

# Comparative sequence analysis suggests a conserved gating mechanism for TRP channels

Eugene Palovcak, Lucie Delemotte, Michael L. Klein, and Vincenzo Carnevale

Institute for Computational Molecular Science, Temple University, Philadelphia, PA 19122

The transient receptor potential (TRP) channel superfamily plays a central role in transducing diverse sensory stimuli in eukaryotes. Although dissimilar in sequence and domain organization, all known TRP channels act as polymodal cellular sensors and form tetrameric assemblies similar to those of their distant relatives, the voltage-gated potassium (Kv) channels. Here, we investigated the related questions of whether the allosteric mechanism underlying polymodal gating is common to all TRP channels, and how this mechanism differs from that underpinning Kv channel voltage sensitivity. To provide insight into these questions, we performed comparative sequence analysis on large, comprehensive ensembles of TRP and Kv channel sequences, contextualizing the patterns of conservation and correlation observed in the TRP channel sequences in light of the well-studied Kv channels. We report sequence features that are specific to TRP channels and, based on insight from recent TRPV1 structures, we suggest a model of TRP channel gating that differs substantially from the one mediating voltage sensitivity in Kv channels. The common mechanism underlying polymodal gating involves the displacement of a defect in the H-bond network of S6 that changes the orientation of the pore-lining residues at the hydrophobic gate.

## INTRODUCTION

A fit cell must perceive and comprehend the conditions of its inner and outer worlds, integrating diverse and transitory physicochemical stimuli into concerted cellular decisions. For this reason, the membranes of even the simplest bacteria are studded with ion channel proteins that detect cellular conditions and translate them into electrochemical information via gated ionic conduction (Hille, 2001).

In eukaryotes, the complexity of cellular life has taken this requirement to its apex. Accordingly, natural selection has elaborated on the ion channel, producing an impressive array of polymodal cellular sensors, the transient receptor potential (TRP) channels (Clapham, 2003). All TRP channels detect multiple physicochemical stimuli, with some overlap among the eight extant TRP subfamilies. However, the response to each stimulus varies substantially from channel to channel, presumably dictated by heterogeneous and subfamily-specific intracellular and extracellular domains (Owsianik et al., 2006). Indeed, since their divergence from the voltage-gated potassium (Kv) channel superfamily over a billion years ago, TRP channel proteins have maintained a tetrameric six-transmembrane (6-TM) architecture and little else (Fig. 1, A–C).

Even so, there is reason to suspect that a common logic underlies TRP channel structure and function

(Su et al., 2007). TRP subunits often form heteromeric channels *in vivo*, even if those subunits come from long-diverged TRP subfamilies (Kobori et al., 2009; Fischer et al., 2014). Also, functional cross-family TRP channel chimeras have been reported (Brauchi et al., 2006). For these heteromers and chimeras to assemble and function, one would expect a common mechanism of pore opening, even if the stimuli that trigger this mechanism differ substantially. These observations suggest that despite heterogeneity in domain organization and gating modality, selection has maintained a set of features in TRP channels underlying their aptitude for polymodal signal transduction.

Discovery of such features necessitates a comparative framework for TRP channels, but high resolution structural and functional data are scant for all but a few known members of the TRP channel superfamily (Liao et al., 2013; Winter et al., 2013). Protein sequences of thousands of TRP channels, on the other hand, are available and have been useful in subfamily-specific analyses of TRP channels (Matsuura et al., 2009; Doñate-Macián and Perálvarez-Marín, 2014). If a large ensemble of TRP channel sequences could be gathered and aligned, statistical features in the resulting multiple sequence alignment (MSA) such as conservation or covariation could be related to selective pressures common to all TRP channels (Halabi et al., 2009; Palovcak et al., 2014).

Correspondence to Vincenzo Carnevale: vincenzo.carnevale@temple.edu

Abbreviations used in this paper: 6-TM, six-transmembrane; aplm-DCA, asymmetric pseudolikelihood-maximization direct coupling analysis algorithm; EC, evolutionary coupling; Kv, voltage-gated potassium; MSA, multiple sequence alignment; TRP, transient receptor potential; VSD, voltage-sensor domain.

© 2015 Palovcak et al. This article is distributed under the terms of an Attribution-Noncommercial-Share Alike-No Mirror Sites license for the first six months after the publication date (see <http://www.rupress.org/terms>). After six months it is available under a Creative Commons License (Attribution-Noncommercial-Share Alike 3.0 Unported license, as described at <http://creativecommons.org/licenses/by-nc-sa/3.0/>).

In this spirit, we searched for sequence–structure–function relationships common to the TRP channel superfamily. We scoured protein sequence databases for an exhaustive set of TRP channel sequences and developed a phylogenetically informed protocol to align the low similarity 6-TM region. Using approaches from information theory and machine learning on a large TRP channel MSA, we characterized the evolutionary pressure exerted on positions throughout the TRP TM region and inferred networks of contacting, coevolving residues running from helix to helix. These observations were contextualized with a parallel analysis on the distantly related Kv channels.

We report sequence features throughout the TRP channel TM core that are not shared with Kv channels, and offer molecular interpretations reasoned on the recently resolved TRPV1 structures (Cao et al., 2013; Liao et al., 2013). Our results establish an evolutionarily informed framework for investigating TRP channel structure and function and suggest a coherent, unified model of TRP channel gating.

## MATERIALS AND METHODS

### Collection and curation of exhaustive TRP and Kv channel MSAs

Annotated TRP channel sequences were collected from the expert-curated SwissProt sequence database and aligned with the fast statistical alignment (FSA) algorithm in the pore domain where most similarity among subfamilies is observed (Bradley et al., 2009). Included sequences spanned the TRPV, TRPM, TRPA, TRPC, TRPP, TRPML, TRPY, and TRPN subfamilies. A profile hidden Markov model (HMM) was constructed on the pore-only TRP MSA and searched against the uncurated UniProt sequence database with HMMER (Eddy, 1998). Thousands of putative TRP channel sequences were collected (E-value of <0.001). Voltage-sensor domain (VSD)-containing Nav, Cav, and BacNav channels were excluded by screening each sequence against an HMM previously constructed to identify VSDs (Palovcak et al., 2014). An approximately maximum-likelihood phylogenetic tree was inferred with FastTree2.1 using the JTT model of amino acid substitution (Price

et al., 2010). Major branches of this tree were identified as known TRP channel subfamilies. Two pore and Catterall channels were also identified as major branches and excluded.

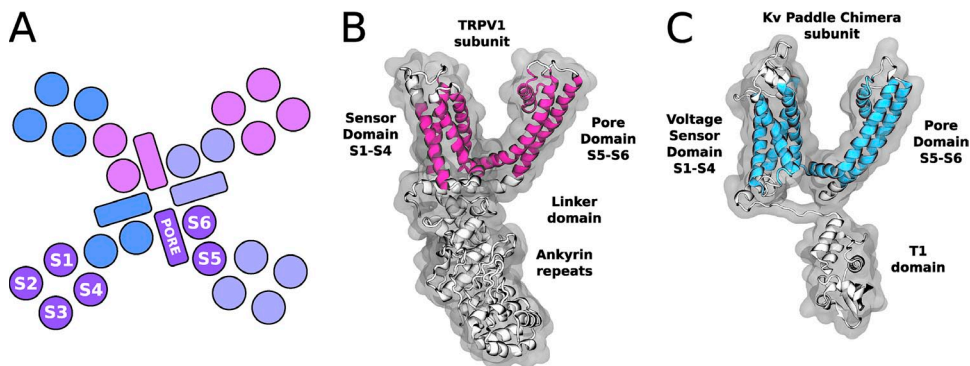
Each identified subfamily was realigned throughout the TM region (S1–S6) with FSA. HMM profiles of each subfamily were constructed and aligned with HAlign (Söding, 2005). Pairs of HMMs were aligned and progressively combined according to the following Newick-formatted phylogenetic tree: (((TRPV,(TRPA,TRPY)),((TRPC,TRPN),TRPM)),TRPP). The TRPML (mucolipins) could not be confidently aligned with our approach (the posterior probability of the profile–profile alignment was <0.50), likely because their sequences were highly similar and provided little information for the profile alignment algorithm. TRPML channels were subsequently excluded. The final TRP channel MSA contained 2,851 sequences. Histograms of pairwise sequence identity were calculated for each TRP subfamily by considering only the aligned 6-TM region, counting the matches for each pair and dividing by the length of the MSA. Sequence logos were generated with Weblogo (Crooks et al., 2004).

A Kv channel MSA was generated similarly. The initial seed MSA for the initial HMM was constructed from rat Kv1-9 sequences in the TM region. Sequences lacking 6-TM helices or the “T[VL]G[FY]GD” potassium selectivity filter sequence were manually removed. Small branches associated with calcium-activated potassium and Kv10-12 channels were also identified and excluded (Yu and Catterall, 2004). A tree inferred from the final Kv MSA contained branches representing Kv1-9 and two large clades of microbial Kv channels. The final Kv channel MSA contained 3,746 sequences.

### Relative entropy calculations for the TRP and Kv superfamilies

To compare structurally equivalent positions, a structure-based alignment of apo TRPV1 to the paddle chimera (Protein Data Bank [PDB] accession nos. 3J5P and 2R9R, respectively) was generated with MultiSeq (Roberts et al., 2006; Long et al., 2007; Liao et al., 2013). We used the apo “closed” structure of TRPV1, as we found alternative conformations of TRPV1 produced identical structure-based sequence alignments. Because of differences in relative orientation between the two structures, the sensor and pore domains were aligned separately. A discrepancy in the structural alignment of the S6 helix was resolved by considering the pore domain of the bacterial voltage-gated sodium channel NavAb structure (PDB accession no. 4EKW), as discussed in the Results (Payandeh et al., 2011).

Amino acid frequency distributions were calculated as described in Palovcak et al. (2014), with sequences with >90% identity down weighted as a single sequence to avoid oversampling of



**Figure 1.** Conservation and divergence of structure in 6-TM cation channels. (A) Conserved architecture of tetrameric 6-TM channels. Four helix-reentrant-loop helix pore domains (S5–S6 helices) assemble like staves of a barrel to form an ion-conducting pore. The S1–S4 helices form isolated sensor domains around the periphery of the pore. Viewed from above, helices are represented as circles and colored to distinguish different subunits. (B) Structure of a representative TRPV1 subunit. Pink helices mark the TM core. TRPV1 contains an intracellular domain and an ankyrin repeat domain, but other TRP channels contain coiled-coils, subfamily-specific regions, and even soluble enzymes. (C) Structure of a Shaker-type Kv channel subunit (the Kv1.2/2.1 paddle chimera). Blue helices mark the TM core. Kv channels feature different intracellular domains than TRP channels.

