1074/jbc.M112.410282


From the §Biochemistry Program, Departments of ¶Internal Medicine and ¶¶Cell Biology and Physiology, Center for Investigation of Membrane Excitability Diseases, **Drug Discovery Program, ¶Department of Pediatrics, and ¶¶Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, Missouri 63110

Background: CLCA proteins activate CaCCs; CLCA proteins have roles in cancer and inflammatory lung diseases, but their mechanism of action is unknown.

Results: CLCA proteins must undergo self-cleavage via their own novel metalloprotease domain in the N terminus to activate CaCCs.

Conclusion: Self-cleavage unmasks the N-terminal fragment, which alone activates CaCCs.

Significance: This work identifies a unique ion channel activation mechanism defining framework to understand CLCA functions in diseases.

The chloride channel calcium-activated (CLCA) family are secreted proteins that regulate both chloride transport and mucin expression, thus controlling the production of mucus in respiratory and other systems. Accordingly, human CLCA1 is a critical mediator of hypersecretory lung diseases, such as asthma, chronic obstructive pulmonary disease, and cystic fibrosis, that manifest mucus obstruction. Despite relevance to homeostasis and disease, the mechanism of CLCA1 function remains largely undefined. We address this void by showing that CLCA proteins contain a consensus proteolytic cleavage site recognized by a novel zinc metalloprotease domain located within the N terminus of CLCA itself. CLCA1 mutations that inhibit self-cleavage prevent activation of calcium-activated chloride channel (CaCC)-mediated chloride transport. CaCC activation requires cleavage to unmask the N-terminal fragment of CLCA1, which can independently gate CaCCs. Gating of CaCCs mediated by CLCA1 does not appear to involve proteolytic cleavage of the channel because a mutant N-terminal fragment deficient in proteolytic activity is able to induce currents comparable with that of the native fragment. These data provide both a mechanistic basis for CLCA1 self-cleavage and a novel mechanism for regulation of chloride channel activity specific to the mucosal interface.

The chloride channel calcium-activated (CLCA) proteins are a complex family targeted for a role in cancer (1, 2) and inflammatory diseases (3) but are poorly understood in terms of molecular structure and function. The original annotation of this family as calcium-activated chloride channels (CaCCs) was based on the observation that overexpression of several different CLCA paralogues from various species all induced chloride current in response to cytosolic calcium flux (4, 5). However, bioinformatic (3) and experimental data (6–8) indicate that CLCA proteins generally lack essential features to form ion channels by themselves as they either contain only a single transmembrane anchor or are fully released in soluble form. Indeed, a recent study provides strong evidence that human CLCA1 functions as a secreted factor that increases the activity of other proteins that act as endogenous CaCCs (9).

Significant interest in CLCA proteins stems from their association with human disease. CLCA1 has been linked to the pathogenesis of human asthma and chronic obstructive pulmonary disease; CLCA1 expression is significantly increased in the airways of these types of patients (6), and polymorphisms in the CLCA1 gene have been reported in a subset of patients with asthma (10) and chronic obstructive pulmonary disease (11). In animal models of these diseases, the mouse and horse homologues of CLCA1 were shown to be necessary and sufficient for driving increased mucus production (12–15). Most importantly, there is evidence that CLCA1 stimulates an increase in mucus production by initiating a MAPK signaling pathway to express mucus (the major protein component of mucus) in...
Humans, a pathway that is highly activated in humans with chronic obstructive pulmonary disease (16). Understanding the mechanism of CLCA function in signaling for mucus overproduction could lead to effective anti-mucus therapies.

Other studies suggest that CLCA proteins are also associated with the pathogenesis of cystic fibrosis (CF); mutations in the cystic fibrosis transmembrane conductance regulator (CFTR), a chloride channel found in the apical membranes of mucosal epithelial cells, disrupt normal chloride transport, resulting in insufficiently salted and hydrated mucus. Fatal intestinal disease found in CFTR-deficient mice, however, is corrected by overexpression of native mouse CLCA3 (an orthologue of human CLCA1) (17). Consistent with this observation, a mutation in the CLCA1 gene is found in a subset of CF patients with more severe intestinal disease (18). These studies indicate that CLCA1 may function to alleviate CFTR deficiency symptoms by increasing endogenous CaCC activity and compensating for defective CFTR-mediated chloride transport. To that end, an understanding of the mechanism of activation could also be exploited to produce effective CF therapies.

In the present study, we aimed to better understand CLCA1 function with an analysis of CLCA processing. We recognized that proteolytic processing is critical to signaling function for other proteins (19) and that CLCA proteins are uniformly subjected to proteolytic cleavage within the secretory pathway to yield N- and C-terminal CLCA fragments of ~70 and 38 kDa, respectively (3). Here, using a combination of sequence analysis, structure prediction, and biochemical, biophysical, and electrophysiological techniques, we demonstrate that all CLCA proteins contain a consensus cleavage motif, which is recognized by a novel zinc metalloprotease domain located within the N terminus of CLCA itself. In addition to self-proteolysis, we also demonstrate that CLCA paralogues are capable of cross-proteolysis. Finally, we demonstrate that this self-cleavage event is a required step for CLCA1-based activation of CaCCs, which is mediated solely through the N-terminal fragment. Taken together, these data support a paradigm of CLCA activation through self-proteolysis to unmask an N-terminal fragment capable of gating CaCCs.

**EXPERIMENTAL PROCEDURES**

**Expression Constructs**— Constructs were generated using standard PCR and molecular biology techniques. All constructs were verified by DNA sequencing. For determination of proteolytic cleavage sites, full-length mature-form CLCA proteins (i.e., without their endogenous signal sequences) were cloned into the pHLsec vector (20) containing an optimized signal sequence and C-terminal hexahistidine tag. Soluble CLCA1s were designed by omitting C-terminal membrane anchors from those CLCA1s that contain them (i.e., human CLCA2, human CLCA4, mouse CLCA4). Specifically, constructs contained the following residues: CLCA1 22–914; CLCA2 32–899; CLCA4 22–876; mouse CLCA3 22–913; mouse CLCA4 22–897. Mutations were introduced into the full-length CLCA1 pHLsec construct using Phusion mutagenesis (New England Biolabs) for mutational analysis of the cleavage site and predicted active site residues. Full-length CLCA1 and select mutants containing the optimized signal sequences were subcloned from pHLsec constructs into the pCDNA3.1 expression vector (Invitrogen). A dual-tagged CLCA1 construct in pCDNA3.1 was made by inserting a FLAG tag directly after the signal sequence and a hexahistidine tag at the C terminus of the protein. Mammalian cell expression constructs encompassing the protease and VWA domain of CLCA1 (22–477), CLCA2 (32–455), and CLCA4 (22–459) as well as the substrate region (285–915) of CLCA1 were cloned into pHLsec with C-terminal hexahistidine tags. A tag-free bacterial expression construct composed of the protease and VWA domain of CLCA1 (22–477) was cloned into pET23b. Mammalian expression constructs of CLCA1 N-terminal fragment (22–695) and C-terminal fragment (696–915) were cloned into the pHLsec vector containing an optimized signal sequence and C-terminal hexahistidine tag (20).

**Protein Expression and Purification**— Proteins in mammalian expression vectors were expressed by transient transfection of FreeStyle 293F cells using 293Fectin cultured in serum-free FreeStyle 293 media (Invitrogen). Culture supernatants were collected 72 h after transfection, and proteins were purified to homogeneity using nickel affinity chromatography. CLCA1 22–477 was expressed in *Escherichia coli* as insoluble protein that was recovered from inclusion bodies, denatured in 6 M guanidine hydrochloride, and refolded by rapid dilution into buffer consisting of 50 mM Tris, pH 8.5, 400 mM arginine, 5 mM reduced glutathione, 0.5 mM oxidized glutathione, 10 mM CaCl$_2$, 0.1 mM ZnCl$_2$. The resulting soluble protein was purified by gel filtration followed by ion exchange chromatography. The protein was well folded as assessed by circular dichroism.

**Determination of Proteolytic Cleavage Sites by Edman Degradation**— Full-length CLCA proteins were expressed by transient transfection of FreeStyle 293F cells as described above. Culture supernatants were collected 72 h after transfection, and C-terminal fragments were captured and purified using nickel affinity chromatography. C-terminal fragments were isolated by SDS-PAGE and membrane transfer and then analyzed by N-terminal sequencing (Edman degradation) to determine the site of proteolytic cleavage. The first 5 residues were identified in each case. Additionally, intracellular processing of CLCA1 was assessed by immunoprecipitating the C-terminal fragment from the lysates of washed lung epithelial H292 cells transfected with a CLCA1-expressing adenovirus. Isolation and analysis were performed as described above.

**Sequence Analysis**— CLCA sequences and naming were based on recent nomenclature as recently reviewed (3).

**Analysis of Proteolytic Processing of Wild-type and Mutant CLCA1**— Mutants were expressed in 293F cells as described above. Culture supernatants and cell lysates were collected 72 h after transfection and analyzed by Western blot using antibodies that recognize either the N-terminal fragment (anti-N-CLCA1 mAb 8D3; epitope region: 477–695) or the C-terminal fragment (anti-His$_6$, antibody, Bethyl Laboratories). Processing of the dual-tagged hCLCA1 construct was analyzed using anti-FLAG antibody (M2 mAb, Sigma) in addition to the other antibodies. To test the CLCA1 tobacco etch virus (TEV) mutant for cleavage by exogenously added TEV protease, transfection was carried out as above followed by the addition of TEV protease to the culture medium (to a concentration of 0.2 mg/ml) 48 h after...
transfection. Samples were collected after 72 h after transfection and analyzed as above.

**Proteolytic Digestion Assays**—Purified protease (CLCA1 22–477, CLCA2 32–455, and CLCA4 22–459) and substrate (CLCA1 285–915) were produced as described above. Digestion experiments were carried out by incubating substrate (0.5 μM) and protease (2.0 μM) in 100 μl of digestion buffer (20 mM Hepes, pH 7.5, 150 mM NaCl, 10 mM CaCl₂, 10 μM ZnCl₂) at 37 °C for 18 h with samples taken at discrete time points and analyzed by SDS-PAGE/Western blot using the anti-CLCA1 mAb 8D3. Digestion experiments in the presence of inhibitors were carried out with GM-6001 (Millipore), Marimastat (Tocris), Batimastat (Tocris), or Zeynepstat001 at 40 μM or Halt (Thermo) at 1× concentration. All inhibitor stocks were dissolved in dimethyl sulfoxide (DMSO).

**Fluorogenic CLCA1Peptide Digestion Assays**—A fluorogenic peptide substrate consisting of a donor-acceptor FRET pair conjugated to the human CLCA1 cleavage sequence was synthesized by AnaSpec (DABCYL-QQSGALYIPG-EDANS). Experiments were carried out at 37 °C in the digestion buffer described above. Reaction progress was monitored using a BioTek plate fluorometer (excitation, 340 nm; emission, 490 nm). Refolded CLCA1 (22–477) was used as the protease. For reactions in the presence of inhibitors, [protease] = 10 μM, [substrate] = 3.25 μM, and [inhibitor] = 20 or 40 μM, respectively. Experiments were conducted in triplicate.

**Chemical Synthesis of a Custom MMP Inhibitor**—Synthesis of Zeynepstat001 is described in the supplemental material.

**Heterologous Expression for Electrophysiology**—Mutations predicted to interfere with the metalloprotease activity of CLCA1 (H156A, E157Q) or a disrupted cleavage site (contra) were generated using the pCDNA3.1 expression vector; cDNAs encoding the N-terminal or C-terminal fragments of CLCA1 were cloned into pHSec, as described above. HEK293T cells were plated in 6-well dishes and cultured in Dulbecco’s modified Eagle’s medium (Invitrogen) supplemented with 10% fetal bovine serum, 10° units/liter penicillin, and 100 mg/liter streptomycin. At 80% confluency, cells were transfected with the relevant plasmids using 293Fectin transfection reagent at a 1:tomycin. At 80% confluency, cells were transfected with the 42140 CaCl₂, Hepes and 10 mM CaCl₂, 10 μM ZnCl₂ at 37 °C for 18 h with samples taken at discrete time points and analyzed by SDS-PAGE/Western blot using the anti-CLCA1 mAb 8D3. Digestion experiments in the presence of inhibitors were carried out with GM-6001 (Millipore), Marimastat (Tocris), Batimastat (Tocris), or Zeynepstat001 at 40 μM or Halt (Thermo) at 1× concentration. All inhibitor stocks were dissolved in dimethyl sulfoxide (DMSO).

**Fluorogenic CLCA1 Peptide Digestion Assays**—A fluorogenic peptide substrate consisting of a donor-acceptor FRET pair conjugated to the human CLCA1 cleavage sequence was synthesized by AnaSpec (DABCYL-QQSGALYIPG-EDANS). Experiments were carried out at 37 °C in the digestion buffer described above. Reaction progress was monitored using a BioTek plate fluorometer (excitation, 340 nm; emission, 490 nm). Refolded CLCA1 (22–477) was used as the protease. For reactions in the presence of inhibitors, [protease] = 10 μM, [substrate] = 3.25 μM, and [inhibitor] = 20 or 40 μM, respectively. Experiments were conducted in triplicate.

**Chemical Synthesis of a Custom MMP Inhibitor**—Synthesis of Zeynepstat001 is described in the supplemental material.

**Fluorogenic CLCA1 Peptide Digestion Assays**—A fluorogenic peptide substrate consisting of a donor-acceptor FRET pair conjugated to the human CLCA1 cleavage sequence was synthesized by AnaSpec (DABCYL-QQSGALYIPG-EDANS). Experiments were carried out at 37 °C in the digestion buffer described above. Reaction progress was monitored using a BioTek plate fluorometer (excitation, 340 nm; emission, 490 nm). Refolded CLCA1 (22–477) was used as the protease. For reactions in the presence of inhibitors, [protease] = 10 μM, [substrate] = 3.25 μM, and [inhibitor] = 20 or 40 μM, respectively. Experiments were conducted in triplicate.

**Chemical Synthesis of a Custom MMP Inhibitor**—Synthesis of Zeynepstat001 is described in the supplemental material.

**Fluorogenic CLCA1 Peptide Digestion Assays**—A fluorogenic peptide substrate consisting of a donor-acceptor FRET pair conjugated to the human CLCA1 cleavage sequence was synthesized by AnaSpec (DABCYL-QQSGALYIPG-EDANS). Experiments were carried out at 37 °C in the digestion buffer described above. Reaction progress was monitored using a BioTek plate fluorometer (excitation, 340 nm; emission, 490 nm). Refolded CLCA1 (22–477) was used as the protease. For reactions in the presence of inhibitors, [protease] = 10 μM, [substrate] = 3.25 μM, and [inhibitor] = 20 or 40 μM, respectively. Experiments were conducted in triplicate.

**Chemical Synthesis of a Custom MMP Inhibitor**—Synthesis of Zeynepstat001 is described in the supplemental material.

**Fluorogenic CLCA1 Peptide Digestion Assays**—A fluorogenic peptide substrate consisting of a donor-acceptor FRET pair conjugated to the human CLCA1 cleavage sequence was synthesized by AnaSpec (DABCYL-QQSGALYIPG-EDANS). Experiments were carried out at 37 °C in the digestion buffer described above. Reaction progress was monitored using a BioTek plate fluorometer (excitation, 340 nm; emission, 490 nm). Refolded CLCA1 (22–477) was used as the protease. For reactions in the presence of inhibitors, [protease] = 10 μM, [substrate] = 3.25 μM, and [inhibitor] = 20 or 40 μM, respectively. Experiments were conducted in triplicate.

**Chemical Synthesis of a Custom MMP Inhibitor**—Synthesis of Zeynepstat001 is described in the supplemental material.

**Fluorogenic CLCA1 Peptide Digestion Assays**—A fluorogenic peptide substrate consisting of a donor-acceptor FRET pair conjugated to the human CLCA1 cleavage sequence was synthesized by AnaSpec (DABCYL-QQSGALYIPG-EDANS). Experiments were carried out at 37 °C in the digestion buffer described above. Reaction progress was monitored using a BioTek plate fluorometer (excitation, 340 nm; emission, 490 nm). Refolded CLCA1 (22–477) was used as the protease. For reactions in the presence of inhibitors, [protease] = 10 μM, [substrate] = 3.25 μM, and [inhibitor] = 20 or 40 μM, respectively. Experiments were conducted in triplicate.

**Chemical Synthesis of a Custom MMP Inhibitor**—Synthesis of Zeynepstat001 is described in the supplemental material.
sequence on proteolytic processing of CLCA1 (Fig. 2A). A point mutation (A696P) or variations that matched the cleavage sequence in other CLCAs (CLCA2 and CLCA4) did not affect proteolytic cleavage of CLCA1; in contrast, variations that significantly altered the consensus sequence (contra, TEV) abolished proteolytic processing (Fig. 2B). These mutations did not grossly affect protein folding as all variant proteins were robustly secreted into the media supernatants. Furthermore, TEV protease added to cultures expressing the TEV CLCA1 variant, in which the TEV protease cleavage site replaced the consensus site, was not cleaved either (Fig. 2C), suggesting that the cleavage site is not surface-exposed to exogenous proteases and may only be accessible to the internal metalloprotease domain in the natively folded full-length protein. Taken together, these results pinpoint the location of the proteolytic cleavage site in CLCA proteins and imply that a common protease with specific access to the site is responsible for cleavage of all human CLCA proteins.

**Proteolytic Processing of CLCA Requires Catalytic Residues in a Predicted Zinbin Metalloprotease Domain**—Extensive bioinformatic searching of the MEROPS protease database (23) using the CLCA consensus cleavage sequence did not provide any reasonable candidates for the protease responsible for cleaving CLCA family members. Thus, we decided to revisit the untested hypothesis that the N-terminal CLCA-N domain itself houses a metalloprotease domain similar to that of matrix metalloproteases (MMPs) (24). Indeed, when various CLCA-N domain sequences were submitted to the updated PHYRE2 fold prediction server (25), most predictions were high-confidence hits to zinbin metalloprotease catalytic domains.

The CLCA-N domain contains 8 cysteine residues that are invariant across all family members (3). With 2 of these cysteines lying in the predicted catalytic region (amino acids 22–199) (Fig. 3C) and 6 in an adjacent cysteine-rich domain (amino acids 200–283) (26) (Fig. 3A), it is most likely that the CLCA-N domain is architecturally similar to the ADAM or ADAMTS family of zinbin metalloproteases, rather than MMPs. Like the CLCA-N domain, ADAM and ADAMTS contain disulfide bonds in their catalytic regions (Fig. 3A) (27). The catalytic domains of zinbin metalloproteases contain a common HEXXHXXXXXX motif, which includes the catalytically required glutamate (bold) and zinc-chelating residues (3 histidine residues) (27). CLCA sequences in the predicted catalytic region are almost completely invariant and reveal a similar but unique catalytic motif: HEXXXXXXX (Fig. 3C). Previous studies have reported that mutations of Glu-157 in human CLCA1 (24), murine CLCA3 (28), and murine CLCA6 (29) to Glu all produce mutant proteins that are not proteolytically processed. However, the significance of these results is difficult to interpret as these studies probed the protein from crude cell lysates. The observed lack of cleavage could have been due to loss of the catalytically required residue or to gross misfolding of the protein caused by the mutation, which would also result in detecting only full-length proteins in cell lysates. To unambiguously assess the catalytic importance of these predicted residues, we performed site-directed mutagenesis of the consensus cleavage site in CLCAs. S and P labels refer to protease subsite and pocket designations, respectively.
active site residues, we analyzed proteolytic processing of CLCA1 point mutants fully secreted into media supernatants, i.e. proteins that have passed the cellular quality control machinery and are correctly folded. Mutations to all predicted active site residues (H156A, E157Q, H160A, D167A, E168A) produced only uncleaved, full-length proteins, whereas wild type and a mutation outside the active site (Q150A) displayed proteolytic processing (Fig. 3B). The data strongly suggest that the CLCA proteins contain an N-terminal zinc metalloprotease domain and that they are capable of self-cleavage.

**CLCA Metalloprotease Domains Process CLCA Proteins and Are Blocked by Specific Inhibitors**—To definitively confirm the self-proteolysis activity of CLCA1 by its metalloprotease domain, we devised a biochemical digestion assay utilizing purified proteins and monoclonal antibodies. This was carried out with separately expressed and purified CLCA1 protease and CLCA1 substrate proteins produced in mammalian cells (Fig. 4A). Substrate consisted of the full CLCA1 protein excluding the metalloprotease domain. Protease consisted of the CLCA1 metalloprotease and VWA domains as this construct was stable and robustly expressed. These proteins were mixed, incubated at 37 °C, and then analyzed at various time points for accumulation of cleavage product by Western blot with an antibody that specifically recognized an epitope found only in the substrate. Substrate alone displayed no degradation, whereas the addition of the CLCA1 protease resulted in the appearance of a cleavage product of appropriate size over time (Fig. 4B). The addition of the divalent metal cation chelators EDTA and 1,10-phenanthroline completely abrogated proteolytic activity as they sequester the catalytically required zinc. In contrast, the addition of a commercial mixture containing protease inhibitors to all classes except metalloproteases (Halt, Thermo) had no effect on proteolysis of the substrate. Taken together with the previously presented mutational analysis, these data unambiguously demonstrate that the CLCA proteins contain an N-terminal metalloprotease domain responsible for self-cleavage.

We next assessed the ability of various commercial and custom MMP inhibitors to block proteolysis in this assay (Fig. 4C). These molecules belong to the class of hydroxamate-based inhibitors, which consist of the moiety attached to a peptide mimetic chain. These inhibitors act by chelating the zinc through the hydroxamate group with the peptide mimicking side chains binding to the S1 and S2 pockets (supplemental Fig. S1) (30). The inhibitors displayed differing levels of potency, with Batimastat being most effective, followed by GM-6001 and Marimastat. In addition, we synthesized a custom inhibitor consisting of a hydroxamate moiety fused to the P1 and P2 residues of the CLCA1 cleavage sequence (termed Zeynepstat001). This inhibitor was also effective in reducing the activity of the CLCA1 protease.
**Biophysical Characterization of CLCA Protease Activity**—

We then developed a quantitative biophysical assay to characterize CLCA1 metalloprotease activity. This assay employed a fluorogenic peptide consisting of the CLCA1 cleavage sequence conjugated to a donor-quencher FRET pair for the substrate and CLCA1 protease refolded from bacterially expressed inclusion bodies. This allowed determination of the kinetics of protease activity ($K_m = 1.10 \times 10^{-6} \text{M}$, $V_{max} = 0.403$ fluorescence units/s) (Fig. 4, D and E). To further address metal dependence, experiments in the presence of chelators or excess divalent metal ions were carried out. The addition of EDTA abolished protease activity, whereas the addition of excess Zn$^{2+}$ partially restored activity (Fig. 4D). In contrast, the addition of Ca$^{2+}$ did not restore activity (data not shown). We used this assay to perform a more quantitative assessment of the MMP inhibitors. These experiments indicated that Batimastat was most effective, almost completely inhibiting protease activity at high concentrations, followed by GM-6001, Zeynepstat001, and Marimastat (Fig. 4F). We also tested the inhibitory activity of acetohydroxamic acid, which consists of only the zinc-binding moiety found in this class of compounds.

**FIGURE 3.** Mutation of predicted catalytic residues blocks proteolytic processing of hCLCA1. **A**, schematic of hCLCA1 highlighting the revised view of the CLCA-N domain as a metalloprotease. Labels denote the following: SS, signal sequence; CAT, metalloprotease catalytic domain; CYS, Cys-rich domain. *Inset* displays the predicted catalytic site of hCLCA1 threaded onto the structure of ADAMTS-1 (Protein Data Bank (PDB) ID: 2V4B) (50). Catalytic and zinc-chelating residues are highlighted. **B**, Western blot analysis of media supernatants from 293F cells expressing hCLCA1 variants. Blot was developed with a His$_6$ antibody (anti-6His), which recognizes full-length and hCLCA1 C-terminal fragment. C, sequence alignment of CLCA family members in the region of the catalytic site. Color coding is as follows: magenta, invariant residues; yellow, conservation score of 5 or greater as determined by ALSCRIPT (49).
of inhibitors. It displayed weak inhibitory activity (data not shown). Taken together, the data definitively show that CLCA1 is a Zn$^{2+}$-dependent metalloprotease and that inhibitors of CLCA1 protease activity can be rationally designed.

**CLCA Proteases Can Cross-cleave CLCA Substrates**—As multiple CLCA proteins are expressed at mucosal surfaces (3), we examined the ability of other CLCA proteases to process CLCA1 substrate. The CLCA1 substrate was incubated with...
proteases, from either CLCA2 or CLCA4. Both proteases cleaved the CLCA1 substrate, although at a reduced rate (Fig. 4G). Taken together with the data demonstrating that self-cleavage of CLCA1 is observed in variants where the cleavage sequence is mutated to that of CLCA2 or CLCA4 (Fig. 2B), this result suggests that the members of the CLCA family are capable of self-cleavage as well as cross-cleavage of other family members. As multiple CLCA proteins are expressed in the same tissues, this suggests that CLCA proteins might cross-cleave each other in vivo; however, the physiological relevance is unknown at this time.

CLCA Proteases Do Not Contain N-terminal Prodomains—A key aspect of zincc metalloproteases is their intrinsic capacity for self-regulation, and the most common mode of regulation is removal of a prodomain. For example, MMPs and ADAM family proteases both contain an N-terminal prodomain, consisting of about 80 amino acids at the N terminus of the protein, that blocks access to the protease active site. This domain must be proteolytically removed for substrate access (27, 31). Structure predictions of the CLCA-N domain metalloprotease region indicate that the characteristic catalytic fold would begin around residue 45, which would leave a short stretch of ~20 amino acids that may act as a prodomain. To probe whether CLCA1 contains an N-terminal prodomain that regulates self-cleavage, we examined the processing of a dual-tagged hCLCA1 construct (Fig. 4H). This construct contains a FLAG tag directly after the signal sequence so that detection of the FLAG tag on the fully processed protein would preclude cleavage of any N-terminal prodomain. The secreted CLCA1 protein was processed normally as detected by antibodies that recognized the CLCA1 N-terminal and C-terminal fragments. In addition, Western blotting for the FLAG tag revealed an intact N terminus. Collectively, these observations indicate that CLCA1, and by implication CLCA proteins in general, do not contain N-terminal prodomains.

