Predicting lipid and ligand binding sites in TRPV1 channel by molecular dynamics simulation and machine learning

Wenjun Zheng and Han Wen

Department of Physics, State University of New York at Buffalo, Buffalo, NY 14260

Running title: small-molecule binding sites in TRPV1

Address: 239 Fronczak Hall, Buffalo, NY 14260

Phone: (716) 645-2947

Fax: (716) 645-2507

E-mail: wjzheng@buffalo.edu

Abstract

As a key cellular sensor, the TRPV1 channel undergoes a gating transition from a closed state to an open state in response to many physical and chemical stimuli. This transition is regulated by small-molecule ligands including lipids and various agonists/antagonists, but the underlying molecular mechanisms remain obscure. Thanks to recent revolution in cryo-electron microscopy, a growing list of new structures of TRPV1 and other TRPV channels have been solved in complex with various ligands including lipids. Toward elucidating how ligand binding correlates with TRPV1 gating, we have performed extensive molecular dynamics simulations (with cumulative time of 20 μs), starting from high-resolution structures of TRPV1 in both the closed and open states. By comparing between the open and closed state ensembles, we have identified state-dependent binding sites for small-molecule ligands in general and lipids in particular. We further use machine learning to predict top ligand-binding sites as important features to classify the closed vs open states. The predicted binding sites are thoroughly validated by matching homologous sites in all structures of TRPV channels bound to lipids and other ligands, and with previous functional/mutational studies of ligand binding in TRPV1. Taken together, this study has integrated rich structural, dynamic, and functional data to inform future design of small-molecular drugs targeting TRPV1.

Key words: closed state, gating, ligand binding, lipid binding, machine learning, molecular dynamics, open state, random forest, TRPV1 channel

Introduction

The transient receptor potential (TRP) channels are a superfamily of cation channels, which are activated by various physical and chemical stimuli such as heat [1,2], cold [3-5], voltage [6], acid [7,8], force [9-11], and small-molecule agonists (e.g., capsaicin [2]). The TRP channels make promising drug targets [12,13] because of their involvements in many signaling pathways linked to various diseases [14,15]. As a prototype TRP channel, TRPV1 forms a homo-tetramer, with each subunit comprised of a six-helix (S1-S6) transmembrane domain (TMD) and an intracellular domain (ICD) (see Fig 1). Like canonical voltage-gated ion channels [16,17], the TMD of TRPV1 consists of two structural domains — the S1-S4 domain on the channel periphery and the S5-S6 pore domain enclosing a central pore (see Fig 1). However, TRPV1 lacks a charged S4 helix so its S1-S4 domain serves a different function than a voltage sensor. The pore domain features a channel pore with two constrictions/gates at G643 and I679 (see Fig 1). The N-terminal portion of ICD forms an ankyrin repeats domain (ARD) which can bind various proteins and ligands [18]. The C-terminal domain (CTD) of ICD contains a highly conserved alpha helix known as the TRP helix [19], which is implicated in the coupling of stimulus sensing to channel gating [20] and interactions with other proteins and ligands [21,22]. At the TMD-ICD interface is a membrane proximal domain (MPD, including a linker of residues 400-415, see Fig 1), which may serve as a mediator between ARD and TMD and a heat sensor [23]. The above functional domains are connected by several flexible linkers (see Fig 1), including the S2-S3 linker involved in vanilloid binding [24], and the S4-S5 linker which may mechanically control the channel gate in voltage-gated ion channels [25] and TRPV1 [24].

High-resolution structures of full-length TRP channels are required to understand their gating and regulation mechanisms at the molecular level. In 2013, the labs of Julius and Cheng used cryo-EM to determine

the first structures of a minimal functional construct of the rat TRPV1 in three distinct forms at 3-4 Å resolutions [24,26] — a closed apo structure, an open structure bound with double-knot toxin (DkTx, a peptide toxin) and resiniferatoxin (RTX, a vanilloid agonist), and a partially open structure bound with capsaicin (a vanilloid agonist). Improved TRPV1 structures were later solved in lipid nanodiscs, revealing details of protein-lipid/ligand interactions [27] . Recently, a growing list of high-resolution structures have been solved for TRPV2, V3, V4, V5, and V6 [28-47], paving the way for quantitative structure-driven studies of the TRP-channel gating mechanism.

Molecular Dynamics (MD) simulation is the method of choice for investigating protein dynamics and energetics under physiological conditions (e.g. in the presence of water, lipids, and ions) with atomic details [48], which has been applied to various ion channels [49-54]. However, MD simulation is highly expensive, demanding the use of a massively parallelized or special-purpose supercomputer (such as Anton [52]). Thanks to recent developments in computing hardware and software (e.g., the use of graphics processing units to accelerate MD simulation [55]), one can now routinely simulate a large biomolecular system (with $\geq 10^5$ atoms) at a speed of 1-10 ns/day on a single computer node. However, it remains challenging for MD to access the μs - ms time scales relevant to many biomolecular transitions (including the gating transition in TRPV1). Among alternative strategies to overcome the time-scale limit of MD simulation, coarse-grained modeling (e.g., the elastic network model [56-58]) has been developed using reduced protein representations and simplified force fields [59,60], and applied to dynamic analysis of ion channels [61-64].

In our recent studies [62,65-67], we utilized all-atom MD simulation to probe the heat activation mechanism of TRPV1. To this end, we have accumulated extensive MD trajectories starting from both the closed and open structures of TRPV1 at different temperatures. Using standard structural and energetic analysis tools for MD data

[64,67-69], we have identified key residues involved in heat activation, but other activation mechanisms (e.g., by ligand binding) remain unexplored.

In the present study, we perform longer (microseconds) MD simulation of TRPV1 in the closed (C) and the open (O) state, and use machine learning to identify residue-specific features that distinguish the C-state and the O-state ensembles. We focus on specific sites accessible to binding of small-molecule ligands in a state-dependent manner to regulate the thermodynamic equilibrium of TRPV1 gating. The underlying hypothesis is: those binding pockets specifically formed in the C (O) state make promising targets for ligand antagonists (agonists) to decrease (increase) TRPV1 activity. We first focus on lipid-binding sites in TRPV1, and then extend to ligand-binding sites in general. We thoroughly validate the predicted ligand- and lipid-binding sites in comparison with a growing list of new cryo-EM structures of TRPV channels, and with functional/mutational data from the literature.

Materials and Methods

MD simulation setup

We downloaded pre-oriented PDB files of the closed and open structures of TRPV1 (PDB ids: 5irz and 5irx) from the OPM database [68]. Following ref [66], we completed these structures by adding the ARD residues (residues 110-334) of the full-length structures of TRPV1 (PDB ids: 3j5p and 3j5q) after superimposing along residues 335-384. We used the Membrane Builder function [69] of the CHARMM-GUI webserver [70] to embed TRPV1 in a bilayer of 1-palmitoyl-2-oleoyl phosphatidylcholine (POPC) lipids surrounded by a box of water and ions (with a buffer distance of 15 Å). K^+ and Cl^- ions were added to ensure 0.15 M ionic concentration and zero net charge. The entire system contains ~265,000 atoms. Following our previous study [67], to maintain the key inter-subunit contacts between each ARD and the MPD of an adjacent subunit, we added inter-subunit harmonic restraints between the C_α atoms of residue pairs 376–245 and 375–210 with a spring constant of 500 kJ/mol/nm^2. After energy minimization, six steps of equilibration were performed (with gradually reduced harmonic restraints applied to protein, lipids, water, and ions). Finally we conducted five independent production runs of 2000 ns in the NPT ensemble. The Nose–Hoover method [71] was used with temperature T = 30°C. The Parrinello-Rahman method [72] was used for pressure coupling. A 10-Å switching distance and a 12-Å cutoff distance were used for nonbonded interactions. The particle mesh Ewald method [73] was used for electrostatics calculations. The LINCS algorithm [74] was used to constrain the hydrogen-containing bond lengths, which allowed a 2-fs time step for MD simulation. The energy minimization and MD simulation were carried out with the GROMACS program [75] version 5.0.3 using the CHARMM36 force field [76] and the TIP3P water model [77].

Lipid-binding residues detection and scoring

For a given residue position n and an MD trajectory, we calculate the fraction f_{nm} of time when it forms polar atomic contacts (with minimal oxygen/nitrogen atomic distance < 5 Å) with a lipid m. The sum of $f_n = \sum_m f_{nm}$ is then averaged over all five trajectories of the C/O-state ensemble, and added up for four symmetrically related residues of the tetramer. By focusing on polar atomic contacts, we aim to detect specific electrostatic interactions between each residue and the lipid head groups, while excluding nonspecific hydrophobic contacts with lipid tails. Electrostatic interactions are particularly relevant to anionic phospholipids like PIP2 [78]. By computing differences in average f_n between the C and O state ensembles, we identify lipid-binding sites that differ most in lipid accessibility between the two states.