TRP channel subunit (TRPV1). Pink helices mark the TM core. TRPV1 contains an intracellular domain and an ankyrin repeat domain, but other TRP channels contain coiled-coils, subfamily-specific regions, and even soluble enzymes. (C) Structure of a Shaker-type Kv channel subunit (the Kv1.2/2.1 paddle chimera). Blue helices mark the TM core. Kv channels feature different intracellular domains than TRP channels.

highly similar sequences. Relative entropies for TRP and Kv MSAs were calculated according to:

$$H_i(\text{TRP} \parallel \text{Kv}) = \sum_{a \in q} \text{TRP}_i(a) \ln \left( \frac{\text{TRP}_i(a)}{\text{Kv}_i(a)} \right) \quad (1)$$

$$H_i(x \parallel Q_{\text{ref}}) = \sum_{a \in q} X_i(a) \ln \left( \frac{X_i(a)}{Q_{\text{ref}}(a)} \right), \quad (2)$$

where  $X$  represents the TRP or Kv MSA, the subscript  $i$  denotes the position in the MSA, and  $a$  is each amino acid in the 20-amino acid alphabet  $q$ .  $Q_{\text{ref}}$  is a previously calculated amino acid reference distribution for membrane proteins (Palovcak et al., 2014).

#### Average polarity calculations for the TRP- and Kv-sensor domains

For each position in the sensor domains of TRP and Kv channels, the average polarity was calculated as the expected value of the amino acid frequency distributions for that position (normalized to exclude gaps) and a biological amino acid hydrophobicity scale, as described in Chowdhury et al. (2014). In this scale, greater values indicate a larger free energy associated with inserting a mono-substituted TM helix into a membrane; more polar residues have greater values, given in kilocalorie per mole.

#### Inferring networks of evolutionary coupled residues and estimating their significance

The asymmetric pseudolikelihood-maximization direct coupling analysis algorithm (aplmDCA) was used to estimate the parameters of the maximum entropy probabilistic model consistent with selected MSA statistics (univariate and bivariate frequency distributions). Default parameters were used for field and coupling regularization and sequence reweighting ( $\lambda_h = 0.01$ ,  $\lambda_j = 0.01$ ,  $\theta = 0.1$ ). The optimal parameters are used to calculate an “evolutionary coupling score” (EC score) for each pair of positions given by the quadratic mean of all coupling parameters involving those positions (Ekeberg et al., 2013). Hereafter, we will call the matrix of all ECs between all pairs of positions the “EC map.”

In a fully connected model, the number of parameters increases quadratically with sequence length. Specifically, the number is  $400((N(N-1))/2) + 20N$ , where  $N$  is the MSA length. To reduce the number of parameters to be estimated, we calculated the EC maps separately for the sensor domain (S1–S4), pore domain (S5–pore–S6), and the interface (S1, S5, S5–pore–S6) in both the TRP and Kv superfamily. Still, with so many parameters, overfitting is a serious concern. To assess the robustness of EC maps with respect to a different choice of the training dataset used in the statistical inference, we performed an error analysis to assign statistical significance to EC scores.

#### Error analysis

Our goal is to calculate the standard error of the maximum-likelihood estimator used in the statistical inference step. Ultimately, to gauge the “strength” of a coupling, we would like to take into account not only its value but also the statistical fluctuations resulting from the finiteness of the sample. Expressing each coupling in units of its standard error (in the form of a Z-score) allows us to assess its statistical significance and provides an expedient metric to appreciate, simultaneously, the strength and degree of robustness of each evolutionary-coupled pair. Moreover, the statistical significance enables us to associate a p-value to each threshold chosen to select the top ranking couplings. Calculation of the standard error of a maximum-likelihood estimator is, in general, a complex task. A conceptually simple scheme to achieve

this goal, the so-called jackknife approach, is based on the use of appropriately chosen subsets of the original dataset (Reeds, 1978). Specifically, the jackknife estimator of a parameter  $\theta$  is obtained by calculating the former for all subsamples resulting from the removal of a single observation without replacement, i.e., for each so-called pseudo-observation (Quenouille, 1949; Shao and Wu, 1989). Here, we have used a variation of this approach, the delete- $d$  jackknife, which is obtained by omitting  $d$  observations each time, because it is statistically consistent for a wider class of estimators compared with the delete-one version (Efron and Tibshirani, 1993). In particular, the variance of the estimator  $\hat{\theta}$  is:

$$\text{Var}(\hat{\theta}) = \frac{N-d}{d} \frac{\sum_{i=1}^{\binom{N}{d}} (\hat{\theta}_{-d_i} - \hat{\theta})^2}{\binom{N}{d}}, \quad (3)$$

where  $N$  is the sample size, and  $d$  is the number of omitted observations. The sum runs over all the distinct subsets of size  $d$  that can be extracted without replacement from the original dataset. For each of them, a pseudo-observation is built, and the parameter  $\hat{\theta}_{-d_i}$  is recalculated. Finally,  $\hat{\theta}$  is the average of all the  $\hat{\theta}_{-d_i}$  terms; namely,

$$\hat{\theta} = \frac{\sum_{i=1}^{\binom{N}{d}} \hat{\theta}_{-d_i}}{\binom{N}{d}}. \quad (4)$$

Note how the prefactor multiplying the sum in Eq. 3 compensates for the increased mean-square fluctuation as  $d$  increases. This estimator of the variance is consistent, i.e., it converges in probability to the true variance as the number points increase, as long as  $d$  is comprised between  $\sqrt{N}$  and  $N$  (Shao and Tu, 1995). An important byproduct of this analysis is a bias-corrected estimator of the parameter  $\hat{\theta}^*$  (i.e., with bias terms proportional to  $1/N$  removed):

$$\hat{\theta}^* = \frac{[N\hat{\theta}^0 - (N-d)\hat{\theta}]}{d}, \quad (5)$$

which enables comparisons between datasets of different size (here,  $\hat{\theta}^0$  is the parameter calculated on the entire dataset) (Narsky and Porter, 2014). A convenient choice for  $d$  is  $N/2$ , which makes the prefactor of Eq. 3 equal to 1. In practice, exhaustive enumeration of all the possible pseudo-observations is not computationally feasible; thus, we approximated the mean-square fluctuation term of Eq. 3 by stochastically sampling  $K$  pseudo-observations (Gentle, 2009).

The parameters,  $\hat{\theta}$  in the equations above, that we are interested in here are the EC scores provided by the aPLM DCA estimator. Following the approach outlined above, we computed the EC scores 100 times by omitting half of the sequences from the MSA each time. Thus, the  $EC_z$  or Z-score of the EC is defined, for each  $(i,j)$  pair as:

$$EC_z(i,j) = \frac{\langle \widetilde{EC}(i,j) \rangle_k}{\left[ \left\langle \left( \widetilde{EC}(i,j) - \langle \widetilde{EC}(i,j) \rangle_k \right)^2 \right\rangle_k \right]^{1/2}}, \quad (6)$$

where  $\langle \cdots \rangle_k$  denotes the average over the  $K$  resampled MSAs ( $K = 100$  in this case), and  $\widetilde{EC}(i,j)$  is the standardized EC score:

$$\widetilde{EC}(i,j) = \frac{EC(i,j) - \langle EC(i,j) \rangle_{i,j}}{\left[ \left\langle \left( EC(i,j) - \langle EC(i,j) \rangle_{i,j} \right)^2 \right\rangle_{i,j} \right]^{1/2}}. \quad (7)$$

In Eq. 7,  $\langle \dots \rangle_{i,j}$  indicates the average over the distinct residue pairs. The standardization step is performed to homogenize all the pseudo-observations by removing constant terms and scale factors. In the following analysis, we will consider  $EC_Z$  scores above 4 as corresponding to a one-tailed p-value of  $3.2 \times 10^{-5}$ .

### Online supplemental material

Fig. S1 shows a small MSA of representative TRP channel sequences aligned in the TM region, with the residues considered for the analysis highlighted by red shading. Fig. S2 shows the relative entropy of the TRP channel TM region with respect to Kv channels. Mapping of the relative entropy on the structure of TRPV1 highlights the regions where the distributions of residues are most similar: positions with low relative entropy primarily putatively face the lipids, whereas the core-facing regions of both channel families have distinct residue distributions except for the indicated lower portions of S3 and S5. The online supplemental material is available at <http://www.jgp.org/cgi/content/full/jgp.201411329/DC1>.

## RESULTS

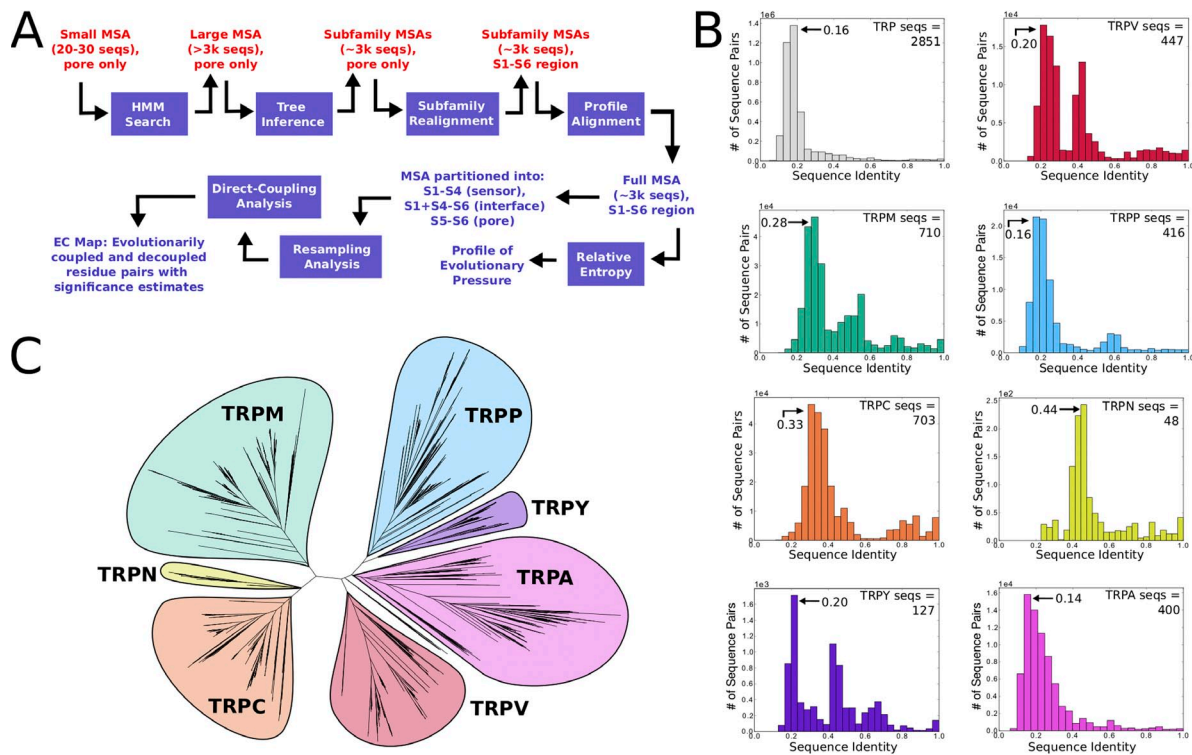
### Profiles of TRP subfamilies enable alignment in the TM region

A reliable alignment of the TRP channel TM region is needed to reflect the evolutionary events of amino acid substitution, insertion, or deletion that have shaped TRP sequences. Additionally, insights obtained from such a

dataset can only be generalized to the superfamily if the MSA is comprehensive. However, in the majority of the TM region, TRP channel sequences are highly heterogeneous, precluding sequence collection and curation by familiar methods. To overcome this obstacle, we developed a custom alignment protocol for the TRP superfamily.