Self-cleavage of CLCA1 Is Required to Modulate CaCCs—To assess the functional role of CLCA1 self-cleavage in regulating CaCC activity, we designed mutations that generate either an inactive metalloprotease catalytic site (H156A and E157Q) or an impaired cleavage site (contra) (Fig. 5). When no Ca\(^{2+}\) was added to the pipette buffer, Cl\(^{-}\) current density was 2.7 ± 1.5 pA/pF in vector-transfected cells and 3.5 ± 1.2 pA/pF in cells transfected with WT-hCLCA1 (n = 6 cells each) (Fig. 5C). These data are in agreement with reported biophysical properties of Ca\(^{2+}\)-activated Cl\(^{-}\) channels in native cells and heterologous expression systems (4, 32–35) and are consistent with a previous report that CLCA1 modulates the activity of endogenous CaCCs in HEK293T cells (9). In cells transfected with CLCA1 variants, in which the metalloprotease activity was abolished (H156A, E157Q) or in which the cleavage site was disrupted (contra), the gluconate-sensitive currents were markedly decreased (Fig. 5C), and on average, the density of Cl\(^{-}\) currents was comparable with that measured in vectortransfected cells (Fig. 5D). Expression levels of all the variants were similar to the WT (Fig. 5B), indicating that the observed effects were due to impaired proteolytic processing, rather than altered synthesis, trafficking, or secretion. These data demonstrate that self-processing of CLCA1 by its novel zincc metalloprotease domain is required for its activation of CaCCs.

The N-terminal Fragment of CLCA1 Is Sufficient to Modulate CaCC Currents—To further delineate the role of CLCA self-cleavage in activation of CaCCs, we assessed the ability of each of the cleavage fragments to induce Ca\(^{2+}\)-activated currents (Fig. 5E). At +80 mV, Cl\(^{-}\) current density was 4.0 ± 0.7 pA/pF in cells transfected with empty pHSec vector and increased to 35 ± 5 pA/pF in cells expressing full-length CLCA1. Current density in cells expressing the CLCA1 C terminus was as low as in cells transfected with empty pHSec (4.7 ± 0.6 pA/pF). In contrast, in cells expressing the N-terminal fragment, currents were similar to those measured in cells transfected with the full-length construct (32 ± 5 pA/pF). These results demonstrate that the N-terminal fragment of CLCA1 is necessary and sufficient to regulate Ca\(^{2+}\)-activated Cl\(^{-}\) currents and that self-cleavage is required to release an inhibitory C-terminal frag-

![FIGURE 4. Self-cleavage of HCLCA1. A, schematic of protein constructs used in purified protein proteolysis assays. Labels denote the following: SS, signal sequence; CAT, metalloprotease catalytic domain; CYS, Cys-rich domain; 6His, hexahistidine tag. B, proteolysis of hCLCA1 substrate by hCLCA1 protease. Samples were incubated at 37 °C with samples taken at time points as noted. Reactions are labeled as follows: A, substrate only; B, substrate (0.5 μM) + protease (2 μM); C, substrate + protease + 15 mM EDTA; D, substrate + protease + 15 mM 1,10 phenanthroline; E, protease + substrate; F, protease + substrate + HALT (1 × C). C, effect of commercial and custom MMP inhibitors on hCLCA1 proteolytic cleavage. Conditions were the same as the above with the following inhibitors added to digestion reactions at 40 μM: B, GM-6001; C, Marimastat; D, Batimatost; E, ZY39; F, activity of purified refolded hCLCA1 protease on a fluorogenic peptide corresponding to the hCLCA1 cleavage sequence (DABCA-CLQQSALSYS-EDANS). Protease and substrate were incubated together for 60 min, and product formation was measured over time. Reactions were carried out in triplicate and averaged with standard deviations shown as error bars. Color coding is as follows: green, protease (10 μM) + substrate (3.25 μM) + red, protease (10 μM) + substrate (3.25 μM) + 15 mM EDTA; gray, protease (10 μM) + substrate (3.25 μM) + 15 mM EDTA + 20 μM ZnCl\(_{2}\); E, enzyme velocity versus substrate concentration plot for refolded human CLCA1 protease (22–473) (at 4 μM) using the human CLCA1 fluorogenic reporter peptide as a substrate. Three replicates were performed for each concentration. E, effect of commercial and custom MMP inhibitors on hCLCA1 protease activity in the fluorogenic peptide digestion assay. Reaction conditions were the same as the above with inhibitors added at 20 and 40 μM. Activity was reported as the percentage of reaction velocity in the absence of inhibitors. Reactions were carried out in triplicate. G, proteolysis of hCLCA1 substrate by hCLCA2 and hCLCA4 protease. Reactions are labeled as follows: A, substrate only; B, substrate + hCLCA2 protease; C, substrate + hCLCA4 protease. H, expression of a dual-tagged hCLCA1 protein in 293F cells. Top: schematic of the construct containing an N-terminal FLAG tag and C-terminal His\(_{6}\) tag. The region containing the BD3 mAb epitope is highlighted. Bottom: Western of supernatants from 293F cells expressing dual-tagged hCLCA1. Left: anti-FLAG and BD3 mAb blot. Right: anti-His\(_{6}\), blot. N-term, N terminus; C-term, C terminus.](https://www.jbc.org/content/10.1074/jbc.M111.327884)

---

**CLCA Self-cleavage Is Required for Gating CaCCs**

![Image](https://www.jbc.org/content/10.1074/jbc.M111.327884)

Volume 287 | Number 50 | December 7, 2012 | www.jbc.org | JOURNAL OF BIOLOGICAL CHEMISTRY | 42145

**FIGURE 4. Self-cleavage of hCLCA1. A, schematic of protein constructs used in purified protein proteolysis assays. Labels denote the following: SS, signal sequence; CAT, metalloprotease catalytic domain; CYS, Cys-rich domain; 6His, hexahistidine tag. B, proteolysis of hCLCA1 substrate by hCLCA1 protease. Samples were incubated at 37 °C with samples taken at time points as noted. Reactions are labeled as follows: A, substrate only; B, substrate (0.5 μM) + protease (2 μM); C, substrate + protease + 15 mM EDTA; D, substrate + protease + 15 mM 1,10 phenanthroline; E, protease + substrate; F, protease + substrate + HALT (1 × C). E, effect of commercial and custom MMP inhibitors on hCLCA1 proteolytic cleavage. Conditions were the same as the above with the following inhibitors added to digestion reactions at 40 μM: B, GM-6001; C, Marimastat; D, Batimatost; E, ZY39; F, activity of purified refolded hCLCA1 protease on a fluorogenic peptide corresponding to the hCLCA1 cleavage sequence (DABCA-CLQQSALSYS-EDANS). Protease and substrate were incubated together for 60 min, and product formation was measured over time. Reactions were carried out in triplicate and averaged with standard deviations shown as error bars. Color coding is as follows: green, protease (10 μM) + substrate (3.25 μM) + red, protease (10 μM) + substrate (3.25 μM) + 15 mM EDTA; gray, protease (10 μM) + substrate (3.25 μM) + 15 mM EDTA + 20 μM ZnCl\(_{2}\); E, enzyme velocity versus substrate concentration plot for refolded human CLCA1 protease (22–473) (at 4 μM) using the human CLCA1 fluorogenic reporter peptide as a substrate. Three replicates were performed for each concentration. E, effect of commercial and custom MMP inhibitors on hCLCA1 protease activity in the fluorogenic peptide digestion assay. Reaction conditions were the same as the above with inhibitors added at 20 and 40 μM. Activity was reported as the percentage of reaction velocity in the absence of inhibitors. Reactions were carried out in triplicate. G, proteolysis of hCLCA1 substrate by hCLCA2 and hCLCA4 protease. Reactions are labeled as follows: A, substrate only; B, substrate + hCLCA2 protease; C, substrate + hCLCA4 protease. H, expression of a dual-tagged hCLCA1 protein in 293F cells. Top: schematic of the construct containing an N-terminal FLAG tag and C-terminal His\(_{6}\) tag. The region containing the BD3 mAb epitope is highlighted. Bottom: Western of supernatants from 293F cells expressing dual-tagged hCLCA1. Left: anti-FLAG and BD3 mAb blot. Right: anti-His\(_{6}\), blot. N-term, N terminus; C-term, C terminus.**
ment and unmask the N-terminal fragment for interaction with the channel.

**Proteolytic Activity of the CLCA1 N-terminal Fragment Is Not Required to Modulate CaCC Currents**—To address whether the proteolytic activity of the N-terminal fragment was required for activation of CaCCs, we assessed the ability of the E157Q variant of this fragment to induce Ca^{2+}-dependent currents in HEK293T cells. In cells transfected with this variant, Cl^{-} current density was close to that of the WT N-terminal fragment (30 ± 6 pA/pF, Fig. 5E). This demonstrates that gating of CaCCs by the N-terminal fragment of CLCA1 does not involve direct proteolytic clipping of the channel and suggests a direct interaction that induces gating.

**DISCUSSION**

**CLCA1 Is a Self-cleaving Zincin Metalloprotease**—Since their discovery nearly two decades ago (36), there has been much controversy regarding the specific functions of CLCA proteins and their connection to CaCC activity (9). In the present study, we demonstrate that CLCA proteins, specifically CLCA1, utilize a distinct self-processing mechanism in regulating CaCC activity. This discovery is supported by several novel findings.
First, we have demonstrated that CLCA proteins represent a novel class of zinc metalloproteases capable of self- and cross-cleavage. It should be noted that previous studies have attempted to address the issue of CLCA metalloprotease activity (28, 29); however, these studies were based on experiments utilizing CLCA proteins found in crude membrane fractions. The authors claimed that purified CLCA proteins could not be produced. Here, using purified proteins, peptides, and a comprehensive functional analysis, we unambiguously demonstrate that CLCA proteins are novel zinc metalloproteases that self-cleave. Up to the present study, the recognized secreted mammalian zinc endopeptidases consisted of MMP, membrane-bound MMP (MTMMP), ADAM, and ADAMTS families (26). These metalloproteases all contain a HEXXHXXGXX(H/D) catalytic motif, with zinc-binding histidines (the third histidine is sometimes replaced by aspartate) and the catalytically required base/acid glutamate (bold) (27). Mutational and sequence analysis of the novel CLCA metalloprotease domain reveals a related HEXXHXXXGXXDE catalytic motif. Aside from histidine, aspartate is the most common residue in the third chelating position. However, aspartate is not strictly conserved at that position within the CLCA family, whereas the adjacent (and chemically similar) glutamate is invariant (Fig. 3C). Mutational analysis indicates that both residues are structurally required for self-processing, although based on the current data, we cannot conclude whether the aspartate or glutamate constitutes the third zinc-binding residue. Regardless of this identity, the CLCA catalytic motif appears to be unique among secreted mammalian zinc metalloproteases.

A second key aspect of our findings relates to the regulation of metalloprotease activity. This is typically controlled at four levels: expression; compartmentalization; pro-enzyme activation; and inactivation (usually by inhibitors). MMP and ADAM proteins are secreted as zymogens that contain an N-terminal prodomain of around 80–100 amino acids, which folds against and blocks the catalytic active site (37). Removal of the prodomain is required for substrate molecules to access the active site cleft. However, a similar prodomain arrangement does not appear to be present in the CLCA proteins as the beginning of the predicted metalloprotease catalytic domain is within 10–20 residues of the N terminus of the mature protein. Furthermore, the processing of a dual-tagged CLCA1 protein indicates that an N-terminal prodomain is not present. Thus, the regulation of CLCA metalloproteases has distinct features that suggest the regulation of specific CLCA proteins will be unique and specific to their function. The observation that the addition of a TEV protease sequence into CLCA1 in the presence of TEV protease indicates that the cleavage site may be buried in the natively folded full-length protein, suggesting that regulation may be achieved by conformational change. The trigger for cleavage is uncertain, but it is possible that the shift in pH that occurs along the secretory pathway (starting from pH 7.4 in the ER to 5.5 in secretory endosomes) could control the compartmental location of cleavage. Additionally, from the current data, it is unclear whether self-cleavage predominantly occurs intra- or intermolecularly. Future studies will be required to define this mechanism and evaluate the role of endogenous metalloprotease inhibitors, such as the tissue inhibitor of metalloprotease (TIMP) proteins (38), in the regulation of CLCA protease activity.

A third unique aspect revealed by our analysis of CLCA metalloprotease activity is the distinct nature of the consensus cleavage site in CLCA proteins. The sequence for this site shows extreme conservation on the prime (’) side of the scissors bond (Fig. 1B), suggesting that the substrate-binding cleft in the CLCA metalloproteases is similar in feature to others in that its specificity is mainly built into the prime side, with the unprimed side being rather flat and featureless (27). Sequence analysis of the mammalian CLCA proteins suggests the presence of a cysteine-rich domain (containing six invariant cysteines) adjacent to the catalytic domain (Fig. 3A). Cys-rich domains are found in a number of ADAM and ADAMTS family members; they control substrate selectivity and access to the catalytic site (26). MMP (39) and ADAM family (26, 31) proteases play key roles as modulators of inflammation and innate immunity through activation or inactivation of cytokines, chemokines, or other proteins. The major substrate of the CLCA metalloprotease domains appears to be the CLCA protein itself, but given the central involvement of CLCAs in chronic inflammatory airway diseases, additional CLCA protease substrates should be also considered.

Self-cleavage Is a Required Feature of CaCC Activation—We demonstrate that CLCA proteins are distinct in their mode of modifying ion channel activity. In particular, the observation that self-cleavage of CLCA1 is required for regulation of CaCC activity introduces a novel mechanism for controlling ion channel gating. The CLCA mechanism exhibits some similarity to the \( \alpha_{\delta} \) subunit of voltage-gated calcium channels, which are also secreted, proteolytically cleaved, and contain a VWA domain (40). Proteolysis cleaves this protein into two pieces and also relieves the \( \alpha_{\delta} \) subunit from a transmembrane domain, and these events are required for enhanced surface expression of voltage-gated calcium channels (41). However, although \( \alpha_{\delta} \) subunits do contain an uncharacterized N-terminal domain, they do not appear to contain features consistent with being a metalloprotease, nor do structure prediction algorithms detect the presence of one.

Our data also indicate details of a mechanism of how CLCA1, and CLCA proteins in general, activate CaCCs. It has been proposed that CLCA1 increases CaCC conductance by directly affecting the permeation pathway, rather than via enhanced trafficking or surface expression of endogenous channels (9). Our results demonstrate that proteolytic cleavage is required for CLCA1-mediated activation of CaCCs, and the N-terminal fragment alone is sufficient for activation. Additionally, the fact that the N-terminal fragment E157Q variant is able to activate CaCCs just as effectively as WT demonstrates that the proteolytic activity of this fragment is not required to activate the channel. This suggests that the CLCA1 N-terminal fragment activates CaCCs by direct interaction with the channel. This mechanism is thus distinct from the manner by which secreted proteases activate epithelial sodium channels (ENaCs) in the airway through proteolysis of the channel (42). We propose that the C-terminal fragment of CLCA masks the N-terminal region in the full-length protein and that self-cleavage is required to expose the N-terminal fragment, which can then interact with
the channel (Fig. 6). The precise nature of these interactions remains to be addressed.

Consequences for Airway Disease Mechanisms and Therapies—Although the precise identity of the CaCC activated by CLCAs is uncertain, a primary candidate is the anoctamin (also called TMEM16) family of proteins, which are the only presently identified CaCC proteins (43). Within this family, only Ano1 and Ano2 have been shown to be surface-expressed and create Ca\(^{2+}\)-activated Cl\(^{-}\) currents (44, 45). Studies in Ano1\(^{-/-}\) mice indicate that Ano1-mediated Cl\(^{-}\) secretion is necessary for normal airway surface liquid homeostasis (46), implying a similar functional role to CFTR in the airway. Future studies will be required to determine whether this potential functional association between CLCA and anoctamin family members is based on direct physical association.

The identification of self-processing of CLCA1 as a requirement for modulation of CaCC activity has significant implications for airway disease. In CF, CLCA1 activity appears to produce improvement of the CF phenotype by activation of compensating Cl\(^{-}\) channels (17, 47). In asthma and chronic obstructive pulmonary disease, increased CLCA1 activity is linked to increased mucus production (16), but any connection to CaCC activity still needs to be defined. Our study is the first to provide experimental approaches aimed at dissecting these two seemingly conflicting functional roles of CLCAs by biochemically characterizing mutations that prevent CaCC activation and identifying the fragment responsible for activation. Our findings also raise the intriguing possibility of developing CLCA1 protease inhibitors as research tools and therapeutic agents. The observed differences in effectiveness among MMP-like inhibitors at blocking CLCA1 metalloprotease activity provides useful insights for the design of potent and specific CLCA1 inhibitors. The main structural difference between these inhibitors is side chain size for the predicted S2’ moiety (indole, isopropyl, phenyl, and isobutyl for GM-6001, Marinastat, Batimastat, and Zynepstat001, respectively) (supplemental Fig. S1). Our results suggest that a larger hydrophobic residue in the S2’ pocket produces a more effective inhibitor, possibly via strengthening noncovalent interactions. These findings may be useful in guiding the design of more potent and specific CLCA metalloprotease inhibitors to prevent excess CLCA1 activity found in airway disease and perhaps cancer as well.

A final implication of our findings derives from the differences in function between human CLCA proteins. Of the three CLCA proteins expressed at mucosal surfaces (CLCA1, CLCA2, and CLCA4), only CLCA1 appears to regulate mucin gene expression and consequent mucus production (16), although CLCA1 and CLCA2 can both regulate CaCC activity (9, 48) (CLCA4 has not been tested). A thorough mechanistic understanding of how CLCA2 (or CLCA4) is able to activate CaCCs without triggering mucus production could lead to the development of novel selective CF therapeutics that exploit these mechanisms.

Acknowledgments—We thank Jen Alexander-Brett and Christopher Nelson for valuable experimental suggestions and Gregory Goldberg for valuable suggestions on the manuscript.

REFERENCES

tional analyses of two new calcium-activated chloride channel family members from mouse eye and intestine. J. Biol. Chem. 279, 41792–41800
SUPPLEMENTARY MATERIAL

Self-cleavage of CLCA1 by a novel internal metalloprotease domain controls calcium-activated chloride channel activation


Figure S1. Chemical structures of the MMP inhibitors used in this study. Arrows denote presumed metalloprotease structural features targeted by particular moieties of the inhibitors.
Preparation of Zeynepstat001 (Compound 9): scheme

1. \( \text{AcCl} + \text{NH} \rightarrow \text{THF} \) at -20 °C
2. \( n-\text{BuLi} \) at -78 °C
3. \( \text{LiHMDS} \) at -78 °C
4. \( \text{THF} \) at -20 °C
5. \( \text{EDC/DMAP} \) and \( \text{MeNH}_2 \) HCl/dioxane
6. \( 10\% \text{Pd/C} \)
7. \( \text{H}_2, \text{EtOAc} \)
8. \( \text{EDC/HOBt/DIEA} \)
9. \( \text{BOC-Leu-OH} \)
Synthesis:

(4S)-4-Benzyl-3-propanoyloxazolidin-2-one (1)

(S)-Benzyl-2-oxazolidinone (5.12 g, 28.8 mmol) was dissolved in THF (35 mL) under a nitrogen atmosphere and the reaction was cooled to -20 °C. A solution of n-buLi (1.6 M in hexane) (18 mL, 28.8 mmol) was added and the mixture was stirred for 1 hr at -20 °C. Propionyl chloride (2.45 g, 26.5 mmol) was added and the reaction was stirred at room temperature for 20 minutes and quenched with saturated aqueous sodium bicarbonate. The THF was evaporated and the reaction was partitioned between ethyl acetate and aqueous sodium bicarbonate. The organic layer was evaporated to give 6.62 g of a clear oil. The material was chromatographed on 90 g of silica using a linear gradient from 5%-40% ethyl acetate/hexane to give 5.38 g of product as a clear oil (87% yield), which provided a single peak by LC-MS analysis (M+H = 234.1).

1H NMR (400 MHz, CDCl3) δ: 7.36-7.20 (m, 5H), 4.71-4.65 (m, 1H), 4.23-4.15 (m, 2H), 3.31 (dd, J = 13.6, 3.2 Hz, 1H), 3.05-2.88 (m, 2H), 2.77 (dd, J = 13.2, 9.2 Hz, 1H), 1.21 (t, J = 7.2 Hz, 3H).

Benzyl (3R)-4-(4S)-4-benzyl-2-oxo-oxazolidin-3-yl]-3-methyl-4-oxo-butanoate (2)

LiHMDS (1M in THF) (27.7 mL, 27.7 mmol) was added to a solution of 1 (5.38 g, 23.06 mmol) in THF (105 mL) at -78 °C and the mixture was stirred at 0 °C for 20 minutes. Benzyl bromoacetate (6.87 g, 30 mmol) in THF (70 mL) was added dropwise to this solution over 90 minutes. The reaction was stirred for 30 minutes at 0 °C and then room temperature over night. The reaction was quenched using saturated aqueous sodium bicarbonate, the THF was evaporated, and the reaction was partitioned between ethyl acetate and aqueous sodium bicarbonate. The organic layer was evaporated to give 10.57 g of an orange oil. The material was chromatographed on 90 g of silica using a linear gradient from 5%-40% ethyl acetate/hexane to give 5.33 g of product as a viscous yellow oil (60% yield), which provided a single peak by LC-MS analysis (M+H = 382.1).

1H NMR (400 MHz, CDCl3) δ: 7.36-7.20 (m, 10H), 5.11 (s, 2H), 4.66-4.60 (m, 1H), 4.27-4.07 (m, 3H), 3.23 (dd, J = 13.2, 2.8 Hz, 1H), 3.02 (dd, J = 13.2, 9.2 Hz, 1H), 2.56-2.49 (m, 2H), 1.23 (d, J = 6.8 Hz, 3H).

(2R)-4-benzyloxy-2-methyl-4-oxo-butanoic acid (3)

Solutions of 30% hydrogen peroxide (8.5 mL) and lithium hydroxide-hydrate (0.82 g) in water (25 mL) were added to a solution of 2 (5.33 g, 13.97 mmol) in THF/water (100 mL/25 mL). The mixture was stirred for 2 h at 0°C and 2 h at room temperature, then acidified with 2 N aqueous hydrochloric acid. The THF was evaporated and the aqueous layer was extracted using ethyl acetate and concentrated to give 5.48 g of a clear oil. The material was chromatographed on 90 g of silica using a linear gradient from 0%-60% ethyl acetate/hexane to give 1.58 g of product as a clear oil (51% yield), which provided a single peak by LC-MS analysis (M+H = 223.1).

1H NMR (400 MHz, CDCl3) δ: 7.38-7.31 (m, 5H), 5.14 (s, 2H), 3.02-2.94 (m, 1H), 2.80 (dd, J = 16.8, 8.0, 1H), 2.48 (dd, J = 16.6, 5.6 Hz, 1H), 1.26 (d, J = 6.8 Hz, 3H).

tert-Butyl N-[(1S)-3-methyl-1-(methylcarbamoyl)butyl]carbamate (4)

BOC-Leu-OH (5 g, 21.6 mmol) was dissolved in dichoromethane (200 mL). EDC HCl (4.97 g, 25.9 mmol) and DMAP (cat.) were added to the solution. Methylamine (2 M in
THF) (13 mL, 26 mmol) was then added and the mixture was stirred for 2 h under N₂. The reaction mixture was washed with 1N aqueous hydrochloric acid (2 x 100 mL), saturated sodium bicarbonate (2 x 100 mL) and saturated brine (100 mL). The organic layer was dried over anhydrous sodium sulfate and concentrated to give 4.06 g of product as a white solid (77% yield). This material does not fly in the LC-MS. ¹H NMR (400 MHz, CDCl₃) δ: 6.14 (br s, 1H), 4.86 (br s, 1H), 4.07 (br s, 1H), 2.80 (s, 3H), 1.71-1.59 (m, 3H), 1.41 (s, 9H), 0.94 (s, 6H).

(2S)-2-amino-N,4-dimethyl-pentanamide (5)
Compound 4 (1 g, 4.09 mmol) was stirred in 4 M hydrochloric acid/dioxane (2 mL, 8 mmol). A small quantity of methanol was added to increase solubility. After 2-1/2 h the reaction was evaporated and the resulting solid was triturated in diethyl ether and filtered to give 0.68 g of product as the hydrochloride salt. LC-MS gave a single peak (M+H = 145.1). ¹H NMR (400 MHz, DMSO) δ: 8.52 (br s, 1H), 8.2 (br s, 2H), 3.64 (br s, 1H), 3.52 (s, 1H), 2.61 (d, J = 4 Hz, 3H), 1.58-1.48 (m, 3H), 0.85 (d, J = 6.4 Hz, 6H).

Benzyl (3R)-3-methyl-4-[(1S)-3-methyl-1-(methylcarbamoyl)butyl]amino]-4-oxo-butanoate (6)
Compound 5 (1.26 g, 6.97 mmol), HOBt (containing 20% water) (1.08 g, 6.38 mmol), EDC·HCl (1.45 g, 7.56 mmol) and DIEA (0.825 g, 6.38 mmol) were added to a solution of compound 3 (1.29 g, 5.80 mmol) in dichloromethane (100 mL) and stirred for 2 h at room temperature. The reaction was quenched using excess 0.5 M aqueous hydrochloric acid, layers were separated and the organic layer dried over sodium sulfate after which it was filtered and evaporated to give 2.06 g of white solid. The material was chromatographed on 24 g of silica using a linear gradient from 20% ethyl acetate/hexane to 100% ethyl acetate to give 1.37 g of product as a white solid (68% yield), which provided a single peak by LC-MS analysis (M+H = 349.1). ¹H NMR (400 MHz, CDCl₃) δ: 7.38-7.28 (m, 5H), 6.54 (br s, 1H), 6.6.35 (br s, 1H), 5.11 (s, 2H), 4.45-4.39 (m, 1H), 2.79-2.70 (m, 5H), 2.53-2.46 (m, 1H), 1.72-1.52 (m, 3H), 1.26-1.18 (m, 3H), 0.97-0.89 (m, 6H).

(3R)-3-methyl-4-[(1S)-3-methyl-1-(methylcarbamoyl)butyl]amino]-4-oxo-butanoic acid (7)
Compound 6 (1.54 g, 4.42 mmol) was dissolved in ethyl acetate (65 mL), 10% Pd/C (0.200 g) was added and the mixture was stirred under hydrogen (initial pressure 50 p.s.i.) for 1 h. The mixture was filtered through diatomaceous earth and evaporated to give 1.16 g of product as a white foam (100%). LC-MS gave a single peak (M+H = 259.1). ¹H NMR (400 MHz, CDCl₃) δ: 7.38-7.28 (m, 5H), 6.54 (br s, 1H), 6.6.35 (br s, 1H), 5.11 (s, 2H), 4.52 (d, J = 8.0 Hz, 2H), 2.84-2.70 (m, 5H), 2.47 (dd, J = 16.4, 4.4 Hz, 1H), 1.66-1.55 (m, 3H), 1.25-1.18 (m, 3H), 0.93-0.89 m, 6H).