Ligand-binding pocket detection and scoring

The TRPV1 conformations generated by MD are pooled into a C/O-state ensemble for binding pocket detection by two state-of-the-art programs (fpocket and concavity, downloaded from: http://fpocket.sourceforge.net/ and http://compbio.cs.princeton.edu/concavity/). Both programs [79,80] combine physical, chemical, and geometrical features to identify potential ligand-binding pockets in a given protein conformation. Following Cimermancic et al. [81], we define the fpocket score for a given residue position as the maximum druggability score among the alpha spheres within 5 Å of that residue, and the concavity score is generated by concavity directly on a per-residue basis. Each per-residue pocket score s_n (ranging between 0 and 1) is then averaged over the conformational ensemble, and then added up for four symmetrically related residues of the tetramer.

Machine learning by the random forest algorithm

Random forest (RF) is an ensemble learning method for classification that operates by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes of the individual trees [82]. Using the RandomForest package of R (https://cran.r-project.org/web/packages/randomForest/randomForest.pdf), we set the hyper-parameters as follows: ntree = 1000, mtry ~ sqrt(p), where p = number of residue-specific features. The training set for ML consists of 500 snapshots of MD simulations from the C/O-state ensemble, each represented by a feature vector of p dimensions. Each vector element stores the number of lipids in polar atomic contact with a given residue position in a snapshot. Five-fold cross-validation was used to ensure no over-fitting. Then all training-set data was used to generate a final RF classifier. The relative importance scores of all features were computed by the Mean Decrease in Gini. We chose RF owning to its robustness in selecting important features based on noisy MD data.

Results and discussion

Microseconds MD simulation of TRPV1 in the C and O state

In our previous studies [66,67], we explored the conformational dynamics of TRPV1 in the C/O state by conducting multiple 200-ns MD simulations of TRPV1 in the presence of a lipid bilayer, water, and ions. The MD trajectories were then combined to form a structural ensemble of the C/O state for comparative analysis. These simulations were relatively short compared to the sub-millisecond gating dynamics of TRPV1, and the C and O states were simulated at different temperature (30°C for C and 60°C for O), making it not straightforward to compare their dynamics. To address these caveats, we have performed longer MD simulations of TRPV1 in the C/O state at the same temperature (30°C) using the latest cryo-EM TRPV1 structures [27] (see Methods). Five 2000-ns trajectories were produced in each state, giving significantly more extensive conformational sampling than our previous studies [66,67].

To assess the conformational stability of TRPV1 in MD simulation, we calculated the root mean square deviation (RMSD) of Cα atoms relative to the initial structure for each state (see Fig S1). In our earlier study [67], we found high RMSD in the TRPV1 tetramer in both states owning to its highly flexible ARD [24], which is consistent with the notion that the heat-sensitive TRPV1 channel possesses high entropy (particularly in the O state) [83]. However, such high flexibility in ARD obscures the dynamics of TMD. Therefore, we limit RMSD calculation to the core TRPV1 structure (denoted the expanded TMD, including residues 400-710 comprising the MPD linker, S1-S6 helices, and the TRP helix, see Fig 1a). We observed relatively low RMSD in the C state (2.1±0.1 Å, see Fig S1) and higher RMSD in the O state (2.7±0.2 Å, see Fig S1), which are comparable to our previous shorter MD simulations [67] .

The higher RMSD in the O-state simulation could be attributed to the following two causes: 1. The removal of DkTx and RTX could cause TRPV1 to undergo further structural relaxation during the simulation. 2. TRPV1 could undergo a transition toward the C state because the O state is thermostatically less stable than the C state at 30˚C [83]. We measured RMSD for the O-state trajectories relative to the closed structure (PDB id: 5irz), which does not decrease as a function of time (see Fig S1). Therefore, the larger structural changes in the O-state simulation are likely due to relaxation in the absence of DkTx and RTX instead of the backward O-to-C transition.

The O state ensemble exhibits greater flexibility than the C state ensemble

The thermodynamics of heat activation requires TRPV1 to have higher enthalpy and entropy in the O state than the C state [83]. To see this, we assess the conformational flexibility at individual residue positions in the

expanded TMD by calculating the root mean square fluctuation (RMSF) of Cα atoms in the C/O-state ensemble [67]. For the C-state ensemble at 30°C, the RMSF profile exhibits pronounced peaks at the MPD linker, the S1-S2 linker, the S2-S3 linker, the outer pore, and the C-terminus of the TRP helix (see Fig 2a). For the O-state ensemble at 30°C, the RMSF profile features higher peaks in the above regions (see Fig 2a and 2b). This confirms our previous finding of a more flexible O-state ensemble than the C-state ensemble at different temperatures [67]. This finding provides a dynamic explanation for high entropy in the O state of TRPV1 essential to its heat activation. Additionally, the above flexible regions may also be targeted by ligand binding to modulate TRPV1 gating (see below).

TRPV1 undergoes modest conformational changes under physiological conditions

A high-quality MD simulation is expected to both maintain the global integrity of the initial experimental structure and explore conformational changes in response to a change of environmental conditions (e.g., removal of ligands, higher temperature, etc). To show this for our MD simulation, we calculated the average structures for the C-state and O-state ensembles, and superimposed them onto the initial cryo-EM structures (see Fig 3). The averaged structures of the TRPV1 expanded TMD showed modest changes both relative to the initial structures (RMSD = 1.4 Å and 1.9 Å for C state and O state, respectively) and between the two states (RMSD = 1.9 Å).

In the C state, we observed the following structural changes relative to the initial structure (see Fig 3a): the outer-pore region moving inward, the S2-S3 linkers also moving inward (toward the S4-S5 linkers), and the MPD linkers moving toward the TRP helices. In contrast, the rest of S1-S4 domain and pore domain showed little changes. Taken together, these small changes support the overall stability of the closed structure captured by cryo-EM [27], and hint for increasing couplings between the above functional domains under physiological conditions.

In the O state, we observed greater changes from the initial structure (see Fig 3b): the outer-pore region undergoing local restructuring (in response to removal of DkTx), the S2 C-terminus and S2-S3 linkers moving

inward (toward the S4-S5 linkers), and the MPD linkers moving closer to the TRP helices. Some of these changes resemble the C state (see Fig 3a), which may result from the change of environmental conditions from cryo-EM sample to MD system with water/membrane/ions.

Between the C and O states (see Fig 3c), we observed coupled inward motions of the MPD linkers, S2-S3 linkers, and TRP helices, along with small outward motions of S5/S6 helices as expected from pore opening. Surprisingly, instead of pulling the TRP helices (and the adjoined lower S6 helices) outward to open the lower gate, these motions cause a narrowing of the channel at the intracellular side of TMD [67].

In sum, our structural comparison of C/O-state ensembles revealed new modest conformational changes that deviate from the cryo-EM structures, supporting the value of MD simulation in exploring functional motions of proteins under physiological conditions. It is natural to ask if the above small C-to-O conformational changes are functionally significant in modulating the relative stability of C vs O state. To partially address this issue, we focus on their impact on the binding accessibility of small-molecule ligands (and lipids in particular).

State-specific lipid-binding sites distinguish between the C and O states

TRPV1 gating is regulated by various lipids [78] including Cholesterol [84] and Phosphatidylinositol 4,5-bisphosphate (PIP2) [85,86]. Although bound lipids have been resolved by cryo-EM in recent structures of TRPV1[27] and other TRPV channels [29,39,46], the functional roles of these lipid-binding sites remain obscure. Our extensive MD simulations of TRPV1 explore dynamic interactions between TRPV1 and surrounding lipids in both the C and O states. This rich MD data allows us to identify key sites involved in lipid binding [87] in a state-dependent fashion. To probe specific protein-lipid interactions relevant to lipid regulation, we calculated the fraction of time f_n when a given residue n forms polar atomic contacts with lipids (see Methods) in the C or O state ensembles (see Fig 4a, 4b).