Most simply, we leverage knowledge of TRP channel phylogeny to construct statistical profiles (profile hidden Markov models, or HMMs) of each TRP subfamily (Eddy, 1998; Price et al., 2010). These subfamily profiles are then aligned in pairs, reconstructing the full superfamily branch by branch (Söding, 2005). In this way, we were able to generate a plausible MSA of thousands of TRP channel sequences throughout their TM region. This method works because the statistical profiles of two protein families (for instance, the TRPV and TRPA subfamilies) contain much more information for a pairwise alignment than two representative sequences (such as human TRPV1 and TRPA1) (Martinsen et al., 2004). A detailed description of this workflow can be found in the top row of Fig. 2 A and in Materials and methods.

To compare the sequence features of the TRP channel superfamily with those of the well-studied Kv channels, we used the same protocol to construct a Kv channel



**Figure 2.** Sequence diversity in the TRP channel superfamily. (A) The sequence analysis workflow followed in this study. The top row shows the protocol for sequence collection and curation (script in red), whereas the bottom row shows the analyses performed on the final full MSA (script in purple). This workflow was applied both to the TRP and Kv channel superfamily. Details of procedures are available in Materials and methods. (B) Histograms of pairwise sequence identity for the TRP channel superfamily. Only the aligned TM region was used and the identity was calculated for all pairs of sequences in each subfamily. (C) Phylogenetic tree of the full TRP channel MSA inferred with FastTree2. Annotated sequences were used to identify branches corresponding to known TRP subfamilies.

MSA of similar size (described in Materials and methods). The final TRP MSA contained 2,851 sequences, whereas the Kv channel MSA contained 3,746 sequences.

The TRP channel MSA captures the expected diversity of the TRP superfamily

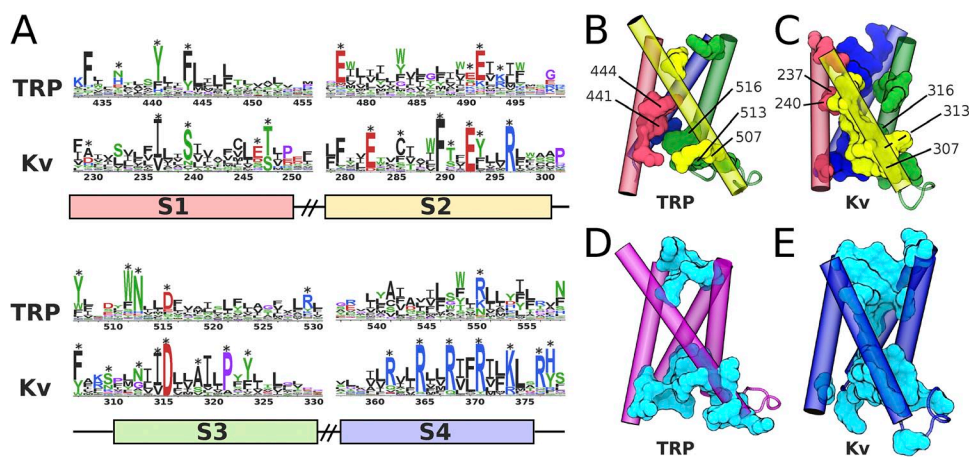
From the phylogenetic tree inferred from the TRP MSA, we identified eight subfamilies varying in size and diversity. However, the small and homogeneous TRPML (mucopolin) subfamily could not be confidently aligned with our protocol and was subsequently excluded from the analysis (see Materials and methods). For the remaining seven subfamilies, histograms of sequence identity showed major modes between 14 and 44% identity, with the major mode of the full MSA showing 16% identity (Fig. 2 B). The multimodal nature of the subfamily histograms mirrors the finer phylogenetic structure observed in the inferred tree (Fig. 2 C). Consistent with recent phylogenetic analyses on smaller trees of TRP channels, we find homologues of all TRP channel subfamilies (except TRPN) in single-celled eukaryotes and a highly diversified and expanded TRPA subfamily in insects (Cai, 2008; Peng et al., 2015).

Although the TRPY and TRPN subfamilies were relatively small (127 and 48 sequences, respectively), TRPCs, TRPVs, TRPAs, TRPPs, and TRPMs each contained hundreds of putative members, implying that the TRP channel MSA is heterogeneous and does not over-represent one or two TRP subfamilies. We therefore expect that insights generated from this dataset should be general to the TRP superfamily. This strategy has been validated by the structural determination of another member of the TRP family, namely TRPA1 (Paulsen et al., 2015). Structural superposition of TRPV1 and TRPA1 induces a sequence alignment that is completely consistent with the one inferred from the profile–profile matching approach. We are thus confident that the alignment is accurate across the entire TRP family.

Relative entropy can serve as a proxy for evolutionary pressure

Guided by a structural alignment of TRPV1 to the Shaker-type paddle chimera Kv structure (Long et al., 2007), we were able to directly compare the distributions of amino acids observed at equivalent positions in the TM region of each superfamily (the colored regions in Fig. 1, B and C). In large, diverse protein families, the nature of the amino acid distribution observed at a certain position is indicative of a common evolutionary pressure acting throughout the family. Visual inspection of aligned sequence logos of the sensor domain show that despite substantial structural similarity, very few positions share similar amino acid distributions between TRP and Kv channels (Fig. 3 A). To make this observation more quantitative, we calculated the relative entropy for each equivalent position in the MSAs.

The relative entropy (or Kullback–Leibler divergence) is a nonsymmetric measure of the difference between two probability distributions (Thomas, 2006). For each structurally equivalent position between two MSAs, the relative entropy quantitatively describes how different the distributions of amino acids at those positions are (see Materials and methods). The relative entropy between TRP and Kv channels is lowest for lipid-facing residues and much higher within the core of the sensor domain, throughout the ion conduction pathway in the pore domain, and at the intersubunit interface (Fig. S2 A). The S3 and S5 helices are exceptions, showing low relative entropy between TRP and Kv channels (Fig. S2 B). In general, two limiting scenarios can give rise to the observed low mutual entropy: both the distributions are similar to the uniform distribution (high entropy), or they deviate from it in a similar fashion, i.e., they are both low entropy with coincident modes. Specifically, although we expect the distribution of the lipid-facing residues to be random-like, the low mutual entropy of



**Figure 3.** Single-site statistics of the TRP- and Kv-sensor domain. (A) Aligned sequence logos for the TRP and the Kv channel superfamily in the sensor domain (S1–S4). Rat TRPV1 numbering is used for the TRP logo, and Shaker numbering is used for the Kv logo. Positions marked with an asterisk have  $H_{\text{rel}}(X(i)||Q_{\text{ref}}) > 1.25$ , where  $X$  is the TRP or Kv MSA position  $i$ . These high relative entropy positions are mapped onto the TRP (B)- and Kv (C)-sensor domains as molecular surfaces. Helices are colored as follows: red, S1; yellow, S2; green, S3; blue, S4. (D and E)

Average polarity in the TRP (D)- and Kv (E)-sensor domains. Average polarity is calculated as the expected value of a position's amino acid frequencies (normalized to remove gaps) and a biological hydrophobicity scale. Positions with an average polarity >1.0 kcal/mol (slightly more polar than serine) are shown as a blue surface.

S3 and S5 is potentially the signature of shared (although loose) patterns of conservation. To distinguish between these two alternative hypotheses, we calculated the mutual entropy with respect to a common reference distribution  $Q_{\text{ref}}$  for the two MSAs separately, with the ultimate goal of detecting conserved patterns of residues. In fact, the low mutual entropy of S3 and S5 turns out to reflect conserved patterns. The functional implications of this conserved sequence feature are discussed in details in the following sections concerning the sensor and pore domains. Technical details of this analysis are described in Materials and methods section.

#### Few common sequence motifs are observed in the TRP- and Kv-sensor domains

It is unclear what role the sensor domain plays in the structure and function of TRP channels (Kalia and Swartz, 2013). Certain positions at the base of S3 show similar distributions of amino acids in both TRP and Kv channels (TRP positions 507, 513, and 516; Kv positions 307, 313, and 316). This motif is conserved in most sensor domains (Kumánovics et al., 2002; Palovcak et al., 2014) and has recently been implicated in the gating of TRPM8 and TRPM2 channels (Winking et al., 2012).

Other positions have high relative entropy in both superfamilies but different amino acid compositions, implying that the position is important in both families but for different reasons. For example, TRP S1 positions 441 and 444 are conserved aromatic residues, whereas the equivalent Kv positions 237 and 240 are conserved hydrophobic and polar residues, respectively. In the voltage-sensitive Kv channels, positions 237 and 240 occupy the gating pore through which the positively charged S4 gating charges are shuttled (Long et al., 2007). Their mutation strongly affects the voltage sensitivity and kinetics of gating charge transport (Lacroix et al., 2014). To our knowledge, positions 441 and 444 have not been mutated in TRPV1, but in TRPM8, mutations of the equivalent Y745 (erroneously described as an S2 residue in Bandell et al., 2006) specifically decrease menthol sensitivity (Winter et al., 2013).

To visualize the spatial distribution of evolutionary pressure through the sensor domains, we map relative entropy onto the TRP and Kv structures (Fig. 3, B and C). Both feature high relative entropy positions at the intracellular interior of the sensor, whereas in Kv channels, high relative entropy positions are also found at the extracellular interior (Kv 247, 248, 283, 286, 322, 324, 362, and 365) and correspond mostly to residues such as the acidic countercharges and the basic gating charges. As expected, no putatively lipid-facing positions in either family had relative entropy greater than our chosen threshold (1.25 nats).