(2R)-N'-benzyloxy-2-methyl-N-[(1S)-3-methyl-1-(methylcarbamoyl)butyl]butanediamide (8)
O-Benzylhydroxylamine hydrochloride (0.84 g, 5.26 mmol), HOBt (containing 20% water) (0.82 g, 4.88 mmol), EDC·HCl (1.1 g, 5.74 mmol) and DIEA (0.63 g, 4.88 mmol) were added to a solution of compound 7 (1.14 g, 4.41 mmol) in dichloromethane (80 mL)
and stirred for 2 h at room temperature. The reaction was quenched using excess 0.5 M aqueous hydrochloric acid and the product precipitated. The organic and aqueous layers were filtered to give 1.16 g of product as a white solid. The solids were triturated in dichloromethane (40 mL) and filtered to give 0.803 g of product as a white solid (50% yield), which provided a single peak by LC-MS analysis (M+H = 364.2). 1H NMR (400 MHz, DMSO) δ: 10.97 (s, 1H), 7.85 (d, J = 8 Hz, 1H), 7.72 (d, J = 4 Hz, 1H), 7.33 (m, 5H), 4.72 (s, 2H), 4.18-4.12 (m, 1H), 2.74-2.68 (m, 1H), 2.52 (m, 3H), 2.21-2.15 (m, 1H), 1.96-1.90 (m, 1H), 1.52-1.49 (m, 1H), 1.45-1.32 (m, 2H), 0.98-0.92 (m, 3H), 0.82-0.75 (m, 6H).

(2S)-2-[(2R)-4-(hydroxyamino)-2-methyl-4-oxo-butanoylamino]-N,4-dimethylpentanamide (9) [Zeynepstat001]

Compound 8 (0.768 g, 2.11 mmol) was dissolved in methanol (80 mL), 10% Pd/C (0.08 g) was added and the mixture was stirred under hydrogen (initial pressure 56 psi) for 45 min. The mixture was filtered through diatomaceous earth and evaporated to give 0.57 g of product as a light tan solid. The material was triturated in dichloromethane (40 mL), filtered and dried in a vacuum oven to give 0.513 g of product as a tan solid. This material was re-crystallized from ethanol/ether (10 mL/10 mL) to give 0.264 g of product as an off-white solid. LC-MS gave a single peak (M+H = 274.1). 1H NMR (400 MHz, DMSO) δ: 10.34 (s, 1H), 8.68 (s, 1H), 7.84 (d, J = 8.4 Hz, 1H), 7.75 (d, J = 4.4 Hz, 1H), 4.16-4.10 (m, 1H), 2.70-2.64 (m, 1H), 2.52 (m, 3H), 2.20-2.15 (m, 1H), 1.96-1.90 (m, 1H), 1.53-1.48 (m, 1H), 1.42-1.37 (m, 2H), 0.94-0.92 (m, 3H), 0.83-0.76 (m, 6H).

References:
Secreted CLCA1 modulates TMEM16A to activate Ca\(^{2+}\)-dependent chloride currents in human cells

Monica Sala-Rabanal\(^{1,2}\), Zeynep Yurtsever\(^{2,3,4,5}\), Colin G Nichols\(^{1,2}\), Tom J Brett\(^{1,2,4,5,6}\)

\(^{1}\)Department of Cell Biology and Physiology, Washington University School of Medicine, St Louis, United States; \(^{2}\)Center for the Investigation of Membrane Excitability Diseases, Washington University School of Medicine, St Louis, United States; \(^{3}\)Biochemistry Program, Washington University School of Medicine, St Louis, United States; \(^{4}\)Department of Internal Medicine, Washington University School of Medicine, St Louis, United States; \(^{5}\)Drug Discovery Program in Pulmonary and Critical Care Medicine, Washington University School of Medicine, St Louis, United States; \(^{6}\)Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St Louis, United States

Abstract
Calcium-activated chloride channel regulator 1 (CLCA1) activates calcium-dependent chloride currents; neither the target, nor mechanism, is known. We demonstrate that secreted CLCA1 activates calcium-dependent chloride currents in HEK293T cells in a paracrine fashion, and endogenous TMEM16A/Anoctamin1 conducts the currents. Exposure to exogenous CLCA1 increases cell surface levels of TMEM16A and cellular binding experiments indicate CLCA1 engages TMEM16A on the surface of these cells. Altogether, our data suggest that CLCA1 stabilizes TMEM16A on the cell surface, thus increasing surface expression, which results in increased calcium-dependent chloride currents. Our results identify the first Cl\(^{-}\) channel target of the CLCA family of proteins and establish CLCA1 as the first secreted direct modifier of TMEM16A activity, delineating a unique mechanism to increase currents. These results suggest cooperative roles for CLCA and TMEM16 proteins in influencing the physiology of multiple tissues, and the pathology of multiple diseases, including asthma, COPD, cystic fibrosis, and certain cancers.

DOI: 10.7554/eLife.05875.001

Introduction
The calcium-activated chloride channel regulator (CLCA- previously known as chloride channel calcium activated) proteins (Cunningham et al., 1995) are a family of secreted self-cleaving metalloproteases that activate calcium-dependent chloride currents (I\(_{\text{CaCC}}\)) in mammalian cells (Yurtsever et al., 2012). CLCA family members are highly expressed in mucosal epithelia where they play important roles in mucus homeostasis and related diseases (Patel et al., 2009). For example, human CLCA1 plays a central role in interleukin (IL-) 13-induced mucus cell metaplasia, the main source of inflammatory mucus overproduction in chronic obstructive airway diseases, such as asthma and COPD (Alevy et al., 2012). Both clinical and animal model studies suggest a compensatory role for CLCAs in the context of cystic fibrosis (CF): the fatal intestinal disease, meconium ileus, arising in CFTR-deficient mice is corrected by overexpression of mCLCA3 (an orthologue of human CLCA1) (Young et al., 2007) and, correspondingly, mutations in CLCA1 are found in a subset of CF patients with aggravated intestinal disease (van der Doef et al., 2010). At the cellular level, overexpression of
CLCA proteins leads to activation of calcium-dependent chloride currents (Gandhi et al., 1998; Britton et al., 2002; Elble et al., 2002; Greenwood et al., 2002), and this functional observation had caused CLCAs to be initially misidentified as calcium-activated chloride channels (CaCCs) themselves (Cunningham et al., 1995). However, further bioinformatic and biochemical studies have demonstrated that CLCA proteins are secreted, soluble proteins and that they act to modulate CaCCs that are endogenous to mammalian cells (Gibson et al., 2005; Hamann et al., 2009; Yurtsever et al., 2012). The molecular identity of these channels, the mechanism of CLCA activation, and their potential roles in CLCA-mediated diseases, remain unknown.

TMEM16A (also known as Anoctamin1/DOG1) was recently identified as the first genuine CaCC in mammals by three independent groups (Caputo et al., 2008; Schroeder et al., 2008; Yang et al., 2008). 10 members of the TMEM16/Anoctamin family have been identified (TMEM16A-K, or Ano1-10); these proteins, predicted to be transmembrane proteins with eight membrane-spanning helices, have been found to function predominantly as CaCCs that are endogenous to mammalian cells (Gibson et al., 2005; Hamann et al., 2009; Yurtsever et al., 2012). The molecular identity of these channels, the mechanism of CLCA activation, and their potential roles in CLCA-mediated diseases, remain unknown.

TMEM16A was discovered several years ago, but remains the only calcium-dependent chloride channel known in mammals. Sala-Rabanal, Yurtsever et al. showed that adding CLCA1 to cells caused more TMEM16A channels to appear in the cell surface membrane and thereby increased the flow of chloride ions. The CLCA protein also physically interacted with the chloride channel in the membrane to stabilize it; no other protein has been shown to regulate ion channels in this way before.

The findings of Sala-Rabanal, Yurtsever et al. provide a much clearer understanding of how the CLCA protein and the chloride channel work. Both of these proteins are known to contribute to excess mucus production in airway diseases; and both have been linked to cardiovascular diseases and certain cancers. These new findings may therefore also help researchers to target these proteins and develop treatments for these diseases.

DOI: 10.7554/eLife.05875.002
secreted regulators of TMEM16A activity have been identified as of yet. Here we report that secreted CLCA1 modulates TMEM16A-dependent calcium-activated chloride currents, and that this activation can occur in a paracrine fashion. Furthermore, we show that CLCA1 and TMEM16A co-localize and physically interact on the surface of mammalian cells, and that CLCA1 increases the level of TMEM16A protein at the cell surface, representing a novel mechanism of channel regulation by a secreted protein. We thus demonstrate a first downstream target of CLCA proteins and provide the first example of a secreted protein modulator of TMEM16A activity. These findings have significant implications for the roles of CLCA1 and TMEM16A proteins as cooperative partners, not only in the physiology and pathophysiology of the airways, but also in those of other tissues and organs.

Results

Secreted CLCA1 can activate Ca^{2+}-dependent chloride currents in a paracrine fashion

We previously demonstrated that I_{CaCC} are activated in HEK293T (293T) cells overexpressing human CLCA1 (Yurtsever et al., 2012). Given that CLCA1 proteins are cleaved and secreted from these cells, we hypothesized that exogenous CLCA1 may activate I_{CaCC}. In a first set of experiments to test this idea, GFP-expressing cells that had been co-cultured overnight with cells transfected with CLCA1-pHLsec plasmid (CLCA1) or with empty pHlsec vector (pHLsec) were tested for I_{CaCC} by means of whole-cell patch clamp electrophysiology (Figure 1A). In the presence of 10 μM intracellular Ca^{2+} and physiological concentrations of extracellular Cl^−, robust, slightly outward rectifying currents were activated in cells co-cultured with CLCA1-transfected cells, but only substantially smaller currents were detected in cells co-cultured with vector-transfected cells (Figure 1B–D). In a complementary experiment, whole-cell I_{CaCC} were measured in untransfected cells that had been cultured in medium obtained from CLCA1- or from pHlsec-transfected cells (Figure 2A). We observed activation of large currents in cells exposed to CLCA1-conditioned medium that had the same Ca^{2+}- and voltage-dependence properties as those induced in cells co-cultured with CLCA1-expressing cells (Figure 2B–D). As shown in Figure 2B–C, outward rectification of the CLCA1-activated current decreases at higher Ca^{2+} concentrations. In addition, current reversal potential shifts positive upon lowering extracellular Cl^−. These features are in agreement with the properties of the Ca^{2+}-dependent Cl^− conductance observed in CLCA1-expressing 293T cells (Hamann et al., 2009; Yurtsever et al., 2012), and consistent with those of CaCCs in native cells and heterologous expression systems (Jeong et al., 2005; Yamazaki et al., 2005; Xiao et al., 2011). These data indicate that secreted CLCA1 can activate I_{CaCC} in a paracrine fashion.

CLCA1-dependent I_{CaCC} are carried by TMEM16A

We next focused on identifying the CaCC responsible for carrying the CLCA1-mediated currents. The CLCA1-modulated I_{CaCC} in 293T cells are Ca^{2+}-dependent, moderately outward rectifying in the presence of μM concentrations of intracellular Ca^{2+}, Cl^−-selective, and blocked by gluconate (Hamann et al., 2009; Yurtsever et al., 2012) (Figures 1, 2), closely resembling the biophysical characteristics of those observed for TMEM16A currents in heterologous expression systems (Schoeder et al., 2008; Yang et al., 2008; Xiao et al., 2011), proteoliposomes (Terashima et al., 2013), and native tissues (Caputo et al., 2008).

Given the biophysical and pathophysiological parallels between TMEM16A currents, and those activated by CLCA1, we hypothesized that CLCA1-activated currents may be carried by TMEM16A. Consistent with this idea, 293T cells were transfected with either TMEM16A siRNA or with non-specific, scrambled RNA (siControl), and cultured in CLCA1-conditioned medium. Exposure to secreted CLCA1 led to the activation of I_{CaCC} in siControl-transfected cells (Figure 3A,B) that were comparable to the activation recorded in untransfected cells (Figure 2B,D), but these CLCA1-dependent currents were knocked down to essentially background levels in TMEM16A siRNA-transfected cells (Figure 3A,B). The TMEM16A siRNA significantly decreased expression of TMEM16A protein, assessed by Western blot (Figure 3C). These results demonstrate that CLCA1-dependent I_{CaCC} in 293T cells are indeed carried by TMEM16A.
CLCA1 colocalizes with and increases cell surface levels of TMEM16A protein

Next, we used immunohistochemistry and confocal microscopy to examine CLCA1 and TMEM16A localization in non-permeabilized HEK293T cells. Cells transfected with pHLSec vector alone did not display noticeable staining for either CLCA1 or TMEM16A (Figure 4A–D), consistent with lack of endogenous expression of CLCA1 and low endogenous levels of TMEM16A in these cells (Kunzelmann et al., 2009; Pritchard et al., 2014). However, cells transfected with CLCA1 stained strongly both for CLCA1 and, surprisingly, for TMEM16A (Figure 4E–H), suggesting that CLCA1 increases TMEM16A protein levels. Furthermore, signal for both proteins clearly overlapped with
**Figure 2.** Activation of calcium-dependent chloride currents by secreted CLCA1. (A) Untransfected cells were cultured in medium from pHLsec- or CLCA1-expressing cells, and assayed by patch clamp electrophysiology. (B–D) Whole-cell currents measured in cells from experiments as in (A), superfused with standard ([154 mM Cl]_out) or reduced Cl^- ([14 mM Cl]_out) extracellular solution; and in the absence or presence of 1 μM or 10 μM free Ca^{2+} in the pipette (respectively, [0 μM Ca^{2+}]_in, [1 μM Ca^{2+}]_in, or [10 μM Ca^{2+}]_in). (B) Representative current traces obtained with the same pulse protocol and displayed as in Figure 1B. Membrane capacitance was similar in all cases at ~25 pF. (C) Current-voltage relationships at the end of the 600-ms voltage steps. Membrane potential values were corrected off-line for the calculated liquid junction potentials, respectively -~5.5 mV ([0 μM Ca^{2+}]_in) and -~6.0 mV ([1 μM Ca^{2+}]_in) or [10 μM Ca^{2+}]_in for the experiments in [154 mM Cl^-]_out; and -~20 mV for the experiments in [14 mM Cl^-]_out. Data are presented as means ± S.E. (n = 5–20). Inset, CLCA1-mediated currents right-shifted ~ +15 mV upon reduction of extracellular Cl^-; symbols have been removed for clarity. (D) Current density at +100 mV, from the same experiments as in (C). Symbols represent data from individual patches; bars indicate the means ± S.E. of all experiments. *p < 0.01 (one-way ANOVA, F = 10.4 and p = 2.1 x 10^{-8}; followed by Tukey test). DOI: 10.7554/eLife.05875.004
the membrane stain (WGA), consistent with a model in which CLCA1 and TMEM16A associate with and stabilize one another on the cell surface. Since secreted CLCA1 can activate TMEM16A-mediated I_{CaCC} in a paracrine manner (Figures 1–3), we carried out similar imaging experiments to determine whether exogenously applied secreted CLCA1 also increased TMEM16A surface expression. Cells cultured in media from pHLsec-transfected cells again displayed no detectable staining for either CLCA1 or TMEM16A (Figure 4I–L), but cells exposed to secreted CLCA1 displayed robust staining for TMEM16A. Signal for CLCA1 was also detected in a few cells, overlapping with TMEM16A and WGA staining (Figure 4M–P). Surprisingly, although TMEM16A surface levels increased after exposure to CLCA1, total TMEM16A in cells did not change (Figure 3C). These results indicate that exogenous secreted CLCA1 colocalizes with and enhances the fraction of TMEM16A located at the cell surface.

CLCA1 associates with cell surface TMEM16A

To investigate whether CLCA1 and TMEM16A associate directly with one another on the cell surface, we adapted an assay commonly used to identify immunological receptor-ligand pairs (Altman et al., 1996). We previously demonstrated that CLCA1 is cut into two fragments by self-cleavage and that the N-terminal fragment is necessary and sufficient to activate CaCCs in HEK293T cells (Yurtsever et al., 2012). Thus, for these assays we developed a CLCA1 cell-staining reagent composed of the N-terminal fragment of CLCA1 (N-CLCA1) containing a specific biotinylation motif on the C-terminus (Figure 5A). Biotinylated N-CLCA1 was coupled to SA-PE (streptavidin conjugated to phycoerythrin) to produce a tetrameric fluorescent reagent with enhanced avidity toward its ligand. Cell-binding assays were carried out in the presence or absence of an anti-TMEM16A antibody raised against epitopes in the last extracellular loop and then

Figure 3. Genetic knockdown of TMEM16A inhibits CLCA1-mediated calcium-dependent chloride currents. (A–B) HEK293T cells transfected with RNAi negative control (siControl) or TMEM16A siRNA were incubated in CLCA1-conditioned medium and assayed by patch-clamp electrophysiology, in standard extracellular solution ([154 mM Cl^{-}] _{out} ) and 10 μM free Ca^{2+} in the pipette ([10 μM Ca^{2+}] _{in} ). (A) Representative current traces obtained with the same pulse protocol and displayed as in Figure 1B. Membrane capacitance was similar in all cases at ~25 pF. (B) Current density at +100 mV. Symbols represent data from individual patches (n = 14); bars indicate the means ± S.E. of all experiments. *p < 0.01 (unpaired Student’s t test). (C) Effect of CLCA1 and/or TMEM16A siRNA treatment on TMEM16A protein expression. Upper panel: top, TMEM16A; and bottom, actin (loading control) Western blot from solubilized HEK293T cells. Lanes are labeled as follows: pHLsec-t, pHLsec transfected cells; CLCA1-t, CLCA1-transfected cells; CLCA1-c, cells treated with CLCA1-conditioned medium; TMEM16A siRNA, cells transfected with TMEM16A siRNA; TMEM16A siRNA + CLCA1, cells transfected with TMEM16A siRNA and treated with CLCA1-conditioned medium. Bar graph: quantitation of TMEM16A band intensity normalized to actin band intensity using ImageJ (NIH).

DOI: 10.7554/eLife.05875.005
analyzed by flow cytometry. The tetramerized N-CLCA1 displayed robust binding to intact HEK293T cells compared to background, and this binding was significantly reduced by pre-incubating the cells with the anti-TMEM16A antibody (Figure 5B). Two control antibodies, one raised against an intracellular epitope of TMEM16A and the other an isotype control, did not affect N-CLCA1 binding (Figure 5C). In order to validate that the biotinylation of N-CLCA1 did not adversely affect function, we carried out whole-cell patch clamp experiments where either purified N-CLCA1 or purified biotinylated N-CLCA1 was exogenously applied to HEK293T cells. We found that both of these proteins were able to robustly activate the observed currents (Figure 5D,E). These results indicate that N-CLCA1 engages TMEM16A on the surface of HEK293T cells, and suggests that the enhanced TMEM16A level is a consequence of stabilization by CLCA1.

Figure 4. CLCA1 colocalizes with TMEM16A and increases TMEM16A surface expression. (A–D) Membrane (WGA) or immunostaining of HEK293T cells transfected with pHLsec vector; (E–H), or with CLCA1. Surface TMEM16A is greatly increased by expression of CLCA1. (I–L) Membrane (WGA) or immunostaining of HEK293T cells cultured in conditioned media from cells transfected with pHLsec vector; (M–P) or cells cultured in conditioned media from cells transfected with CLCA1. TMEM16A surface expression is greatly enhanced after cells are exposed to secreted CLCA1.

DOI: 10.7554/eLife.05875.006
Discussion

It has been demonstrated that purified TMEM16A protein reconstituted in proteoliposomes recapitulates the permeation, pharmacological, voltage- and Ca\(^{2+}\)-dependence properties of TMEM16A channels characterized in heterologous expression systems and native cell models (Terashima et al., 2013), which implies that TMEM16A does not require association with other proteins for CaCC activity. However, a recent proteomics approach has identified a large number of endogenous proteins implicated in protein trafficking, surface expression, folding and stability that interact with TMEM16A, including SNAREs such as syntaxins and syntaxin-binding proteins, and the ezrin-radixin-moesin (ERM) scaffolding complex (Perez-Cornejo et al., 2012). Our data identify CLCA1 as the first secreted direct regulator of TMEM16A, and our findings suggest that CLCA1 may modulate TMEM16A channel activity by stabilizing it at the cell surface, much like the SNARE and ERM protein networks of the TMEM16A interactome.

So how does CLCA1 modulate TMEM16A currents? We observe that CLCA1 increases TMEM16A surface expression without increasing expression of the protein (Figure 3C). A model consistent with our data and the current literature is that CLCA1 engages and stabilizes dimeric TMEM16A on the surface of the cell. Previous studies have shown that TMEM16A can exist as a dimer (Fallah et al., 2011; Sheridan et al., 2011), dimerization being mediated by an

Figure 5. N-CLCA1 engages TMEM16A on the cell surface. (A) Schematic of CLCA1 N-terminal fragment (N-CLCA1) construct with specific biotinylation site and resultant tetrameric cell-staining reagent created after complexation with SA-PE. (B) Flow cytometry of intact HEK293T cells stained with SA-PE alone (black line), N-CLCA1/SA-PE (green line), or N-CLCA1/SA-PE in the presence of anti-TMEM16A antibody S-20 (red line). (C) Flow cytometry of intact HEK293T cells either stained with SA-PE alone (black line), N-CLCA1/SA-PE (green line), N-CLCA1/SA-PE in the presence of anti-TMEM16A antibody C-5 (raised against an intracellular TMEM16A epitope; orange line), or N-CLCA1/SA-PE in the presence of anti-Aquaporin5 antibody G-19 (blue line). (D–E) Cells were incubated in the absence ([-]) or presence ([+]) of purified N-terminal (N-term) CLCA1 protein before (N-CLCA1) or after biotinylation (N-CLCA1biotin), and assayed by patch-clamp electrophysiology, in standard extracellular solution ([154 mM Cl\(^{-}\)]_out and 10 \(\mu\)M free Ca\(^{2+}\) in the pipette ([10 \(\mu\)M Ca\(^{2+}\)]_in). (D) Representative current traces obtained with the same pulse protocol and displayed as in Figure 1B. Membrane capacitance was similar in all cases at \(\sim\)25 pF. (E) Current density at +100 mV. Symbols represent data from individual patches (n = 8–11); bars indicate the means \(\pm\) S.E. of all experiments. *p < 0.05 (unpaired Student’s t test).

DOI: 10.7554/eLife.05875.007
intracellular region in the N-terminus of TMEM16A (Tien et al., 2013). Mutations to this region abolish dimerization, prevent protein trafficking to the plasma membrane, and, consequently, ablate channel activity (Tien et al., 2013). These observations indicate that dimerization regulates TMEM16A trafficking and channel activity. It is possible that TMEM16A dynamically shuttles between the cell surface and intracellular compartments. One possibility is that CLCA1 engages monomeric TMEM16A and drives dimerization; alternatively, CLCA1 may engage and stabilize dimeric TMEM16A on the cell surface, thereby preventing its removal and, consequently, increasing calcium-dependent chloride currents (Figure 6). This highlights an unprecedented mechanism for regulating ion channel currents by a secreted protein.

The identification of CLCA1 as a modulator of TMEM16A activity raises the possibility of functional associations between other CLCA and TMEM16 family members. Four to eight CLCA (Patel et al., 2009) and ten TMEM16 family members (Pedemonte and Galietta, 2014) are expressed in mammalian tissues. A number of these TMEM16 proteins have poorly defined functions and do not obviously traffic to the cell surface when expressed alone (Duran et al., 2012). Future studies will be needed to determine whether other CLCA proteins can associate with other TMEM16 proteins and influence their function. CLCA1 (Yang et al., 2013), CLCA2 (Sasaki et al., 2012; Walia et al., 2012), and CLCA4 (Yu et al., 2013) have all been implicated in various cancers as have a number of TMEM16 proteins (West et al., 2004; Dutertre et al., 2010; Duvvuri et al., 2012; Liu et al., 2012; Qu et al., 2014), and such studies could have tremendous implications for cooperative CLCA/TMEM16 roles in cancer and other diseases.

Here we report that secreted CLCA1 modulates TMEM16A-dependent I_{CaCC}, and that this activation can occur in a paracrine fashion. Furthermore, we show that CLCA1 and TMEM16A colocalize and physically interact on the surface of mammalian cells, with CLCA1 increasing the level of TMEM16A protein at the cell surface. We thus demonstrate a first downstream target of CLCA proteins, solving the 20-year-old mystery regarding how CLCA proteins activate I_{CaCC}, and provide the first example of a secreted protein modulator of TMEM16A activity. CLCA1 (Alevy et al., 2012) and TMEM16A (Huang et al., 2012; Scudieri et al., 2012) have been separately observed to play critical roles in chronic inflammatory airway disease models. Our findings have significant implications for the roles of CLCA1 and TMEM16A proteins as cooperative partners, not only in the physiology and pathophysiology of the airways, but also in those of other tissues and organs.

Figure 6. Model for CLCA1 modulation of TMEM16A-mediated calcium-dependent chloride currents. Following secretion and self-cleavage of CLCA1, the N-terminal fragment (N-CLCA1) acts in paracrine fashion (1). Dimerization appears to regulate surface trafficking of TMEM16A. N-CLCA1 engages TMEM16A on the cell surface (2), stabilizing TMEM16A dimers, preventing internalization (3) and in turn, results in increased TMEM16A surface expression and calcium-dependent chloride current density.

DOI: 10.7554/eLife.05875.008
Materials and methods

Reagents

The following commercial antibodies were used according to the manufacturer’s specifications:
mouse-anti-human-TMEM16A monoclonal antibody C-5 (Santa Cruz Biotechnology, Dallas, TX);
goat-anti-human-TMEM16A polyclonal antibody S-20 (Santa Cruz Biotechnology); mouse-anti-actin monoclonal antibody C4 (Millipore, Billerica, MA); goat-anti-human-Aquaporin5 polyclonal antibody G-19 (Santa Cruz Biotechnology); rabbit anti-6-His-antibody-HRP conjugate (Bethyl Laboratories, Montgomery, TX), goat anti-mouse IgG antibody-HRP conjugate (Santa Cruz Biotechnology); wheat germ agglutinin (WGA)-Alexa Fluor 633 conjugate (Life Technologies, Carlsbad, CA); donkey anti-goat IgG-Alexa Fluor 488 conjugate (Life Technologies); donkey anti-rabbit IgG-Alexa Fluor 594 conjugate (Life Technologies); and rabbit anti-human CLCA1 polyclonal antibody 1228 (Biosystems, Rockford, IL). Mouse anti-human CLCA1 monoclonal antibody 8D3 was produced in-house and used as previously described (Alevy et al., 2012; Yurtsever et al., 2012). Streptavidin conjugated to phycoerythrin (SA-PE) was purchased from BD Biosciences (San Jose, CA). Hype-5 transfection reagent was purchased from OZ Biosciences (San Diego, CA).