We found numerous lipid-binding residues (with average f_n >0.5 in C or O state) in the following regions (see Fig 4a, 4b): 405 (in the MPD linker), 427-437, 453-472 (in the S1-S2 linker), 487-508 (in the S2-S3 linker), 530-536 (in the S3-S4 linker), 559-579 (in the S4-S5 linker), 628-632 and 652-657 (in the outer pore), 707 and 710 (in the TRP helix). Most of these regions (e.g. the MPD linker, the S2-S3 linker, the S4-S5 linker, the outer pore, and the TRP helix) are involved in the C-to-O conformational changes (see Fig 3c) and state-dependent interactions observed in our previous MD studies [66,67], allowing lipid binding to modulate the gating transition. For validation of the predicted lipid-binding residues, we compare them with the lipid-binding sites observed in TRPV structures (Table S1) which were identified using the same criterion of polar atomic contacts (see Methods). The lipid-binding sites in non-TRPV1 structures were mapped back to TRPV1 by sequence alignment (Table S1). Among 64 known lipid-binding residues, 50% exactly match the predicted lipid-binding residues (see Fig 4a, 4b).

To functionally annotate the above lipid-binding sites, we next focus on a small subset of them distinctly formed in the C and O states. To this end, we subtract the average f_n between the C and O states (see Fig 4c), and select top 10 residues with the most positive / negative differences (E405, R455, L465, N467, Y472, P501, L506, Y530, K535, R575, see Fig 4e):

E405 (in the MPD linker) and R575 (in the S4-S5 linker) have lower f_n in the O state than the C state, which is consistent with our previous finding that E405 forms hydrogen bond with R575 in the O state [67]. If the coupling between MPD linker and S4-S5 linker via E405 and R575 is essential to gating, then lipid binding to these residues can potentially inhibit gating. In several earlier studies, R575 was proposed to mediate PIP2 binding and associated TRPV1 activation [88-91]. However, R575 was also observed to bind lipid in a closed structure of TRPV3 (PDB id: 6uw4)[29]. Our finding favors its role in allowing lipid binding to stabilize the C state.

P501 and L506 (in the S2-S3 linker) have higher f_n in the O state than the C state, thus allowing lipid binding to stabilize the O state. The C-to-O transition features the S2-S3 linker moving inward (toward the S4-

S5 linker, see Fig 3c). Therefore, lipid binding to the S2-S3 linker can enable activation via this motion. In support of this idea, a recent cryo-EM study of TRPV5 (PDB id: 6dmu)[34] observed a PIP2-binding site at residues 416-419 (corresponding to residues 503-506 of TRPV1).

The remaining six lipid-binding residues are located in the S1-S2 linker (R455, L465, N467, Y472) and the S3-S4 linker (Y530, K535) at the extracellular side of TMD. They may explain the extracellular PIP2-binding sites [85]. Interestingly, these residues are also involved in non-lipid ligand binding: R455 and Y530 correspond to a 2-APB binding site in a TRPV3 structure (PDB id: 6dvz)[41]; K535 was implicated in vanilloid binding by a mutational study [92].

The above analysis of ensemble-averaged f_n differences does not take into account the dynamics of lipid binding (i.e. residue-lipid contacts forming and breaking during MD simulation) in the C/O state. To remedy this caveat, we employ machine learning (ML) to train a classifier to distinguish the C and O state ensembles based on their differences in residue-lipid contacts. We choose the random forest algorithm for ML to select key residue-specific features for classification (see Methods). After training, we evaluate the importance of all residues (see Fig 4d) and select top 10 features corresponding to key residues where lipid binding changes significantly between the C and O states (see Fig 4e).

The ML-identified residues (W426, V457, P461, N467, P501, L506, Y530, K535, K571, R575) mostly overlap with the residues found by the analysis of average f_n differences (see Fig 4e). Additional sites like K571 and W426 were uncovered by ML: the K571 site was found to bind lipid in three cryo-EM structures of TRPV1 (PDB id: 5irz)[27], TRPV5 (PDB id: 6dmu)[34], and TRPV6 (PDB id: 6bo8)[36]; W426 is critically positioned to interact with the TRP helix, so lipid binding at this site may modulate gating by moving the TRP helix (see Fig 3c).

In sum, by comparing lipid-binding residues between the C and O state ensembles, we identified a promising list of lipid-binding residues linked to lipid regulation of the equilibrium between the C and O states. The predicted residues are located in well-known functional domains including the MPD linker, the S2-S3 linker, and the S4-S5 linker, and less studied regions like S1-S2 and S3-S4 linkers (see Fig 4e). Future functional/mutational studies targeting these sites will test their roles in lipid regulation of TRPV1 gating.

State-dependent ligand-binding pockets differentiate between the C and O states

Going beyond lipid-binding sites, we next search for possible binding pockets for any small-molecule ligands in TRPV1. A key requirement for successful drug design is to identify potential binding sites for small-molecule ligands to modulate target protein functions. These sites are often located in exposed and concave pockets with certain structural, dynamical, and physiochemical characteristics that favor molecular interactions. When such pockets are already formed in a ligand-unbound protein structure, various computational methods are available to detect them with reasonable accuracy (such as fpocket and concavity) [79,80]. However, in many cases, such pockets are cryptic in the absence of ligands [93], and extensive conformational sampling is needed to capture the conformation with the desired pocket properly formed and detectable [81]. Our extensive MD simulations of TRPV1 explore conformational changes in protein surface affecting its accessibility to ligand binding in both the C and O states. This rich MD data may allow us to identify key sites for ligand binding in a state-dependent fashion which can tune the C-to-O thermodynamic equilibrium to regulate channel activity. To accurately find possible ligand-binding pockets, we have applied two state-of-the-art pocket-finding methods (fpocket and concavity [79,80]) to individual MD snapshots in the C or O state. The two methods are based on distinct algorithms with fpocket favoring nonpolar sites while concavity considering both polar and nonpolar sites [79,80]. So their combination allows different types of binding pockets to be discovered. Either method generates a summary

pocket score between 0 and 1 at each residue position (see Methods), which is then averaged over the C/O-state ensemble, and added for four symmetric residue positions of the tetramer. A cutoff average score of 0.5 is used to select pocket residues.

Consistent with the observation of small C-to-O conformational changes (see Fig 3c), the predicted pocket residues are similarly distributed in the C and O states (see Fig 5 and Fig 6). The fpocket-predicted pocket residues are concentrated at the extracellular side of S1-S4 domain (see Fig 5a, 5b), including F448, A451, A452, R455, V475, G477, E478, I479, S481, V482, M523, S526, V527, Y530, F531, Y537, M541. They are dominated by nonpolar residues as favored by the fpocket algorithm. In contrast, the concavity-predicted pocket residues are more widely distributed at the intracellular side of TMD, involving the following regions (see Fig 6a, 6b): the ARD (E293, A295, D296, N297, T298, V299, S343, G344, K345, I346), the MPD linker (S402, S403, S404, E405, T406, P407, N408, R409, D411, L414, E416), the S2-S3 linker (V508, D509), the S4-S5 linker (Q561, K571, R575), the TRP helix (E684, N687, K688, I689, A690, Q691, E692, S693, K694, N695, L699), and the CTD (N748, W749, T750, T751). Notably, most concavity-predicted pocket residues are near the central intracellular cavity (see Fig 1a) where water, lipids, and other ligands can readily access to modulate the energetics of gating.

To functionally characterize the above pocket residues, we focus on a subset of them preferably formed for ligand binding in the C or O state. To this end, we first calculate the C-to-O differences in average pocket scores and select top 10 residues with the most positive or negative differences, then we use ML to identify top 10 residue-specific features that distinguish between the C and O states. To validate the predicted pocket residues' role in ligand binding, we analyze all TRPV structures bound with various biologically relevant ligands (see Table

S2) and check if the observed ligand-binding residues overlap with our predicted pocket residues (based on sequence alignment between TRPV1 and other TRPV channels).

We first discuss the results of fpocket. The top 10 residues with maximal C-to-O change in pocket score are F448, A451, A452, R455, I479, S481, M523, S526, F531, M541, all of which exhibit negative change in pocket score (see Fig 5c), suggesting that they are more accessible to ligand binding in the C state than the O state. To incorporate dynamic fluctuations in pocket score, we have performed ML using the RF algorithm to select top 10 features to distinguish the C state from the O state (see Methods). The top 10 residues with maximal feature importance are F448, A451, A452, P461, S481, V482, M523, V527, Y537, M541 (see Fig 5d). They mostly overlap with the residues identified based on the C-to-O change in pocket score, confirming the importance of the extracellular S1-S4 site in state-dependent ligand binding. In support of the functional importance of this site, a new TRPV3 structure (PDB id: 6dvz) [41] was solved with 2-APB bound to this site (contacting residues corresponding to A451, R455, I479, S481, V482, V527, M541 of TRPV1). In TRPV3, this site (site 4 in ref [41]) was bound to 2-APB in the O state but not the C state, which suggests that ligand binding at this site may stabilize the O state. Further supporting the involvement of these residues in ligand binding, mutations F522L, M523L, K535E, and E536L/W ablated the vanilloid sensitivity [94]; mutations V538A/G/I/L/T and A539P ablated proton activation/response [95].