Solvation of the TRP sensor is limited to the intracellular side. In the Kv channel–sensor domain, polar residues inside the sensor act to solvate the gating pore, with

upper and lower crevices forming an hourglass-shaped vestibule of water, which focuses the membrane electric field onto the gating charges, allowing efficient transport of charge (Starace and Bezanilla, 2004; Krepkov et al., 2009). In the TRPV1 structure, the lower vestibule is splayed open and occupied by the polar face of the C-terminal TRP helix, suggesting that it is solvated (Liao et al., 2013). The upper portion of the vestibule, on the other hand, is densely packed with bulky aromatic and hydrophobic residues and is therefore likely unsolvated.

It is not clear whether this asymmetric solvation is unique to TRPV1. To test this, we measured the average polarity in the TRP- and Kv-sensor domains (following a protocol described in Chowdhury et al., 2014; see Materials and methods) and mapped positions with conserved high polarity (on average, more polar than serine) onto the TRP- and Kv-sensor domain structures (Fig. 3, D and E). Although the Kv-sensor domain shows the expected hourglass shape, the TRP sensor recapitulates the pattern of polar residues observed in the TRPV1 structure, suggesting that an asymmetrically solvated sensor domain is a conserved feature of TRP channels.

#### Networks of evolutionarily coupled residues can be revealed through statistical inference

Although the relative entropy view of evolutionary pressure is intuitive, it misses a great deal of the information that can be extracted from such large MSAs. Evolution may act on more variable positions by strictly tuning their covariation with other residues (Süel et al., 2003; Halabi et al., 2009). This is especially true for residues that make physical contact in the native protein structure (Göbel et al., 1994; Morcos et al., 2011). In general, directly measuring this covariance with statistics such as the mutual information does not produce easily interpretable results, as transitive interactions between chains of residues can yield indirect correlations (Weigt et al., 2009). In Weigt et al. (2009), the authors showed that a probabilistic model assuming a direct interaction between coevolving residue pairs can explain the observed frequency distribution of amino acids in a large protein family MSA. Strong interactions, or “evolutionary couplings,” are highly predictive of native protein contacts (Ekeberg et al., 2013).

Only a subset of contacting residue pairs in a protein are detected as ECs. We recently examined the network of ECs present in VSDs and found that the structure of this network was highly informative about the general mechanism by which these molecular machines sense voltage (Palovcak et al., 2014).

We calculated EC networks for the TRP and the Kv channel superfamily using the aplmDCA (Ekeberg et al., 2013). We also performed an error analysis to assess the strength and robustness of the ECs we inferred. In our framework, the statistical significance of an evolutionarily coupled pair of amino acids is given by a Z-score

( $EC_z$  score). Only ECs exceeding a chosen significance threshold of  $EC_z > 4$  are reported in this study; considering a single-tailed test and correcting for the multiple comparison, this corresponds to a Bonferroni-corrected p-value of  $\sim 0.1$ . Technical aspects of the evolutionary network inference with aplmDCA and the error analysis are described in Materials and methods.

Before delving into the details of the TRP and Kv channel EC maps, we cautiously note the conceptual differences between ECs and energetic couplings in proteins. Energetic couplings, as measured by double-cycle mutant analysis, correspond to the nonadditive energetic effects of double mutants on the stability or functional activity of a macromolecule (Yifrach and MacKinnon, 2002). ECs are instead meant to describe how knowledge of the identity of a residue at a given position affects our expectations about a different one. Although it is reasonable to expect that an energetically favorable pairwise interaction is conserved along evolution through compensatory mutations, the “strength” of the associated statistical signal may vary: interactions that are of nonpolar nature can arguably rely on a larger repertoire of residue pairs compared with polar H-bonding ones that require a more precise match.

Chains of such pairwise interactions may underlie allosteric signal communication (Yifrach and MacKinnon, 2002; Süel et al., 2003). However, unlike protein structure, which is consistently conserved among homologues, the conservation of discrete allosteric pathways has not been demonstrated to the same extent (Fodor and Aldrich, 2004).

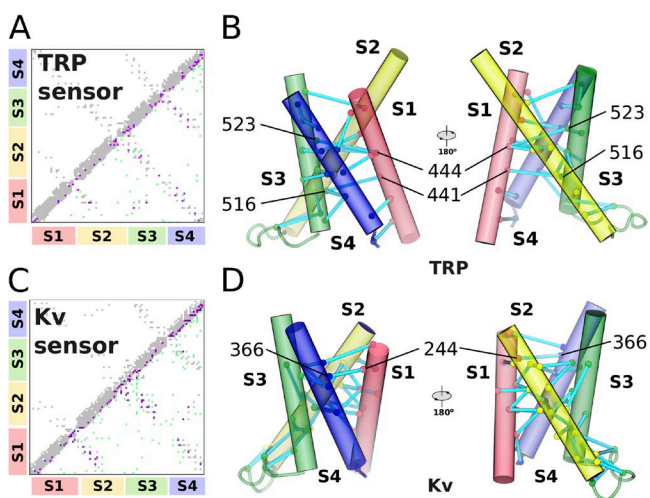
Here, we use the presence of ECs to draw attention to physically contacting residues that have coevolved unambiguously. Coevolution implies that those physical interactions have been important during evolution and are likely to be conserved among members of the superfamily; however, we cannot make a claim on the energetics of specific molecular processes. Finally, a comment is in order concerning coevolving pairs that are not in contact in the available experimental structures analyzed in this paper: in analogy with our previous investigation of the VSD (Palovcak et al., 2014), the coevolution observed for some of these pairs could be the signature of residue–residue contacts that exist only in channel conformations distinct from those captured by the experiments so far. Exploring this fascinating hypothesis amounts to distinguishing with accuracy bona fide coevolution signals from false positives, a task that is beyond the scope of the current investigation.

**ECs involving TRP S4 are consistent with a static helix**  
Mapping the strong and robust ECs ( $EC_z > 4$ ) onto the structures of the TRPV1- and paddle chimera–sensor domains, we observe that the upper S1–S2 and lower S2–S3 interfaces exhibit numerous ECs in both families, although the couplings in Kv channels are more

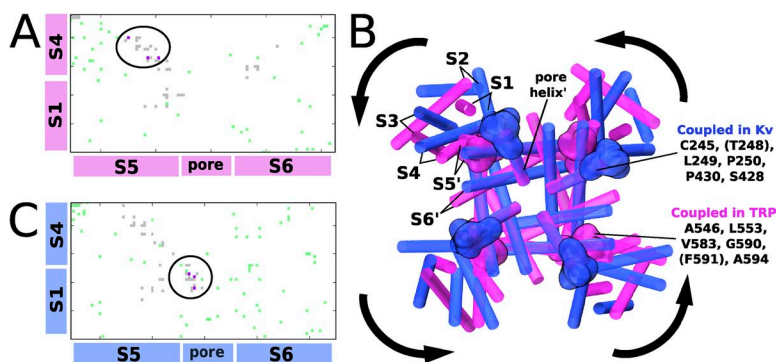
extensive and run further down the S1–S2 interface (Fig. 4, B and D, right).

In the TRP sensor, several couplings are also observed between the aforementioned aromatic TRP S1 positions 441 and 444 and bulky residues on S3 (TRP 516 and 523; Fig. 4 A). Such interactions would obstruct the gating pore of a Kv channel VSD and are observed neither as residue contacts nor ECs in the Kv channels.

Most significantly, we find differences in the ECs of the S4 helices of TRP and Kv channels. In previous work, we demonstrated that VSDs in general do not feature significant coevolution between the hydrophobic contact interfaces of S4 and the neighboring S1 or S3 helices but do show coevolution between the gating charges and acidic or polar residues in the gating pore (Palovcak et al., 2014). These results are consistent with a recent investigation revealing that coevolution analysis can detect residue contacts unique to distinct configurations (Morcos et al., 2013). Here, the pattern of ECs is shown experimentally by the portability of S3b–S4 paddle motif among chimeric voltage-gated cation channels, where the gating charges are conserved among the transposed paddle motifs, but the hydrophobic interfaces are not (Alabi et al., 2007). In analogy with these results, we interpret an absence of extensive networks of coevolving pairs as the signature of the mobility of the structural elements involved. This association between coevolution and structural dynamics is, at this level, a



**Figure 4.** Networks of ECs in the TRP- and Kv-sensor domains. (A) The EC map for the TRP-sensor domain overlaid on the contact map of the apo TRPV1-sensor domain. In all EC maps, the axes represent the primary structure of the sensor domain. In the lower triangle, contacting noncoupled positions are gray, noncontacting coupled positions are green, and contacting coupled positions are red (these are of greatest interest to the present study). (B) Cartoon representation of the TRP sensor with contacting, coupled pairs connected with bars. (C) The EC map for the Kv-sensor domain overlaid on the contact map of the paddle chimera VSD. (D) Cartoon representation of the Kv sensor with coupled contacts mapped.



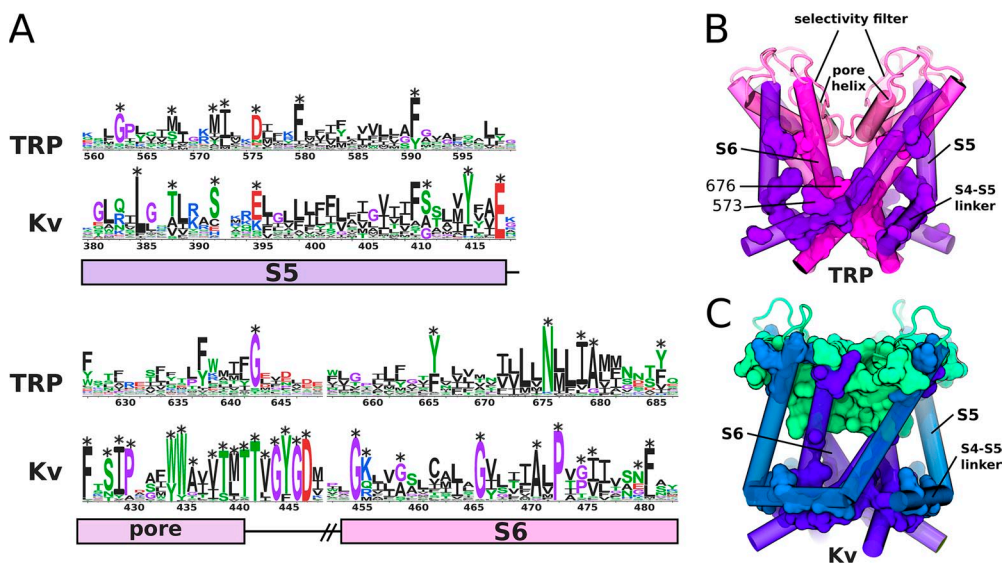
**Figure 5.** Intersubunit coevolution between the sensor and domain of TRP and Kv channels. (A) EC map of the TRP intersubunit interface overlaid on the contact map TRPV1. To minimize the number of parameters, only the contacting S1 and S4 helices and the pore domain were used in the model of the interface. The EC map is colored as in Fig. 3. (B) Cartoon representation of the structurally aligned TM region of TRPV1 (magenta) and the paddle chimera (blue). Arrows illustrate the counterclockwise shift in the TRP-sensor domain with respect to the paddle chimera. Residues making up the coevolving intersubunit interfaces in both channels are represented as colored surfaces. (C) EC map of the Kv intersubunit interface overlaid on the contact map of the paddle chimera.

working hypothesis. We argue that because the dynamic remodeling of the contact map associated with a conformational change entails the interaction of each residue with more than one partner, correlations among residues are potentially of higher order than the pairwise ones considered here. Accessing these higher order correlations and thus testing this hypothesis is challenging because of the limited number of sequences available, and is beyond the scope of this type of analysis.