Heterologous expression of CLCA1

Full length human CLCA1 (22–914) (CLCA1) cloned into pHLsec vector was used throughout (Yurtsever et al., 2012). HEK293T cells were cultured in 6-well dishes in Dulbecco’s Modified Eagle Medium (Life Technologies) supplemented with 10% fetal bovine serum, 10^5 units/l penicillin and 100 mg/l streptomycin, at 37°C and 5% CO2. Cells were transfected at 80% confluence using 293fectin transfection reagent (Life Technologies) at a 1:2 ratio (μg DNA: μl 293fectin), using 1 μg of plasmid DNA per 1 million cells. Experiments were conducted in cells that were transiently transfected with CLCA1, or in cells that were exposed to exogenous CLCA1 protein via means of two different experimental approaches: either co-culture with CLCA1 transfected cells; or treatment with CLCA1-conditioned medium. For co-culture experiments, cells transfected with CLCA1, empty pHLsec vector (pHLsec), or EGFP-pCDNA3.1 plasmid (GFP) were trypsinized 24 hr post-transfection, and GFP-expressing cells were mixed at a 1:1 ratio with either CLCA1 or pHLsec-transfected cells, and replated at low density on UV-sterilized, 8 mm round German glass coverslips (Electron Microscopy Sciences, Hatfield, PA). Following trypsin treatment, all transfected cells were pelleted by centrifugation and washed with sterile PBS prior to replating to prevent carry-over of transfection medium. After 24 hr, the GFP-expressing cells were assayed for Ca^{2+}-dependent Cl− currents by patch clamp electrophysiology. For conditioned medium experiments, cells were transfected with either CLCA1 or pHLsec for 6 hr, then transfection medium was removed, cells were washed with sterile PBS, and fresh medium was applied; following overnight incubation, medium from these cells was harvested and centrifuged gently (1500×g, 5 min) to remove non-adherent cells. Untransfected cells were plated onto round coverslips and incubated for 24 hr in 2 ml of cleared CLCA1- or pHLsec-conditioned medium supernatants.

Recombinant expression of CLCA1 and in vitro biotinylation

The N-terminal fragment of CLCA1 (22–694; N-CLCA1) was cloned into pHL-Avitag3 vector (Aricescu et al., 2006), which contains a BirA biotin ligase recognition motif and hexahistidine tag at the C-terminus. This secreted protein was expressed in 293F cells via transient transfection using Hype-5 at 1:1.5 μl ratio (μg DNA: μl Hype-5), using 1 μg of plasmid DNA per 1 million cells. Media supernatants were harvested after 72 hr. Protein was purified from media supernatant using Ni-NTA chromatography and eluted in 5 ml buffer A (50 mM K2HPO4 pH 8, 300 mM NaCl and 250 mM imidazole). Purified N-CLCA1 was concentrated to a final volume of 300 μl in a centrifuge concentrator and protein concentration was calculated from absorbance at 280 nm. For the experiments reported in Figure 5D,E, protein was added onto the untransfected cells at 10 μg/ml and incubated for 24 hr prior to whole-cell patch clamp experiments. The same volume of buffer A was added onto cells as buffer control. For in vitro biotinylation, N-CLCA1 containing the specific biotinylation tag at the C-terminus was exchanged into buffer B (100 mM Tris pH 7.5, 200 mM NaCl, and 5 mM MgCl2) and specifically biotinylated by addition of biotin and Escherichia coli BirA ligase (produced and purified in-house) at 4°C overnight. Excess biotin was removed using a 2 ml desalting column.
Biotinylated N-CLCA1 was added onto the untransfected cells at 10–50 μg/ml and incubated for 24 hr prior to whole-cell patch clamp experiments. The same volume of buffer B was added onto cells as buffer control.

**siRNA knockdown of TMEM16A**

To investigate the molecular identity of CLCA1-modulated CaCCs, a targeted approach was taken focusing on TMEM16A. For siRNA knockdown of TMEM16A, cells plated in 48-well plates were transfected with either 200 nM TMEM16A siRNA (HSS123904; 5′-AAG UUA GUG AGG UAG GCU GGG AAC C-3′, Life Technologies) or 200 nM medium GC-content Stealth RNAi negative control (12935300; 5′-GGU UCC CAG CCU ACC UCA ACU U-3′, Life Technologies) using Lipofectamine 2000 (Life Technologies) at a 20:2 ratio (pmol siRNA: μl Lipofectamine 2000); 24 hr later, cells were plated onto round coverslips and incubated for an additional 24 hr in CLCA1- or pHLSec-conditioned medium as described above. TMEM16A knockdown was estimated at 60–70% as assayed by qPCR.

**Whole-cell patch clamp recordings**

Experiments were performed at 25°C, 24 hr after co-culture or incubation in conditioned medium. Micropipettes were prepared from non-heparinized hematocrit glass (Kimble-Chase, Vineland, NJ) on a horizontal puller (Sutter Instruments, Novato, CA), and filled to a typical electrode resistance of 2 megaohms with pipette solution containing 150 mM N-methyl-D-glucamine (NMDG) chloride, 10 mM Hepes, 2 mM MgCl₂, 8 mM HEDTA, and 5.8 mM CaCl₂ to attain 10 μM free Ca²⁺, as calculated by means of the CaBuf program (available through Katholieke Universiteit Leuven). Selected experiments were performed with a pipette solution containing (mM) 150 NMDG chloride, 10 Hepes and 2 MgCl₂, in the absence ([0 μM Ca²⁺]₀) or presence of 5 mM EGTA and 4 mM CaCl₂ to attain 1 μM free Ca²⁺ ([1 μM Ca²⁺]₀). The pH of all pipette solutions was adjusted to 7.1 with Tris. The bath solution was 10 mM Hepes, 1 mM CaCl₂ and 1 mM MgCl₂; plus 150 mM NaCl (standard extracellular), [154 mM Cl⁻]₀), or 140 mM Na-gluconate and 10 mM NaCl (reduced extracellular Cl⁻; [14 mM Cl⁻]₀), and adjusted to pH 7.4 with Tris. After formation of a gigahm seal and establishment of whole-cell configuration, cells were voltage-clamped at 0 mV. A pulse protocol was applied in which membrane potential was held at 0 mV for 50 ms and stepped to a test value for 600 ms before returning to the holding potential for an additional 400 ms. The test potential varied from −100 to +100 mV in 20 mV increments. Membrane capacitance was calculated from the integral of the current transient in response to 10 mV depolarizing pulses, and was monitored for stability throughout the experiment. Data were filtered at 2 kHz, and signals were digitized at 5 kHz with Liquid junction potentials were calculated using Clampex JPCalc software and command voltages were corrected a posteriori as specified in the figure legends. Results are presented as mean ± S.E., differences between two groups were assessed by unpaired Student’s t test with Welch’s correction, and differences between more than two groups were evaluated by one-way ANOVA and post-hoc Tukey test (Prism 5.0c, GraphPad Software, San Diego, CA).

**Immunohistochemistry**

For staining experiments, cells were either transfected or exposed to conditioned medium as described above. Following 24 hr incubation, cells were fixed on glass slides with 4% paraformaldehyde (PFA) in PBS for 5 min and washed twice with PBS. Cells were blocked for 1 hr at room temperature with 1% blocking solution in PBS (Life Technologies) and then incubated with primary antibodies (rabbit anti-human CLCA1 polyclonal antibody 1228 at 1:100 dilution and goat-anti-human-TMEM16A polyclonal antibody S-20 at 1:50 dilution) overnight at 4°C. Slides were washed and incubated with WGA-Alexa Flour 633 conjugate (5 μg/ml) for 30 min at room temperature, followed by secondary antibodies (donkey anti-rabbit IgG-Alexa Fluor 594 conjugate at 1:250 dilution and donkey anti-goat IgG-Alexa Fluor 488 conjugate at 1:200 dilution) for 2 hr at room temperature. Washed slides were then mounted in VECTASHIELD H-1200 Mounting Medium with DAPI (Vector Laboratories, Burlingame, CA). Confocal microscopy was carried out using a Zeiss LSM 510 META Confocal Laser Scanning Microscope.
Flow cytometry binding assay

Human CLCA1 was assayed for binding to cell-surface TMEM16A using a flow cytometry-based binding assay. Prior to staining, intact HEK293T cells were treated with human FcR blocking reagent at 1:100 dilution (Miltenyi Biotec, San Diego, CA) for 15 min. The biotinylated N-CLCA1 was pre-incubated with SA-PE at a 4:1 molar ratio for 15 min at room temperature to produce fluorescently labeled tetramers of N-CLCA1 (N-CLCA1/SA-PE). Cells (4 x 10^5 cells/sample) were either stained with SA-PE alone (1:50) or N-CLCA1/SA-PE (1:50) diluted in PBS containing 1% BSA (FACS buffer) at 4˚C. In order to validate specific binding of N-CLCA1/SA-PE to cell surface TMEM16A, goat-anti-human-TMEM16A polyclonal antibody S-20 (1:10), which was raised against a 15–20 amino acid peptide within residues 820–870 (corresponding to the last extracellular loop; UniProt Q5XXA6), was added prior to addition N-CLCA1/SA-PE. A goat polyclonal IgG antibody for human Aquaporin5 was used (1:10) as an isotype control for the blocking antibody S-20. Mouse-anti-human-TMEM16A monoclonal antibody C5, which binds a cytosolic epitope, was used as a second control antibody (1:10). Following staining, cells were washed with FACS buffer, and then analyzed by flow cytometry (BD FACScan). Data analysis was performed using FlowJo (Tree Star, Ashland, OR).

Western blotting and densitometric quantitation

HEK293T cells were pelleted, lysed in lysis buffer (1.5 mM KH₂PO₄, 2.7 mM KCl, 4.3 mM Na₂HPO₄, 137 mM NaCl and 1% Triton X-100 in deionized water), and then diluted 1:2 in 2x SDS containing 2-mercaptoethanol. Samples were boiled for 5 min, and then loaded on a 4–12% bis-tris Nupage gel (Life Technologies). The proteins were transferred to nitrocellulose membranes using an iBlot Gel Transfer Device (Life Technologies). Membranes were blocked by 0.5% nonfat milk in PBS with 0.1% TWEEN. Primary antibodies (mouse α-human CLCA1 8D3, 1:4000; mouse-anti-human-TMEM16A monoclonal antibody C-5, 1:1000; mouse-anti-actin monoclonal antibody C4, 1:5000) in blocking buffer were incubated on the membrane for 15 min. Following three washes with PBS-TWEEN, secondary antibodies (goat-anti-mouse IgG-HRP conjugate 1:5000) in blocking buffer were applied for 15 min. After three PBS-TWEEN washes, signal was detected using Pierce ECL Western Blotting Substrate (Thermo Fisher Scientific, Rockford, IL). Developed films were scanned and converted to 8-bit tiff files. Protein bands were processed equally and the pixel intensities were quantified with ImageJ 1.48 (http://imagej.nih.gov/ij).

Acknowledgements

This work was supported by NIH R01-HL119813 (to TJB) and R01-HL54171 (to CGN), American Lung Association RG-196051 (to TJB), a CIMED Pilot and Feasibility grant (to TJB), and American Heart Association Predoctoral Fellowship PRE19970008 (to ZY).

Additional information

<table>
<thead>
<tr>
<th>Funder</th>
<th>Grant reference number</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Heart, Lung, and Blood Institute (NHBLI)</td>
<td>R01-HL119813</td>
<td>Tom J Brett</td>
</tr>
<tr>
<td>American Heart Association (AHA)</td>
<td>PRE19970008</td>
<td>Zeynep Yurtsever</td>
</tr>
<tr>
<td>American Lung Association</td>
<td>RG-196051</td>
<td>Tom J Brett</td>
</tr>
<tr>
<td>National Heart, Lung, and Blood Institute (NHBLI)</td>
<td>R01-HL54171</td>
<td>Colin G Nichols</td>
</tr>
<tr>
<td>Washington University in St. Louis Pilot and Feasibility Grant</td>
<td></td>
<td>Tom J Brett</td>
</tr>
</tbody>
</table>

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.
Author contributions
MS-R, ZY, Conception and design, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article; CGN, TJB, Conception and design, Analysis and interpretation of data, Drafting or revising the article

Author ORCIDs
Tom J Brett, http://orcid.org/0000-0002-6871-6676

References


DOI: 10.7554/eLife.05875.
Video Article

Efficient Mammalian Cell Expression and Single-step Purification of Extracellular Glycoproteins for Crystallization

Daniel L. Kober1,2,3, Zeynep Yurtseser2,3,4,5, Thomas J. Brett2,3,5,6,7

1Molecular Microbiology and Microbial Pathogenesis Program, Washington University School of Medicine
2Department of Internal Medicine, Washington University School of Medicine
3Drug Discovery Program in Pulmonary and Critical Care Medicine, Washington University School of Medicine
4Biochemistry Program, Washington University School of Medicine
5Center for the Investigation of Membrane Excitability Diseases, Washington University School of Medicine
6Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine
7Department of Cell Biology and Physiology, Washington University School of Medicine

* These authors contributed equally

Correspondence to: Thomas J. Brett at TBRET@dom.wustl.edu

URL: http://www.jove.com/video/53445
DOI: doi:10.3791/53445

Keywords: mammalian protein expression, nickel affinity chromatography, glycosylation, macromolecular crystallography

Date Published: 11/16/2015


Abstract

Production of secreted mammalian proteins for structural and biophysical studies can be challenging, time intensive, and costly. Here described is a time and cost efficient protocol for secreted protein expression in mammalian cells and one step purification using nickel affinity chromatography. The system is based on large scale transient transfection of mammalian cells in suspension, which greatly decreases the time to produce protein, as it eliminates steps, such as developing expression viruses or generating stable expressing cell lines. This protocol utilizes cheap transfection agents, which can be easily made by simple chemical modification, or moderately priced transfection agents, which increase yield through increased transfection efficiency and decreased cytotoxicity. Careful monitoring and maintaining of media glucose levels increases protein yield. Controlling the maturation of native glycans at the expression step increases the final yield of properly folded and functional mammalian proteins, which are ideal properties to pursue X-ray crystallography. In some cases, single step purification produces protein of sufficient purity for crystallization, which is demonstrated here as an example case.

Video Link

The video component of this article can be found at http://www.jove.com/video/53445/

Introduction

Understanding protein structure at an atomic level is key to uncovering the molecular basis of biological pathways and diseases. X-ray protein crystallography is the most widely used/applicable method for determining macromolecular structures. The main challenge of this method is obtaining sufficient amounts of properly folded, pure protein. This becomes an issue particularly when working with secreted mammalian proteins, which undergo specific post-translational modifications.

Bacterially-expressed proteins are the primary source of crystallized proteins deposited in the Protein Data Bank1. Bacterial expression systems are largely preferred because they are fast, inexpensive and typically produce high yields of protein. However, extracellular domains of mammalian proteins expressed in bacteria are often not properly folded, in which case refolding and extensive purification steps are required for obtaining homogeneously folded protein. Additionally, many mammalian proteins require post-translational glycosylation to achieve proper folding2. Although expression and glycosylation in yeast or insect cells can overcome the folding problem, post-translational modifications, including glycosylation, differ significantly from those of mammalian cells3, yielding proteins with incorrect or non-homogeneous modifications.

Mammalian cells express all the required molecular machinery to ensure proper post-translational modifications and folding; however, these expression systems are not typically preferred by most labs, due to limited yields and high costs of reagents and consumables. Polyethyleneimine (PEI), a standard transfection reagent is relatively cheap but imposes considerable cytotoxicity and low transfection efficiency, resulting in increased costs in cell media, DNA, and culturing equipment. Many alternatives to PEI are prohibitively expensive. We address these issues by describing a combination of improved cell culture tools and chemically modified PEI for the quick and relatively inexpensive method for the expression of secreted mammalian proteins, followed by single-step purification. This robust method gives sufficient yields for functional and biochemical studies4, and in some cases, results in protein amenable to crystallization without further purification.
This protocol describes several techniques to maximize expression and yield for secreted mammalian proteins in human embryonic kidney (HEK) 293F cells grown in suspension. Transfection efficiency (and cost), protein production and purification are all greatly enhanced by following this protocol. PEI modified by the addition of carbamates through a single-step ring-opening reaction (PEI-TMC-25, synthesis and properties described in detail in ref5) greatly improves transfection efficiency, reduces the cytotoxicity from cationic membrane disruption and accordingly reduces experiment costs. Furthermore, cell viability and protein expression are greatly improved with the addition of culture supplements to supply glucose and vitamins. Importantly for the production of glycosylated proteins, treatment with kifunensine, a non-toxic chemical inhibitor of Mannosidase I, produces proteins with defined, immature glycans, which can be removed by the endoglycosidase EndoHf to yield proteins with a single N-acetylgalcosamine in place of a full-length N-linked glycan. Finally, the secretion of proteins into a serum-free, chemically defined medium allows rapid and facile purification for structural and biochemical studies. Single-step nickel-nitrilotriacetic acid (Ni-NTA) resin purification removes the majority of contaminating species in the supernatant and, in some cases, can yield protein of sufficient purity for crystallization.

## Protocol

### 1. Production of Milligram Quantities of Plasmid DNA for Large-scale Transient Transfection

1. Clone the protein of interest into a high copy number mammalian expression vector using restriction site cloning, or other appropriate technique.
   1. For optimal results, use pHILsec \(^7\) vector, which has a built-in C-terminal 6His-tag, a strong promoter Kozak sequence and an optimized secretion signal.

2. Transform the plasmid onto competent cells.
   1. Add 20 µl of competent E. coli cells onto 1 µg of plasmid DNA and incubate on ice for 30 minutes.
   2. Heat shock cells at 42 °C for 35 sec, then incubate on ice for 2 min.
   3. Add 300 µl of microbial growth medium (SOC) and incubate at 37 °C for 45 mins, shaking at 220 rpm.
   4. Plate cells on agar plate with appropriate antibiotic selection.
      1. Use 100 µg/ml carbenicillin if the plasmid is in the pHILsec vector.

3. Culture colonies in 250 ml of LB Media supplemented with 100 µg/ml antibiotic (carbenicillin) overnight at 37 °C, shaking at 220 rpm.

4. Purify DNA from culture using Hi-Speed Plasmid Maxi Kit according to manufacturer's protocol.
   1. Elute DNA in buffer EB (10 mM Tris-Cl, pH 8.5), instead of buffer TE.
   2. Aliquot the purified plasmid at amount needed for transfection and store at -20 °C.

### 2. Large-scale Culture and Transient Transfection of 293F Cells

1. Supplement 1 L 293F media with 10 ml of glucose and 5 ml Pen/Strep (both 100x). Store at 4 °C. 5 ml Pen/Strep is a sufficient strength in serum-free conditions and the reduced antibiotic concentration improves cell viability during transfection, which improves protein yields.

2. Culture 293F cells in 300 ml media in 1 L polycarbonate baffled Erlenmeyer flasks with vented caps at 37 °C with 8% CO\(_2\), while shaking in a standard tissue culture incubator.

3. Dilute cells to 5 x 10^5/ml density one day before transfection.

4. On the day of transfection, supplement culture medium by adding 10% volume of 2% w/v Cell Boost in 293F media.
   1. Measure glucose concentration using a glucose monitor according to the manufacturer's instructions and use supplements as needed to achieve a glucose concentration of 500 mg/dL.

5. Add kifunensine (1 µg/ml final concentration) at this step to control protein glycosylation.

6. Calculate volume of DNA required for 1 µg plasmid per 1 x 10^6 cells. Under sterile conditions, dilute DNA in 5 ml serum-free medium.

7. Calculate volume of transfection reagent required for 1 µg plasmid per 2 µl transfection reagent. Under sterile conditions, dilute transfection reagent (PEI-TMC-25) in 5 ml serum-free medium.

8. Add transfection reagent into DNA solution in 1 ml increments, mixing gently. Incubate for 30 minutes at room temperature for reagent-DNA complexes to form. Then add the solution onto the cells in a drop-wise fashion.

9. Allow transfected cells to express protein for 72-96 hr. Supplement with ~10% volume Cell Boost Media daily, or as necessary to keep glucose reading 400-600 mg/dl.

### 3. Purification

1. Decant culture into a centrifuge flask, centrifuge for 20 min at 1,300 x g to pellet cells and then collect the supernatant. If necessary, spin a second time and/or use 0.22 µm filter to clarify supernatant.

2. Add 10% volume 10x Ni-NTA binding buffer (1.5 M NaCl, 0.5 M K2HPO4, 0.1 M Tris pH 8.5, 50 mM imidazole).

3. Prepare a gravity column by adding 2 ml of Ni-NTA slurry in a column and equilibrating with 10 column volumes (CV) of 1x binding buffer. If possible, do all column steps in a 4 °C room. Alternatively, chill protein and all buffers on ice before column step, and keep protein and collected flow-through on ice.

   Note: Ni-NTA slurry is 50% resin by volume and the manufacture's stated binding capacity is 50 mg/ml. Ni-NTA beads can be re-charged for multiple uses.
4. Flow the supernatant over the resin and collect flow-through. Repeat this step.
5. Wash with 10 CV of wash buffer (300 mM NaCl, 50 mM K₂HPO₄, 20 mM imidazole pH 8).
6. Elute the protein in 5 CV of elution buffer (300 mM NaCl, 50 mM K₂HPO₄, 250 mM imidazole pH 8).
7. If deglycosylation is required:
   1. For a final volume of 0.5 ml, concentrate eluate to 0.43 ml using a centrifugation concentrator. If precipitates form, pellet any debris by centrifugation at 16,000 x g and 4 °C.
   2. Add 50 µl of 500 mM Na-Citrate pH 5.5.
   3. Add 20 µl of EndoHf (1 x 10⁶ U/ml). Incubate at room temperature for 2 hours.
      Note: The enzyme works optimally at 37 °C, which may cause the concentrated protein to aggregate. Extend the room temperature incubation, if deglycosylation, assessed by SDS-PAGE or immunoblotting, is incomplete. The enzyme does not have activity at 4 °C.
   4. To remove EndoHf: Wash Amylose Resin 3x in phosphate buffered saline (PBS) or final storage buffer. Incubate protein with resin for 1 hr at 4 °C. Spin 5 min at 1,000 x g to pellet beads and collect the supernatant.
   5. Concentrate protein using appropriate molecular weight cutoff centrifugation filter and buffer exchange into storage buffer (150 mM NaCl and 20 mM HEPES pH 7.5).

Representative Results

Herein follows the results of this expression system applied to a secreted 13 kDa immunoglobulin (Ig) domain from the human protein triggering receptor expressed on myeloid cells 2 (hTREM2, residues 19-132). TREM2 is a type I transmembrane protein containing a single extracellular Ig domain that has two disulfide bonds and two N-linked glycosylation sites. Unlike many other Ig domain proteins, TREM2 was not amenable to refolding from bacterial inclusion bodies. Subsequent mutagenesis confirmed N-linked glycans are required for proper expression and folding. To facilitate structural and functional studies, TREM2 was introduced into the pHLsec vector with a C-terminal 6His-tag using standard molecular biology techniques. Transient expression in HEK293F cells treated with kifunensine yielded protein that was purified by Ni-NTA chromatography (Figure 1B, lane 2) and the sample was then deglycosylated to produce natively folded, homogenous protein (Figure 1B, lane 3). 293F expression of TREM2 produced 5-10 mg/L (5-10μg/million transfected cells). After buffer exchanging into the storage buffer, this protein was crystallized (Figure 1C). Despite the final purity of about 80%, these crystals reliably reproduced, diffracted, and were shown by silver-stain to only contain TREM2 protein (Figure 1D). This observation suggests the biochemical homogeneity of the protein (i.e. folding and post-translational modifications) can, in some cases, be more critical than overall purity for crystallization success.

In addition to crystallization, this system offers a robust tool for structural and functional studies. It is exploited to produce natively glycosylated protein and achieve >95% purity by size-exclusion chromatography (Figure 1E, F). This additional purification step, which shows the solution behavior of the protein, is also an ideal way to monitor protein quality. The purified protein should elute at a volume corresponding to its molecular weight and should be the most abundant species in the sample. Abnormally large amounts of aggregated protein may indicate unstable protein and provide clues to optimize protein production and purification. Furthermore, this final purified protein is superior for use in quantitative biophysical experiments such as circular dichroism spectroscopy, thermal stability, and investigating protein-ligand interactions. Lastly, controlling the extent of glycosylation provides a robust tool to study glycan-dependent functions.
HEK 293F cells offer robust production of proteins requiring post-translational modifications. This system allows rapid and scalable expression of natively folded proteins containing disulfides, glycosylation, and phosphorylation that would otherwise be absent using more routine expression tools. In addition, this system can be used for the expression and purification of multi-protein complexes simply by co-transfection of multiple plasmids. Besides TREM2, this system has been extensively used for functional studies with other proteins of interest in the lab\textsuperscript{10,11}. Mammalian cells also offer endotoxin-free protein expression optimal for \textit{in vivo} experiments or for production of natively folded antigens to generate conformation-dependent antibodies.

Optimal cell viability and transfection conditions are the most crucial for efficient protein production. For better cell viability, low-passage cell cultures are boosted with cell culture supplements and the antibiotic concentration in the growth media is reduced. The modified PEI-TMC-25 has suitable transfection efficiency with most proteins; however, there are other options available at moderate price that offer increased transfection efficiency and reduced cytotoxicity. Hype 5 (Oz Bioscience) increases yields compared to PEI or other reagents, at a considerably reduced cost compared to more expensive transfection reagents, such as 293fectin. The reagent type and DNA: reagent ratios can be optimized for individual expressions and the needs of the desired experiment.

The method described here has several advantages over other methods of protein production. Mammalian cells offer native chaperone folding and post-translational modifications unavailable to bacteria and yeast. The reduced time for plasmid construction and preparation, along the use of media, the pH of which can be directly modified, make it superior to baculovirus-based production in insect cells; and finally, the serum-free media offers facile scaling (higher cell density) and purification not attainable to adherent 293T cells. The chief limitation of this system is only the time and scale the individual researcher is able to commit to protein expression.

Included below are troubleshooting options for the most common issues encountered.

Troubleshooting

If there is low protein expression, avoid passaging 293F cells longer than 2-3 months. Prolonged passaging results in decreased cell health reflected in slower doubling times and significantly reduced protein expression. Cell growth should be monitored daily and cultures no longer doubling daily should be discarded. Cell viability can be improved by using culture supplements while passaging cells. Monitor glucose to ensure a culture concentration of 400-600 mg/dl Cell density should not exceed 2 million/ml and cell viability should be ≥98% by trypan blue exclusion. Dilute dense cells to 0.5 million/ml and allow to double overnight before transfection. Optimize transfection using different ratios of transfection reagent: DNA (within the range of 1 μg DNA: 1.5-6 μl reagent). This can be done using 2 ml cultures in a 6-well plate and the output measured by immunoblotting. Doing 12 hr time points can indicate if the protein is being expressed and degraded rapidly; check the cell lysate if protein is not observed in the supernatant fraction. Misfolded proteins, which fail to be secreted, will still be apparent in the cell lysate by immunoblotting.