We next discuss the results of concavity. The top 10 residues with maximal C-to-O change in pocket score are S402, S403, S404, E405, T406, R409, D411, E416, K571, S693 (see Fig 6c), most of which (except D411 and E416) show negative change in pocket score, suggesting that they are more accessible to ligand binding in the C state than the O state. A majority of them are in the MPD linker, while K571 is in the S4-S5 linker and S693 is in the TRP helix (see Fig 6e). To incorporate dynamic fluctuations in pocket score, we performed ML using RF to select top 10 residues with maximal feature importance (see Fig 6d, 6e): K238, S403, T406, R409, L503,

F507, V567, A680, M682, S693, which are more widely distributed in the ARD (K238), the MPD linker (S403, T406, R409), the S2-S3 linker (L503, F507), the S4-S5 linker (V567), the S6 helix (A680, M682), and the TRP helix (S693). The above predicted pocket residues are supported by evidence from ligand-bound TRPV structures and other functional studies (see Table S2) summarized as follows:

In the MPD linker, A400 corresponds to V297 which is bound with an inhibitor in a TRPV5 structure (PDB id: 6pbe) [32]; P407 and H410 correspond to R302 and R305 which are bound with PIP2 in a TRPV5 structure (PDB id: 6dmu) [34]; H410, L413, and V415 correspond to H417, L420, and L422 which are bound with 2-APB in two TRPV3 structures (PDB ids: 6dvz and 6ot5) [41].

In the S2-S3 linker, L503 corresponds to F416 bound with PIP2 in a TRPV5 structure (PDB id: 6dmu) [34]; F507 is at a site bound with multiple ligands (including lipid, RTX, inhibitor, 2-APB) in several TRPV structures (PDB ids: 6lgp of TRPV3 [39] , 5irx of TRPV1 [27], 6pbe of TRPV5 [32] , 6d7x of TRPV6 [42]). Additionally, S502 is a key phosphorylation site for TRPV1 regulation [92] .

In the S4-S5 linker, V567 corresponds to V577 bound with lipid in a TRPV3 structure (PDB id: 6lgp) [39], and V525 bound with RTX in a TRPV2 structure (PDB id: 6bwj) [47]; K571 is at a site bound with lipid in three structures of TRPV1 (PDB id: 5irz) [27], TRPV5 (PDB id: 6dmu) [34] and TRPV6 (PDB id: 6bo8) [36]. Additionally, the K571E mutation was found to disrupt specific binding for 2-APB [92].

In the channel pore are A680 and M682 which are adjacent to the lower gate at I679, and may serve as a binding site for channel blockers. From functional studies, the M682A mutation caused impaired capsaicin and heat activation [96] , the E684G/V mutations caused gain of function and strong toxicity when expressed in yeast [97] .

In the TRP helix, K694 was implicated in PIP2 binding in several computational and functional studies [88-91].

In sum, by predicting pocket residues that differ between the C and O state ensembles, we identified possible ligand-binding sites potentially involved in the regulation of the C-to-O equilibrium of TRPV1. The predicted pocket residues are located in key functional domains including the MPD linker, the S2-S3 linker, the S4-S5 linker, the pore, the TRP helix, and the extracellular site of S1-S4 domain (see Fig 5e and 6e). These predictions will inform future efforts for drug design by targeting these sites for positive/negative regulation of TRPV1 gating.

Conclusion

In conclusion, we have performed, to our knowledge, the most extensive MD simulations (with cumulative time of 20 μs) of TRPV1 in both the closed and open states. Detailed comparison between the closed and open state ensembles confirmed our previous findings of distinct structure and dynamics between the two states based on shorter simulations [62,65-67]. Furthermore, by comparing ligand binding accessibility between the two ensembles, we identified state-dependent binding sites for small-molecule ligands in general and lipids in particular. We further used machine learning to predict top ligand-binding sites as important features to classify the closed vs open states. The predicted binding sites are thoroughly validated by matching corresponding sites in all structures of TRPV channels bound to lipids and other ligands [28-47], and with previous functional/mutational studies of ligand binding in TRPV1 [92]. Taken together, this comprehensive study has integrated structural, dynamic, and functional data to inform future design of small-molecular drugs targeting TRPV1.

Figure Legends

Figure 1. Structural architecture of TRPV1:

(a) Side view of TRPV1 structure: a representative subunit is colored by domains as follows — the ARD (red), the MPD (green), the MPD linker (light green), the S1-S4 helices (cyan), the S2-S3 linker (light cyan), the S4-S5 linker (light purple), the pore domain (purple, including the outer pore, the S5 helix, and the S6 helix), the TRP helix (blue). The remaining subunits are colored in silver. Residues G643 and I679 at the upper and the lower gate are shown as spheres colored in orange and yellow, respectively. The expanded TMD is boxed. (b) Top view of TRPV1 structure with the same color coding as (a).

Figure 2. RMSF analysis of TRPV1 flexibility: (a) RMSF of the expanded TMD of TRPV1 in the C state (blue) and the O state (red); (b) Two diagonally opposite subunits of TRPV1 colored by RMSF of the O state (blue/red for low/high RMSF). In (a), the following key regions are marked by horizontal bars and colored as follows: the MPD linker (green), the S1-S2 linker (yellow), the S2-S3 linker (cyan), the S4-S5 linker (magenta), the outer pore (red), and the TRP helix (blue). RMSF is averaged over four symmetrically related residue positions of the TRPV1 tetramer.

Figure 3. State-dependent conformational changes in the expanded TMD of TRPV1: (a) Average structure of the C-state ensemble (blue) superimposed with the initial closed structure from cryo-EM (cyan); (b) Average structure of the O-state ensemble (red) superimposed with the initial open structure from cryo-EM (yellow); (c) Average structure of the C-state ensemble (blue) superimposed with the average structure of the O-state ensemble (red). Major domain motions are marked by arrows. For clarity, only two diagonally opposite subunits of TRPV1 are shown.

Figure 4. Lipid-binding residues predicted by analysis of lipid-binding fraction f_n: (a) Average f_n as a function of the residue position in the C state; (b) Average f_n as a function of the residue position in the O state; (c) C-to-O change in average f_n as a function of the residue position; (d) feature importance of ML (random forest); (e) Top 10 lipid-binding residues shown as spheres on the average structures of C state (blue) and O state (red). The black dots in (a) and (b) mark the positions of known lipid-binding sites in all TRPV structures (see Table S1).

Figure 5. Pocket residues predicted by fpocket: (a) Average pocket score s_n as a function of the residue position in the C state; (b) Average s_n as a function of the residue position in the O state; (c) C-to-O change in average s_n as a function of the residue position; (d) feature importance of ML (random forest); (e) Top 10 pocket residues

shown as spheres on the average structures of C state (blue) and O state (red). The black dots in (a) and (b) mark the positions of known ligand-binding residues in all TRPV structures (see Table S2).

Figure 6. Pocket residues predicted by concavity: (a) Average pocket score s_n as a function of the residue position in the C state; (b) Average s_n as a function of the residue position in the O state; (c) C-to-O change in average s_n as a function of the residue position; (d) feature importance of ML (random forest); (e) Top 10 pocket residues shown as spheres on the average structures of C state (blue) and O state (red). The black dots in (a) and (b) mark the positions of known ligand-binding residues in all TRPV structures (see Table S2).

Acknowledgement

We thank funding support from American Heart Association (17GRNT33690009). Computational support was provided by the Center for Computational Research at the University at Buffalo.

References

1. Caterina MJ, Rosen TA, Tominaga M, Brake AJ, Julius D. A capsaicin-receptor homologue with a high threshold for noxious heat. Nature 1999;398(6726):436-441.
2. Caterina MJ, Schumacher MA, Tominaga M, Rosen TA, Levine JD, Julius D. The capsaicin receptor: a heat-activated ion channel in the pain pathway. Nature 1997;389(6653):816-824.
3. Story GM, Peier AM, Reeve AJ, Eid SR, Mosbacher J, Hricik TR, Earley TJ, Hergarden AC, Andersson DA, Hwang SW, McIntyre P, Jegla T, Bevan S, Patapoutian A. ANKTM1, a TRP-like channel expressed in nociceptive neurons, is activated by cold temperatures. Cell 2003;112(6):819-829.
4. Nilius B, Appendino G, Owsianik G. The transient receptor potential channel TRPA1: from gene to pathophysiology. Pflugers Arch 2012;464(5):425-458.
5. Karashima Y, Talavera K, Everaerts W, Janssens A, Kwan KY, Vennekens R, Nilius B, Voets T. TRPA1 acts as a cold sensor in vitro and in vivo. Proc Natl Acad Sci U S A 2009;106(4):1273-1278.
6. Voets T, Droogmans G, Wissenbach U, Janssens A, Flockerzi V, Nilius B. The principle of temperature-dependent gating in cold- and heat-sensitive TRP channels. Nature 2004;430(7001):748-754.

