Protein semisynthesis underscores role of a conserved lysine in activation and desensitization of acid-sensing ion channels

Debayan Sarkar1§, Iacopo Galleano1,2§, Sofie Yuewei Ou1§, Gül Refika Uzun1§, Keith K. Khoo1, Gerbrand Jan van der Heden van Noort2, Stephan Alexander Pless1*

1Department of Drug Design and Pharmacology, University of Copenhagen, 2100 Copenhagen, Denmark
2Department of Cell and Chemical Biology, Leiden University Medical Centre, Leiden, The Netherlands

§§ These authors contributed equally

* Co-corresponding author:
Stephan Pless
Department of Drug Design and Pharmacology
University of Copenhagen
Jagtvej 160, 2100 Copenhagen, Denmark
Mail: stephan.pless@sund.ku.dk
SUMMARY

Acid-sensing ion channels (ASICs) are trimeric ligand-gated ion channels that open a cation-conducting pore in response to proton-binding to their extracellular domain (ECD). This activation is typically short-lived, as channel opening is followed by a rapid and spontaneous pore closure due to a process termed fast desensitization. This intrinsic channel closing mechanism is important, because excessive activation of ASICs during periods of prolonged acidosis in conditions such as inflammation and ischemia, has been linked to several pathologies, including pain and stroke. Structural and functional data suggest that a conserved lysine in the ECD (Lys211 in mASIC1a) plays a role in receptor function. However, the precise physico-chemical contributions to activation and fast desensitization are difficult to dissect with conventional mutagenesis, as replacement of Lys211 with naturally occurring side chains invariably changes multiple parameters, such as size, charge, and H-bonding pattern simultaneously.

Here, we study the contribution of Lys211 to activation and fast desensitization of mASIC1a using tandem protein-trans splicing (tPTS) to incorporate non-canonical lysine analogs. We conduct extensive optimization efforts to improve apparent splicing yields and subsequently interrogate the function of semisynthetic mASIC1a using two-electrode voltage clamp. Our data show that both charge and side chain length of Lys211 are crucial to activation and fast desensitization of mASIC1a, with the latter being sensitive to even the most subtle mutational disruptions. Our work emphasizes that tPTS can enable atomic-scale interrogations of how side chains contribute to the function of complex membrane proteins.
INTRODUCTION

Acid-sensing ion channels (ASICs) are trimeric ligand-gated ion channels that open a cation-selective pore in response to proton binding to their extracellular domain (ECD). They contribute to fast synaptic neurotransmission and have been associated with pathological conditions of both the peripheral nervous system (e.g. pain, diabetic neuropathy) and the central nervous system (e.g. stroke, fear, neurodegenerative diseases). Consequently, ASICs are emerging as possible drug targets for both small molecules and peptide-/animal toxin-derived lead molecules.

Structural evidence has confirmed that each ASIC subunit contains intracellular N- and C-terminal domains and a transmembrane domain (TMD) encompassing two transmembrane helices that are linked to the large ECD. Binding of protons to the ECD triggers a series of conformational changes that lead to opening of the channel pore. This transient pore opening is typically followed by a process termed fast desensitization, during which the channel enters a non-conductive state in the continued presence of high proton concentrations. Alternatively, exposure to less elevated proton levels can evoke conformational changes that evoke a non-conductive state without prior pore opening. This process is termed steady-state desensitization (SSD).

Previous functional and structural work has identified Lys211 (mASIC1a numbering) to play an important role in channel gating in different ASIC isoforms, as well as across ASICs from different species, likely because it forms a H-bond network with Leu351 and Asp355 of the neighboring subunit (Figure 1A). Additionally, Lys211 has been suggested to contribute to a state-dependent binding site of Cl- ions. Together with Arg309 and Glu313 from the adjacent subunit (Figure 1A), this anion binding site was implicated in the process of fast desensitization and had been suggested to affect acid-induced cell death.

However, conventionally employed mutagenic approaches, such as side chain deletion or substitution by alanine, are likely to cause global disruptions and can therefore complicate data interpretation. This is especially true for side chains situated at the functionally important subunit interface of ASICs or the related epithelial sodium channels, ENaCs. A high-resolution dissection of a possible contribution to channel activation and fast desensitization of Lys211 therefore requires the incorporation of Lys analogs with subtle changes in side chain length or pKa.