If Ni-NTA retention is low, \textit{i.e.} protein remains in flow-through after column binding, use fresh resin. Increasing pH of sample to 8.5 will result in stronger binding to the resin, although this will also increase non-specific binding. Concentrating the sample and batch binding to resin overnight...
Review Article

Novel Roles for Chloride Channels, Exchangers, and Regulators in Chronic Inflammatory Airway Diseases

Monica Sala-Rabanal,1,2 Zeynep Yurtsever,2,3,4,5 Kayla N. Berry,4,6 and Tom J. Brett1,2,4,5,7

1Department of Cell Biology and Physiology, Washington University School of Medicine, St. Louis, MO 63110, USA
2Center for the Investigation of Membrane Excitability Diseases, Washington University School of Medicine, St. Louis, MO 63110, USA
3Biochemistry Program, Washington University School of Medicine, St. Louis, MO 63110, USA
4Department of Internal Medicine, Washington University School of Medicine, St. Louis, MO 63110, USA
5Drug Discovery Program in Pulmonary and Critical Care Medicine, Washington University School of Medicine, St. Louis, MO 63110, USA
6Medical Scientist Training Program, Washington University School of Medicine, St. Louis, MO 63110, USA
7Department of Biochemistry and Molecular Biophysics, Washington University School of Medicine, St. Louis, MO 63110, USA

Correspondence should be addressed to Tom J. Brett; tbrett@wustl.edu

Received 24 August 2015; Accepted 13 October 2015

Academic Editor: Mauricio Retamal

Copyright © 2015 Monica Sala-Rabanal et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Chloride transport proteins play critical roles in inflammatory airway diseases, contributing to the detrimental aspects of mucus overproduction, mucus secretion, and airway constriction. However, they also play crucial roles in contributing to the innate immune properties of mucus and mucociliary clearance. In this review, we focus on the emerging novel roles for a chloride channel regulator (CLCA1), a calcium-activated chloride channel (TMEM16A), and two chloride exchangers (SLC26A4/pendrin and SLC26A9) in chronic inflammatory airway diseases.

1. Introduction

The chronic inflammatory airway disease, asthma, and chronic obstructive pulmonary disease (COPD) are significant causes of morbidity and mortality in children and adults. Asthma affects over 300 million people worldwide, and the prevalence is increasing among all demographics. COPD is currently the third leading cause of death in USA. These diseases are hallmarked by a Th2-mediated inflammatory response which drives the three pathologies that contribute to airway obstruction in these diseases: chronic inflammation; airway muscle constriction due to airway hyperreactivity (AHR); and mucus overproduction due to mucous cell metaplasia (MCM). A central feature of these diseases is production of the inflammatory cytokines IL-4 and IL-13, which drive MCM and contribute to AHR.

The inflammatory signaling upregulates the expression of hundreds of proteins in the airway epithelia. A number of these proteins have roles in anion transport across membranes, including chloride channels, channel regulators, and transporters. The identity, function, and elucidated mechanism of action of these proteins have lagged behind their cation channel counterparts. However, recent advances in several technologies, including high throughput screening, have made it possible to consider the development of specific inhibitors and activators for these classes of proteins [1]. The development of such therapeutics, however, requires an intimate knowledge of the roles these proteins play in airway homeostasis and mucociliary clearance. Anion channels play very crucial roles in mucus function. Mucus is composed of 97% water and 3% solids, with the main solid component being the mucin proteins [2]. Mucin proteins are secreted in a dehydrated form and require anion channel activity to instill chloride and bicarbonate ions that ensure proper salination, hydration, and pH of the mucus gel layer. Proper control of this is crucial as is exemplified by the disease cystic fibrosis.
(CF), which is caused by loss of function mutations to the chloride channel cystic fibrosis transmembrane conductance regulator (CFTR) that produces thick, sticky mucus deficient in mucociliary clearance or innate antimicrobial properties [3].

Here we discuss what is currently known about the function of four exciting, new, and emerging proteins affecting anion channel activity in inflammatory airway epithelia: a chloride channel regulator (CLCA1), a calcium-activated chloride channel (TMEM16A), and two chloride exchangers (SLC26A4/pendrin and SLC26A9). In particular, we focus on recently uncovered contributions to airway diseases and mucus function, in order to answer whether they can be targeted by inhibitors or activators and whether they should be.

2. The CLCA Family of Chloride Channel Regulators

The CLCA family of proteins was originally misidentified as calcium-activated chloride channels and has long been associated with chronic inflammatory airway diseases. Their evolving functional identity and the possible role they play in these diseases have only recently been elucidated.

2.1. CLCAs: Association with Chronic Inflammatory Airway Diseases. Asthmatic inflammation results from a Th2-mediated mechanism, where the cytokines IL-4 and IL-13 bind their receptors and activate the transcription factor STAT6 to drive inflammation and mucus overproduction in the airways [4, 5]. In mouse models of both allergic and respiratory virus induced-asthma, CLCA1 (previously known as mCLCA3 or gob-5) expression has been solidly linked to IL-13 driven MCM [6, 7] and controversially linked to AHR [6], both hallmarks of asthma and COPD. Similar results were observed in vitro with the human pulmonary mucopidermoid cell line NCI-H292, in which expression of the protein significantly increased mucin gene MUC5AC expression and subsequent mucus production [8, 9], implying that CLCA1 can drive MCM. Studies using Clcat−/− mice, however, have failed to show reduced response to IL-13 stimulation, as these mice showed the same phenotype as wild-type (WT) mice [7, 10]. Other members of the family, particularly CLCA2 and potentially CLCA4A and CLCA4B, have also been observed to be upregulated and to induce and colocalize with the inflammatory mucin protein, MUC5AC [7]. These results indicate a possible functional redundancy between members of the mouse CLCA family, which is unlikely to translate to human biology, as there are only three human CLCA proteins, compared to the seven mouse CLCA proteins [II] (Figure 1).

Of the three human CLCA proteins, CLCA1 has been identified as a potential biomarker of inflammation and MCM in the airways [12] and suggested as a potential drug target for treatment of asthma and COPD. Similar to mouse CLCA1, overexpression of human CLCA1 in NCI-H292 cells increases MUC5AC expression and mucus levels [6, 8, 9]. Its expression is upregulated in primary cell models of IL-13 driven MCM and siRNA-mediated knockdown of CLCA1 prevents IL-13-driven mucus overproduction [9]. These experimental observations suggest a central role for CLCA1 function in IL-13-mediated MCM (Figure 2). Additionally, it is highly overexpressed in the airway epithelia of asthmatic patients [8] and can be found in the bronchoalveolar lavage fluid (BALF) at high levels [13]. In contrast, other members of the human CLCA family are not upregulated in response to IL-13 [8, 9], suggesting that CLCA1 is the sole family member with an essential role in MCM in human airways.

2.2. The Conceptual Evolution of CLCA Proteins from “Channels” into “Channel Regulators”. The family of CLCA proteins were first cloned from bovine and murine samples, where overexpression of these proteins increased Ca2+-sensitive Cl− conductance, which led to their initial misannotation as Ca2+-activated Cl− channels (CaCCs) [14–16]. In addition, observations that nonspecific chloride channel blockers, such as niflumic acid and DIDS, seemed to reduce both currents and the mucus production [17] erroneously supported this hypothesis. However, modern bioinformatics algorithms and experimental approaches have definitively demonstrated that CLCA proteins are soluble, secreted proteins that do not constitute ion channels themselves [13, 18], and subsequent studies demonstrated that CLCA proteins activate currents through an endogenous CaCC [18–20]. As a consequence, the CLCA nomenclature has been updated and the family is now recognized as Ca2+-activated Cl− channel regulator proteins. All members of this family (with the exception of the likely pseudogenes human CLCA3 and mouse CLCA4C) are synthesized as full-length proteins and proteolytically cleaved into two fragments. It has been demonstrated that, for some of the family members, the cleavage is carried out by a zinc-dependent matrix metalloprotease-like (MMP-like) domain located in the N-terminus of the protein [18, 21] (Figure 1). However, sequence analysis has revealed that all CLCA family members contain the required MMP active site motif [18]; thus self-cleavage is likely a conserved feature of all CLCAs. Self-cleavage is required to unmask the N-terminal fragment CLCA1, which then interacts with the CaCC [18] (Figure 2). The molecular identity of this previously unknown CaCC has been shown to be TMEM16A and the direct interaction between the channel and the N-terminal fragment stabilizes and increases the cell surface expression of the channel, thereby increasing currents [20].

2.3. CLCA1 as a Potential Regulator of Cytokine Expression. Upstream of CLCA1 expression, the involvement of Th2 cytokines IL-4 and IL-13 has been shown in cellular and animal models as discussed above. The relationship between CLCA1 and downstream cytokine signaling, however, is still under investigation. The few articles that do exist report contradictory observations regarding the role of CLCA1 as a signaling molecule for cytokine expression. Challenging Clcat−/− mice with lipopolysaccharide (LPS), Long et al. observed increased KC (keratinocyte-chemoattractant) levels in BALF, but no change in MIP-2 (macrophage inflammatory protein 2) or IL-17 levels in the knockout mice compared
Figure 1: Domain architecture schematic for human and mouse CLCA proteins. Each row contains the corresponding human and mouse homologs. Mouse CLCA proteins are labeled according to the recently updated naming commissioned by the Mouse Gene Nomenclature Committee (MGNC) in order to align the numbering established by the Human Gene Nomenclature Committee and the Rat Genome Database. The previously used names for the mouse proteins are shown below the current names and are in italics. Scissors denote the experimentally determined location of proteolytic cleavage sites [18]. Human CLCA3 and mouse CLCA4C are likely pseudogenes because they contain premature stop codons. Labels denote the following domains: CAT: matrix-metalloprotease-like catalytic domain; CYS: matrix-metalloprotease-like cysteine rich domain; vWA: von Willebrand factor type A domain; FnIII: fibronectin type III domain; TM: transmembrane domain; GPI: glycosylphosphatidylinositol anchor.

Figure 2: Schematic of CLCA1-driven MCM in human airways based on the current literature. IL-13 induces CLCA1 gene expression through activated STAT6. CLCA1 protein is expressed, is secreted, and undergoes proteolytic self-cleavage to yield two fragments (N-CLCA1: N-terminal fragment; C-CLCA1: C-terminal fragment). N-CLCA1 engages and activates the CaCC TMEM16A. Downstream, a signaling pathway is activated through MAPK13 which leads to induction of the inflammatory mucin MUC5AC, followed by goblet cell differentiation and subsequent MCM. It is currently unknown whether or how the steps highlighted in the dashed ellipse (CLCA1 cleavage and activation of TMEM16A) contribute to the activation of MAPK13.
to WT mice [22]. In contrast, Dietert et al. observed significantly decreased levels of IL-17 and CXCL-1 in BALF from the CclαT−/− mice infected with Staphylococcus aureus [23]. Using a cellular model of inflammation, Ching et al. showed that CLCA1-conditioned media increased proinflammatory cytokine (IL-6, IL-8, IL-1β, and TNFα) mRNA levels in monocyte cell line U-937 and primary porcine alveolar macrophages. Immunopurified CLCA1 protein only increased IL-8 and IL-1β levels significantly [24]. If such a regulatory mechanism exists for cytokine expression, modulation of CLCA1 function with small molecules to treat mucus cell metaplasia might also alter the inflammatory response in the airways.

3. TMEM16: The First Family of CaCCs

While CaCC conductance was a long-observed phenomenon in the airways and could be separated from CFTR currents, the molecular identity of the channels responsible for these currents remained elusive until the late 2000s. The TMEM16/Anoctamin family was identified in 2008 as the first bona fide CaCCs [25–27]. However, based on their electrical and pharmacological characterization, only two of the ten family members, TMEM16A and TMEM16B, displayed properties previously observed for CaCCs in the airways [28,29], whereas most of the other members function as lipid scramblases. Of these two, TMEM16A expression has been verified in airway epithelium and airway smooth muscle cells [30].

3.1. TMEM16A Is Linked to Chronic Inflammatory Airway Diseases

The predicted topology for TMEM16 family members is based on the recent landmark crystal structure of the fungal Nectria haematococca TMEM16 (nhTMEM16) which has 10 transmembrane domains instead of the previously predicted 8 (Figure 3) [31]. The purified and reconstituted protein, which was shown to be a homodimer [32], constitutes a channel on its own and does not require other proteins to be active [33]. While the Ca2+ sensitivity of the channel is well documented [31,34,35] and the protein directly binds Ca2+, the possible involvement and mechanism of interaction with calmodulin as a calcium sensor and binding partner are still controversial [35–37].

Much like CLCA1, expression of TMEM16A is upregulated by IL-4/IL-13 stimulation [38–41]. Upon upregulation, TMEM16A colocalizes to the apical plasma membrane of goblet cells along with the mucin MUC5AC [41–43]. It is expressed in airway smooth muscle cells and has been shown to play a role in AHR [42]. Additionally, inhibitors of TMEM16A have been shown to block mucus secretion [42,44] whereas small molecule activators increase secretion [45].

3.2. The Potential for Targeting CLCA1 and TMEM16A in Chronic Inflammatory Airway Diseases

CLCA1 is a promising therapeutic target for asthma and COPD, as it is the only member of its family to be upregulated in models of IL-13 mediated mucus overproduction [9], is a secreted protein [18], is expressed in goblet cells [7,9,16], and associates with mucin granules [46,47]. Overexpression of CLCA1 in airway
epithelium induces MUC5AC expression via a signaling cascade involving the kinase MAPK13, and siRNA-knockdown of CLCA1 blocks mucus production in these models [9], implying a critical role for CLCA1 function in inflammatory mucus production. Consistent with these observations, DNA vaccines [48] and antibodies targeting mouse CLCA1 [49] have displayed some effectiveness in reducing airway inflammation and MUC5AC levels in mouse asthma models. Similarly, as mentioned above, TMEM16A inhibitors block ATP-dependent mucus secretion, suggesting a central role for this pair in inflammatory mucus overproduction. However, it should be noted that these small molecules have relatively low potency and questionable selectivity and these findings need to be supported by additional experiments to determine the role TMEM16A plays in mucus secretion [50]. An important question that currently remains unanswered is what role this pair of molecules plays in mucociliary clearance. It is well known that anion channel activity is required for secreted mucins (MUC5AC and MUC5B) to function properly in a mucosal immunity and mucociliary clearance capacity [51]. Mucin proteins are secreted in dense, dehydrated granules and require anion channel passage of chloride and bicarbonate ions to ensure proper hydration, salination, and pH control [2]. Thus, any therapy targeting anion channel activity in airway diseases should proceed with caution to avoid any potential detrimental impacts to mucociliary clearance and innate mucosal immunological properties. Along these lines, exploiting the mechanism CLCA1-mediated regulation of TMEM16A action might be a possible therapeutic route for CF, utilizing a potential compensatory channel to make up for loss of CFTR activity [52]. Further along these lines, TMEM16A has been shown to carry not only chloride ions but also bicarbonate ions [36]; thus activation of TMEM16A in the setting of CF could be beneficial to adjust mucus hydration and pH.

4. SLC26: An Ancient Family with Unexpected New Roles

Over the last decade the anion exchanger pendrin (PDS, SLC26A4), once thought to be limited mainly to the inner ear, kidney, and thyroid, has been found to be upregulated by inflammatory cytokines in the bronchial epithelium, where it contributes to the pathogenesis of inflammatory airway diseases [33–35] and also to the host response to bacterial infections [56, 57]. Another member of the family, SLC26A9, is prominently expressed in the airway epithelia, where it interacts with CFTR to modulate mucus production [58]. The discovery of these crucial roles in lung physiology and pathophysiology makes these anion transporters intriguing new biomarkers for airway disease and promising novel pharmacological targets.

4.1. Pendrin: An Anion Exchanger with Critical Roles in Ear, Kidney, and Lung Physiology. Pendrin (PDS, SLC26A4) is a member of the SLC26 family of multifunctional anion transporters and channels [59, 60]. The eleven mammalian SLC26 genes encode proteins with cytoplasmic N-termini and C-termini flanking a transmembrane core of unknown structure, predicted to span the lipid bilayer 10 to 14 times (Figure 4). Mutagenesis, homology modeling, and molecular dynamics simulation data are consistent with the hypothesis that the SLC26 transmembrane fold consists of two nesting, inverted repeats of 5–7 helices, resembling that of the CLC Cl−/H+ antiporter channel proteins [61, 62] and the recently solved three-dimensional structure of SLC26Dg, a bacterial H+–coupled fumarate symporter, has clarified this [63]. It has been suggested that SLC26 proteins organize in functional homodimers or homotetramers [64], though each subunit is thought to constitute an independent translocation pathway. The C-terminal cytoplasmic region of all SLC26 proteins includes a sulfate transporter and antisigma factor antagonist (STAS) domain (Figure 4), which has been implicated in nucleotide binding and hydrolysis [65]. SLC26A4, or pendrin, functions as an electroneutral exchanger of Cl−, HCO3−, I−, NO3−, formate, SCN−, and other monovalent anions. It is expressed in cochlear epithelial cells of the spiral prominence, in root cells, in spindle cells of the stria vascularis, in epithelial cells of the endolymphatic sac, and in epithelial cells surrounding the hair cells of the saccule, utricle, and ampulla [66]. Additionally, pendrin is expressed in the apical membrane of thyrocytes [67, 68], renal collecting duct Type B intercalated cells [69], salivary gland cells [70], and airway epithelia [53].

Pendrin function is important in several settings. In the inner ear, pendrin helps maintain Cl− and HCO3− homeostasis, which is crucial for normal hearing and for the development of bony structures such as the cochlea and the vestibular aqueduct [71]. In the thyroid gland, pendrin contributes I− to the follicle for thyroxine biosynthesis [72], and in the cortical collecting duct the transporter is implicated in Cl− reabsorption through functional coupling with the epithelial Na+ channel ENaC and the Na+–dependent Cl−/HCO3− exchanger NDCBE/SLC4A8 [73]. Most interestingly, it has been shown that in the bronchial epithelium pendrin mediates an increase in Cl−/SCN− exchange in response to IL-4 stimulation [74] to provide SCN− substrate to lactoperoxidase for the synthesis of hypoiodocyanate (OSCN−), a molecule with antimicrobial properties [75], and this underscores the emerging role of pendrin in innate airway defense mechanisms (Figure 5).

4.2. Pendrin and the Pathogenesis of Inflammatory Lung Disease: Too Much of a Good Thing? Pendrin was first identified by positional cloning as the disease gene for Pendred syndrome (OMIM number 247600), an autosomal recessive condition characterized by deafness with enlargement of the vestibular aqueduct, complex abnormalities in cochlear structure, and variably penetrant euthyroid goiter [76–78]. Pendrin is also implicated in DFNB4, an autosomal recessive form of nonsyndromic deafness [79]. SLC26A4 mutations that are clinically associated with Pendred syndrome cause complete loss of transport function when studied in heterologous expression systems, mostly due to retention in various intracellular compartments, whereas those exclusively associated with DFNB4 have residual transport activity [80]. In recent years, pendrin gain of function, mainly due to increased surface expression, has been linked to respiratory
diseases including bronchial asthma, COPD, and rhinovirus infection, rhinitis, and chronic rhinosinusitis [54, 55, 81–86].

The association between pendrin and inflammatory airway disease was first proposed in 2005, when it was observed that pendrin expression was upregulated in three different murine asthma models, including transgenic overexpression of IL-13 in lung [83]. Later, it was reported that induction of asthma or COPD in mice by inhalation of ovalbumin or elastase, respectively, resulted in increased pendrin expression; direct overexpression of pendrin in the lung led to increased mucus production and secretion and neutrophilic infiltration [86]. In subsequent works, the link between inflammatory cytokines, in particular IL-4 and IL-13, and pendrin overexpression has been cemented [81, 85, 87], and in a recent study SLC26A4 was identified as the most upregulated gene in human asthmatic bronchi [88]. A major downstream effect of IL-4 and IL-13 signaling is the activation of the signal transducer and activator of transcription 6 (STAT6). Following ligand-receptor binding, associated Janus kinases (JAKs) activate the receptor, allowing STAT6 to then be recruited and activated by phosphorylation. Once phosphorylated, STAT6 homodimerizes and translocates to the nucleus where it regulates the transcription of target genes via binding to N4 interferon-γ activated sequences (N4 GAS) in the promoter region [89]. The pendrin promoter contains at least one N4 GAS motif, and STAT6 has been shown to bind this sequence in vitro, thus suggesting that increases in pendrin promoter activity via STAT6 represent at least one mechanism by which IL-4 and IL-13 increase pendrin activity [87]. Cytokines other than IL-4 and IL-13 may be responsible for increases in pendrin expression; IL-1β, a macrophage-secreted cytokine involved in the immunopathogenesis of asthma and COPD, has also been shown to increase pendrin levels in rodent and human bronchial epithelial cells [74, 90].

Signaling through IL-4/IL-13 mediates airway hyperresponsiveness, eosinophilic inflammation, mucus cell metaplasia and mucus overproduction, subepithelial fibrosis, and increased viscosity of the airway surface liquid (ASL), all of which are common to bronchial asthma and COPD [91]. Pendrin may play a major role in the pathogenesis of asthma or COPD by regulating some of these responses, in particular ASL thickness and mucus production (Figure 5). In lung epithelial cells, reabsorptive Na+ transport through ENaC is suppressed whereas secretory Cl− transport through CFTR and CaCCs is stimulated, which collectively results in a net secretory phenotype whereby water osmotically flows into the lumen and ASL viscosity decreases. On the other hand, pendrin imports Cl− in exchange for other anions, and thus an IL-4/IL-13-mediated increase in pendrin activity may shift the equilibrium towards a reabsorptive phenotype, resulting in the osmotic flow of water into the interstitium and the thinning of the ASL [85]. In asthmatic mice, mucus overproduction is accompanied by an increased pendrin expression at the apical surface of bronchial epithelial cells,

Figure 4: Domain architecture schematic for human SLC26 family proteins discussed here, SLC26A4 (pendrin) and SLC26A9, based on the crystal structure of SLC26Dg (PDB ID 5DA0). Upper inset shows general schematic while the lower insets show details of the C-terminal cytoplasmic region for each protein. Labels denote the following domains: STAS: sulfate transporter and antisigma factor antagonist domain; PDZ: PSD95-Dlg1-ZO-1 domain. Locations of SLC26A9 interaction with CFTR R domains and WNK kinases are denoted. The location of residues of a crystallographically determined ligand binding site is highlighted with stars.
and in pendrin overexpression cell models, production of MUC5AC, a major mucus protein in asthma and COPD patients, is increased [86]. In mice, pendrin overexpression is also accompanied by neutrophil-dominant inflammation, suggesting that, in this system, mucus production may be induced not only by a direct effect of pendrin on airway epithelial cells, but also by an indirect effect of pendrin by recruiting inflammatory neutrophils [86].

In bronchial epithelial cells, IL-4/IL-13 signaling upregulates the expression of CLCA1 [9], CFTR [92], and pendrin [87], whereas it downregulates the expression of the β and γ subunits of ENaC. Because upregulation of certain proteins, such as pendrin, might aggravate asthma or COPD symptoms, whereas the downregulation of ENaC and the downregulation of CLCA1 might be protective (Figure 5), it is unsurprising that pharmacological strategies aimed at the blocking of the IL-4/IL-13 pathway are not as successful as they were anticipated to be [93]. Selective inhibition of pendrin could be an intriguing new strategy for asthma/COPD therapy, but as noted above, pendrin contributes to the secretion of SCN\(^{−}\), a substrate of lactoperoxidase for the production of the protective, antimicrobial OSCN\(^{−}\) [74], and this should be taken into account when exploring novel treatment avenues.

### 4.3. The Next Frontier: Pendrin and Infectious Lung Disease

Most recently, pendrin has been implicated in the IL-17A-dependent host inflammatory response to bacterial airway infections (Figure 5) [56, 57]. IL-17A is critical for the immune response of the lung to infectious bacteria, for example, *Haemophilus influenzae*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, and *Bordetella pertussis* [94]. The latter is the etiologic agent of whooping cough, or pertussis disease, which is a resurgent condition of great clinical concern as it can progress to pulmonary inflammation and death in infants and for which there is no effective treatment [95]. Pertussis toxin (PT), the virulence factor of *B. pertussis*, undermines the host immune system by inhibiting macrophage and neutrophil
responses, suppressing the production of antibodies against the bacteria, and inducing proinflammatory cytokines, in particular IL-17A [96]. One of the most highly upregulated genes in association with PT activity is SLC26A4 [97, 98], and pendrin is upregulated in human bronchial epithelial cells exposed to IL-17A [56]. In the lungs of B. pertussis-infected mice there is an increase in pendrin levels that is concomitant with an increase in IL-17A but not in IL-4/IL-13 levels, and pendrin upregulation is significantly hampered in II-17A-null mice [57]. Other host factors may be involved in PT-dependent upregulation of pendrin, such as IL-1β and IFN-γ, as both are upregulated during B. pertussis infection and have been linked with pendrin upregulation [85, 98, 99]. Taken together, these recent advances suggest that the upregulation of pendrin, with its associated inflammatory pathology, is a major mechanism of virulence for the pertussis toxin and position pendrin as a potential novel therapeutic target for the treatment of whooping cough.

4.4. SLC26A9: A Novel CFTR Regulator. SLC26A9 is another member of the SLC26 family of anion exchangers and channels. It is robustly expressed in apical airway epithelia [100] and gastric parietal cells [101] and to a lesser extent in the kidney, brain, and reproductive tracts [59, 102–104]. The function of SLC26A9 is still unclear: it has been described as a Cl\(^-\) channel with a small degree of bicarbonate transport [104], a Cl\(^-\)/HCO\(_3\)\(^-\) channel or an HCO\(_3\)\(^-\) exchanger [105, 106]. One group reported increased Cl\(^-\) conductance particularly in high bicarbonate conditions [107], and others have found that SLC26A9 activity is coupled to Na\(^+\) transport [58, 103, 108]. Like other members of the family, SLC26A9 is predicted to span the membrane 14 times, and it contains a STAS domain followed by a PSD95-Dlg1-Zo-1 (PDZ) domain in the C-terminal cytoplasmic region (Figure 4) [59, 109]. Two groups propose that SLC26A9 proteins, including SLC26A9, interact with CFTR initially via their PDZ domains [110, 111]. This is followed by a stronger interaction between the CFTR R domain and the SLC26 STAS domain, which is enhanced by PKA-dependent phosphorylation of the R domain [111–113]. The interaction between SLC26A9 and CFTR has been described in multiple studies. However, whether the interaction is stimulatory or inhibitory is still controversial and may be cell type-dependent [108, 114]. Not much is known about the regulation of SLC26A9, but WNK kinases, also known to regulate other transporters and channels involved in osmoregulation, have been shown to inhibit SLC26A9 activity via interaction with the STAS domain [109]. Though the influence of SLC26A9 on CFTR has been reported, the reciprocal interaction is less clear. Multiple groups have described CFTR regulation of SLC26A9 activity and expression, but the results are not consistent [58, 114–118]. Evidence of CFTR and SLC26A9 coexpression has been found in the lung, trachea, stomach, and sweat gland [119].

Due to its high expression in the lung, numerous studies have investigated the role of SLC26A9 in lung disease. Anagnostopoulou and colleagues [120] first reported that SLC26A9 activity is responsible for increased constitutive Cl\(^-\) current under Th2 inflammatory conditions, but not in normal physiology. The authors also found that SLC26A9 prevents airway mucous obstruction after stimulation with IL-13. These changes were due to increased SLC26A9 activity, which may be due to changes in regulation by WNK kinases [104]. Unlike CFTR and TMEM16A, SLC26A9 is downregulated in patients with allergic asthma (Figure 4). Further investigation revealed that a SNP in the 3′ UTR of the SLC26A9 gene likely reduced expression levels in these patients, possibly through enhanced binding of hsa-miR-632 [120].