These observations are largely recapitulated when considering just Kv channel VSDs, although we see more of the gating charges coevolving with gating pore residues, and we do see one hydrophobic interaction between S4 and S1 (Kv 244–366) (Fig. 4, B and D). In contrast, the S4 helix of the TRP-sensor domain is extensively coupled at the hydrophobic faces of the S1 and S3 helices, stabilized into a putatively stationary conformation by specifically tuned interactions (Fig. 4, A and C, left). Consistently, chimeragenesis studies show that the corresponding paddle motif is apparently not portable in TRP channels (Kalia and Swartz, 2013).

### The TRP sensor–pore interface has shifted through evolution

In Lee et al. (2009), sequence analysis and electrophysiology illustrated the functional importance of a coevolved contacting interface between the VSD and pore of Kv channels. This interface involved residues at the top of the S1 helix and the pore helix of a neighboring subunit. Its disruption yields channels that either do not traffic to the membrane or are defective in the final concerted step of activation gating (Lee et al., 2009; Shem-Ad et al., 2013). Despite using a dataset of Kv channels with a different size (ours is an order of magnitude larger), phylogenetic scope (ours is more confined), and algorithm for coevolution detection (theirs is a perturbation-based approach, whereas ours infers structured probabilistic graphical models), we similarly observe strong ECs at this interface in Kv channels (Fig. 5, B and C). Additionally, Kv position 248, the residue most sensitive to mutation at this interface, shows up with high relative entropy, suggesting a specific evolutionary pressure that constrains this position to small polar residues (Fig. 3 A).



**Figure 6.** Single-site statistics of the TRP and Kv pore domain. (A) Aligned sequence logos of the TRP and Kv pore domain. Positions with  $H_{rel}(X(i)||Q_{ref}) > 1.25$  are marked with an asterisk. High relative entropy ( $>1.25$  nats) positions are displayed on the TRPV1 pore domain structure (B) and paddle chimera pore domain structure (C) in space-filling representation. The following color code has been used for representing different structural elements of the channels: purple, S4–S5 linker and S5; radiant orchid, pore helix; magenta, S6 for TRPV1; steel blue, S4–S5 linker and S5; green, pore helix; indigo, S6 for paddle chimera.



In TRPV1, this sensor–pore domain interface has shifted (Liao et al., 2013). With respect to the paddle-chimera structure, the TRPV1-sensor domain is rotated counter-clockwise, breaking most of the S1–pore helix interactions (Fig. 5 B). Primary contact between the sensor and the pore is made instead through the S4 helix and the S5 helix of a neighboring subunit. This rolling of the sensor domain is similarly observed in the bacterial sodium channel NavAb (Payandeh et al., 2011). Accordingly, we see no positions with high relative entropy on the top of the TRP S1 analogous to Kv 248, and we see strong EC of S4 positions to contacting S5 positions (Figs. 2 A and 5, A and B). Interestingly, the structure of TRPA1 reveals an analogous shift of the interface, thereby confirming the generality of our coevolution analysis (Paulsen et al., 2015).

The pore architecture is more variable in TRP channels than in Kv channels

Aligning the sequence logos of the TRP and Kv pore domains, we see again that despite structural similarity, the underlying distributions of amino acids observed at the same structural position greatly differ in the two subfamilies (Fig. 6 A). This difference is perhaps most striking around the pore helix and selectivity filter (TRP 627–649 and Kv 427–448). Kv channels have the common constraint of selective potassium ion conduction, which is echoed by the high relative entropy of the pore architecture responsible for this function (Fig. 6 C, the region in green). TRP channels have more variable cation selectivity depending on the channel and may exploit the pore architecture for channel-specific modulation of gating and permeation (Owsianik et al., 2006). TRP channels have comparatively low relative entropy in this region (Fig. 6 B).

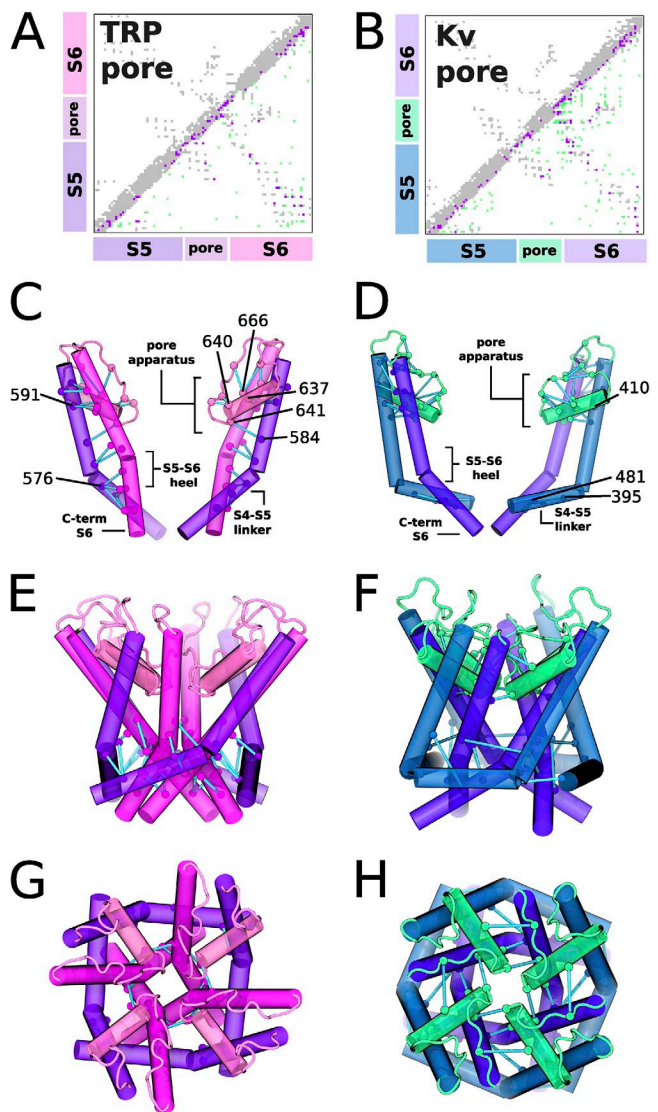
In both superfamilies, the pore itself is made up by the S6 helices of the four subunits. However, when comparing TRP and Kv channels, the relative entropy of the distribution of residues in the S6 helix is high. TRP channels feature a conserved asparagine (TRP 676) common to voltage-gated sodium and calcium channels, whereas Kv channels contain a characteristic “P[VT][PG]” motif (Guda et al., 2007). As observed previously in a smaller dataset, unlike almost all voltage-gated cation channels, TRP channels do not possess a gating hinge glycine at TRP position 670 (Kv 466) (Zhao et al., 2004).

The interface between the base of S5 (TRP 560–574 and Kv 380–394, the so-called “S4–S5 linker”) and the C-terminal region of S6 shows high relative entropy in both TRP and Kv channels, although in neither case are the residues strictly conserved. This interface has been shown to be crucial for VSD–pore coupling in Kv channels and allosteric modulation of pore opening in TRP channels (Lu et al., 2002; Labro et al., 2008; Boukalova et al., 2010). Curiously, TRP position 573 in this region represents an insert present only in TRPV channels (Fig. S1).

A coevolving network at the base of S5–S6 is observed in TRP but not Kv channels

A network of ECs bridges the top interface between the S5 and S6 helices of both families (Fig. 7, C and D, left in both). This network centers on a sensor-facing aromatic residue (TRP 591 and Kv 410) but also runs behind the pore helix and throughout the pore architecture (Fig. 7, C and D). In Kv channels, these couplings mount the pore helix to S5 and S6, the selectivity filter to the pore helix, and the pore loops to each other. In TRP channels, similar features are seen but with less density.

Similarly, the S4–S5 linkers are coupled to the C-terminal portion of S6 in both TRP and Kv channels.



**Figure 7.** EC in the TRP and Kv pore domains. (A) EC maps of the TRP pore domain, colored as in Fig. 3. (C) Intrasubunit ECs mapped onto a TRPV1 monomer. (E and G) Intersubunit ECs mapped onto the tetrameric TRPV1 pore from the side (E) and from above (G). (B, D, F, and H) Analogous figures for the Kv channels, mapped onto the structure of the paddle chimera pore.

As mentioned above, this interface plays a substantial role in coupling the voltage-dependent motion of S4 to pore opening (Lu et al., 2002; Labro et al., 2008).

In contrast, the bent region in the middle of the pore domain has several very strong intrasubunit ECs in TRP channels and none in Kv channels. One mostly acidic position, TRP 576 (Kv 395), has three such ECs. In Kv channels, 395 coevolves with one nearby aromatic residue (Kv 481) on an adjacent subunit, but this intersubunit coupling is not in physical contact per se and does not correspond to an EC in TRP channels. A recent study also noted this coevolution between Kv 395–481 and investigated the interaction in detail (Pless et al., 2013). However, the equivalent positions (and their very different interacting partners) have not been characterized in TRP channels.