A widely applied approach to incorporate such non-canonical amino acid (ncAA) analogs into membrane proteins is by genetic code expansion. However, protein yields, non-specific incorporation or lack of sufficiently effective discrimination by the amino-acyl tRNA synthetase between ncAA side chains with similar physico-chemical properties can be challenging. In some cases, these issues can be overcome by alternative approaches such as protein semisynthesis using protein trans-splicing (PTS). PTS involves formation of a peptide bond between two proteins or protein fragments (termed as exteins) by means of traceless native chemical ligation mediated by a split intein. Split inteins consist of an N-terminal part (IntN) and a C-terminal part (IntC), which spontaneously recombine to form an active intein.
the protein, the approach can be expanded to tandem protein trans-splicing (tPTS), in which two orthogonal split inteins are used to insert a synthetic peptide into a protein.19,28,29 This strategy has been successfully used to incorporate ncAAs in the ECD of the ATP-gated P2X2 receptor, in addition to incorporating post-translational modifications in intracellular linkers of the cardiac sodium channel NaV1.5.29,30 tPTS requires the N- and C-terminal fragments of the ion channel to be expressed in a cell by recombinant means, while the synthetically produced peptide (Pept X^{SYN}) can be introduced into the cell by micro-injection or cell squeezing, depending on cell type.29 Here, we aim to insert a recombinant (Pept X^{REC}) or a Pept X^{SYN} into the mASIC1a ECD using the strategy outlined in Figure 1B. In the recombinantly expressed N-terminal and C-terminal transmembrane constructs (N-TMD and C-TMD), split intein Int_{N}-A is placed at the C-terminus of the N-TMD fragment, whereas the orthogonal Int_{C}-B is positioned at the N-terminus of the C-TMD fragment. The latter contains a faux transmembrane helix (faux-TMD) that is introduced to preserve the correct topology of the C-TMD fragment in the membrane during protein trafficking (28,29, Figure 1B). To reconstitute full-length mASIC1a using tPTS, Int_{C}-A is situated at the N-terminus of peptide X, while Int_{N}-B is attached to its C-terminus.

With a goal to establish and optimize tPTS to incorporate ncAA Lys analogs at position 211 in mASIC1a, we systematically screened 11 orthogonal split intein pairs across two different splice sites, conducted extein sequence optimization, and evaluated the potential of covalently linking N-TMD and C-TMD constructs. Our results, based on reconstitution of WT mASIC1a (i.e. using Pept X^{REC}), establish the orthogonal CfuDnaESspDnaB^{M86} split intein pair as ideal for tPTS in the mASIC1a ECD. We then generated semisynthetic mASIC1a variants by inserting four ncAA analogs of Lys via Pept X^{SYN} at position 211 and provide direct evidence for the importance of side chain charge and length in determining both activation and fast desensitization properties of mASIC1a. Overall, we demonstrate the potential of tPTS for the evaluation of functional properties of ASICs with atomic resolution.

Figure 1: A semisynthetic strategy to probe the contributions of Lys211 to the Cl-binding site in the ASIC ECD. (A) Side view of a cASIC1 channel (PDB:2qts). Subunit in front is shown in violet cartoon representation with the Lys211 side chain highlighted as salmon-colored spheres and the two remaining subunits are depicted with light grey and light green surfaces.
The top inset shows the proposed H-bond network between the Lys211 amine and the main chain carbonyl oxygen of Leu351 and side chain carboxylate of Asp355 \(^{6,11}\). The bottom inset shows the Cl\(^{-}\)-binding pocket in cASIC1 (PDB: 2qts), in which the Cl\(^{-}\) ion is predicted to be coordinated by the side chains of Lys211, Arg309 and Glu313. (B) Schematic overview of the strategy to reconstitute full-length mASIC1a. This is achieved by recombinantly expressing two separate mASIC1a fragments comprising the first transmembrane helix and the N-terminal part of the ECD (N-TMD) and the C-terminal part of the ECD and the second transmembrane helix (C-TMD; note the faux transmembrane domain (Faux TMD) at its N-terminus to ensure correct topology) in the presence of a recombinantly expressed or synthetically prepared peptide X in Xenopus laevis oocytes (Pept X\(^{REC}\) and Pept X\(^{SYN}\), respectively). Tandem protein trans-splicing in the presence of peptide X results in formation of full-length mASIC1a monomers that trimerize to form functional ion channels. Inteins A and B are indicated by square and round symbols, respectively.
RESULTS

Optimizing splice site location and choice of split intein pairs. Previous studies with orthogonal split inteins used for protein trans-splicing have indicated that the splicing efficiency and yield depend on several factors. These include presence of nucleophilic Cys, Ser or Thr side chains, location of splice site, amino acids sequence present in close proximity to the splice sites, and orthogonality among split intein pairs. We therefore set out to identify the optimal combination of splice site, orthogonal split intein pair and sequence context in order to replace Lys211 in mASIC1a.