7. Jordt SE, Tominaga M, Julius D. Acid potentiation of the capsaicin receptor determined by a key extracellular site. Proc Natl Acad Sci U S A 2000;97(14):8134-8139.
8. Tominaga M, Caterina MJ, Malmberg AB, Rosen TA, Gilbert H, Skinner K, Raumann BE, Basbaum AI, Julius D. The cloned capsaicin receptor integrates multiple pain-producing stimuli. Neuron 1998;21(3):531-543.
9. Howard J, Bechstedt S. Hypothesis: a helix of ankyrin repeats of the NOMPC-TRP ion channel is the gating spring of mechanoreceptors. Curr Biol 2004;14(6):R224-226.
10. Sotomayor M, Corey DP, Schulten K. In search of the hair-cell gating spring elastic properties of ankyrin and cadherin repeats. Structure 2005;13(4):669-682.
11. Yin J, Kuebler WM. Mechanotransduction by TRP channels: general concepts and specific role in the vasculature. Cell Biochem Biophys 2010;56(1):1-18.
12. Gunthorpe MJ, Szallasi A. Peripheral TRPV1 receptors as targets for drug development: new molecules and mechanisms. Curr Pharm Des 2008;14(1):32-41.
13. Nilius B. Transient receptor potential TRP channels as therapeutic drug targets: next round! Curr Top Med Chem 2013;13(3):244-246.
14. Nilius B, Voets T, Peters J. TRP channels in disease. Sci STKE 2005;2005(295):re8.
15. Nilius B. TRP channels in disease. Biochim Biophys Acta 2007;1772(8):805-812.
16. Catterall WA. Ion channel voltage sensors: structure, function, and pathophysiology. Neuron 2010;67(6):915-928.
17. Catterall WA. Voltage-gated sodium channels at 60: structure, function and pathophysiology. J Physiol 2012;590(Pt 11):2577-2589.
18. Gaudet R. A primer on ankyrin repeat function in TRP channels and beyond. Molecular bioSystems 2008;4(5):372-379.
19. Montell C. Physiology, phylogeny, and functions of the TRP superfamily of cation channels. Sci STKE 2001;2001(90):re1.
20. Garcia-Sanz N, Valente P, Gomis A, Fernandez-Carvajal A, Fernandez-Ballester G, Viana F, Belmonte C, Ferrer-Montiel A. A role of the transient receptor potential domain of vanilloid receptor I in channel gating. The Journal of neuroscience : the official journal of the Society for Neuroscience 2007;27(43):11641-11650.
21. Numazaki M, Tominaga T, Takeuchi K, Murayama N, Toyooka H, Tominaga M. Structural determinant of TRPV1 desensitization interacts with calmodulin. Proc Natl Acad Sci U S A 2003;100(13):8002-8006.
22. Prescott ED, Julius D. A modular PIP2 binding site as a determinant of capsaicin receptor sensitivity. Science 2003;300(5623):1284-1288.
23. Yao J, Liu B, Qin F. Modular thermal sensors in temperature-gated transient receptor potential (TRP) channels. Proc Natl Acad Sci U S A 2011;108(27):11109-11114.
24. Cao E, Liao M, Cheng Y, Julius D. TRPV1 structures in distinct conformations reveal activation mechanisms. Nature 2013;504(7478):113-118.
25. Bagneris C, Naylor CE, McCusker EC, Wallace BA. Structural model of the open-closed-inactivated cycle of prokaryotic voltage-gated sodium channels. The Journal of general physiology 2015;145(1):5-16.
26. Liao M, Cao E, Julius D, Cheng Y. Structure of the TRPV1 ion channel determined by electron cryo-microscopy. Nature 2013;504(7478):107-112.
27. Gao Y, Cao E, Julius D, Cheng Y. TRPV1 structures in nanodiscs reveal mechanisms of ligand and lipid action. Nature 2016;534(7607):347-351.
28. Dang S, van Goor MK, Asarnow D, Wang Y, Julius D, Cheng Y, van der Wijst J. Structural insight into TRPV5 channel function and modulation. Proceedings of the National Academy of Sciences of the United States of America 2019;116(18):8869-8878.
29. Deng Z, Maksaev G, Rau M, Xie Z, Hu H, Fitzpatrick JAJ, Yuan P. Gating of human TRPV3 in a lipid bilayer. Nature structural & molecular biology 2020;27(7):635-644.

30. Deng Z, Paknejad N, Maksaev G, Sala-Rabanal M, Nichols CG, Hite RK, Yuan P. Cryo-EM and X-ray structures of TRPV4 reveal insight into ion permeation and gating mechanisms. Nature structural & molecular biology 2018;25(3):252-260.
31. Dosey TL, Wang Z, Fan G, Zhang Z, Serysheva, II, Chiu W, Wensel TG. Structures of TRPV2 in distinct conformations provide insight into role of the pore turret. Nature structural & molecular biology 2019;26(1):40-49.
32. Hughes TE, Del Rosario JS, Kapoor A, Yazici AT, Yudin Y, Fluck EC, 3rd, Filizola M, Rohacs T, Moiseenkova-Bell VY. Structure-based characterization of novel TRPV5 inhibitors. eLife 2019;8.
33. Hughes TET, Lodowski DT, Huynh KW, Yazici A, Del Rosario J, Kapoor A, Basak S, Samanta A, Han X, Chakrapani S, Zhou ZH, Filizola M, Rohacs T, Han S, Moiseenkova-Bell VY. Structural basis of TRPV5 channel inhibition by econazole revealed by cryo-EM. Nature structural & molecular biology 2018;25(1):53-60.
34. Hughes TET, Pumroy RA, Yazici AT, Kasimova MA, Fluck EC, Huynh KW, Samanta A, Molugu SK, Zhou ZH, Carnevale V, Rohacs T, Moiseenkova-Bell VY. Structural insights on TRPV5 gating by endogenous modulators. Nature communications 2018;9(1):4198.
35. Huynh KW, Cohen MR, Jiang J, Samanta A, Lodowski DT, Zhou ZH, Moiseenkova-Bell VY. Structure of the full-length TRPV2 channel by cryo-EM. Nature communications 2016;7:11130.
36. McGoldrick LL, Singh AK, Saotome K, Yelshanskaya MV, Twomey EC, Grassucci RA, Sobolevsky AI. Opening of the human epithelial calcium channel TRPV6. Nature 2018;553(7687):233-237.
37. Pumroy RA, Samanta A, Liu Y, Hughes TE, Zhao S, Yudin Y, Rohacs T, Han S, Moiseenkova-Bell VY. Molecular mechanism of TRPV2 channel modulation by cannabidiol. eLife 2019;8.
38. Saotome K, Singh AK, Sobolevsky AI. Determining the Crystal Structure of TRPV6. In: Kozak JA, Putney JW, Jr., editors. Calcium Entry Channels in Non-Excitable Cells. Boca Raton (FL); 2018. p 275-292.
39. Shimada H, Kusakizako T, Dung Nguyen TH, Nishizawa T, Hino T, Tominaga M, Nureki O. The structure of lipid nanodisc-reconstituted TRPV3 reveals the gating mechanism. Nature structural & molecular biology 2020;27(7):645-652.
40. Singh AK, McGoldrick LL, Demirkhanyan L, Leslie M, Zakharian E, Sobolevsky AI. Structural basis of temperature sensation by the TRP channel TRPV3. Nature structural & molecular biology 2019;26(11):994-998.
41. Singh AK, McGoldrick LL, Sobolevsky AI. Structure and gating mechanism of the transient receptor potential channel TRPV3. Nature structural & molecular biology 2018;25(9):805-813.
42. Singh AK, Saotome K, McGoldrick LL, Sobolevsky AI. Structural bases of TRP channel TRPV6 allosteric modulation by 2-APB. Nature communications 2018;9(1):2465.
43. Zubcevic L, Borschel WF, Hsu AL, Borgnia MJ, Lee SY. Regulatory switch at the cytoplasmic interface controls TRPV channel gating. eLife 2019;8.
44. Zubcevic L, Herzik MA, Jr., Chung BC, Liu Z, Lander GC, Lee SY. Cryo-electron microscopy structure of the TRPV2 ion channel. Nature structural & molecular biology 2016;23(2):180-186.
45. Zubcevic L, Herzik MA, Jr., Wu M, Borschel WF, Hirschi M, Song AS, Lander GC, Lee SY. Conformational ensemble of the human TRPV3 ion channel. Nature communications 2018;9(1):4773.
46. Zubcevic L, Hsu AL, Borgnia MJ, Lee SY. Symmetry transitions during gating of the TRPV2 ion channel in lipid membranes. eLife 2019;8.
47. Zubcevic L, Le S, Yang H, Lee SY. Conformational plasticity in the selectivity filter of the TRPV2 ion channel. Nature structural & molecular biology 2018;25(5):405-415.
48. Karplus M, McCammon JA. Molecular dynamics simulations of biomolecules. Nat Struct Biol 2002;9(9):646-652.
49. Nury H, Poitevin F, Van Renterghem C, Changeux JP, Corringer PJ, Delarue M, Baaden M. One-microsecond molecular dynamics simulation of channel gating in a nicotinic receptor homologue. Proc Natl Acad Sci U S A 2010;107(14):6275-6280.