No intersubunit ECs are observed within the upper pore of TRP channels

Several ECs in the pore domain were found to be in contact across subunits. In both families, numerous intersubunit ECs are observed at the lower region of the pore, connecting N- to C-terminal regions of S4–S5 linkers and the S6 helices (Fig. 7, E–H). In TRP channels, several

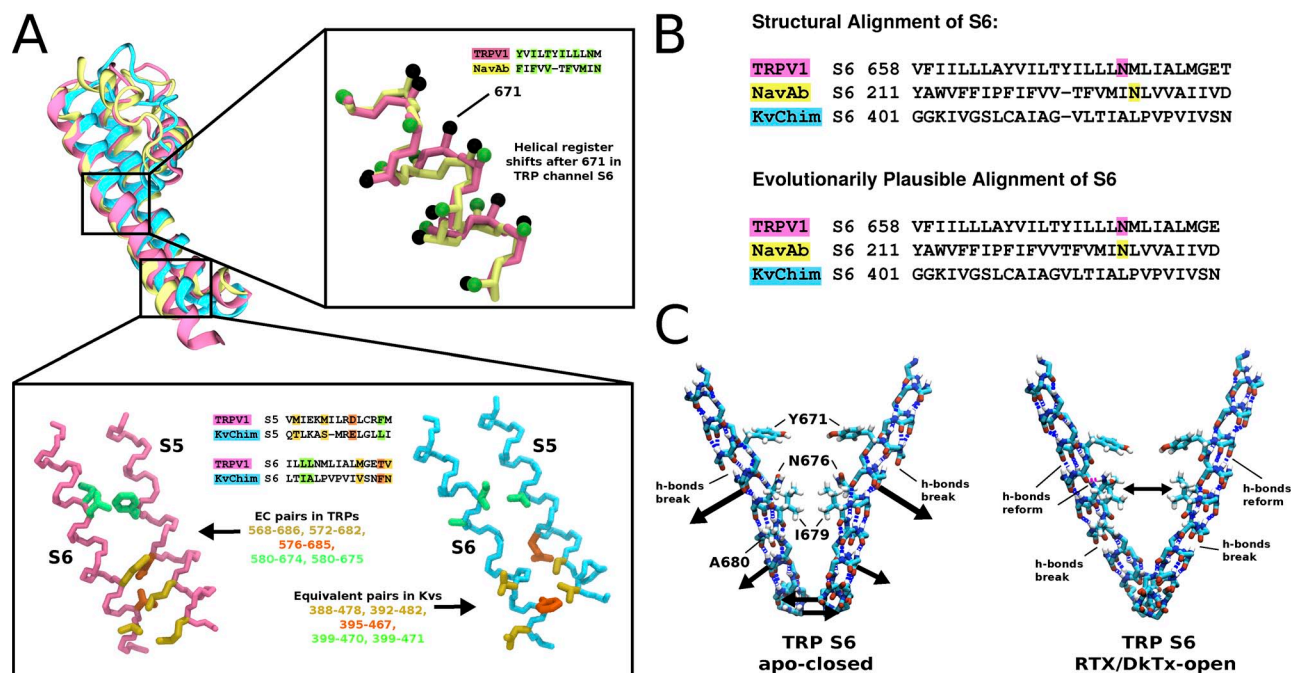
strong intersubunit couplings were also observed between the conserved hydrophobic residues of the lower S6 helix.

Intersubunit ECs are observed at the upper region of the pore in Kv channels. Most of these couplings link the pore helix to the S6 helix of a neighboring subunit. In TRP channels, there are no such intersubunit ECs, even though the intersubunit interface occupies a similar contact area.

Evolution has encoded a register-shifting helical distortion into the S6 helix of TRP channels

Close examination of the structure of TRPV1 reveals a helical distortion in the middle of S6, where Y671 appears to have popped out from the helix and shifted the helical register of the TRPV1 S6 (the helical repeat goes from  $i + 4$  to  $i + 5$  after Y671; see Fig. 8 A, top inset). Because the S6 helices line the ionic permeation pathway in TRP and Kv channels, and presumably contain the gate, subtle structural features are worth investigating in detail.

Structural alignments of the TRPV1 S6 helix to the paddle chimera S6 and the bacterial sodium channel NavAb S6 indicate that position Y671 is an insertion. However,



**Figure 8.** A helical distortion in the TRP channel pore. (A) Structural alignment of a pore domain subunit of TRPV1 (pink), NavAb (yellow), and the paddle chimera (blue). (Top inset) Y671 of TRPV1 has no aligned residue in NavAb, in effect shifting the register of the helix (at the top of S6, the green and black C- $\beta$  atoms are aligned, whereas after 671, they are staggered). (Bottom inset) Several pairs corresponding to contacting ECs in TRPV1 (left) are not coupled nor even in contact in the equivalent pairs of the Kv paddle chimera (right). (B; top) The “structural alignment” of TRPV1, NavAb, and the paddle chimera S6 helices, which suggests that TRP 671 is an insert picked up in TRP channels. (Bottom) An “evolutionarily plausible” alignment of TRPV1, NavAb, and the paddle chimera S6 in which the conserved asparagine observed in Nav, Cav, and TRP channels is aligned. (C) Gating motions of TRPV1 as determined by cryoelectron microscopy. On the left, two S6 helices of the apo “closed” TRPV1 pore domain are shown. On the right, the same “open” pore domain helices from the RTX/DkTx-bound TRPV1 structure are shown. Broken hydrogen bonds beneath Y671 reform as the channel shifts from closed to open, but only at the expense of new broken hydrogen bonds above A680, caused by the helical distortion.

considering the sequence alignment of TRPV1 to NavAb rules this out: if TRP 671 is an insertion, it would cause the conserved N676 to be misaligned from the obviously equivalent position in NavAb (N211) (Fig. 8 B). Instead, the observed bulging residue may be caused by an inherent geometrical distortion of the TRP S6 helix. Consistent with this hypothesis, we recall the aforementioned contacting couplings at the S5–S6 interface in TRP channels. Not only are the equivalent pairs not coevolving in Kv channels, most are not even in contact (Fig. 8 A, bottom inset). It is precisely the register shift of S6 that enables these evolutionarily coupled residues to make contact in TRPV1. Although the distortion of the S6 helix is unambiguous from the TRPV1 structure and its representative electron density map, this featured was not discussed previously (Liao et al., 2013). We hypothesize that these residues have been properly selected such that favorable, highly specific interactions pin the S6 helix to the S5 helix out of register, distorting the helix and breaking the backbone hydrogen-bonding network (Fig. 8 C). This helical defect has been recently observed in TRPA1, another member of the TRP family (Paulsen et al., 2015). Strikingly, a recently determined structure of the ryanodine receptor RyR1 (PDB accession no. 3J8H) shows a similar distortion of S6 with the bulging residue (Val4926) in the corresponding helix turn (Yan et al., 2015), thereby suggesting that this mechanism could be shared with other evolutionarily related members of the 6-TM superfamily.

## DISCUSSION

Comparative sequence analysis of the TRP channel superfamily has illuminated subtle and unanticipated sequence and structural features that differentiate the TRPV1 structures from those of Kv channels, including the asymmetric solvation of the TRP sensor, the shift of the intersubunit sensor–pore domain interface away from the pore helix, and the register-shifting helical distortion in S6. Such structural differences are robustly represented in the ensemble statistics of the TRP and Kv sequences, implying their significance not only in TRPV1 but also throughout the TRP superfamily. Collectively, these results suggest that TRPV1 is a good representative of the TRP TM region. This insight, coupled with our nearly comprehensive MSA of TRP channel sequences, will facilitate the accurate comparative modeling of other TRP channels while more detailed structural information is collected.

TRP channels have diverged in sequence and structure for at least a billion years; common sequence motifs and patterns of amino acid covariation present in the ensemble of their extant sequences should therefore represent the action of common, constraining evolutionary pressures. We now attempt to rationalize these observations with previously conducted experiments in a model of TRP channel structure and function.

First, we note that despite structural similarity, the TRP sensor is not likely acting as a canonical voltage sensor in any TRP homologue. Unlike Kv channels, TRP channels feature asymmetric solvation of the sensor and a gating pore plugged by conserved aromatic residues. The primary dynamic gating element of Kv channels is the positively charged S4 helix (Aggarwal and MacKinnon, 1996). The requirement for voltage-dependent mobility in the Kv S4 has largely relaxed the pressure for highly tuned hydrophobic helical interaction with the contacting S1 and S3 helices, and we see no EC there (Fig. 3, D and E). On the other hand, in weakly voltage-sensitive TRP channels, only one positive position is found on S4 (TRP 551), and several positions of the S4 helix are evolutionarily coupled to positions in the S1 and S3 helices, consistent with a stationary, structural role for the TRP S4 (Fig. 4, A and B).

Also, Kv channels possess an S1–pore helix interface whose disruption yields nontrafficking or substantially modulated channels (Lee et al., 2009; Shem-Ad et al., 2013). This interface is almost entirely absent in the TRPV1 structure, as the sensor has shifted with respect to the pore domain (Fig. 4 B). The equivalent S1 positions are significantly more variable in TRP channels, and mutation of S629A in TRPV1 produces functional channels with normal capsaicin sensitivity and weakened but not abolished proton sensitivity, consistent with a more relaxed pressure on this region in TRP channels (Ryu et al., 2007). Most significantly, strong ECs observed at this interface in Kv channels are absent in TRP channels, appearing instead at the more substantial S4–S5 intersubunit interface (Fig. 4 A).

Why shift this interface away from the pore helix? If the lack of ECs on the hydrophobic interfaces of the S4 helix in Kv channels is indeed an imprint of its dynamic role in channel gating, our data suggest that the pore helix has taken up its role in TRP channels. Although the extensive coupling of TRP S4 to the S1, S3, and S5 helices strongly suggests that it is not making large movements during gating, the pore helix of TRP channels is largely decoupled with respect to the Kv pore helix, losing intersubunit ECs to the aforementioned S1 helix and the contacting S6 helix of a neighboring subunit. Two intrasubunit couplings mount the pore helix to S5 and S6 with hydrogen-bonded interactions in TRPV1 (TRP 584–641 and 637–666), and disruption of these interactions in TRPV1 (T641S) or yeast TRPY1 (where TRP 666 corresponds to TRPY1/YVC1 Y442C) yields constitutively active or completely nonfunctional channels, respectively (Su et al., 2007; Myers et al., 2008). Mutation of a conserved aromatic residue making a ternary contact with both these pairs (TRP 640) also yields a constitutively active channel (Myers et al., 2008). Additionally, structures of TRPV1 in three conformations show a different location of the pore helix in each conformation, suggesting that it is a highly dynamic element (Cao et al.,

2013). Several important gating modalities including proton potentiation, tarantula toxin binding, and possibly thermal gating localize to the TRPV1 pore architecture (Ryu et al., 2007; Bohlen et al., 2010; Grandl et al., 2010), and it has also been recently suggested that the weak voltage sensitivity of TRPV1 may be caused by the movement of the intrinsic dipole of the pore helix in response to membrane depolarization (Cao et al., 2013). Considering this body of evidence and the EC data, we speculate that the dynamic movement of the pore helix, mediated by the conserved and coevolving networks of residues described above, is a fundamental feature of gating in most, if not all, TRP channels.

TRP channels lack a glycine gating hinge (Kv 466, TRP 670) that is conserved in Nav, Cav, and Kv channels (Zhao et al., 2004). Even so, the resolved conformations of TRPV1 show the S6 helix pivoting in this region to open and close the intracellular gate (Cao et al., 2013). We traced the flexibility of the TRP S6 helix to a register-shifting helical distortion around TRP 671 and identified strong, TRP-specific ECs at the joint in S5–S6 that pin S6 into this distorted conformation. Like the half-twist of a Möbius strip, the broken backbone hydrogen bonds induced by this distortion can be shifted but not fully ironed out.