First, we evaluated the sequence context around Lys211 for the presence of nucleophilic Cys, Ser or Thr side chains that would enable splicing with no or minimal alterations of the native mASIC1a sequence, depending on the nucleophile preference of different split inteins. This led to the identification of two similar but distinct splice site combinations: For splice site combination #1, we chose the N-TMD, peptide X (including Lys211) and C-TMD fragments to span mASIC1a residues 1-193, 194-213 and 214-526, respectively; those chosen for splice site combination #2 spanned residues 1-179, 180-213, and 214-526, respectively (Figure 2A). This enabled us to take advantage of the presence of either a Cys (Cys194) (splice site combination #1) or a Ser (Ser180) (splice site combination #2) at the N terminus of peptide X, to accommodate split inteins with either Cys- and Ser-dependent splicing mechanisms. The C-terminal splice junction at Thr214 was kept identical for both splice site combination #1 and #2 because the next Ser/Cys residues in mASIC1a sequence are situated over 25 amino acids further downstream, which would have necessitated a much longer peptide X. However, due to concerns over cross-reactivity with Thr-based split inteins such as TvoVM33, this required us to mutate Thr214 to Cys or Ser (depending on the split intein in use) to ensure efficient splicing.

Next, we generated mASIC1a constructs with four different split intein pairs at splice site combination #1 (CfaDnaE-SspDnaBM86, MjaKlbA-SspDnaBM86, NpuDnaE-SspDnaBM86, and MjaKlbA-VidaL) and seven different split intein pairs at splice site combination #2 (NrdJ1-SspDnaBM86, NrdJ1-SspDnaX, NrdJ1-VidaL, gp41-1-SspDnaX, gp41-1-CatInt, IMPDH-1-SspDnaX, and SspGyrB-SspDnaX). In all cases, we used the native mASIC1a sequence as the extein sequence context, except for a Thr to Ser/Cys mutation at position 214 position of the C-TMD fragments (see Supplemental Table T1 for sequence details). We then translated the corresponding N-TMD, Pept XREC and C-TMD fragments into mRNA and injected them into Xenopus laevis oocytes for recombinant expression. This was followed by measurements of whole-cell mASIC1a currents elicited by application of pH 6.0 over a period of 6 days using two-electrode voltage clamp (Supplemental Figure S1). For these recordings we made the following assumptions (see also 29): the absence of current would indicate no or incomplete splicing, while the presence of pH-induced currents would originate from fully reconstituted mASIC1a constructs, where the current amplitude is roughly indicative of the splicing efficiency.

As expected, expression of just the N-TMD and C-TMD fragments did not yield currents for any of the constructs (Figure 2B). This confirms that the N-TMD and C-TMD fragments do not yield functional channels in absence of Pept XREC. But when co-expressed with Pept XREC mRNA, we observed robust
currents for two of the split intein pairs at splice site combination #1 (CfaDnaE as split intein A, SspDnaB<sup>M86</sup> as split intein B; NpuDnaE as split intein A, SspDnaB<sup>M86</sup> as split intein B, respectively) and two of the split intein pairs at splice site combination #2 (NrdJ1 as split intein A, SspDnaB<sup>M86</sup> as split intein B and gp41-1 as split intein A, SspDnaX as split intein B, respectively) (Figure 2B). However, even for the most promising split intein combination (CfaDnaE as split intein A, SspDnaB<sup>M86</sup> as split intein B), currents rarely exceeded amplitudes of 10 µA, even after multiple days of incubation. This is significantly less than what can be achieved from expression of full-length WT mASIC1a, which typically yields >20 µA within about 24 hrs. We therefore sought to improve splicing yields and/or efficiency by optimizing the sequence context of the different split inteins tested here to more closely resemble their natural sequence context.
**Figure 2:** Evaluation of different split intein combinations and sequence context to optimize for highest current amplitudes. (A) Left panel: Structure of cASIC1 (same color-coding as in Figure 1A), with the part of the ECD to be inserted as Pept X highlighted in purple and Lys211 in salmon spheres. Right panel: Two splice site combinations were used in this study. The sequence margins of the N- and C-TMD fragments (violet) and the Pept XREC (purple) shown along with the intein A (square symbols) and B (round symbols) parts linked to them. (B) Left panel: Representative traces of mASIC1a currents from five different split intein combinations. Note that the extein sequence has not been altered beyond the Thr214 Ser/Cys substitution for the C-TMD constructs. Black bars indicate application of pH 6.0 to oocytes expressing only N- and C-TMD fragments (N + C, left) or in the presence of a recombinantly expressed Pept XREC using the indicated split intein combination (N + X + C, right). Scale bars: x, 10 s; y, 1 μA. Right panel: List of all split inteins combinations tested here categorized based on splice site combination, along with resulting maximal mASIC1a-mediated currents in *Xenopus laevis* oocytes elicited by application of pH 6.0 (n=6-10 for each combination). (C) Schematic of the trans-splicing reaction with the relevant extein sequence residues in N-TMD, C-TMD and peptide X denoted with -1, -2, -3, +1, +2, +3, which corresponded to positions that were altered to optimize splicing. In this example, the split intein sequences corresponding to CfaDnaE have been shown, where sequence optimization was required only for peptide X. The native (mASIC1a) sequence is shown in purple and optimized sequence is depicted in violet (left panel). (D) Daily progression of reconstituted mASIC1a current levels in oocytes compared between native (mASIC1a sequence) and optimized (*) extein sequences for two different combinations of split inteins (Intein A: CfaDnaE or NrdJ1; Intein B: SspDnaB<sup>M86</sup> (both)) (right panel) (n=6-11 for both combinations). Floating bars depict the mean along with minimum and maximum current margins.