50. Nury H, Van Renterghem C, Weng Y, Tran A, Baaden M, Dufresne V, Changeux JP, Sonner JM, Delarue M, Corringer PJ. X-ray structures of general anaesthetics bound to a pentameric ligand-gated ion channel. Nature 2011;469(7330):428-431.
51. Zhu F, Hummer G. Pore opening and closing of a pentameric ligand-gated ion channel. Proc Natl Acad Sci U S A 2010;107(46):19814-19819.
52. Jensen MO, Jogini V, Borhani DW, Leffler AE, Dror RO, Shaw DE. Mechanism of voltage gating in potassium channels. Science 2012;336(6078):229-233.
53. Zheng W, Wen H. Investigating dual Ca(2+) modulation of the ryanodine receptor 1 by molecular dynamics simulation. Proteins 2020.
54. Zheng W, Liu Z. Investigating the inter-subunit/subdomain interactions and motions relevant to disease mutations in the N-terminal domain of ryanodine receptors by molecular dynamics simulation. Proteins 2017;85(9):1633-1644.
55. Stone JE, Hardy DJ, Ufimtsev IS, Schulten K. GPU-accelerated molecular modeling coming of age. Journal of Molecular Graphics & Modelling 2010;29(2):116-125.
56. Atilgan AR, Durell SR, Jernigan RL, Demirel MC, Keskin O, Bahar I. Anisotropy of fluctuation dynamics of proteins with an elastic network model. Biophysical journal 2001;80(1):505-515.
57. Tama F, Sanejouand YH. Conformational change of proteins arising from normal mode calculations. Protein Eng 2001;14(1):1-6.
58. Zheng W, Doniach S. A comparative study of motor-protein motions by using a simple elastic-network model. Proc Natl Acad Sci U S A 2003;100(23):13253-13258.
59. Tozzini V. Coarse-grained models for proteins. Curr Opin Struct Biol 2005;15(2):144-150.
60. Tozzini V. Minimalist models for proteins: a comparative analysis. Q Rev Biophys 2010;43(3):333-371.
61. Zheng W, Wen H, Iacobucci GJ, Popescu GK. Probing the Structural Dynamics of the NMDA Receptor Activation by Coarse-Grained Modeling. Biophysical journal 2017;112(12):2589-2601.
62. Zheng W, Qin F. A combined coarse-grained and all-atom simulation of TRPV1 channel gating and heat activation. The Journal of general physiology 2015;145(5):443-456.
63. Zheng W. Toward decrypting the allosteric mechanism of the ryanodine receptor based on coarse-grained structural and dynamic modeling. Proteins 2015;83(12):2307-2318.
64. Zheng W, Sachs F. Investigating the structural dynamics of the PIEZO1 channel activation and inactivation by coarse-grained modeling. Proteins 2017;85(12):2198-2208.
65. Zheng W, Wen H. Heat activation mechanism of TRPV1: New insights from molecular dynamics simulation. Temperature 2019;6(2):120-131.
66. Wen H, Zheng W. Decrypting the Heat Activation Mechanism of TRPV1 Channel by Molecular Dynamics Simulation. Biophysical journal 2018;114(1):40-52.
67. Wen H, Qin F, Zheng W. Toward elucidating the heat activation mechanism of the TRPV1 channel gating by molecular dynamics simulation. Proteins 2016;84(12):1938-1949.
68. Lomize MA, Lomize AL, Pogozheva ID, Mosberg HI. OPM: orientations of proteins in membranes database. Bioinformatics 2006;22(5):623-625.
69. Jo S, Kim T, Im W. Automated Builder and Database of Protein/Membrane Complexes for Molecular Dynamics Simulations. PloS one 2007;2(9).
70. Jo S, Kim T, Iyer VG, Im W. CHARMM-GUI: a web-based graphical user interface for CHARMM. Journal of computational chemistry 2008;29(11):1859-1865.
71. Martyna GJ, Hughes A, Tuckerman ME. Molecular dynamics algorithms for path integrals at constant pressure. J Chem Phys 1999;110(7):3275-3290.
72. Parrinello M, and A. Rahman. Polymorphic transitions in single crystals: a new molecular dynamics method. J Appl Phys 1981;52:7182– 7190.

73. Darden T, York D, Pedersen L. Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems. The Journal of Chemical Physics 1993;98(12):10089-10092.
74. Hess B, H. Bekker, ., J. G. E. M. Fraaije. LINCS: a linear constraint solver for molecular simulations. . J Comput Chem 1997;18: 1463–1472.
75. Pronk S, Pall S, Schulz R, Larsson P, Bjelkmar P, Apostolov R, Shirts MR, Smith JC, Kasson PM, van der Spoel D, Hess B, Lindahl E. GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. Bioinformatics 2013;29(7):845-854.
76. Huang J, MacKerell AD, Jr. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. Journal of computational chemistry 2013;34(25):2135-2145.
77. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. The Journal of Chemical Physics 1983;79(2):926-935.
78. Rohacs T. Phosphoinositide regulation of TRPV1 revisited. Pflugers Archiv : European journal of physiology 2015;467(9):1851-1869.
79. Capra JA, Laskowski RA, Thornton JM, Singh M, Funkhouser TA. Predicting protein ligand binding sites by combining evolutionary sequence conservation and 3D structure. PLoS computational biology 2009;5(12):e1000585.
80. Le Guilloux V, Schmidtke P, Tuffery P. Fpocket: an open source platform for ligand pocket detection. BMC bioinformatics 2009;10:168.
81. Cimermancic P, Weinkam P, Rettenmaier TJ, Bichmann L, Keedy DA, Woldeyes RA, Schneidman-Duhovny D, Demerdash ON, Mitchell JC, Wells JA, Fraser JS, Sali A. CryptoSite: Expanding the Druggable Proteome by Characterization and Prediction of Cryptic Binding Sites. Journal of molecular biology 2016;428(4):709-719.
82. Tin Kam H. The random subspace method for constructing decision forests. IEEE Transactions on Pattern Analysis and Machine Intelligence 1998;20(8):832-844.
83. Yao J, Liu B, Qin F. Kinetic and energetic analysis of thermally activated TRPV1 channels. Biophysical journal 2010;99(6):1743-1753.
84. Picazo-Juarez G, Romero-Suarez S, Nieto-Posadas A, Llorente I, Jara-Oseguera A, Briggs M, McIntosh TJ, Simon SA, Ladron-de-Guevara E, Islas LD, Rosenbaum T. Identification of a binding motif in the S5 helix that confers cholesterol sensitivity to the TRPV1 ion channel. The Journal of biological chemistry 2011;286(28):24966-24976.
85. Senning EN, Collins MD, Stratiievska A, Ufret-Vincenty CA, Gordon SE. Regulation of TRPV1 ion channel by phosphoinositide (4,5)-bisphosphate: the role of membrane asymmetry. The Journal of biological chemistry 2014;289(16):10999-11006.
86. Klein RM, Ufret-Vincenty CA, Hua L, Gordon SE. Determinants of molecular specificity in phosphoinositide regulation. Phosphatidylinositol (4,5)-bisphosphate (PI(4,5)P2) is the endogenous lipid regulating TRPV1. The Journal of biological chemistry 2008;283(38):26208-26216.
87. Muller MP, Jiang T, Sun C, Lihan M, Pant S, Mahinthichaichan P, Trifan A, Tajkhorshid E. Characterization of Lipid-Protein Interactions and Lipid-Mediated Modulation of Membrane Protein Function through Molecular Simulation. Chemical reviews 2019;119(9):6086-6161.
88. Brauchi S, Orta G, Mascayano C, Salazar M, Raddatz N, Urbina H, Rosenmann E, Gonzalez-Nilo F, Latorre R. Dissection of the components for PIP2 activation and thermosensation in TRP channels. Proceedings of the National Academy of Sciences of the United States of America 2007;104(24):10246-10251.
89. Grycova L, Holendova B, Bumba L, Bily J, Jirku M, Lansky Z, Teisinger J. Integrative binding sites within intracellular termini of TRPV1 receptor. PloS one 2012;7(10):e48437.
90. Poblete H, Oyarzun I, Olivero P, Comer J, Zuniga M, Sepulveda RV, Baez-Nieto D, Gonzalez Leon C, Gonzalez-Nilo F, Latorre R. Molecular determinants of phosphatidylinositol 4,5-bisphosphate (PI(4,5)P2) binding to transient receptor potential V1 (TRPV1) channels. The Journal of biological chemistry 2015;290(4):2086-2098.