Free energy changes associated with shifts in frustrated backbone hydrogen bond networks have been estimated to be very small (<1 kcal/mol), consistent with the modest free energies measured for TRP channel activation (Brauchi et al., 2004; Yao et al., 2010; Cao and Bowie, 2012). We propose that the aforementioned motion of the pore helix is translated into pore gating by enhancing the relative stability of S6 backbone hydrogen bond networks that correspond to an open intracellular gate (Fig. 8 C). Importantly, the recently reported structure of another member of the TRP family, TRPA1, reveals an analogous distortion of S6 (Paulsen et al., 2015). This observation strongly supports our fundamental hypothesis concerning a common gating mechanism and lends credibility to the specific mechanism envisioned here. For both TRPV1 and TRPA1, binding of agonists or antagonists is able to induce subtle conformational transitions in the region surrounding the helical defect, thereby promoting a displacement of some of the pore-lining residues (Cao et al., 2013; Paulsen et al., 2015). Equivalently, this conformational equilibrium could be shifted by perturbations propagating from intracellular cytosolic domains. In this model, signals from distant domains detecting diverse stimuli could be integrated into one allosteric pore-opening step. Because this argument comes largely from the ensemble statistics of TRP channel sequences and not the specific features of a single channel, we expect this mechanism may underlie poly-modal gating throughout the superfamily.

Although the features identified in this study are robust measurements of the statistical properties of sequences

in the TRP superfamily, this model of TRP channel structure and gating is, of course, speculative. It should be seen as an early hypothesis-generating step in a full explanation of TRP channel gating rather than its culmination. Even so, the usefulness of the approximate answers obtainable through exploratory data analysis should not be trivialized (Tukey, 1962). We expect that this analysis will serve as a solid platform for the design and interpretation of experiments conducted across the diverse and daunting TRP channel superfamily.

We thank Sandipan Chowdhury and the members of the laboratory of Kenton Swartz for insightful critiques.

This work was supported by the National Institutes of Health via National Institute for General Medical Sciences (grant P01 GM55876 to M.L. Klein) and by the National Science Foundation (grant ACI 1440059 to M.L. Klein and V. Carnevale). L. Delemotte receives funding from European Union Seventh Framework Program “Voltsens” (grant PIOF-GA-2012-329534).

The authors declare no competing financial interests.

Kenton J. Swartz served as editor.

Submitted: 12 November 2014

Accepted: 11 May 2015

## REFERENCES

- Aggarwal, S.K., and R. MacKinnon. 1996. Contribution of the S4 segment to gating charge in the Shaker K<sup>+</sup> channel. *Neuron*. 16: 1169–1177. [http://dx.doi.org/10.1016/S0896-6273\(00\)80143-9](http://dx.doi.org/10.1016/S0896-6273(00)80143-9)
- Alabi, A.A., M.I. Bahamonde, H.J. Jung, J.I. Kim, and K.J. Swartz. 2007. Portability of paddle motif function and pharmacology in voltage sensors. *Nature*. 450:370–375. <http://dx.doi.org/10.1038/nature06266>
- Bandell, M., A.E. Dubin, M.J. Petrus, A. Orth, J. Mathur, S.W. Hwang, and A. Patapoutian. 2006. High-throughput random mutagenesis screen reveals TRPM8 residues specifically required for activation by menthol. *Nat. Neurosci.* 9:493–500. <http://dx.doi.org/10.1038/nn1665>
- Bohlen, C.J., A. Priel, S. Zhou, D. King, J. Siemens, and D. Julius. 2010. A bivalent tarantula toxin activates the capsaicin receptor, TRPV1, by targeting the outer pore domain. *Cell*. 141:834–845. <http://dx.doi.org/10.1016/j.cell.2010.03.052>
- Boukalova, S., L. Marsakova, J. Teisinger, and V. Vlachova. 2010. Conserved residues within the putative S4-S5 region serve distinct functions among thermosensitive vanilloid transient receptor potential (TRPV) channels. *J. Biol. Chem.* 285:41455–41462. <http://dx.doi.org/10.1074/jbc.M110.145466>
- Bradley, R.K., A. Roberts, M. Smoot, S. Juvekar, J. Do, C. Dewey, I. Holmes, and L. Pachter. 2009. Fast statistical alignment. *PLoS Comput. Biol.* 5:e1000392. <http://dx.doi.org/10.1371/journal.pcbi.1000392>
- Brauchi, S., P. Orio, and R. Latorre. 2004. Clues to understanding cold sensation: Thermodynamics and electrophysiological analysis of the cold receptor TRPM8. *Proc. Natl. Acad. Sci. USA*. 101:15494–15499. <http://dx.doi.org/10.1073/pnas.0406773101>
- Brauchi, S., G. Orta, M. Salazar, E. Rosenman, and R. Latorre. 2006. A hot-sensing cold receptor: C-terminal domain determines thermosensation in transient receptor potential channels. *J. Neurosci.* 26:4835–4840. <http://dx.doi.org/10.1523/JNEUROSCI.5080-05.2006>
- Cai, X. 2008. Unicellular Ca<sup>2+</sup> signaling ‘toolkit’ at the origin of metazoa. *Mol. Biol. Evol.* 25:1357–1361. <http://dx.doi.org/10.1093/molbev/msn077>

- Cao, E., M. Liao, Y. Cheng, and D. Julius. 2013. TRPV1 structures in distinct conformations reveal activation mechanisms. *Nature*. 504:113–118. <http://dx.doi.org/10.1038/nature12823>
- Cao, Z., and J.U. Bowie. 2012. Shifting hydrogen bonds may produce flexible transmembrane helices. *Proc. Natl. Acad. Sci. USA*. 109:8121–8126. <http://dx.doi.org/10.1073/pnas.1201298109>
- Chowdhury, S., B.W. Jarecki, and B. Chanda. 2014. A molecular framework for temperature-dependent gating of ion channels. *Cell*. 158:1148–1158. <http://dx.doi.org/10.1016/j.cell.2014.07.026>
- Clapham, D.E. 2003. TRP channels as cellular sensors. *Nature*. 426:517–524. <http://dx.doi.org/10.1038/nature02196>
- Crooks, G.E., G. Hon, J.-M. Chandonia, and S.E. Brenner. 2004. WebLogo: A sequence logo generator. *Genome Res*. 14:1188–1190. <http://dx.doi.org/10.1101/gr.849004>
- Doñate-Macián, P., and A. Perálvarez-Marín. 2014. Dissecting domain-specific evolutionary pressure profiles of transient receptor potential vanilloid subfamily members 1 to 4. *PLoS ONE*. 9:e110715. <http://dx.doi.org/10.1371/journal.pone.0110715>
- Eddy, S.R. 1998. Profile hidden Markov models. *Bioinformatics*. 14:755–763. <http://dx.doi.org/10.1093/bioinformatics/14.9.755>
- Efron, B., and R.J. Tibshirani. 1993. An Introduction to the Bootstrap. Chapman and Hall/CRC, Boca Raton, FL. 456 pp. <http://dx.doi.org/10.1007/978-1-4899-4541-9>
- Ekeberg, M., C. Lövkvist, Y. Lan, M. Weigt, and E. Aurell. 2013. Improved contact prediction in proteins: Using pseudolikelihoods to infer Potts models. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys*. 87:012707. <http://dx.doi.org/10.1103/PhysRevE.87.012707>
- Fischer, M.J., D. Balasuriya, P. Jeggle, T.A. Goetze, P.A. McNaughton, P.W. Reeh, and J.M. Edwardson. 2014. Direct evidence for functional TRPV1/TRPA1 heteromers. *Pflugers Arch*. 466:2229–2241. <http://dx.doi.org/10.1007/s00424-014-1497-z>
- Fodor, A.A., and R.W. Aldrich. 2004. On evolutionary conservation of thermodynamic coupling in proteins. *J. Biol. Chem*. 279:19046–19050. <http://dx.doi.org/10.1074/jbc.M402560200>
- Gentle, J.E. 2009. Computational Statistics. Springer, Berlin. 728 pp. <http://dx.doi.org/10.1007/978-0-387-98144-4>
- Göbel, U., C. Sander, R. Schneider, and A. Valencia. 1994. Correlated mutations and residue contacts in proteins. *Proteins*. 18:309–317. <http://dx.doi.org/10.1002/prot.340180402>
- Grandl, J., S.E. Kim, V. Uzzell, B. Bursulaya, M. Petrus, M. Bandell, and A. Patapoutian. 2010. Temperature-induced opening of TRPV1 ion channel is stabilized by the pore domain. *Nat. Neurosci*. 13:708–714. <http://dx.doi.org/10.1038/nn.2552>
- Guda, P., P.E. Bourne, and C. Guda. 2007. Conserved motifs in voltage-sensing and pore-forming modules of voltage-gated ion channel proteins. *Biochem. Biophys. Res. Commun*. 352:292–298. <http://dx.doi.org/10.1016/j.bbrc.2006.10.190>
- Halabi, N., O. Rivoire, S. Leibler, and R. Ranganathan. 2009. Protein sectors: Evolutionary units of three-dimensional structure. *Cell*. 138:774–786. <http://dx.doi.org/10.1016/j.cell.2009.07.038>
- Hille, B. 2001. Ion Channels of Excitable Membranes. Third edition. Sinauer Associates, Inc., Sunderland, MA. 814 pp.
- Kalia, J., and K.J. Swartz. 2013. Exploring structure-function relationships between TRP and Kv channels. *Sci Rep*. 3:1523. <http://dx.doi.org/10.1038/srep01523>
- Kobori, T., G.D. Smith, R. Sandford, and J.M. Edwardson. 2009. The transient receptor potential channels TRPP2 and TRPC1 form a heterotetramer with a 2:2 stoichiometry and an alternating subunit arrangement. *J. Biol. Chem*. 284:35507–35513. <http://dx.doi.org/10.1074/jbc.M109.060228>
- Krepkiy, D., M. Mihailescu, J.A. Freitas, E.V. Schow, D.L. Worcester, K. Gawrisch, D.J. Tobias, S.H. White, and K.J. Swartz. 2009. Structure and hydration of membranes embedded with voltage-sensing domains. *Nature*. 462:473–479. <http://dx.doi.org/10.1038/nature08542>
- Kumánovics, A., G. Levin, and P. Blount. 2002. Family ties of gated pores: evolution of the sensor module. *FASEB J*. 16:1623–1629. <http://dx.doi.org/10.1096/fj.02-0238hyp>
- Labro, A.J., A.L. Raes, A. Grottesi, D. Van Hoorick, M.S.P. Sansom, and D.J. Snyders. 2008. Kv channel gating requires a compatible S4-S5 linker and bottom part of S6, constrained by non-interacting residues. *J. Gen. Physiol*. 132:667–680. <http://dx.doi.org/10.1085/jgp.200810048>
- Lacroix, J.J., H.C. Hyde, F.V. Campos, and F. Bezanilla. 2014. Moving gating charges through the gating pore in a Kv channel voltage sensor. *Proc. Natl. Acad. Sci. USA*. 111:E1950–E1959. <http://dx.doi.org/10.1073/pnas.1406161111>
- Lee, S.-Y., A. Banerjee, and R. MacKinnon. 2009. Two separate interfaces between the voltage sensor and pore are required for the function of voltage-dependent K<sup>+</sup> channels. *PLoS Biol*. 7:e47. <http://dx.doi.org/10.1371/journal.pbio.1000047>
- Liao, M., E. Cao, D. Julius, and Y. Cheng. 2013. Structure of the TRPV1 ion channel determined by electron cryo-microscopy. *Nature*. 504:107–112. <http://dx.doi.org/10.1038/nature12822>
- Long, S.B., X. Tao, E.B. Campbell, and R. MacKinnon. 2007. Atomic structure of a voltage-dependent K<sup>+</sup> channel in a lipid membrane-like environment. *Nature*. 450:376–382. <http://dx.doi.org/10.1038/nature06265>
- Lu, Z., A.M. Klem, and Y. Ramu. 2002. Coupling between voltage sensors and activation gate in voltage-gated K<sup>+</sup> channels. *J. Gen. Physiol*. 120:663–676. <http://dx.doi.org/10.1085/jgp.20028696>
- Marti-Renom, M.A., M.S. Madhusudhan, and A. Sali. 2004. Alignment of protein sequences by their profiles. *Protein Sci*. 13:1071–1087. <http://dx.doi.org/10.1110/ps.03379804>
- Matsuura, H., T. Sokabe, K. Kohno, M. Tominaga, and T. Kadowaki. 2009. Evolutionary conservation and changes in insect TRP channels. *BMC Evol. Biol*. 9:228. <http://dx.doi.org/10.1186/1471-2148-9-228>
- Morcos, F., A. Pagnani, B. Lunt, A. Bertolino, D.S. Marks, C. Sander, R. Zecchina, J.N. Onuchic, T. Hwa, and M. Weigt. 2011. Direct-coupling analysis of residue coevolution captures native contacts across many protein families. *Proc. Natl. Acad. Sci. USA*. 108:E1293–E1301. <http://dx.doi.org/10.1073/pnas.1111471108>
- Morcos, F., B. Jana, T. Hwa, and J.N. Onuchic. 2013. Coevolutionary signals across protein lineages help capture multiple protein conformations. *Proc. Natl. Acad. Sci. USA*. 110:20533–20538. <http://dx.doi.org/10.1073/pnas.1315625110>
- Myers, B.R., C.J. Bohlen, and D. Julius. 2008. A yeast genetic screen reveals a critical role for the pore helix domain in TRP channel gating. *Neuron*. 58:362–373. <http://dx.doi.org/10.1016/j.neuron.2008.04.012>
- Narsky, I., and F.C. Porter. 2014. Statistical Analysis Techniques in Particle Physics: Fits, Density Estimation and Supervised Learning. Wiley, Hoboken, NJ. 459 pp.
- Owsianik, G., D. D’hoedt, T. Voets, and B. Nilius. 2006. Structure–function relationship of the TRP channel superfamily. In *Reviews of Physiology, Biochemistry and Pharmacology*. S.G. Amara, E. Bamberg, S. Grinstein, S.C. Hebert, R. Jahn, W.J. Lederer, R. Lill, A. Miyajima, H. Murer, S. Offermanns, et al., editors. Springer, Berlin. 61–90. <http://dx.doi.org/10.1007/s10254-005-0006-0>
- Palovcak, E., L. Delemotte, M.L. Klein, and V. Carnevale. 2014. Evolutionary imprint of activation: The design principles of VSDs. *J. Gen. Physiol*. 143:145–156. <http://dx.doi.org/10.1085/jgp.201311103>
- Paulsen, C.E., J.-P. Armache, Y. Gao, Y. Cheng, and D. Julius. 2015. Structure of the TRPA1 ion channel suggests regulatory mechanisms. *Nature*. 520:511–517. <http://dx.doi.org/10.1038/nature14367>
- Payandeh, J., T. Scheuer, N. Zheng, and W.A. Catterall. 2011. The crystal structure of a voltage-gated sodium channel. *Nature*. 475:353–358. <http://dx.doi.org/10.1038/nature10238>