**Extein sequence optimization did not enhance current amplitudes.** Several previous studies have demonstrated that the sequence contexts of the exteins, especially the -1, -2, -3, +1, +2, +3 positions around the split inteins can influence the efficiency and yields of splicing reactions of split inteins (see Figure 2C, left panel).<sup>21,20,33-36</sup> It is thus conceivable that the extein sequence context of the mASIC1a sequence around the splice sites impacts splicing efficiency. We aimed to address this potential issue by testing if mutating the (native) mASIC1a sequence to that of the split intein in its original context would improve apparent splicing yields.

To this end, we therefore introduced mutations in the N-TMD, Pept X<sup>REC</sup> or C-TMD constructs for all 11 split intein combinations to reflect the natural extein sequence context of each tested split intein, which we refer to as optimized sequences. Supplemental Tables T2 and T3 outline the mutations incorporated for each split intein combination. Notably, extein sequence optimization failed to elicit pH-induced currents in all seven split intein combinations that had already yielded either negligible or no mASIC1a currents with their natural sequence context (*MjaKlbA – SspDnaB<sup>M86</sup>*, gp41-1 – CatInt, *MjaKlbA – VidaL*, *NrdJ1 – SspDnaX*, IMPDH1 – *SspDnaX*, *NrdJ11 – VidaL* and *SspGyrB – SspDnaX*; see Supplemental Figure S1). Further, extein sequence-optimized constructs with the *NpuDnaE – SspDnaB<sup>M86</sup>* and gp41-1 – *SspDnaX* split intein combinations no longer elicited detectable pH-induced currents, suggesting a lack of functional channel reconstitution after sequence optimization. By contrast, sequence optimization of the *CfaDnaE - SspDnaB<sup>M86</sup>* split intein pair increased current amplitudes on day 2, although maximally observed current amplitudes (around day 5) were not affected (Figure 2D). Finally, the *NrdJ1 - SspDnaB<sup>M86</sup>* split intein pair showed increased current amplitudes on day 6, yet not exceeding current amplitudes obtained with the *CfaDnaE - SspDnaB<sup>M86</sup>* split intein pair.
In conclusion, our sequence optimization attempts proved largely unsuccessful, even detrimental, to splicing, as assessed by current amplitudes. However, the robust currents observed with the native mASIC1a extein sequence context of the CfaDnaE - SspDnaB<sup>M86</sup> split intein pair (except for the Thr214Ser mutation), made this particular split intein combination the most promising candidate for additional optimization attempts.

**Covalently linking N-TMD and C-TMD did not increase current amplitudes.** We next sought to test if a forced spatial proximity of the recombinantly expressed N-TMD and C-TMD fragments would increase splicing yields and thereby current amplitudes. To this end, we designed a construct in which N-TMD and C-TMD were linked by a 30-amino acid flexible linker, consisting of 6 repeats of a GlyGlyGlyGlyGlySer sequence stretch (Figure 3A; note that this design eliminates the need for a *Faux* TMD). Injecting mRNA for this construct (NC linked) and the corresponding peptide X<sup>REC</sup> fragment yielded pH-sensitive currents, but we did not observe currents in control experiments where only the mRNA for NC linked was injected (Figure 3B). However, maximal current amplitudes for NC linked + peptide X<sup>REC</sup> remained consistently lower than for those observed for the separately expressed N-TMD + X<sup>REC</sup> + C-TMD constructs (Figure 3C). We therefore chose the latter to proceed for a more detailed characterization, followed by insertion of synthetic peptides.