91. Ufret-Vincenty CA, Klein RM, Collins MD, Rosasco MG, Martinez GQ, Gordon SE. Mechanism for phosphoinositide selectivity and activation of TRPV1 ion channels. The Journal of general physiology 2015;145(5):431-442.
92. Winter Z, Buhala A, Otvos F, Josvay K, Vizler C, Dombi G, Szakonyi G, Olah Z. Functionally important amino acid residues in the transient receptor potential vanilloid 1 (TRPV1) ion channel--an overview of the current mutational data. Molecular pain 2013;9:30.
93. Vajda S, Beglov D, Wakefield AE, Egbert M, Whitty A. Cryptic binding sites on proteins: definition, detection, and druggability. Current opinion in chemical biology 2018;44:1-8.
94. Jordt SE, Julius D. Molecular basis for species-specific sensitivity to "hot" chili peppers. Cell 2002;108(3):421-430.
95. Ryu S, Liu B, Yao J, Fu Q, Qin F. Uncoupling proton activation of vanilloid receptor TRPV1. The Journal of neuroscience : the official journal of the Society for Neuroscience 2007;27(47):12797-12807.
96. Susankova K, Ettrich R, Vyklicky L, Teisinger J, Vlachova V. Contribution of the putative inner-pore region to the gating of the transient receptor potential vanilloid subtype 1 channel (TRPV1). The Journal of neuroscience : the official journal of the Society for Neuroscience 2007;27(28):7578-7585.
97. Myers BR, Bohlen CJ, Julius D. A yeast genetic screen reveals a critical role for the pore helix domain in TRP channel gating. Neuron 2008;58(3):362-373.

Figures

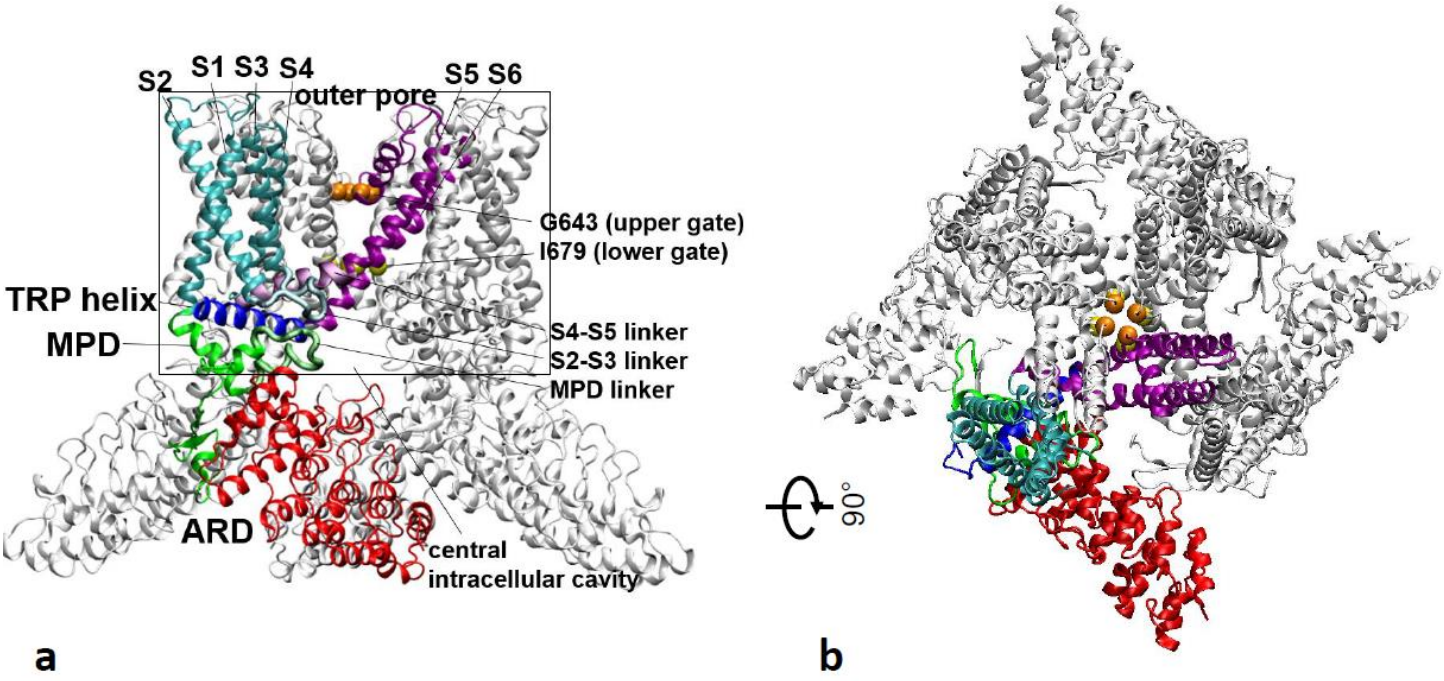


Figure 1

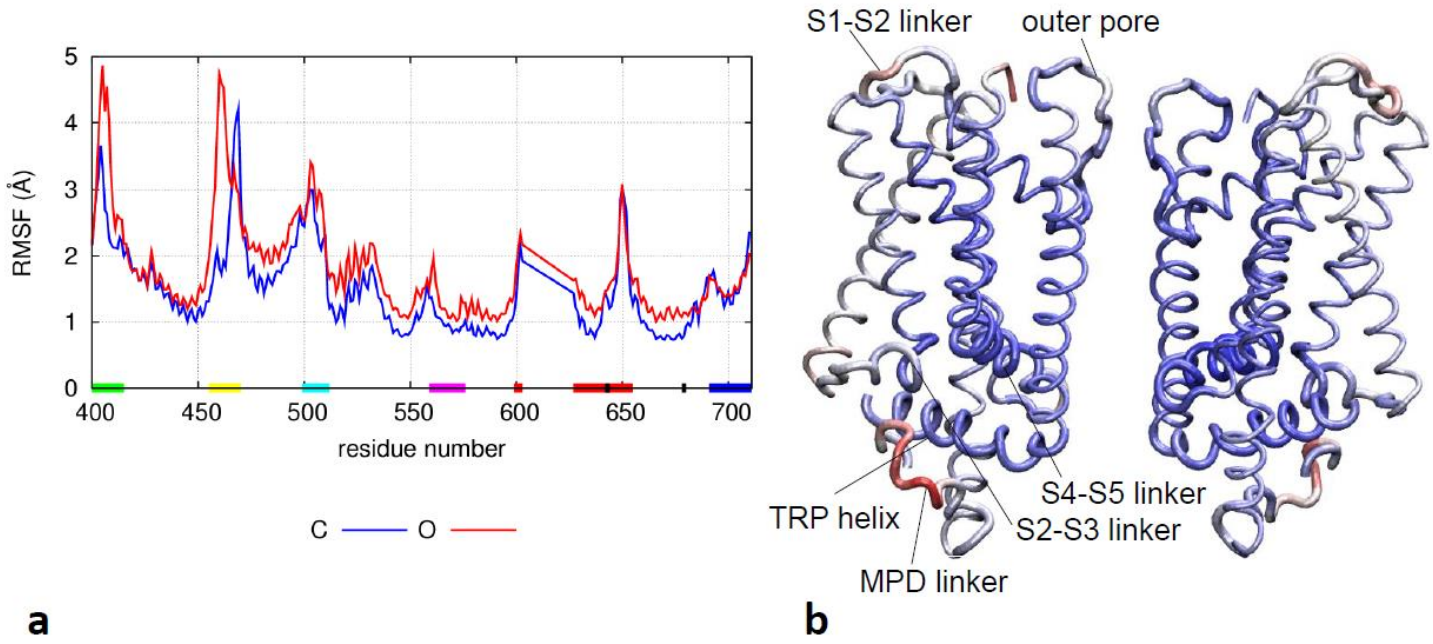

RMSF (Å)
residue number
C
O
a
S1-S2 linker
outer pore
TRP helix
S4-S5 linker
S2-S3 linker
MPD linker
b