SLC26A9 has also been implicated in the pathogenesis of bronchiectasis, the widening of airways frequently due to mucous obstruction, a condition often seen in patients with cystic fibrosis. A recent report [119] identified two patients with diffuse idiopathic bronchiectasis who also had mutations in the SLC26A9 gene. One patient presented with a mutation in a transmembrane domain of SLC26A9 (V486I); the patient’s brother was asymptomatic, though he had the same mutation. The second patient presented with a mutation in the STAS domain of SLC26A9 (R575W) in addition to the Fdel508 mutation in CFTR. Coexpression of both mutants in Xenopus oocytes provided evidence of a decreased interaction between the SLC26A9 STAS domain and the R domain of CFTR. It is thought that wild-type SLC26A9, in conjunction with CFTR loss, may enhance ion conductance and fluid secretion [121]. Thus, loss of SLC26A9, in the setting of CFTR loss, may result in reduced airway surface liquid hydration, mucous blockage, and consequent bronchiectasis.

The authors of the study further narrow the CFTR-SLC26A9 interaction region to a peptide within the STAS domain but do not confirm that the R575W mutation in this peptide disrupts the interaction and activation [119]. In contrast, a second group did not report changes in Cl\(^-\) transport with the R575W mutation [122]. As the second patient’s daughter only carried the Fdel508 mutation and was asymptomatic, the authors speculate that one mutation in CFTR is not sufficient to produce the CF phenotype [119]. However, mutations in modifier genes, such as SLC26A9, may contribute to CF in those heterozygous for CFTR mutations. Thus, SLC26A9 may influence phenotypic expression of heterozygous mutations in ion channels (CFTR, ENaC, and others) involved in airway surface liquid hydration [119].

Further supporting the modifier gene hypothesis, some SLC26A9 mutations have been shown to increase the risk of meconium ileus in patients with CF [123], as well as CF-related diabetes onset [124] and pancreatic disease severity [108, 125]. In the case of CFTR and SLC26A9 double mutations, the exact mechanism causing the phenotype must be further investigated as Cl\(^-\) transport could be due to altered CFTR and/or SLC26A9 function or to impaired regulation of CFTR by SLC26A9 or vice versa [119, 124]. Understanding the various mechanisms of SLC26A9 mutations will be important towards developing therapies that can improve lung diseases such as asthma, CF, and bronchiectasis.

5. Conclusions

Chloride transport proteins play crucial roles in airway health and disease. On one hand, they contribute to proper
mucus function by controlling mucus hydration and pH via controlling chloride and bicarbonate ion transport. On the other hand, they may play a direct role in mucus synthesis, secretion, and AHR. Recent animal model studies have emphasized the crucial role that mucus and mucin proteins play in innate mucosal immunity. Deletion of MUC5B (the main secreted mucin protein produced in the airway under homeostatic conditions) results in impaired mucociliary clearance and increased microbial infection [126]. Knockout of CFTR in pig results in airway mucus that is more acidic and deficient in antimicrobial activity due to loss of defensin function at low pH [3, 127] and is also deficient in mucociliary clearance as it remains tethered to secreting cells [128]. This could be due to improper proteolytic processing of mucin proteins, since loss of CFTR function in the intestine impairs β-meprin processing and release of secreted mucins in that setting [129]. Thus a complete understanding of how these channels contribute to mucus synthesis, secretion, function, and mucociliary clearance is required to understand the impact of modulating their activity.

Conflict of Interests
The authors declare that there is no conflict of interests regarding publication of this paper.

Authors’ Contribution
Monica Sala-Rabanal and Zeynep Yurtsever contributed equally to this work.

Acknowledgments
This work was supported by NIH R01-HL119813 (to Tom J. Brett), American Heart Association Predoctoral Fellowship PRE19970008 (to Zeynep Yurtsever), and NIH T32GM07200 (to Kayla N. Berry).

References


Mediators of Inflammation


Mediators of Inflammation


The crystal structure of phosphorylated MAPK13 reveals common structural features and differences in p38 MAPK family activation

Zeynep Yurtsever, Suzanne M. Scheaffer, Arthur G. Romero, Michael J. Holtzman and Tom J. Brett

*Biochemistry Program, Washington University School of Medicine, St Louis, MO 63110, USA, bDepartment of Internal Medicine, Washington University School of Medicine, St Louis, MO 63110, USA, cCenter for the Investigation of Membrane Excitability Diseases, Washington University School of Medicine, St Louis, MO 63110, USA, dDrug Discovery Program in Pulmonary and Critical Care Medicine, Washington University School of Medicine, St Louis, MO 63110, USA, eCell Biology and Physiology, Washington University School of Medicine, St Louis, MO 63110, USA, and fBiochemistry and Molecular Biophysics, Washington University School of Medicine, St Louis, MO 63110, USA.

Supporting information: this article has supporting information at journals.iucr.org/d

The p38 MAP kinases (p38 MAPKs) represent an important family centrally involved in mediating extracellular signaling. Recent studies indicate that family members such as MAPK13 (p38γ) display a selective cellular and tissue expression and are therefore involved in specific diseases. Detailed structural studies of all p38 MAPK family members are crucial for the design of specific inhibitors. In order to facilitate such ventures, the structure of MAPK13 was determined in both the inactive (unphosphorylated; MAPK13) and active (dual phosphorylated; MAPK13/pTpY) forms. Here, the first preparation, crystallization and structure determination of MAPK13/pTpY are presented and the structure is compared with the previously reported structure of MAPK13 in order to facilitate studies for structure-based drug design. A comprehensive analysis of inactive versus active structures for the p38 MAPK family is also presented. It is found that MAPK13 undergoes a larger interlobe configurational rearrangement upon activation compared with MAPK14. Surprisingly, the analysis of activated p38 MAPK structures (MAP12/pTpY, MAPK13/pTpY and MAPK14/pTpY) reveals that, despite a high degree of sequence similarity, different side chains are used to coordinate the phosphorylated residues. There are also differences in the rearrangement of the hinge region that occur in MAPK14 compared with MAPK13 which would affect inhibitor binding. A thorough examination of all of the active (phosphorylated) and inactive (unphosphorylated) p38 MAPK family member structures was performed to reveal a common structural basis of activation for the p38 MAP kinase family and to identify structural differences that may be exploited for developing family member-specific inhibitors.

1. Introduction

The mitogen-activated protein kinases (MAPKs) are crucially involved in signal transduction in response to several extracellular stimuli. In mammalian cells, there are four well characterized MAPK families, ERK1/2, ERK5, JNK and p38 MAPK; these are serine/threonine kinases that catalyze the reversible phosphorylation of proteins. Their central involvement in crucial signaling pathways has made them important drug targets.

The p38 MAPKs consist of four family members: p38α (MAPK14), p38β (MAPK11), p38γ (MAPK12) and p38δ (MAPK13). These p38 MAPKs are highly similar in sequence, and all are activated by dual phosphorylation mediated by the MAPK kinases MKK3 and MKK6 (Cuenda & Rousseau,
Within the family, they can be divided into two subgroups based on sequence similarity, cellular expression patterns and substrate specificity: MAPK14 (p38α) and MAPK11 (p38β) represent one subgroup, and MAPK12 (p38γ) and MAPK13 (p38δ) represent the other. MAPK14 (UniProt ID L7RSM2) and MAPK11 (Q15759) display the highest sequence identity (75%). MAPK14 was the first family member to be identified; it is fairly ubiquitously expressed and is centrally involved in pro-inflammatory signaling, thereby making it the focus of several drug-development efforts, while MAPK11 expression is more specific and its function appears to be redundant with that of MAPK14. In contrast, MAPK12 (P53778) and MAPK13 (O15264) are more divergent in sequence (62 and 61% identical to MAPK14, respectively), display more tissue-specific expression patterns and therefore contribute to more restricted functions (Risco & Cuenda, 2012). This tightly controlled expression pattern suggests that distinct biological pathways may be targeted by inhibiting specific isoforms. In this vein, MAPK13 has recently emerged as a potential disease-specific drug target. MAPK13 regulates the insulin secretion and survival of pancreatic β-cells, implying a central role in diabetes (Sumara et al., 2009). MAPK13 is also pivotal in neutrophil chemotaxis pathways, contributing to acute respiratory distress syndrome (ARDS; Ittner et al., 2012). In addition, studies in knockout mice indicate that MAPK13 contributes to pathways activated in chronic inflammatory conditions (Risco et al., 2012), is crucial for the development of arthritis in certain models (Criado et al., 2014) and is centrally involved in tumorigenesis in colitis-associated colon cancer models (Del Reino et al., 2014). Along these lines, we have observed that a signaling pathway progressing through MAPK13 mediates mucus overproduction in chronic inflammatory lung disease, and that MAPK13 inhibitors can block inflammatory mucus production in cell-culture models (Alevy et al., 2012).

Given the importance of MAPK13 as a potential specific drug target in chronic inflammatory lung disease (and possibly in other diseases), we initiated a structure-based drug-design project targeting this kinase by determining the structure of the inactive form of the enzyme and comparing it with two inhibitor-bound complexes (Alevy et al., 2012). In order to develop a more complete understanding of the structural basis for activation of MAPK13, we developed methods to produce activated (dual phosphorylated on Thr180 and Tyr182) MAPK13 (referred to in the following as MAPK13/pTpY), determined the X-ray crystal structure of MAPK13/pTpY and compared it with our previous structure of inactive unphosphorylated MAPK13 (MAPK13). In addition, we compared these with previously determined structures of other inactive (MAPK14 and MAPK11; Wang et al., 1997; Patel et al., 2009) and active (dual phosphorylated; MAPK14 and MAPK12; Bellon et al., 1999; Zhang et al., 2011) p38 MAPKs. These
studies reveal a universal structural basis for activation for the p38 MAPK family.

2. Materials and methods

2.1. Expression constructs

Full-length human MAPK13 consists of 365 amino acids and contains several basic residues at the C-terminus (six of the last 12 residues are either Arg or Lys), and this full-length protein (1–365) would not readily crystallize in initial trials. Thus, for crystallization, a slightly truncated MAPK13 construct (1–352) was designed and cloned into pET-28a as an N-terminally His_6-tagged construct using the NdeI and XhoI restriction-endonuclease sites, as described previously (Alevy et al., 2012). The construct was confirmed by sequencing.

2.2. Protein expression and purification

The cells were transformed into Escherichia coli Rosetta2 (DE3) cells (Stratagene) and colonies were grown on a plate with kanamycin selection. Cultures for protein expression were grown in LB medium using chloramphenicol (20 mg ml⁻¹) and DNAse I followed by sonication. The clarified lysate was subjected to anion-exchange chromatography using a Mono Q column (GE Life Sciences). The protein was bound against a buffer consisting of 20 mM imidazole, 10% glycerol, 10 mM β-mercaptoethanol. The cells were lysed by the addition of 0.5 mg ml⁻¹ lysozyme and DNAsel I followed by sonication. The clarified lysate was passed over Ni–NTA, which was washed with lysis buffer suitable for nickel–nitrilotriacetic acid (Ni–NTA) chromatography (50 mM Tris pH 8.0, 150 mM NaCl, 0.001% NaN₃, 5 mM dithiothreitol, 10% glycerol). The protein (at this point still a mixture of MAPK13 and MAPK13/pTpY) was exchanged into buffer consisting of 50 mM K₂HPO₄ pH 7.5, 150 mM NaCl, 0.001% NaN₃, 1 mM dithiothreitol, 10% glycerol and concentrated using an Amicon spin concentrator (Millipore). MAPK13/pTpY would not crystallize under similar conditions to MAPK13. Therefore, we initiated crystallization trials using broad commercial screens including The JCSG Core I–IV Suites (Qiagen), The PEGs I and II Suites (Qiagen), Crystal Screen (Hampton Research) and Index (Hampton Research), followed by optimization. Crystals were grown at 17°C using the hanging-drop vapour-diffusion method. Hexagonal crystals of MAPK13/pTpY were grown by mixing protein solution (at 10 mg ml⁻¹) with reservoir solution (100 mM bis-tris pH 6.2–6.6, 21% PEG 3350, 200 mM NaCl) in a 1:1 ratio (Fig. 1e). The total drop volume was 2 μl. Orthorhombic crystals of non-phosphorylated MAPK13 (Fig. 1d) have been described previously (Alevy et al., 2012).

2.3. Crystallization of MAPK13/pTpY and MAPK13

MAPK13/pTpY was exchanged into buffer consisting of 20 mM HEPES pH 7.5, 150 mM NaCl, 0.001% NaN₃, 1 mM dithiothreitol, 10% glycerol and concentrated using an Amicon spin concentrator (Millipore). MAPK13/pTpY would not crystallize under similar conditions to MAPK13. Therefore, we initiated crystallization trials using broad commercial screens including The JCSG Core I–IV Suites (Qiagen), The PEGs I and II Suites (Qiagen), Crystal Screen (Hampton Research) and Index (Hampton Research), followed by optimization. Crystals were grown at 17°C using the hanging-drop vapour-diffusion method. Hexagonal crystals of MAPK13/pTpY were grown by mixing protein solution (at 10 mg ml⁻¹) with reservoir solution (100 mM bis-tris pH 6.2–6.6, 21% PEG 3350, 200 mM NaCl) in a 1:1 ratio (Fig. 1e). The total drop volume was 2 μl. Orthorhombic crystals of non-phosphorylated MAPK13 (Fig. 1d) have been described previously (Alevy et al., 2012).

2.4. Data collection, processing and structure determination

X-ray diffraction data were collected on beamline 4.2.2 at the Advanced Light Source (ALS) using a NOIR-1 CCD detector and an 85 × 120 μm beam size. Data were indexed and processed with HKL-2000 (Otwinowski & Minor, 1997). A molecular-replacement solution was obtained using BALBES (Long et al., 2008). The best solution utilized

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Data-collection and refinement statistics for MAPK13/pTpY.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Values in parentheses are for the highest resolution shell.</td>
<td></td>
</tr>
<tr>
<td>Data-collection statistics</td>
<td>PDB 4myg</td>
</tr>
<tr>
<td>Space group</td>
<td>a = 107.23, b = 107.23, c = 125.53</td>
</tr>
<tr>
<td>Unit-cell parameters (Å)</td>
<td>X-ray source ALS 4.2.2</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>1.0070</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>Rmerge (%) 0.088 (0.590)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>97.9 (91.5)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>8.9 (5.5)</td>
</tr>
<tr>
<td>(I/σ(I))</td>
<td>Wilson B factor (Å²) 48.61</td>
</tr>
<tr>
<td>Refinement statistics</td>
<td></td>
</tr>
<tr>
<td>Rwork</td>
<td>0.205 (0.265)</td>
</tr>
<tr>
<td>Rfree</td>
<td>0.255 (0.318)</td>
</tr>
<tr>
<td>No. of amino-acid residues</td>
<td>674</td>
</tr>
<tr>
<td>No. of waters</td>
<td>186</td>
</tr>
<tr>
<td>R.m.s.d., bond lengths (Å)</td>
<td>0.005</td>
</tr>
<tr>
<td>R.m.s.d., angles (°)</td>
<td>1.015</td>
</tr>
<tr>
<td>Average B factor (Å²)</td>
<td>Protein (chain A/chain B) 57.2 (51.2/63.2)</td>
</tr>
<tr>
<td></td>
<td>Water 47.8</td>
</tr>
<tr>
<td>Ramachandran plot</td>
<td>Favorable (%) 93.0</td>
</tr>
<tr>
<td></td>
<td>Allowed (%) 6.1</td>
</tr>
<tr>
<td></td>
<td>Outliers (%) 0.9</td>
</tr>
<tr>
<td></td>
<td>Poor rotamers (No./%) 50.84</td>
</tr>
<tr>
<td></td>
<td>Clashscore (score/percentile) 7.79/99th</td>
</tr>
<tr>
<td></td>
<td>MolProbity score (score/percentile) 1.87/98th</td>
</tr>
<tr>
<td></td>
<td>Luzzati error 0.3067</td>
</tr>
<tr>
<td></td>
<td>PDB code 4myg</td>
</tr>
</tbody>
</table>

6 mg MAPK13 and 4 mg MAPK13/pTpY per litre of culture (i.e. roughly a 3:2 ratio of MAPK13:MAPK13/pTpY).
Figure 2
Structure of MAPK13/pTpY and comparison with MAPK13. (a) Ribbon diagram of MAPK13/pTpY color-coded as follows: N-lobe, green; C-lobe, blue; hinge, red; activation loop, red. Simulated-annealing OMIT $F_{\text{obs}} - F_{\text{calc}}$ electron density contoured at 3.0σ for pThr180 and pTyr182 is shown in the inset. (b) Superposition of MAPK13/pTpY (green with red activation loop and hinge) and MAPK13 (orange with blue activation loop and hinge). MAPK13/pTpY was superposed using the C-lobes. (c, d) Coordination of pThr180 and pTyr182 and comparison of structure of the activation loop between the active and inactive forms. Hydrogen bonds are shown as yellow dashed lines. ‘Warhead’ arginines involved in direct interactions with pThr180 and pTyr182 are shown with grey C atoms. Hydrogen bonds are shown as yellow dashed lines. Note that a major portion of the MAPK13 activation loop is disordered, displays no visible electron density and therefore could not be placed in the inactive structure.
phosphorylated MAPK12 (p38\textsuperscript{Y/Y}\textsuperscript{P}; PDB entry 1cm8; Bellon et al., 1999) as a probe. Solutions were not obtained using unphosphorylated full-length MAPK13 (PDB entry 4exu; Alevy et al., 2012) and could only be obtained by searching separately with the N- and C-lobe domains. This is owing to the large degree of relative motion between the N- and C-lobes that occurs upon phosphorylation. The solution had a \(Q\)-factor of 0.829 and a final \(R_{\text{work}}\) and \(R_{\text{free}}\) of 0.267 and 0.328, respectively. The model was completed in Coot (Emsley et al., 2010) using \(\sigma_A\)-weighted \(2F_o - F_c\) maps, and refinement was carried out in PHENIX (Adams et al., 2010). Because of the moderate resolution, torsion-angle NCS restraints were used throughout. Data-collection and refinement statistics are summarized in Table 1. Following all refinements, model bias in electron-density maps was reduced using the kicked-map feature in PHENIX. The final model for both chains included residues 15–351 (residues 1–14 and 352 could not be located in electron-density maps and were not built). The average error in the coordinates from the Luzzati plot was calculated using SFcheck (Vagune et al., 1999) in the CCP4i interface (Potterson et al., 2003; Winn et al., 2011). LSQ was used within Coot to perform and calculate r.m.s.d.s of \(C^\alpha\) superpositions. Motion between domains upon phosphorylation was analyzed using DynDom using the domain-select mode (Hayward & Berendsen, 1998). All molecular-graphics figures were produced using PyMOL (Schrödinger). All crystallographic software was provided from the latest distributions of the SBGrid (Morin et al., 2013).

3. Results and discussion

3.1. Production, purification and crystallization of active MAPK13/pTpY

Large amounts of pure, active (phosphorylated) forms of MAPKs are required for structural, biophysical and enzymatic studies. For other p38 MAPKs, this has previously been accomplished either by using eukaryotic expression systems capable of post-translational modifications such as insect cells (Pargellis et al., 2002) or by overexpression and purification from bacteria followed by directed phosphorylation by treatment with a constitutively active MKK6 mutant (Bellon et al., 1999). Here, we demonstrate that MAPK13/pTpY is produced by autophosphorylation during expression in \(E.\ coli\) and that the unphosphorylated and phosphorylated MAPK13 can be separated using ion-exchange chromatography (Figs. 1b and 1c), producing protein of sufficient purity for the crystallization of both forms (Figs. 1d and 1e). We have reported previously that \(\lambda\)-phosphatase can be used to dephosphorylate MAPK13 (Alevy et al., 2012); however, this treatment is not very efficient (only increasing the yield of MAPK13 by about 10%) and is unnecessary to produce either MAPK13 or MAPK13/pTpY suitable for crystallization. The expression and purification scheme outlined here is sufficient for these purposes. Previous studies have indicated that activated MAPKs may dimerize and that this dimerization contributes to kinase activity. We observe that MAPK13/pTpY co-elutes with MAPK13 in the context of a gel-filtration chromatography experiment, eluting at a roughly monomeric weight (Fig. 1a). Thus, within the context of a gel-filtration chromatography experiment, MAPK13 appears to be monomeric in both the active and the inactive forms.

3.2. Structure of MAPK13/pTpY and comparison with MAPK13

To gain an understanding of the structural basis for the activation of MAPK13, and a more complete understanding of p38 MAPK family activation in general, we determined the structure of MAPK13/pTpY and compared it with our previously determined structure of inactive MAPK13 (Alevy et al., 2012). In our previously determined MAPK13 structure (PDB entry 4exu) there is one molecule in the asymmetric unit. There are two molecules in the asymmetric unit of the MAPK13/pTpY structure that are nearly identical (the same residues are built; r.m.s.d. on \(C^\alpha\) atoms of 0.94 Å); thus, only the A chain will be discussed and used in structural comparisons throughout this manuscript. While the crystals of MAPK13 diffracted to high resolution (1.70 Å), the crystals of MAPK13/pTpY diffracted to moderate resolution (2.60 Å; see Table 1); however, the phosphorylation sites and covalently bound phosphates, as well as the entire activation loop, were well resolved in the electron-density maps (Fig. 2a). The structures display the standard MAPK topology consisting of N- and C-lobes encompassing a catalytic groove between them. Nearly the entire protein could be placed in electron-density maps for both MAPK13 (2–172 and 181–351) and MAPK13/pTpY (15–351). The main difference between the two structures occurs in the relative orientation of the N- and C-lobes and also in the activation loop (Fig. 2b). The two phosphorylated residues in the activation loop are coordinated by basic side chains from both lobes of the kinase, and this results in a more compact catalytic site and overall structure. The largest difference between the two structures occurs in the activation loop; much of this region is notably absent from electron-density maps and is therefore likely to be disordered in the MAPK13 structure, but it is very well ordered in the MAPK13/pTpY structure (Figs. 2c and 2d). This is in contrast to MAPK14, where the activation loop is clearly observed in both the inactive MAPK14 and activated (MAPK14/pTpY) structures (see, for example, Fig. 5b). We have previously demonstrated that these crystals of inactive MAPK13 are useful for the determination of MAPK13–inhibitor complexes via soaking for the purposes of structure-based drug-design studies (Alevy et al., 2012), since the activation loop is disordered and is not involved in crucial crystal contacts, and therefore is amenable to the conformational adjustments required for inhibitor binding.

3.3. Comparison of activation phosphate coordination in p38 MAPK/pTpY structures

The activation loop is the nexus of conformational differences between unphosphorylated and phosphorylated structures. Phosphorylation of Thr180 and Tyr182 induces
conformational changes in the activation loop in order to stabilize the large negative charges introduced by the phosphate groups. Packing analysis suggests that this conformational change is not greatly influenced by crystal packing. This coordination involves basic residues from both the N- and C-lobes, which in turn drives reorientation of catalytic site residues and, ultimately, closure between the two lobes, producing a conformation that is compatible with substrate binding and/or catalysis. Given the crucial importance of phosphate coordination in producing the catalytically active form, we examined the known p38 MAPK/pTpY structures in order to compare their activation-loop conformations and coordination of phosphate groups. Using superpositions of the C-lobes, MAPK13/pTpY and MAPK12/pTpY displayed nearly identical activation-loop conformations (r.m.s.d. on Cα of 0.23 Å for activation-loop residues), while MAPK14/pTpY displayed a very different activation-loop conformation (r.m.s.d. on Cα of 2.24 Å for activation-loop residues) (Fig. 3). Surprisingly, the basic residues utilized to coordinate the phosphates were not invariant in the structures of the p38 MAPK family, even though nearly all of the residues used are sequence-invariant throughout the family and in proximity (Figs. 3 and 4b). For coordination of the pThr moiety, two invariant arginines (Arg71 and Arg173 in MAPK13) are used in all three structures, while a third (Arg149 in MAPK13) was utilized in two of the structures (Figs. 3b, 3c, 3d and 4b). Additionally, this moiety is coordinated by two unique residues in both MAPK14/pTpY (Arg67; Fig. 3c) and MAPK12/pTpY (Lys69; Fig. 3d). For coordination of the pTyr moiety, one invariant arginine was used by all three of the kinases (Arg186 in MAPK13) as well as an additional invariant arginine in MAPK13/pTpY and MAPK12/pTpY only (Arg189 in MAPK13). A unique histidine (His228), which is not conserved in MAPK13 or MAPK12, was also utilized to coordinate the pTyr phosphate in MAPK14/pTpY (Fig. 3c). Taken altogether, these observations are surprising in that the p38 MAPK family display a high degree of sequence identity (e.g. 61% identity between MAPK13 and MAPK14, the most divergent pair; Fig. 4a) yet achieve stabilization of the active form of the kinase using different structural elements. This is owing to a redundancy of basic residues available in this region and is also likely to be contributed to by the different configurations of N-lobes in the structures. These observations suggest that despite the high degree of sequence identity, structural differences can exist between the active forms of the kinases, and this could contribute to selective activation and possibly be exploited to design specific inhibitor strategies.

3.4. Comparison of structural transition upon activation in MAPK13 versus MAPK14

The only other p38 MAPK family member for which crystal structures of both the inactive unphosphorylated and active pTpY forms have been determined is MAPK14, which is also the least homologous to MAPK13. Thus, we performed a comparison of the structural changes that occur between the

Figure 3
Comparison of activation-loop phosphate-group coordination in p38 MAPK/pTpY crystal structures. (a) Overlay of MAPK13/pTpY (PDB entry 4myg; blue with red activation loop), MAPK14/pTpY (PDB entry 3py3; orange with magenta activation loop) and MAPK12/pTpY (PDB entry 1cm8; green with cyan activation loop) superposed using the C-lobes. Note the similarity in activation-loop conformations in MAPK13/pTpY and MAPK12/pTpY (b, c, d). Coordination of pThr and pTyr phosphate groups. ‘Warhead’ basic residues involved in direct interactions with phosphate groups are colored according to their conserved use in the three structures (black, used in all three; grey, used in two; white, used in one). Hydrogen bonds are shown as yellow dashed lines.
two forms for these family members. Overall, the structural differences between the inactive and active forms of the kinases are not major; however, slightly larger changes occur upon activation of MAPK14. The r.m.s.d. for superposition of MAPK13 compared with MAPK13/pTpY was 2.38 Å (for 329 Cα atoms in common excluding the activation loop) versus 2.55 Å on comparing MAPK14 and MAPK14/pTpY (with the same 329 Cα atoms in common excluding the activation loop). The largest difference between the inactive and activated structures is the conformational rearrangement of the activation loop. This loop is disordered in MAPK13 (residues 172–180 are missing) and thus not built in the MAPK13 structure (see Fig. 2c) and only becomes ordered in the active MAPK13/pTpY structure; however, this loop is ordered in both the MAPK14 and MAPK14/pTpY structures. The coordination of the activation-loop phosphates is carried out by basic residues from both lobes and leads to a more closed and compact structure in the activated state. We compared the relative movement of the two domains in the MAPK13 versus MAPK13/pTpY and MAPK14 versus MAPK14/pTpY structures using DynDom (Hayward & Berendsen, 1998). The difference observed in MAPK13 versus MAPK13/pTpY is most dramatic (25° relative rotation between the N- and C-lobes) compared with the MAPK14 and MAPK14/pTpY pair (only 8° relative rotation) (Figs. 5a and 5b).