**Figure 3:** Forced proximity of N-TMD and C-TMD fragments does not appear to promote splicing. (A) Schematic overview of the strategy to reconstruct full-length mASIC1a from a single recombinantly expressed mASIC1a construct with N- and C-terminal fragments that were connected by a 30 amino acid (GGGGS)<sub>6</sub> linker, termed NC linked. Inteins A (*Cfa*DnaE) and B (*Ssp*DnaB<sup>M86</sup>) are indicated by square and round symbols, respectively. (B) Representative traces of currents from *Xenopus laevis* oocytes expressing only the NC linked construct or NC linked in presence of the Pept X<sup>REC</sup> fragment (NC linked + X<sup>REC</sup>) and subjected to application of pH 6.0. Scale bars: x, 10 s; y, 1 μA. (C) Daily progression of maximal pH 6.0-induced mASIC1a currents in oocytes separately expressing N + C (see Fig 2B) or the NC linked construct in the presence or absence of X<sup>REC</sup> (n=4-8 for each combination). Floating bars depict the mean along with minimum and maximum current margins.
Thr214Ser mutation leads to altered pH sensitivity of activation and SSD. Based on the above results, the CfaDnaE - SspDnaB<sup>M86</sup> split intein pair appeared to display the highest splicing yields. Because this split intein pair required the Thr214Ser mutation at the +1 position of the C-TMD fragment to enable SspDnaB<sup>M86</sup>-mediated splicing (no splicing was observed with the native Thr), we sought to investigate if this sequence alteration caused functional perturbations in mASIC1a function (Supplemental Figure S2). To this end, we first compared the pH sensitivity of activation and steady-state desensitization (SSD) of full-length wildtype mASIC1a channels to full-length mASIC1a channels containing the Thr214Ser mutation. This showed that the mutation itself caused a right-shift in both pH<sub>50</sub> of activation and SSD (6.7 ± 0.1 vs 6.5 ± 0.1 (p=0.0003) and 7.2 ± 0.0 vs 6.9 ± 0.0 (p<0.0001), respectively), suggesting an overall lower sensitivity to protons (Supplemental Figure S2B and Supplemental Table T4). We expected that recombinantly expressed N-TMD+X<sup>REC</sup>+C-TMD (referred to as N+X<sup>REC</sup>+C) constructs (with the CfaDnaE and SspDnaB<sup>M86</sup> split intein pair and hence the Thr214Ser mutation) would yield mASIC1a channels with properties similar to full-length Thr214Ser channels. However, we found the spliced channels to display a minor right-shift in the pH<sub>50</sub> of activation and a slight left-shift in the pH<sub>50</sub> of SSD (6.5 ± 0.1 vs 6.4 ± 0.1 (p=0.006) and 6.9 ± 0.0 vs 7.1 ± 0.0 (p<0.0001) for N+X<sup>REC</sup>+C and full-length Thr214Ser, respectively; see Supplemental Figure S2B and Supplemental Table T4). The reason for these small but significant changes remains unclear. The near-identical properties, however, for both full-length and recombinantly reconstituted mASIC1a channels containing the Thr214Ser mutation and the robust expression of the latter led us to next attempt insertion of synthetic peptides in order to assess the role of Lys211 in mASIC1a function.

ncAA lysine analogs at position 211 reveal importance of side chain length and charge to channel function. In order to assess the contributions of Lys211 side chain charge and length to mASIC1a function, we aimed to incorporate four different ncAA analogs of lysine: ornithine (Orn; protonated amino moiety like Lys, but one carbon shorter), homolysine (hLys; protonated amino moiety like Lys, but one carbon longer), hydroxyl nor-leucine (NleuOH; isosteric to Lys but uncharged) and thialysine (thiaLys; with a decrease in pKa of around one unit and a marginally altered bond geometry (bond angle around 90 degrees, instead of 109.5 degrees) compared to Lys).<sup>37</sup> We used solid phase peptide synthesis (SPPS) to generate synthetic Pept<sub>X</sub> versions (Pept<sub>X</sub><sup>SYN</sup>) containing either Lys, Orn, hLys, NleuOH or thiaLys in position 211 (see Supplemental information for details). For functional characterization of channels containing the Lys analogs, we injected Pept<sub>X</sub><sup>SYN</sup> into <i>Xenopus laevis</i> oocytes that were pre-injected with the N-TMD and C-TMD fragment mRNAs ~24h in advance.<sup>29</sup> Electrophysiological assessments of pH<sub>50</sub> of activation and SSD were performed 24-48 hrs after peptide injection because cells injected with synthetic peptides typically showed increased mortality beyond 48 hrs after peptide injection. Reassuringly, all constructs generated robust currents (Figure 4A) and control experiments in which a Lys-bearing synthetic peptide was inserted (Lys211<sup>SYN</sup>) showed near-identical pH<sub>50</sub> values for activation and SSD when compared to mASIC1a channels reconstituted from only recombinantly
expressed components N+XREC+C (6.4 ± 0.1 and 7.1 ± 0.0, respectively; Figure 4A/B and Supplemental Table T5). This was similar for Orn-containing peptide XSYN (Lys211OrnSYN), indicating that shortening of the side chain has little impact on channel function (activation pH50 = 6.2 ± 0.1; SSD pH50 = 7.1 ± 0.1). By contrast, increasing side chain length (Lys211hLysSYN), altering its pKa (Lys211thiaLysSYN) or removing the positive charge altogether (Lys211NlleuOHSYN) produced a drastic right shift in the activation pH50 (5.2 ± 0.1 for hLys and NlleuOH; 5.3 ± 0.1 for thiaLys), while showing no or relatively minor change in pH50 of SSD (7.0 ± 0.0 for hLys, 7.1 ± 0.1 for thiaLys and 7.2 ± 0.0 for NlleuOH) (Figure 4B and Supplemental Table T4). Together, this indicates that a shorter side chain length at Lys211 has little functional impact but increasing side chain length or altering/removing its charge is detrimental to proton sensitivity of activation and SSD.