- Peng, G., X. Shi, and T. Kadowaki. 2015. Evolution of TRP channels inferred by their classification in diverse animal species. *Mol. Phylogenet. Evol.* 84:145–157. <http://dx.doi.org/10.1016/j.ympev.2014.06.016>
- Pless, S.A., A.P. Niciforovic, J.D. Galpin, J.-J. Nunez, H.T. Kurata, and C.A. Ahern. 2013. A novel mechanism for fine-tuning open-state stability in a voltage-gated potassium channel. *Nat. Commun.* 4:1784. <http://dx.doi.org/10.1038/ncomms2761>
- Price, M.N., P.S. Dehal, and A.P. Arkin. 2010. FastTree 2—Approximately maximum-likelihood trees for large alignments. *PLoS ONE* 5:e9490. <http://dx.doi.org/10.1371/journal.pone.0009490>
- Quenouille, M.H. 1949. Problems in plane sampling. *Ann. Math. Stat.* 20:355–375. <http://dx.doi.org/10.1214/aoms/1177729989>
- Reeds, J.A. 1978. Jackknifing maximum likelihood estimates. *Ann. Stat.* 6:727–739. <http://dx.doi.org/10.1214/aos/1176344248>
- Roberts, E., J. Eargle, D. Wright, and Z. Luthey-Schulten. 2006. MultiSeq: unifying sequence and structure data for evolutionary analysis. *BMC Bioinformatics*. 7:382. <http://dx.doi.org/10.1186/1471-2105-7-382>
- Ryu, S., B. Liu, J. Yao, Q. Fu, and F. Qin. 2007. Uncoupling proton activation of vanilloid receptor TRPV1. *J. Neurosci.* 27:12797–12807. <http://dx.doi.org/10.1523/JNEUROSCI.2324-07.2007>
- Shao, J., and D. Tu. 1995. The Jackknife and Bootstrap. Springer, Berlin. 517 pp. <http://dx.doi.org/10.1007/978-1-4612-0795-5>
- Shao, J., and C.F.J. Wu. 1989. A general theory for jackknife variance estimation. *Ann. Stat.* 17:1176–1197. <http://dx.doi.org/10.1214/aos/1176347263>
- Shem-Ad, T., O. Irit, and O. Yifrach. 2013. Inter-subunit interactions across the upper voltage sensing-pore domain interface contribute to the concerted pore opening transition of Kv channels. *PLoS ONE*. 8:e82253. <http://dx.doi.org/10.1371/journal.pone.0082253>
- Söding, J. 2005. Protein homology detection by HMM-HMM comparison. *Bioinformatics*. 21:951–960. <http://dx.doi.org/10.1093/bioinformatics/bti125>
- Starace, D.M., and F. Bezanilla. 2004. A proton pore in a potassium channel voltage sensor reveals a focused electric field. *Nature*. 427:548–553. <http://dx.doi.org/10.1038/nature02270>
- Su, Z., X. Zhou, W.J. Haynes, S.H. Loukin, A. Anishkin, Y. Saimi, and C. Kung. 2007. Yeast gain-of-function mutations reveal structure-function relationships conserved among different subfamilies of transient receptor potential channels. *Proc. Natl. Acad. Sci. USA*. 104:19607–19612. <http://dx.doi.org/10.1073/pnas.0708584104>
- Süel, G.M., S.W. Lockless, M.A. Wall, and R. Ranganathan. 2003. Evolutionarily conserved networks of residues mediate allosteric communication in proteins. *Nat. Struct. Biol.* 10:59–69. <http://dx.doi.org/10.1038/nsb881>
- Thomas, J.A. 2006. Elements of Information Theory. Second edition. Wiley, Hoboken, NJ. 776 pp.
- Tukey, J.W. 1962. The future of data analysis. *Ann. Math. Stat.* 33:1–67. <http://dx.doi.org/10.1214/aoms/1177704711>
- Weigt, M., R.A. White, H. Szurmant, J.A. Hoch, and T. Hwa. 2009. Identification of direct residue contacts in protein-protein interaction by message passing. *Proc. Natl. Acad. Sci. USA*. 106:67–72. <http://dx.doi.org/10.1073/pnas.0805923106>
- Winking, M., D.C. Hoffmann, C. Kühn, R.-D. Hilgers, A. Lückhoff, and F.J.P. Kühn. 2012. Importance of a conserved sequence motif in transmembrane segment S3 for the gating of human TRPM8 and TRPM2. *PLoS ONE*. 7:e49877.
- Winter, Z., A. Buhala, F. Ötvös, K. Jósavay, C. Vizler, G. Dombi, G. Szakonyi, and Z. Oláh. 2013. Functionally important amino acid residues in the transient receptor potential vanilloid 1 (TRPV1) ion channel—an overview of the current mutational data. *Mol. Pain*. 9:30. <http://dx.doi.org/10.1186/1744-8069-9-30>
- Yan, Z., X.C. Bai, C. Yan, J. Wu, Z. Li, T. Xie, W. Peng, C.C. Yin, X. Li, S.H. Scheres, et al. 2015. Structure of the rabbit ryanodine receptor RyR1 at near-atomic resolution. *Nature*. 517:50–55. <http://dx.doi.org/10.1038/nature14063>
- Yao, J., B. Liu, and F. Qin. 2010. Kinetic and energetic analysis of thermally activated TRPV1 channels. *Biophys. J.* 99:1743–1753. <http://dx.doi.org/10.1016/j.bpj.2010.07.022>
- Yifrach, O., and R. MacKinnon. 2002. Energetics of pore opening in a voltage-gated K<sup>+</sup> channel. *Cell*. 111:231–239. [http://dx.doi.org/10.1016/S0092-8674\(02\)01013-9](http://dx.doi.org/10.1016/S0092-8674(02)01013-9)
- Yu, F.H., and W.A. Catterall. 2004. The VGL-chanome: A protein superfamily specialized for electrical signaling and ionic homeostasis. *Sci. STKE*. 2004:re15.
- Zhao, Y., V. Yarov-Yarovoy, T. Scheuer, and W.A. Catterall. 2004. A gating hinge in Na<sup>+</sup> channels: A molecular switch for electrical signaling. *Neuron*. 41:859–865. [http://dx.doi.org/10.1016/S0896-6273\(04\)00116-3](http://dx.doi.org/10.1016/S0896-6273(04)00116-3)