The activation-loop reconfiguration also induces a large rotation of about 100° in the side chain of Asp168 in the conserved DFG motif in both MAPK13 and MAPK14, which locks the crucial Phe169 side chain into a hydrophobic pocket (Figs. 5c and 5d). Interestingly, there are differences in the ATP-binding site that occur in MAPK14 activation that are not present in MAPK13. In MAPK14, Met109 in the hinge region blocks the ATP-binding site; however, in MAPK14/pTpY Met109 occupies a new pocket and this conformational rearrangement allows a flip of the peptide bond owing to the presence of the adjacent Gly110 (Fig. 5d). The hinge region of MAPK13 does not contain a glycine residue and correspondingly does not undergo large backbone changes in going from the inactive to the active form (Fig. 5c). The arrangement of having a glycine at this position in the hinge region is not frequent, occurring in only about 9% of the human kinome. These observations suggest that targeting this structural transition could be an effective strategy.
to specifically target MAPK14, as has recently been demonstrated with skepinone-L (Koeberle et al., 2012).

### 3.5. Comparison of conserved ‘spine’ structures in p38 MAPKs

Kornev and Taylor recently performed an analysis of the conservation of spatial motifs found in activated but not in inactivated protein kinase structures and uncovered the presence of two crucial hydrophobic elements in activated kinases that they termed ‘spines’ (Taylor & Kornev, 2011). They concluded that all protein kinases contain two contiguous hydrophobic structures that link the N- and C-lobes of the protein in the activated form, that breaking these spines results in an inactive kinase structure, that many kinases

---

![Figure 5](image_url)

**Figure 5**

Comparison of MAPK13 and MAPK14 conformational change upon phosphorylation. (a) MAPK13/pTpY, blue; MAPK13, light blue. (b) MAPK14/pTpY, orange (PDB entry 3py3; Zhang et al., 2011); MAPK14, light orange (PDB entry 1p38; Wang et al., 1997). Note that the activation loop of MAPK14 is ordered in both the inactive and active forms, whereas the MAPK13 activation loop is disordered in the inactive form. (c, d) Comparison of the hinge and ATP-binding pocket region in (c) MAPK13/pTpY and (d) MAPK14/pTpY.
display discontiguous spine structures in their inactive form, and that activation drives the assembly of contiguous spines. These two spines were termed the regulatory spine (R-spine), which consisted of four residues (two from the N-lobe, one from the C-lobe and one from the activation loop), and the catalytic spine (C-spine) consisting of seven residues (two from the N-lobe and five from the C-lobe). The R-spine is usually dynamically assembled upon reconfiguration of the activation loop after phosphorylation, whereas assembly of the R-spine is only completed upon binding of ATP, with the adenine ring docking into and completing the spine structure. Given that structures have now been determined for three p38 MAPK family members in both the activated (MAPK12/pTpY, MAPK13/pTpY and MAPK14/pTpY) and inactivated (MAPK11, MAPK13 and MAPK14) forms, we undertook the first comprehensive comparison of spine structures within this family. All of the R-spine residues are invariant within the p38 MAPK family, as are all of the C-spine residues, with the exception of conservative substitutions at two positions (Fig. 4b). All display a discontiguous R-spine in the inactive forms (Figs. 6b, 6d and 6f), which becomes contiguous after configuration of the activation loop (Figs. 6a, 6c and 6e). This largely comes from the repositioning of Phe169, which takes on varied conformations in the inactive forms (Fig. 6h) but locks into a hydrophobic pocket in the active forms (Fig. 6g). Dislodging of this Phe side chain from this pocket, thereby breaking the C-spine, is the mechanism exploited by so-called type II and III kinase inhibitors (Karcher & Laufer, 2009). Most of the structural transitioning of the C-spine residues comes from pure translation as the kinases take on a more compact structure in the activated form. Thus, the p38 MAPKs display a dynamic assembly of spines in going from the inactive to the active state (similar to that observed in other kinases) and the largest structural change within these spines comes from reconfiguration of the Phe side chain from the DFG motif which completes the R-spine.

4. Conclusion

Here, we report the first production and purification of MAPK13/pTpY suitable for structural and biophysical studies. This allowed us to determine the first structure of the activated form MAPK13/pTpY, to compare it with our previously determined structure of inactive MAPK13 and to perform the first comprehensive analysis of inactive versus active forms of p38 MAPK family members. Including the structure reported in this manuscript, there are currently structures of three p38 MAPK family members in both the inactive state, MAPK11

Figure 6
Comparison of the assembly of ‘spine’ residues in inactive versus active p38 MAPK structures. Catalytic spine residues (C-spine) are shown as yellow CPK spheres and regulatory spine residues (R-spine) are shown as red CPK spheres. (a, b) Comparison of MAPK13/pTpY and MAPK13. (c, d) Comparison of MAPK14/pTpY and MAPK14. (e, f) Comparison of MAPK12/pTpY (PDB entry 1cm8; Bellon et al., 1999) and MAPK11 (PDB entry 3gc9; Patel et al., 2009). (g) Multiple alignment of residues comprising the two spines in p38 MAPK/pTpY structures. Alignment is based on superposition of the C-lobes. (h) Multiple alignment of residues comprising the two spines in p38 MAPK structures. Residue labeling is shown for MAPK13 in both (g) and (h).
although the interlobe rotation is more dramatic for MAPK13 form a more compact structure. This is consistent with what pTyr182. This coordination is achieved using residues from structural changes that occur in comparing the inactive with the active form of MAPK13 occur to stabilize the negative charge introduced by the two phosphates at pThr180 and pTyr182. This coordination is achieved using residues from both the N- and C-lobes of the kinase, which give the active form a more compact structure. This is consistent with what is observed on comparing MAPK14 and MAPK14/pTyr, although the interlobe rotation is more dramatic for MAPK13 (25° for MAPK13 versus 8° for MAPK14) and is closer to that observed for ERK2 (Canagarajah et al., 1997). Another major difference is in the hinge region, in which major backbone rearrangements occur in MAPK14 that are required to allow it to engage the catalytically required ATP. This rearrangement does not occur in MAPK13. This structural transition represents a possible unique feature to target for the development of specific inhibitors.

Stabilization of the major negative charges introduced by dual phosphorylation is achieved using basic residues from both lobes within the p38 MAPK family. A surprising observation from our analysis is that, despite a high degree of sequence similarity in the p38 MAPK family, different residues are utilized to carry out the coordination of the phosphate groups in the different p38 MAPKs. This suggests that the inhibition of specific p38 MAPks may be accomplished by targeting the nonconserved residues that are critical for stabilizing the active form of the kinase.

Among the p38 MAPKs, MAPK13 appears to have a more restricted expression pattern and correspondingly has been linked to specific disease pathways. For example, we have observed that MAPK13 activity is crucial in a model of IL-13-induced mucus overproduction which upregulates the self-cleaving metallotranspeptidase CLCA1 (Aleyev et al., 2012), where self-cleavage is required to activate the CLCA1 protein (Yurtsever et al., 2012). These studies detail the first comprehensive analysis of p38 MAPK structures in the inactive and active (pTyr) forms, thus providing a framework for understanding the structural basis for activation of p38 MAPKs that could be useful in the design of specific inhibitors.

Acknowledgements

This work was supported in part by funding from NIH R01-HL119813 (to TJ), CADET P50-HL107183, CADET II UH2-HL123429 (to MJH) and an American Heart Association Predoctoral Fellowship PRE19970008 (to ZY). The results were derived from work performed at the Advanced Light Source (ALS), Berkeley, California, USA. The ALS is supported by the Office of Basic Energy Sciences of the US DOE (DE-AC02-05CH11231).
First comprehensive structural and biophysical analysis of MAPK13 inhibitors targeting DFG-in and DFG-out binding modes

Zeynep Yurtsever$^{1,2,3}$, Dhara A. Patel$^1$, Daniel L. Kober$^{1,4}$, Alvin Su$^1$, Chantel A. Miller$^1$, Arthur G. Romero$^1$, Michael J. Holtzman$^{1,5}$, and Tom J. Brett$^{1,3,5,6}$

From the Division of Pulmonary and Critical Care Medicine, Department of Medicine$^1$, Biochemistry Program$^2$, Center for the Investigation of Membrane Excitability Diseases$^3$, Microbiology Program$^4$, Department of Cell Biology and Physiology$^5$, and Department of Biochemistry and Molecular Biophysics$^6$

Washington University School of Medicine, St. Louis, Missouri 63110

Address correspondence to: Tom J. Brett, PhD., Campus Box 8052, 660 S. Euclid, St. Louis, MO 63110. Phone: (314) 747-0018; Fax: (314) 362-8987; Email: (tbrett@wustl.edu)
Abstract

The p38 MAP kinases are an important family centrally involved in mediating extracellular signaling in various diseases. While much attention has previously been focused on the ubiquitously expressed family member MAPK14 (p38α), recent studies indicate that family members such as MAPK13 (p38δ) display a more selective cellular and tissue expression and might therefore represent a specific kinase to target in certain diseases. In order to facilitate the design of potent and specific inhibitors, we present here the structural, biophysical, and functional characterization of two new MAPK13-inhibitor complexes, as well as the first comprehensive structural, biophysical, and functional analysis of MAPK13 complexes with four different inhibitor compounds of greatly varying potency. These inhibitors display IC₅₀ values either in the nanomolar range or micromolar range (1000 fold range). The nanomolar inhibitors exhibit much longer ligand-enzyme complex half-lives compared to the micromolar inhibitors as measured by biolayer interferometry. Crystal structures of the MAPK13 inhibitor complexes reveal that the nanomolar inhibitors engage MAPK13 in the DFG-out binding mode, while the micromolar inhibitors are in the DFG-in mode. Detailed structural and computational docking analyses suggest that this difference in binding mode engagement is driven by conformational restraints imposed by the chemical structure of the inhibitors, and may be fortified by an additional hydrogen bond to MAPK13 in the nanomolar inhibitors. These studies provide a structural basis for understanding the differences in potency exhibited by these inhibitors and provide the groundwork for future studies to improve specificity, potency, pharmacodynamics, and pharmacokinetic properties.
1. Introduction

The mitogen activated protein kinase (MAPK) p38α (MAPK14) is the founding and most studied member of the p38 MAPK family. MAPK14 is ubiquitously expressed and centrally involved in pro-inflammatory signaling; it has thus been the focus of numerous drug-development projects for several diseases involving chronic inflammation, including rheumatoid arthritis and inflammatory bowel disease (Feng & Li, 2011, Schieven, 2009, Muller & Knapp, 2010, Schnieders et al., 2012). Recent breakthroughs in structure-based drug design have resulted in highly specific and potent inhibitors to MAPK14 (Koeberle et al., 2012, Azevedo et al., 2012, Millan et al., 2011), but there are remaining concerns as to the practicality of targeting this enzyme, since knockout of this gene in mice results in embryonic lethality (Adams et al., 2000, Mudgett et al., 2000, Tamura et al., 2000), and efforts to develop MAPK14 inhibitors have ended in failed drug trials, usually due to toxicity (Goldstein & Gabriel, 2005). In contrast to MAPK14, other members of the p38 MAPK family display restricted tissue and cellular expression, and therefore may represent viable targets for specific diseases. Along these lines, MAPK11 (p38β) is highly similar to MAPK14 (75% sequence identity), whereas MAPK12 (p38γ) and MAPK13 (p38δ) are more sequence diverse (62% and 61% identical to MAPK14, respectively). Also, the MAPK12/MAPK13 double knockout mouse displays no detrimental phenotypes, only reduced innate inflammatory responses induced by lipopolysaccharide (LPS) (Risco et al., 2012).

MAPK13 has recently emerged as a disease specific p38 MAPK drug target. Compared to MAPK14, MAPK13 displays more tissue-specific expression patterns, and therefore contributes to more restricted functions (Risco & Cuenda, 2012, O'Callaghan et al., 2014). This restricted expression pattern suggests that this kinase may be targeted in specific biological pathways. For example, MAPK13 regulates insulin secretion and survival of pancreatic β-cells, and MAPK13-deficient mice are protected against insulin resistance induced by a high-fat diet and oxidative stress-mediated β-cell
failure, suggesting a role in diabetes (Sumara et al., 2009). MAPK13 is also pivotal in neutrophil chemotaxis pathways, where it contributes to acute respiratory distress syndrome (ARDS) (Ittner et al., 2012). In addition, studies in knockout mice indicate that MAPK13 contributes to pathways activated in chronic inflammatory conditions (Risco et al., 2012), is centrally involved in tumorogenesis in both colitis-associated colon cancer (Del Reino et al., 2014) and squamous cell carcinoma models (Zur et al., 2015), and is crucial for development of arthritis in certain models (Criado et al., 2014). Finally, MAPK13 activation is required for IL-13-dependent airway mucus overproduction which is a characteristic feature of chronic inflammatory lung disease (Alevy et al., 2012).

Given the role of MAPK13 in chronic lung disease (and possibly other diseases), we initiated the first structure-based drug project to develop inhibitors targeting MAPK13 (Alevy et al., 2012). This work also presented the first structures of MAPK13 complexes with inhibitors, which are the only complexes for any p38 MAPK family member other than MAPK14. As noted above, MAPK14 has been the subject of intense structure-based drug design efforts, and numerous structural and biophysical studies of MAPK14-inhibitor complexes have been reported (>170 structures in the PDB as of October 2015). Thus, much is known about the structural requirements for MAPK14 inhibition (Lee & Dominguez, 2005). To develop a more complete understanding of the structural basis for activation of MAPK13, we developed methods to produce both unphosphorylated (inactive) and dually-phosphorylated (active) MAPK13 in quantities for structural and biophysical studies and used these proteins to study complexes with inhibitors for the purposes of targeting this kinase through structure-based drug design (Alevy et al., 2012) (Yurtsever et al., 2015). Here we present a complete functional, biophysical, and structural characterization of two new MAPK13-inhibitor complexes, new biophysical data on two previously reported complexes, and the first comprehensive analysis of MAPK13 in complex with all four structurally characterized inhibitors, which display greatly varying potency. These studies reveal the structural basis for differences in potency for this set of inhibitors.
2. Materials and Methods

2.1. Materials

All inhibitors were synthesized at Washington University as described previously (Alevy et al., 2012).

2.2. Expression constructs

Full-length human MAPK13 is 365 amino acids and contains several basic residues at the C-terminus. For crystallization, a slightly truncated MAPK13 construct (1-352) was designed and cloned into pET28a as a N-terminal 6-His-tagged construct using the Ndel and XhoI restriction endonuclease sites. A full-length human MAPK13 construct was prepared similarly to produce protein for enzyme inhibition and biophysical studies (Alevy et al., 2012). The construct was confirmed by sequencing. A pET28a construct of N-terminal-6-His-tagged lambda-phosphatase was a kind gift from Dima Klenchin (University of Wisconsin, Madison, WI).

2.3. Protein Expression and Purification

 Constructs were transformed into Rosetta2 (DE3) E. coli (Stratagene) and colonies were grown on a plate with kanamycin selection (50 µg/mL). Cultures for protein expression were grown in Luria broth (LB) media using chloramphenicol (40 µg/mL) and kanamycin (50 µg/mL) selection. Both the MAPK13 crystallization construct (1-352) and full-length protein (1-365) were expressed and purified in a similar manner, as outlined here. Protein expression was induced at 30° C by addition of 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and each 1L of media was enriched with 10 ml of saturated glucose solution during protein expression. Protein expression was carried out at 30° C for 4 hours.

Cell pellets were harvested by centrifugation and suspended in lysis buffer suitable for nickel-nitrilotriacetic acid (Ni-NTA) gravity chromatography (50 mM K2HPO4 pH 8.0, 300 mM NaCl, 10 mM imidazole, 10% glycerol, 10 mM 2-mercaptoethanol). Cells were lysed by addition of 0.5 mg/ml lysozyme and DNase I followed by sonication. Clarified lysate was passed over Ni-NTA superflo
resin (Qiagen), which was then washed with lysis buffer containing 20 mM imidazole and then proteins were eluted with 250 mM imidazole. The protein was further purified by gel filtration chromatography (Superdex 200) in a buffer consisting of 20 mM HEPES pH 7.5, 150 mM NaCl, 0.001% NaN₃, 5 mM dithiotreitol (DTT), and 10% glycerol. The protein eluted as a single peak correlating to a monomeric molecular weight (Yurtsever et al., 2015). This peak was harvested and then dephosphorylated in vitro by addition of lambda phosphatase (0.1 mg/ml) and 5 mM MnCl₂ at room temperature for 30 minutes, which was then quenched with 5 mM EDTA. This treatment was not very efficient and only increased the amount of dephosphorylated MAPK13 by about 10%. Following dialysis against a buffer consisting of 20 mM Tris pH 8.0, 10 mM NaCl, 1 mM DTT, and 10% glycerol (buffer A) at 4°C overnight, the protein was then subjected to anion exchange chromatography using a MonoQ column (GE Lifesciences). Protein was injected onto the column using buffer A and then eluted off using a gradient from 0-60% of buffer B (20 mM Tris pH 8.0, 1 M NaCl, 1 mM DTT, and 10% glycerol) over 40 column volumes. This resulted in separation of unphosphorylated and phosphorylated MAPK13 (Yurtsever et al., 2015). The unphosphorylated MAPK13 peak was harvested and used for structural studies with inhibitors. Full-length MAPK13 for enzyme inhibition and biophysical assays was expressed and purified similarly.

2.3. Preparation of MAPK13-inhibitor co-crystals

Crystals of non-phosphorylated MAPK13 were obtained by mixing protein solution (at 10 mg/ml) with reservoir solution (50 mM ammonium tartrate, 18% PEG 3350) in a 4:1 (protein: reservoir) ratio, similar to those previously described by us (Alevy et al., 2012). Co-crystal complexes with inhibitor compounds 58 and 117 were prepared by soaking, similar to our previous work. Inhibitor compound stocks were prepared in DMSO at 100 mM and then added to crystal-containing drops at 1:10 volume, for a final inhibitor concentration of 10 mM in the drop. Crystals were allowed to soak for 10 minutes to 4 hours before cryoprotection in mother liquor supplemented with 25% ethylene glycol, followed by flash-freezing at -160°C in a chilled nitrogen stream.
2.4. Data collection, reduction and structure determination

X-ray diffraction data were collected at either the Advanced Photon Source (beamline 19-ID) or the Advanced Light Source (beamline 4.2.2). Data were indexed and processed with HKL2000 (Otwinowski & Minor, 1997). The phase problem was solved by isostructural replacement using the apo form of MAPK13 (PDB ID: 4YNO) (Alevy et al., 2012) in an initial round of rigid body refinement. Compounds were clearly visible in $F_O - F_C$ difference maps following rigid body refinement (see Fig. 3). Model rebuilding and compound fitting to electron density maps was performed manually using Coot (Emsley & Cowtan, 2004) and refinement was carried out in PHENIX (Adams et al., 2002). Structural analysis of MAPK13-inhibitor contacts was performed in LigPlot+ (Laskowski & Swindells, 2011). All molecular graphics figures were produced using PyMOL (Schrodinger, 2010). All crystallographic software were obtained and maintained by the SBGrid distribution (Morin et al., 2013). Diffraction images were deposited with the Structural Biology Data Grid (Meyer et al., 2016). Data collection and refinement statistics are summarized in Table 1.

2.5. Differential Scanning Fluorimetry

Melting temperature of MAPK13 and MAPK14 was measured in the presence of inhibitors (58, 61, 117, 124) using a Protein Thermal Shift Dye Kit (Life Technologies, CA) according to manufacturer’s specifications. This assay used both the unphosphorylated and the phosphorylated forms of MAPK13 and unphosphorylated MAPK14 at a final concentration of 0.15 mg/ml in the reaction mixture. A titration experiment was carried out first to demonstrate specificity and establish dynamic range. Based on this, single point experiments were carried out on all kinases using a final inhibitor concentration of 25 µM. Each condition was replicated five times. The experiment was carried out on a 7500 Fast Real-Time PCR System (Life Technologies, CA) using the melt curve template. The initial incubation step at 25 °C was modified to be 30 minutes to ensure complex formation between the inhibitors and the proteins. Data were analyzed using Protein Thermal Shift Software v1.1 (Life Technologies).
Additional statistical analysis was carried out in Prism v5.0c. Results are summarized in Figure 6 and Table 3 and presented as mean ± S.E.M.

2.6. MAPK13 Inhibition Assay

Compounds were assayed for inhibition of MAPK13 kinase activity using an immobilized metal affinity polarization (IMAP) as previously described (Alevy et al., 2012). Briefly, the assay contained activated (phosphorylated) MAPK13, a FITC-labeled substrate, and the test compound. Experiments were carried out at concentrations in the linear phase of the rate kinetics. Assay reactions contained 0-100 µM test compound, 5-35 nM (EC₈₀) activated MAPK13, 3 µM (Kₘ,app) ATP, and 100 nM FITC-labeled EGFR peptide substrate (FITC-KRELVERLTPSGEAPNQALLR-NH₂). The IC₅₀ values for each compound were determined from the compound concentration versus fluorescence polarization (mP) plot using nonlinear curve fitting.

2.7. Inhibitor binding kinetics measured by biolayer interferometry

Kinetics of MAPK13 binding to small molecules was assessed using bio-layer interferometry with an Octet (ForteBio, Menlo Park, CA) as described previously (Alevy et al., 2012). MAPK13 was randomly biotinylated in vitro using EZ-link NHS-PEG4-Biotin (Thermo Scientific) at a 1:2 molar ratio (protein: reagent). Super-streptavidin-coated biosensors from ForteBio were used to capture biotinylated MAPK13 onto the surface of the sensor. After reaching baseline, sensors were moved to association step for 300 s and then dissociation for 300 s. Curves were corrected by a double-referencing technique using both biotin-coated pins dipped into the experimental wells and a buffer-only reference. The binding buffer consisted of 20 mM HEPES pH 7.5, 150 mM NaCl, 0.005% Tween 20, and 5% DMSO. MAPK13-inhibitor complex half-lives were calculated from dissociation constants determined from global kinetic analysis (t₁/₂ = ln2/k_dissociation).

2.8. Computational docking
In silico docking studies were carried out in GLIDE from the Schrodinger Suite (Halgren et al., 2004, Friesner et al., 2004). All inhibitors were treated as ligands and prepared from their SMILES strings using LigPrep without changing ionization states. Protein coordinates were imported from pdb files and processed using the Protein Preparation Wizard (PDB ID 4MP9 for MAPK13 DFG-out; 4MP5 for MAPK13 DFG-in; 1KV2 for MAPK14 DFG-out; 1P38 for MAPK14 DFG-in). A rigid grid of 36 Å was generated centered on the DFG residues. Ligands were docked flexibly into the grid in Standard Precision (SP) mode. After the docking simulation was completed, the protein was overlapped with the coordinates from the co-crystal structures. RMSD between the docked coordinates and the crystal coordinates for the inhibitor was calculated using a command-line script, rmsd.py (available from http://www.schrodinger.com/scriptcenter/14). Results are summarized in Table 4.

3. Results and discussion

3.1. A series of inhibitors with greatly varied potencies for MAPK13.

Here we report the steady-state, biophysical, and structural evaluation of a set of MAPK13 inhibitors. The inhibitor design was inspired by analyzing potent MAPK14-inhibitor complex structures, superposing on our unphosphorylated (apo) MAPK13 structure, and realizing smaller substituents in the ATP-binding pocket would be required for optimal binding (Alevy et al., 2012). Specifically, we focused on BIRB-796, which binds to MAPK14 in DFG-out mode and exhibits an IC₅₀ of 4 nM, and inhibits other p38 isoforms, but at much lower potencies (IC₅₀ = 520 nM for MAPK13) (Kuma et al., 2005). Superposition of BIRB-796 into MAPK13 indicated it could sterically clash with the gatekeeper residue, so we designed inhibitors with smaller substituents at this position that we hypothesized would bind in DFG-out mode (Alevy et al., 2012). We began by determining the IC₅₀ values in a steady-state enzyme inhibition assay. These values for 1-(3-tert-butyl-1-methyl-1H-pryzol-5-yl)-3-[4-(pyridin-4-yloxy)phenyl]urea (compound 61) and 2-(morpholin-4-yl)-N-[4-pyridin-4-yloxy)phenyl]pyridine-4-carboxamide (compound 124) were reported previously (Alevy et al., 2012). Here we determine the
IC₅₀ values for 1-(3-tert-butyl-1-methyl-1H-pyrazol-5-yl)-3-[4-(pyridin-4-yl)sulfanyl]phenyl]urea (compound 58) and 3-(4-methyl-1H-imidazol-1-yl)-N-[4-(pyridin-4-xyloxy) phenyl] benzamide (compound 117) (Fig. 1a). Surprisingly, the IC₅₀ values varied over a wide range; while compounds 58 and 61 displayed IC₅₀ values in the nanomolar range, compounds 117 and 124 displayed IC₅₀ values in the micromolar range, over a thousand fold less potent (Figs. 1b and 1c).

3.2. Kinetic and binding analysis of inhibitors with MAPK13.

Inhibitor binding kinetics has become an important drug design criteria. In particular, slow dissociation rates have been associated with improved inhibitor activities due to prolonged target engagement in vivo (Wood et al., 2004). To understand whether inhibitor binding kinetics correlated with potency, we analyzed this property using biolayer interferometry (BLI), in a manner similar to that for compounds 61 and 124 reported previously (Alevy et al., 2012). In these assays, biotinylated MAPK13 was immobilized on the streptavidin-coated BLI pin and association and dissociation kinetics were analyzed. The two compounds that displayed IC₅₀ values in the nanomolar range displayed slower association and dissociation rates with around 100-fold longer complex half-lives (t₁/₂ = 111 to 137 sec) as compared to inhibitor compounds with micromolar IC₅₀ values (t₁/₂ = 1 to 3 sec) (Fig. 2 a-c). These kinetic constants are of similar magnitude as those for MAPK14 complexes with inhibitors measured by surface plasmon resonance (SPR) (Angell et al., 2008). We hypothesized that this difference could correlate with different binding modes or binding pockets utilized by the inhibitors.