**Lys substitutions at position 211 alter kinetics of fast desensitization.** A previous study demonstrated that disrupting the Cl− ion binding site predicted by the X-ray structure of cASIC1 (PDB ID: 2qts,6) (Figure 1A, bottom right panel) accelerates kinetics of fast desensitization: both removing the charge at Lys211 through an Ala mutation and replacement of Cl− by larger anions accelerated cASIC1 fast desensitization kinetics.12 Given the role predicted for Lys211 in Cl− ion binding, we set out to assess the kinetics of fast desensitization in mASIC1a channels containing the above ncAA analogs of Lys in PeptXSYN. To this end, we evaluated the fast desensitization kinetics (τdesensitization) at pH 5.5 (Figure 4C, left panel). Reassuringly, we found that the semisynthetic Lys-containing channels did not show significant differences in kinetics of fast desensitization compared with recombinant channels (N+XREC+C) (see Supplemental Table T5). By contrast, all ncAA analogs of Lys used here (Orn, hLys, thiaLys and NlleuOH) led to a significant acceleration in τdesensitization (Figure 4C, right panel and Supplemental Table T5). Our data demonstrate that even the most subtle side chain manipulations at Lys211 drastically enhanced the process of fast desensitization.
Figure 4: Functional characterization of mASIC1a channels trans-spliced with recombinant N and C terminal fragment and Pept X^{SYN}. (A) Representative traces in the top panel depicts the pH activation profiles on the left and SSD profiles on the right with the chemical structure of the ncAA replacing Lys211 shown at the center. The buffer perfusion protocols are shown on the top left (activation) or top right (SSD) of the panels to depict the duration of each pH steps, and the 20 sec exposure to buffer with activating proton concentrations shown as thickened black bar on the current traces. Time period of exposure to desensitizing buffer conditions is shown in grey bars (last 30 sec of 80 sec). Scale bars: x, 60 sec; y, 0.5 μA. (B) The activation (left) and SSD (right) curves of pH response for all the spliced mASIC1a channels with four different ncAAs at the Lys211 position (n=6-10). (C) Comparison of kinetics of fast desensitization for all five synthetic peptide substitutions with the recombinant spliced construct of the mASIC1a channel, N+X^{REC+C}T214S variant (N+X^{REC+C}). In left panel, N+X^{REC+C} activation trace at pH 5.5 is shown in black and all the traces (normalized to N+X^{REC+C} trace) corresponding to other ncAAs...
has been shown in other colors. Quantitative evaluation of the desensitization time constants for each construct have been shown in the graph on the right panel. Each sphere represents individual data points (n=6-8 for each). The floating bars depict the mean along with minimum and maximum current margins. Scale bars: x, 10 sec.
DISCUSSION

Structure-function studies have proven invaluable to advance our understanding of complex membrane proteins such as ion channels. However, progress regarding the contribution of evolutionarily highly conserved and/or functionally crucial side chains is frequently stymied by the limited chemical space afforded by conventional site-directed mutagenesis. Here we overcome this constraint by employing tPTS to incorporate a range of subtle Lys analogs into position Lys211 in mASIC1a. Our results underscore an important role of this side chain in both channel activation and fast desensitization.

The role of Lys211 in mASIC1a activation and fast desensitization. Previous studies have suggested for Lys211 to be involved in ASIC activation via a H-bonding network with the backbones of Leu351 and Asp355 from the neighboring subunit. This was based on the observations that the distance between Lys211 and Leu351 is decreased in the open state compared with the resting state of chicken ASIC1 (cASIC1) and that deleting Lys211 or covalently bridging the adjacent palm and thumb ECD subdomains by disulfide locking causes a right-shift in proton sensitivity. Here, we find that altering the pKa and altering side-chain geometry or removing the charge of Lys211 (thiaLys and NleuOH, respectively) is roughly as detrimental to mASIC1a activation, as is extending the side chain length by one carbon (hLys). By contrast, a shorter side chain is relatively well tolerated, as Orn incorporation in position 211 does not affect the proton sensitivity of activation.

On the other hand, we observed significantly accelerated desensitization kinetics with all Lys analogs incorporated at position 211. This suggests that the H-bonding network involved in activation is slightly more permissible towards minor disruptions than the process of fast desensitization. This notion is supported by the observation that the conventional Lys211Arg mutation displays near WT-like activation properties despite an increase in fast desensitization kinetics. Our data show that the process of fast desensitization is sensitive to even the most subtle of manipulations. The roughly 3-fold acceleration in $\tau_{\text{desensitization}}$ reported previously for the conventional Lys211Ala mutation is similar to the change we observe for the longer Lys analog hLys or those with altered or removed charge (NleuOH and thiaLys). These effects could be indicative of a disruption of Cl- ions binding near Lys211. Cl- ions have been proposed to bind ASICs in a state-dependent manner as Cl- ions were observed in open and desensitized structures of cASIC1 but not in a resting state at low pH. (Figure 1A). Although Cl- is also coordinated by other side chains, including Arg309 and Glu313, previous work suggested for Lys 211 to make the most functionally relevant contribution in ASIC1a. The physiological relevance of Cl- ions in fast desensitization of various ASICs (incl heterologously expressed ASIC1a, 2a and 3 channels) is underscored by the finding that they affect ASIC currents in native dorsal root ganglia preparations and have been shown to possibly play a role in acid-induced neuronal death.

Our finding that even very minor alterations of side chain length and/or charge significantly affect fast desensitization kinetics suggests that the coordination of Cl- ions is disrupted by even the most subtle
side chain alterations. Such a notion is supported by the observation that proteinaceous ion binding sites tend to be highly sensitive to mutational disruptions, while H-bonding patterns or charge-charge interactions are often more resilient to small perturbations.\textsuperscript{38,39} Our observation that the length of the aliphatic chain in Lys is crucial is further supported by studies demonstrating that longer or shorter Lys analogs are detrimental to ATP recognition in P2X receptors, substrate recognition in Src homology 2 domains, and catalysis by histone lysine methyltransferases and ubiquitin chain-forming enzymes.\textsuperscript{29,40-42}

In summary, we conclude that K211 plays an important role in both ASIC activation and fast desensitization. This notion is supported by the finding that mutations at the adjacent G212 also affect both processes\textsuperscript{15}, thus underlining the broad functional relevance of this protein domain.

**Optimization attempts.** Our initial screening of a range of split intein pairs and two different splice site combinations pointed towards the \textit{CfaDnaE - SspDnaB\textsuperscript{M86}} split intein pair as the most promising candidate. The large span of observed phenotypes, ranging from efficient channel reconstitution (\textit{CfaDnaE - SspDnaB\textsuperscript{M86}}) to intermediate levels (e.g. \textit{gp41-1 – SspDnaX, NpuDnaE - SspDnaB\textsuperscript{M86}} and \textit{NrdJ1 - SspDnaB\textsuperscript{M86}}) or no apparent reconstitution (e.g. \textit{MjaKlbA – SspDnaB\textsuperscript{M86}}, \textit{gp41-1 – CatInt and MjaKlbA – VidaL}) clearly emphasizes the advantages of testing different split intein combinations for a given splice site.\textsuperscript{21,29,33-36} Notably, and despite its highly asymmetric split \textsuperscript{29}, \textit{SspDnaB\textsuperscript{M86}} appears to be highly suitable as intein B: three out of four split intein pair combinations with \textit{SspDnaB\textsuperscript{M86}} as intein B yielded measurable currents. By contrast, only one out of four designs with \textit{SspDnaX} as intein B, and neither of the two designs with VidaL as intein B (Figure 2B) yielded functional mASIC1a constructs. This is not unexpected because the \textit{SspDnaB\textsuperscript{M86}} split intein has undergone optimization from the wild-type \textit{SspDnaB} intein to display greater tolerance toward extein sequences that deviate from its natural sequence context.\textsuperscript{43,44} Similarly, the \textit{CfaDnaE} split intein acquired its favorable splicing properties and stability from a consensus sequence-based design of related DnaE proteins.\textsuperscript{45}