3.3. MAPK13 inhibitors display two different binding modes.

In order to elucidate how our novel inhibitors bound to MAPK13, we determined the X-ray crystal structures of the complexes (Table 1). We had noticed during our structural analysis of the inactive (MAPK13) and dually-phosphorylated active (MAPK13/pTpY) forms of the protein (Yurtsever et al., 2015) that a significant portion of the activation loop (residues 173-180), which is ordered in the dual-phosphorylated protein, is not visible in our high resolution electron density maps of the unphosphorylated form. We therefore reasoned that we should be able to obtain inhibitor complexes
with MAPK13 via soaking inhibitors into crystals of the unphosphorylated form, and indeed we were able to demonstrate the validity of this method previously with the crystal structures of MAPK13 in complex with compounds 61 (PDB ID: 4EYJ) and 124 (4EYM). We observed a strong difference in electron density for the two new compounds using this method (Fig. 3c & e). We found that the compounds displaying nanomolar IC$_{50}$ values all bound similarly and occupied the ATP binding pocket, as well as an adjacent pocket usually occupied by Phe169 (of the DFG motif), displacing it from this pocket (Fig. 3a, c-d, and Table 2). This mode of binding has been termed as “DFG-out” and was first observed in the complexes of c-Abl with Gleevac (Schindler et al., 2000) and MAPK14 with BIRB-796 (Pargellis et al., 2002). In contrast, compounds displaying micromolar IC$_{50}$ values bound in a similar manner, but did not displace Phe169 from the pocket (Fig. 3b,e,f), binding to MAPK13 in what is known as “DFG-in” mode (Table 2). Thus, in the case of the MAPK13 inhibitors presented here, it appears that binding mode creates a marked difference in potency. However, it should be noted that highly potent inhibitors that bind to MAPK14 in DFG-in mode have been identified (e.g., skepinone-L (Koeberle et al., 2012) and SB203580 (Wang et al., 1998)), so binding mode does not always dictate efficacy.


To understand the structural basis for the observed binding modes exhibited by the different MAPK13-inhibitor complexes, we performed a comprehensive analysis of the binding pockets occupied by each inhibitor. The inhibitors analyzed here have similar designs, consisting of a heterocyclic ring that interacts with the hinge region of MAPK13, a central aromatic that occupies the ATP-binding site, and a bulky (mostly hydrophobic) group, with these units spanned with by various polar linkers (Fig. 1a). The binding pockets were very similar, with similar numbers of van Der Waals contacts and hydrogen bonds to MAPK13. We found that, compared to the micromolar inhibitors, the nanomolar inhibitors engaged in one additional hydrogen bond with MAPK13 (3 versus 4) (Fig. 4c). All inhibitors engaged in hydrogen bonds with the backbone amide nitrogens of Met110 (in the hinge) and Asp168 (in the...
DFG motif) (Fig. 4a,b,d,e). The main difference was seen in interactions involving the side-chain carboxylic acid of Glu72. This side-chain engages in a hydrogen bond with the linker between central aromatic and bulky hydrophobic groups in these inhibitors. The nanomolar inhibitors all contain urea linkers whereas the micromolar inhibitors both contain amide linkers. The second hydrogen bond made possible with the urea linkages (red arrow, Fig. 4a,b) assists to lock the bulky hydrophobic groups in relative conformations that require the moieties to bind to the Phe pocket, thus making these inhibitors selective for the DFG-out conformation of the kinase. Biphenyl amide inhibitors similar to compounds 117 and 124 have previously been observed to bind to both DFG-in and DFG-out conformations of MAPK14 and bind in similar modes to those observed in our MAPK13 structures (Angell et al., 2008), thus the inclusion of such groups does not dictate that these inhibitors can only bind in DFG-in modes.

3.5. Sequence comparison of inhibitor binding pockets in mammalian MAPK13 proteins.
Given that numerous animal systems are used for in vitro studies, we carried out analysis of MAPK13 sequences from higher mammals to examine whether there were any gross differences in the MAPK13 binding pockets for the inhibitors characterized here. The MAPK13 sequences for the species analyzed here averaged greater than 90% identity (Fig. 5). Of the numerous residues lining the binding pocket for the inhibitors, only one displayed a very conservative variation. Residue Phe109 (in the hinge region) is a Tyr in chicken MAPK13 (Fig. 5). Structure-based analysis of the inhibitor binding pocket sequence conservation would suggest that the inhibitors analyzed here should bind similarly to MAPK13 proteins in common animal models.

Previous studies have demonstrated the use of differential scanning fluorimetry as a method for identifying potential kinase inhibitors based on the enhanced thermal stability upon complexation (Niesen et al., 2007, Vedadi et al., 2006). A general observation has been that more potent kinase inhibitors lend greater thermal stability to the protein (Fedorov et al., 2012). We decided to analyze our diverse potency inhibitors using this method. We carried out experiments with both the inactive
(unphosphorylated) MAPK13 and active (dually-phosphorylated) MAPK13/pTpY, as well as with MAPK14. First, we carried out a titration on MAPK13 to establish a dynamic range for the inhibitors (Fig. 6). Based on this result, we carried out subsequent experiments with 25 µM inhibitor. The differences in melting temperatures using the two forms of MAPK13 were similar, indicating that the inhibitors effectively bound both forms of the protein (Table 3). The melting point shifts were about a degree higher for MAPK13/pTpY, likely due to the fact that this form is stabilized by contacts involving the phosphorylated activation loop, (Yurtsever et al., 2015) which would be dislodged by the DFG-out inhibitors. We found that the micromolar inhibitors (compounds 117 and 124) induced small increases in melting temperatures of MAPK13 (1-2° C). In contrast, the nanomolar inhibitors all induced large shifts in melting temperatures of MAPK13 (7-8° C). In addition, we found that our inhibitors induced larger T_m shifts in MAPK13 than MAPK14 (by a difference of 2-3 °C). For comparison, we also ran the DSF experiments with the potent MAPK14 inhibitor BIRB-796, which is optimized for MAPK14 (IC_{50} [MAPK14] = 4 nM; IC_{50} [MAPK13] = 520 nM) (Kuma et al., 2005). We found that the melting point shifts corresponded with the known differences in potencies (DTm [MAPK14] = 13.7 °C; [MAPK13] = 5.9 °C). Taken together, these results suggest our best inhibitors characterized here display slight selectivity for MAPK13 versus MAPK14, which should improve with optimization. Thus, the inhibitors which displayed the greatest potency and longest complex half-lives induced the greatest shift in thermal stability, indicating that this could be a robust screening method to presage inhibitor potency for MAPK13, and selectivity versus MAPK14.

3.7. Computational analysis of inhibitor complexes.

We carried out computational docking studies to gain further insight into why subsets of inhibitors bind to the DFG-in conformation versus the DFG-out conformation of MAPK13. In these calculations, the inhibitors were introduced as SMILES strings (with no predefined conformation) and were flexible during docking. The inhibitors were docked to MAPK13 coordinates in both the DFG-in and DFG-out
mode. For compounds 58 and 61, which bind in the DFG-out mode, the calculations reproduced the crystal structures extremely accurately when docked to the MAPK13 DFG-out coordinates (Table 4). However, they showed high variation and poor docking scores when docked to MAPK13 in DFG-in coordinates. In contrast, for compounds 117 and 124, which bind in DFG-in mode, the calculations accurately reproduced the crystal structure coordinates for the inhibitors regardless of whether MAPK13 was in DFG-in or DFG-out mode (Table 4). In other words, regardless of whether or not the Phe pocket was available for inhibitor moiety binding, the DFG-in inhibitors preferred to bind to the same pockets as observed in the crystal structure. These calculations indicate that this set of inhibitor structures can be predicted fairly accurately using GLIDE, and imply that part of the selective driving force is goodness of fit to the available pockets and the inherent structure preferred by inhibitor. To analyze this further, we carried out the same calculation using MAPK14 in both DFG-in and DFG-out mode as the receptor for these inhibitors. With the MAPK14 structures, we obtained similar results as with MAPK13, and the simulations reproduced inhibitors binding configurations observed in the MAPK13 co-crystal structures. Thus, regardless of the receiver kinase, the inhibitors take on configurations observed in the crystal structures, indicating that the major driving forces for this set of inhibitors is chemically-encoded structural constraints of the compound and shape complementarity to available pockets.

4. Conclusions

Here we present the functional, biophysical, and structural characterization of two new inhibitor complexes with MAPK13, as well as a comprehensive functional, biophysical, and structural analysis of all four of our inhibitor complexes. These represent the only four MAPK13 inhibitor complexes structurally characterized to date. The inhibitors examined here displayed widely varying potencies, with IC$_{50}$ values in either the nanomolar range, or, greater than 1000-fold less potent, in the micromolar range. These potencies correlated well with small-molecule binding kinetics measured by BLI. The
nanomolar-potency inhibitors displayed complex half-lives about 100 times longer than those of the micromolar inhibitors. It should be noted that the half-lives measured here using BLI are comparable to those of MAPK14-inhibitor interactions measured by the comparable flow-based method of SPR (Angell et al., 2008). Co-crystal structures of MAPK13 in complex with each of the four inhibitors revealed that the high potency inhibitors all bound to the kinase in DFG-out configuration, whereas the lower potency inhibitors both bound in DFG-in configuration. This is worth noting, since high potency inhibitors can be developed for either mode, for example skepinone-L (Koeberle et al., 2012) and SB203580 (Wang et al., 1998) are highly potent MAPK14 inhibitors that bind in DFG-in mode. Slow inhibitor dissociation rates have emerged as an important parameter for sustained efficacy of kinase inhibitors. In general, slower kinetics have been observed for inhibitors binding to the DFG-out mode of kinases, possibly due to kinetic constraints caused by the necessary structural reconfiguration (Gruenbaum et al., 2009, Pargellis et al., 2002). However, the structural mechanisms that influence inhibitor binding kinetics are poorly understood, thus comprehensive structural and kinetic studies such as those presented here are quite valuable in creating these correlations.

Detailed analysis of the binding pockets for the inhibitors revealed that all inhibitors, regardless of potency, interacted with MAPK13 in a very similar fashion, forming hydrogen bonds with the same residues in all four structures. The high potency inhibitors all engaged in one additional hydrogen bond, but computational docking and chemical reasoning indicate that this is mostly driven by the preferred lowest energy configuration of the inhibitor, which is intrinsic to its structure, and the goodness of fit of that conformation to the available pockets. It is worth noting that one of the development points of the GLIDE software has been to predict accurate kinase complexes with inhibitors (Lyne et al., 2006). The complexes with MAPK13 characterized here were readily predicted and suggest that this software approach could be used effectively for in silico screening to identify new chemical families with the potential to bind MAPK13.
DSF has been shown previously to be an effective method for screening for potential binders and inhibitors of kinases, and the thermal melting shifts usually correlate well with affinity (Fedorov et al., 2012). We found this to be the case with the set of inhibitors here, with the thermal shifts also correlating with binding mode (DFG-in versus DFG-out). These thermal shifts appear to correlate with binding affinities of the inhibitors, as observed in our titration experiments. These experiments do suggest, however, that this method could be utilized effectively to identify new inhibitors for MAPK13 by screening and to prioritize ranking based on magnitude of melting point stabilization. These experiments also suggest this method could be used to screen for inhibitors that selectively target MAPK13 versus MAPK14. The inhibitors analyzed here have not been optimized yet for potency or selectivity, yet they do display some preference for binding MAPK13 in our DSF experiments. Combining our structural studies of inhibitor binding done here with sequence comparison of MAP13 versus MAPK14 (Yurtsever et al., 2015), the most sequence divergent area that contacts the inhibitors is in the hinge region, thus this will be focus of our future iterations of inhibitor design.

Altogether, the functional, structural, and biophysical characterizations presented here provide the first comprehensive analysis of this type for inhibitor complexes with MAPK13, elucidating the structural basis for their mechanism of action, thus providing a knowledge base for the structure-based development of potent, and eventually specific, inhibitors.

Acknowledgements

This work was supported in part by funding from NIH R01-HL119813 (T.J.B), CADET P50-HL107183, CADET II UH2-HL123429 (M.J.H.), American Heart Association Predoctoral Fellowships PRE19970008 (Z.Y.) and PRE22110004 (D.L.K.), and a pre-doctoral training grant T32-GM007067 (D.L.K.). Results were derived from work performed at Argonne National Laboratory (ANL) Structural Biology Center and the Advanced Light Source, Berkeley CA (ALS). ANL is operated by U. Chicago Argonne, LLC, for the U.S. DOE, Office Biological and Environmental Research (DE-AC02-06CH11357). ALS is supported by the Office of Basic Energy Sciences of the
**PDB References:** mitogen-activated protein kinase 13 (p38δ), complex with compound 58 5ekn; complex with compound 117, 5eko.

**Figure 1**
Schematic of MAPK13 inhibitors and their IC_{50} values. (a) Schematic of MAPK13 inhibitor compounds characterized in this manuscript. (b) Data and curve fit used to determine IC_{50} values for compounds 58 and 117. (c) IC_{50} values, 95% confidence range, and coefficients of determination (R^2). * denotes values were previously reported (Alevy et al., 2012).

**Figure 2**
Biolayer interferometry analysis of MAPK13 binding to inhibitors. (a-b) Biolayer interferometry sensorgrams for (a) compound 58 and (b) compound 117. Processed data are shown as black lines with global kinetic fits overlaid as red lines. (c) Table of MAPK13-inhibitor complex half-lives derived from the global kinetic fits of k_{association} and k_{dissociation} \((t_1/2 = \ln 2/k_{dissociation})\). * denotes values were previously reported (Alevy et al., 2012).

**Figure 3**
Structures of MAPK13 complexes with inhibitors. (a) Ribbon diagram of MAPK13 (grey) with DFG-out inhibitors 58 and 61. Hinge and activation loop are both colored red. (b) Ribbon diagram of MAPK13 (grey) with DFG-in inhibitors 117 and 124. Hinge and activation loop are both colored red. (c) Crystal structure of compound 58 (green) in complex with MAPK13 in DFG-out mode. Omit Fo–Fc electron density contoured at 3.0 σ is shown for the compound and 2Fo – Fc electron density contoured at 2.0 σ is shown for Phe169. (d) Our previously reported crystal structure of compound 61 (cyan) in complex with MAPK13 in DFG-out mode for comparison (4EYJ). (e) Crystal structure of compound 117 (yellow) in complex with MAPK13 in DFG-in mode. Omit Fo – Fc electron density contoured at 3.0 σ is shown for the compound and 2Fo – Fc electron density contoured at 2.0 σ is shown for Phe169. (f) Our previously reported crystal structure of compound 124 (magenta) in complex with MAPK13 in DFG-in mode for comparison (4EYM). The dynamic positioning of Phe169 of the DFG motif is shown in all complexes.

**Figure 4**
MAPK13 binding pockets for inhibitor compounds. Binding pockets for (a) compound 58, (b) compound 61, (d) compound 117, and (e) compound 124. All residues that make van der Waals or polar contacts are shown. Hydrogen bonds are shown as thin black lines. One hydrogen bond that is unique to compounds 58 and 61 as compared to 117 and 124 is highlighted with the red arrow. (c) Table of polar and non-polar contacts each compound makes with MAPK13.

**Figure 5**
Sequence alignment of MAPK13 from relevant disease model mammals. Invariant and conserved residues are highlighted with magenta and yellow backgrounds, respectively. Secondary structure of human MAPK13 is shown above the alignment. The hinge region, DFG motif, gatekeeper (G), dual phosphorylation site (P), and activation loop are shown. Residues that form the binding pockets for inhibitor compounds are marked (black I). Figure was generated using ESPript3.0 (Gouet et al., 2003).
Figure 6
Changes in MAPK13 melting temperatures (ΔT_m) induced by binding inhibitors as measured by DSF. Titration of inhibitors is shown while MAPK13 concentration was held at 0.15 mg/ml. The point at each concentration is an average of three measurements. Error bars are SEM. Curve fitting done in PRISM 5 using the one site total binding equation for nonlinear regression. Calculated Kd’s listed in Table 3.
<table>
<thead>
<tr>
<th></th>
<th>Compound</th>
<th>58</th>
<th>117</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection statistics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space Group</td>
<td>P21212</td>
<td>P21212</td>
<td></td>
</tr>
<tr>
<td>Unit Cell (Å)</td>
<td>a=61.0</td>
<td>a=61.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>b=70.2</td>
<td>b=69.7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c=92.9</td>
<td>c=93.1</td>
<td></td>
</tr>
<tr>
<td>Source</td>
<td>ALS 4.2.2</td>
<td>APS 19-ID</td>
<td></td>
</tr>
<tr>
<td>Wavelength(Å)</td>
<td>1.0000</td>
<td>0.9790</td>
<td></td>
</tr>
<tr>
<td>Resolution(Å)</td>
<td>50.0-2.60 (2.69-)</td>
<td>50.0-2.00 (2.07-)</td>
<td></td>
</tr>
<tr>
<td>R&lt;sub&gt;sym&lt;/sub&gt;(%)</td>
<td>0.091(0.508)</td>
<td>0.055(0.369)</td>
<td></td>
</tr>
<tr>
<td>Completeness(%)</td>
<td>88.5 (100)</td>
<td>81.2(87.8)</td>
<td></td>
</tr>
<tr>
<td>Redundancy</td>
<td>7.0(7.3)</td>
<td>4.6(4.4)</td>
<td></td>
</tr>
<tr>
<td>I/σ(I)</td>
<td>20.6(3.6)</td>
<td>21.6(2.5)</td>
<td></td>
</tr>
<tr>
<td><strong>Refinement statistics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R&lt;sub&gt;work&lt;/sub&gt;</td>
<td>0.218</td>
<td>0.200</td>
<td></td>
</tr>
<tr>
<td>R&lt;sub&gt;free&lt;/sub&gt;</td>
<td>0.273</td>
<td>0.241</td>
<td></td>
</tr>
<tr>
<td>Amino Acid Residues(#)</td>
<td>340</td>
<td>340</td>
<td></td>
</tr>
<tr>
<td>Waters (#)</td>
<td>23</td>
<td>213</td>
<td></td>
</tr>
<tr>
<td>RMSD bond length (Å)/angles(°)</td>
<td>0.003/0.667</td>
<td>0.004/0.904</td>
<td></td>
</tr>
<tr>
<td>Wilson B (Å&lt;sup&gt;2&lt;/sup&gt;)</td>
<td>48.4</td>
<td>31.3</td>
<td></td>
</tr>
<tr>
<td>Ave. B. protein</td>
<td>52.2</td>
<td>37.0</td>
<td></td>
</tr>
<tr>
<td>Ave. B inhibitor</td>
<td>40.5</td>
<td>31.3</td>
<td></td>
</tr>
<tr>
<td>Ave. B water</td>
<td>48.1</td>
<td>42.7</td>
<td></td>
</tr>
<tr>
<td>Ramachandran</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>%Favored</td>
<td>87.4</td>
<td>95.3</td>
<td></td>
</tr>
<tr>
<td>%Allowed</td>
<td>12.0</td>
<td>4.4</td>
<td></td>
</tr>
<tr>
<td>%Outliers</td>
<td>0.6</td>
<td>0.3</td>
<td></td>
</tr>
<tr>
<td>Poor rotamers (no. / %)</td>
<td>12 / 3.99%</td>
<td>5 / 1.6%</td>
<td></td>
</tr>
<tr>
<td>Clashscore (score / percentile)</td>
<td>6.72 / 99°</td>
<td>6.10 / 95°</td>
<td></td>
</tr>
<tr>
<td>Molprobity (score / percentile)</td>
<td>2.16 / 94°</td>
<td>1.83 / 85°</td>
<td></td>
</tr>
<tr>
<td>Luzzati Error</td>
<td>0.35</td>
<td>0.24</td>
<td></td>
</tr>
<tr>
<td>PDB ID</td>
<td>5EKN</td>
<td>5EKO</td>
<td></td>
</tr>
</tbody>
</table>

Values in parenthesis are for the highest-resolution shell.
Table 2. Backbone torsion angles that define the conformation of the MAPK13 DFG motif

<table>
<thead>
<tr>
<th>Compound</th>
<th>DFG conformation</th>
<th>Leu 167 $\psi$ (°)</th>
<th>Asp 168 $\psi$ (°)</th>
<th>Phe 169 $\psi$ (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>58</td>
<td>DFG-out</td>
<td>-124.5</td>
<td>170.8</td>
<td>-154.0</td>
</tr>
<tr>
<td>61</td>
<td>DFG-out</td>
<td>-119.6</td>
<td>169.4</td>
<td>-144.7</td>
</tr>
<tr>
<td>117</td>
<td>DFG-in</td>
<td>-107.8</td>
<td>156.4</td>
<td>62.5</td>
</tr>
<tr>
<td>124</td>
<td>DFG-in</td>
<td>-111.3</td>
<td>160.9</td>
<td>55.5</td>
</tr>
</tbody>
</table>

Table 3. Inhibitor-induced thermal stability shifts measured by differential scanning fluorimetry (DSF)

<table>
<thead>
<tr>
<th>Compound</th>
<th>$\Delta T_m$-MAPK13 (°C)</th>
<th>$\Delta T_m$-MAPK13/pTπY (°C)</th>
<th>$\Delta T_m$-MAPK14 (°C)</th>
<th>$K_d$ (µM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>58</td>
<td>8.1 ± 0.3</td>
<td>9.0 ± 0.4</td>
<td>5.2 ± 0.3</td>
<td>1.9 ± 0.3</td>
</tr>
<tr>
<td>61</td>
<td>7.2 ± 0.3</td>
<td>7.8 ± 0.4</td>
<td>5.1 ± 0.3</td>
<td>2.6 ± 0.5</td>
</tr>
<tr>
<td>117</td>
<td>1.8 ± 0.4</td>
<td>2.0 ± 0.5</td>
<td>-1.4 ± 0.4</td>
<td>N.D.***</td>
</tr>
<tr>
<td>124</td>
<td>1.1 ± 0.3</td>
<td>1.2 ± 0.5</td>
<td>-0.8 ± 0.3</td>
<td>N.D</td>
</tr>
<tr>
<td>BIRB-796</td>
<td>5.9 ± 0.2</td>
<td>13.7 ± 0.3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* $\Delta T_m$ are reported at inhibitor concentrations of 25 µM and kinase concentrations at 2.5 µM

** Derived from titrations shown in Fig. 6

***Not determined

Table 4. GLIDE in silico docking results compared to co-crystal structures

<table>
<thead>
<tr>
<th>Compound</th>
<th>Score*</th>
<th>RMSD**</th>
<th>Score</th>
<th>RMSD</th>
<th>Score</th>
<th>RMSD</th>
<th>Score</th>
<th>RMSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAPK13 DFG-in</td>
<td>-5.34</td>
<td>10.15</td>
<td>-5.21</td>
<td>9.99</td>
<td>-9.81</td>
<td>1.07</td>
<td>-6.14</td>
<td>2.00</td>
</tr>
<tr>
<td>MAPK14 DFG-out</td>
<td>-9.25</td>
<td>1.34</td>
<td>-8.70</td>
<td>1.10</td>
<td>-7.96</td>
<td>1.92</td>
<td>-4.20</td>
<td>12.71</td>
</tr>
</tbody>
</table>

* Glide Score in kcal/mol

** RMSD of compound heavy atom positions compared to crystal structure (Å)
Figure 1. Yurtsever, et al.
Figure 2. Yurtsever, et al.
Figure 3. Yurtsever, et al.
Figure 4. Yurtsever, et al.
Figure 5. Yurtsever, et al.
Figure 6. Yurtsever, et al.
References:


Zeynep Yurtsever

Personal phone: 314-882-2378
E-mail: zyurtsever2@gmail.com

LinkedIn profile: https://www.linkedin.com/in/zeynep-yurtsever-9729041b

Professional Profile:


Highlights

• First manuscript detailing dissertation work was selected Paper of the Week.

• Featured as an author in Journal of Biological Chemistry based on the first author publication.

• Received a nationally competitive two-year pre-doctoral fellowship from American Heart Association for Ph.D. dissertation project.

• Developed in vitro and cellular assays to probe protein-protein interactions.

Education

Ph.D. Biochemistry
Washington University in St. Louis, MO
GPA: 3.57
May 2016

B.S. Chemistry (Honors)
Koc University, Istanbul, Turkey
GPA: 3.82
June 2011
**Research Experience**

**Graduate Research Assistant**, 2011-2016  
Washington University in St. Louis  
Ph.D. Dissertation “Mechanism of Calcium-dependent Chloride Channel Activation by the Secreted Regulator CLCA1”.  
Mentors: Dr. Tom Brett, Dr. Colin Nichols

**Undergraduate Research Assistant**, 2010-2011  
Koc University, Istanbul, Turkey  
Characterized hydrogen bonding patterns between aspirin and COX-1 using DFT and MP2 calculations.  
Mentors: Dr. Burak Erman, Dr. Ersin Yurtsever

**Summer Research Intern**, 2010  
University of Heidelberg, Germany  
Carried out microfluidic experiments to assess cell adhesion on self-assembled monolayer coated surfaces.  
Mentor: Dr. Axel Rosenhahn

**Undergraduate Research Assistant**, 2007-2011  
Koc University, Istanbul, Turkey  
Synthesized and characterized properties of CdS/ZnS quantum dots, optimized photoluminescent properties via Mn doping strategies, assessed effect of PEI coating on toxicity and solubility  
Mentor: Dr. Funda Yagci Acar

**Teaching Experience**

**Graduate Teaching Assistant**, 2012  
Washington University in St. Louis  
Course: Chemistry and Physics of Biological Molecules  
Course master: Dr. Katherine Henzler-Wildman  
Responsibilities: problem sessions, hands-on tutorials, review sessions, quiz and exam grading, office hours

**Undergraduate Teaching Assistant**, 2009  
Koc University, Istanbul, Turkey  
Course: Calculus for Science and Engineering Majors  
Course master: Dr. Mehmet Saridereli  
Responsibilities: Problem sessions, review sessions, quiz and exam grading
**Honors and Awards**

American Heart Association Pre-doctoral Fellowship                              2014-2016
Division of Biology and Biomedical Sciences Merit Award                       2014-2016
Vehbi Koc Scholar Award (semi-annually)                                        2007-2011

**Publications  (*co-first author) († Paper of the Week)**

**Yurtsever, Z.** Sala-Rabanal, M. Nichols, C.G, Brett, T.J. “Mechanism of CLCA1-TMEM16A interaction: Mapping the Minimal Interacting Domains”. (manuscript in prep for Molecular Cell)


**Abstracts** (*pre-doctoral poster prize*)


*Yurtsever, Z.* Sala-Rabanal, M. Nichols, C.G, Brett, T.J. “Novel Regulation of TMEM16A by Secreted CLCA1 in Human Cells”. Abstract for poster presentation, Biochemistry Department Retreat 2015, St. Louis, MO.


Yurtsever, Z. “Solving a 20-year old Mystery: How CLCA1 Regulates Ca-activated Chloride Currents in Human Cells”. Abstract for oral presentation, Biochemistry Department Retreat 2014, St. Louis, MO.

*Yurtsever, Z.* Sala-Rabanal, M. Nichols, C.G, Brett, T.J. “Structural and Biochemical Studies of CLCA1: A Novel Mechanism of CaCC Activation by a Secreted Protein”. Abstract for poster presentation, Annual Koster Symposium and CIMED Research Day 2014, St. Louis, MO.


**Memberships**

American Heart Association
American Society of Cell Biology