Somewhat surprisingly, our subsequent optimization attempts failed to improve apparent splicing yields and/or efficiency: neither physically linking the N-TMD/C-TMD fragments nor optimizing the extein sequence context led to a considerable increase in current amplitudes over those observed for the originally tested constructs based on the \textit{CfaDnaE - SspDnaB\textsuperscript{M86}} split intein pair. We conclude that physical proximity is not likely a major limiting factor, and speculate that either sequence optimization was detrimental to channel function or the sequence context might be less important for non-cytosolic splicing events. The latter notion is supported by the fact that extein sequence optimization did not generally improve apparent splicing yields in any of the tested constructs. It is also possible that some of the mutations introduced through the optimization process may be detrimental to ASIC1a function. However, it should be noted that the current amplitudes achieved here with the \textit{CfaDnaE - SspDnaB\textsuperscript{M86}} split intein pair (1 to 6 µA) were significantly greater than those observed previously for tPTS in extracellular domains of structurally related P2X receptors (typically less than 500 nA, see \textsuperscript{29}). The
resulting high signal-to-noise ratio therefore still enabled us to conduct a detailed structure-function analysis of the ncAA-containing mASIC1a variants.

**Limitations of our study.** We observed minor differences in the pH sensitivity of activation and SSD of the full-length channel containing the Thr214Ser mutant compared with the N+XREC+C combination, which also contains Thr214Ser to enable splicing (Supplemental Figure S2). Both constructs should thus comprise the exact same amino acid sequence, the only difference being that the N+XREC+C combination must undergo tPTS to assemble into full-length channels. The observed functional difference could indicate the existence of a subpopulation of channels that contain not fully spliced protomers. We have carefully excluded the possibility of just the N-TMD and C-TMD fragments assembling into functional channels. It is conceivable, however, that in the presence of one or two fully spliced mASIC1a protomers, one or two non- or partially spliced N-TMD and/or C-TMD fragments could assemble to form part of a trimeric channel with altered functional properties. However, this remains speculative and further investigation of this possibility is hampered by the fact that the overall splicing efficiency is exceedingly difficult to quantify, but estimated to be less than 3% (see 29).

While *Xenopus laevis* oocytes are highly amenable to injection of the synthetic peptides required for tPTS28-30, they do not provide the intricate cellular environment afforded by neuronal tissue where ASICs are typically expressed. Additionally, their large physical size results in relatively slow solution exchange. This means we are limited to the study of steady-state measurements, as well as those of the relatively slow kinetics of fast desensitization and we cannot rule out that the $\tau_{\text{desensitization}}$ values reported for the ncAA analogs of Lys may be underestimated due to limited temporal resolution of TEVC. Lastly, we cannot accurately determine channel activation kinetics and are thus unable to assess the effect of ncAA introduction at Lys211 on the activation kinetics of mASIC1a.

**Conclusion.** In this study we demonstrate how subtle analogs of a conserved Lys side chain in mASIC1a can be employed to decipher its contribution to both channel activation and fast desensitization with atomic precision. Our work highlights the potential of tPTS-based approaches for detailed structure-function studies in complex membrane proteins such as ligand-gated ion channels that can be studied with high signal-to-noise methods, making it ideally suited for use with electrophysiology.
MATERIALS AND METHODS

A detailed description of materials and methods is included in the Supplemental Information. In brief, mASIC1a variants were generated by site-directed mutagenesis on wildtype mASIC1a clone.\textsuperscript{46} The split intein-containing clones were obtained from a commercial source (Twist Biosciences). Recombinant, as well as semisynthetic mASIC1a constructs were expressed in \textit{Xenopus laevis} oocytes, and proton-induced currents were measured using TEVC. First, a screen for recombinant reconstitution of mASIC1a channels using 11 different split intein combinations was performed. This was followed by different approaches to improve splicing yields (as measured by maximal current amplitudes). We finally conducted a detailed analysis of proton sensitivity of activation, steady-state desensitization properties of both recombinant and semisynthetic channels, including comparison of fast desensitization kinetics.
ACKNOWLEDGEMENTS

We would like to thank Drs Valeria Kalienkova and Timothy Lynagh for suggestions regarding the experimental design and obtaining preliminary data, respectively. We would also like to thank Dr Mette H. Poulsen and Hendrik Harms for helpful comments on the manuscript. Our work was supported by the Lundbeck Foundation (R303-2018-2900 to IG) and the Independent Research Fund Denmark (9039-00335B to SAP).

AUTHOR CONTRIBUTIONS

D.S., I.G., K.K.K. and S.A.P. designed the research. D.S., I.G., S.Y.O. and G.R.U. performed the experiments and analyzed the data. I.G. and G.J.V.D.H.V.N. designed the synthesis of the peptides. D.S., I.G. and S.A.P. wrote the manuscript with input from all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.
REFERENCES