Figure 2

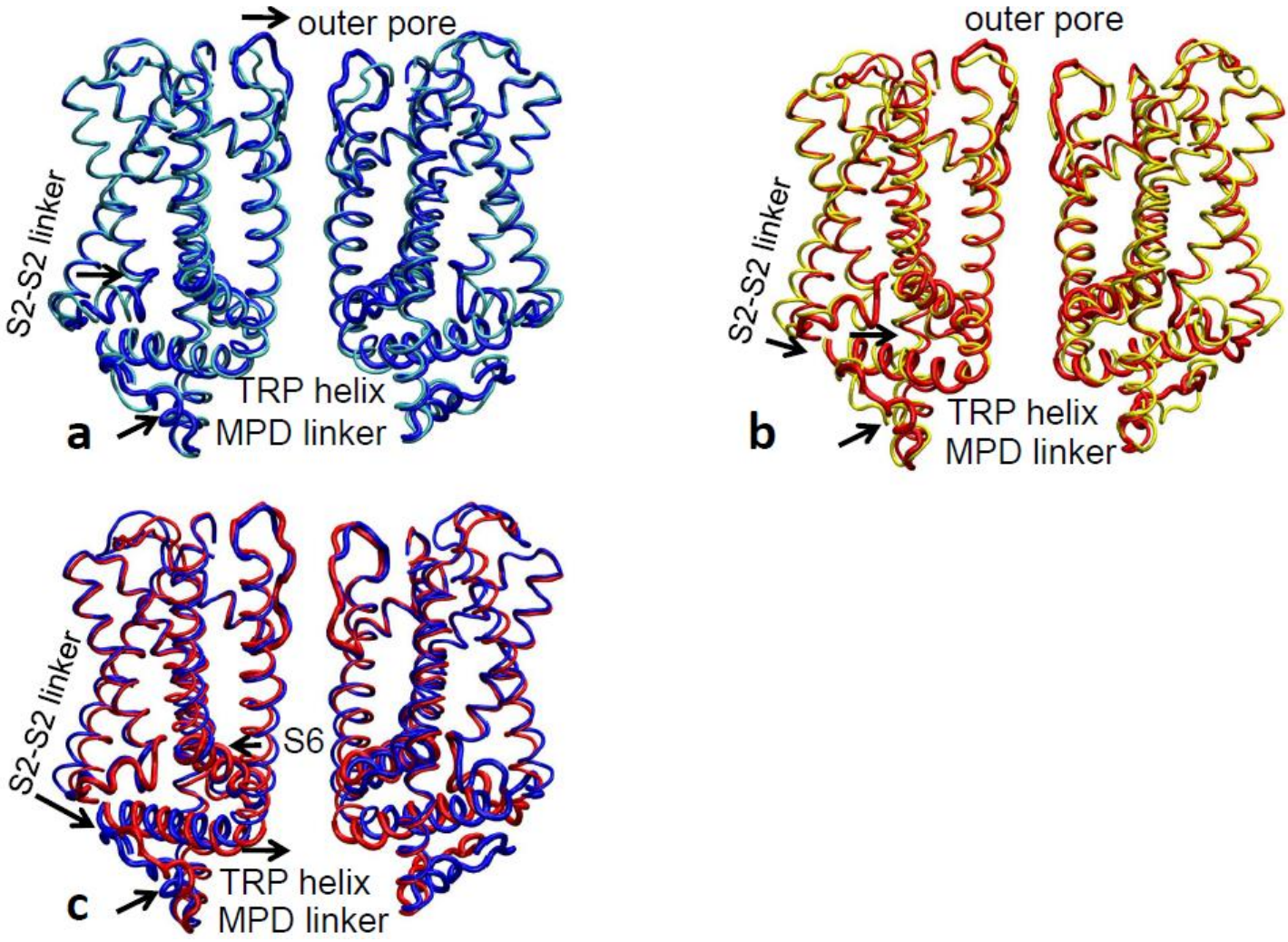

outer pore
S2-S2 linker
TRP helix
MPD linker
a
outer pore
S2-S2 linker
TRP helix
MPD linker
b
S2-S2 linker
S6
TRP helix
MPD linker
c

Figure 3

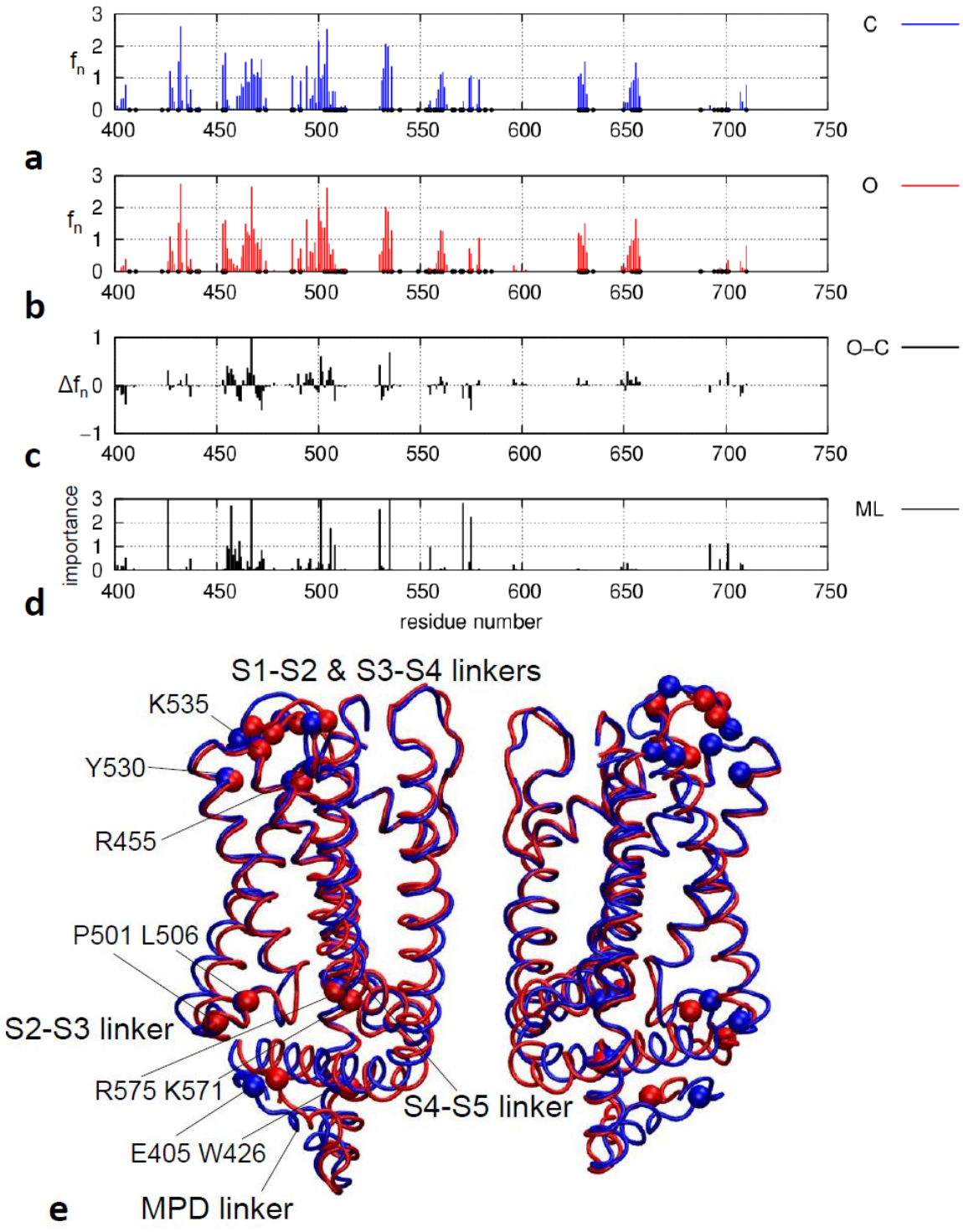

a
b
c
d
e
C
O
O–C
ML
f_n
Δf_n
importance
residue number
S1-S2 & S3-S4 linkers
K535
Y530
R455
P501 L506
S2-S3 linker
R575 K571
E405 W426
MPD linker
S4-S5 linker

Figure 4

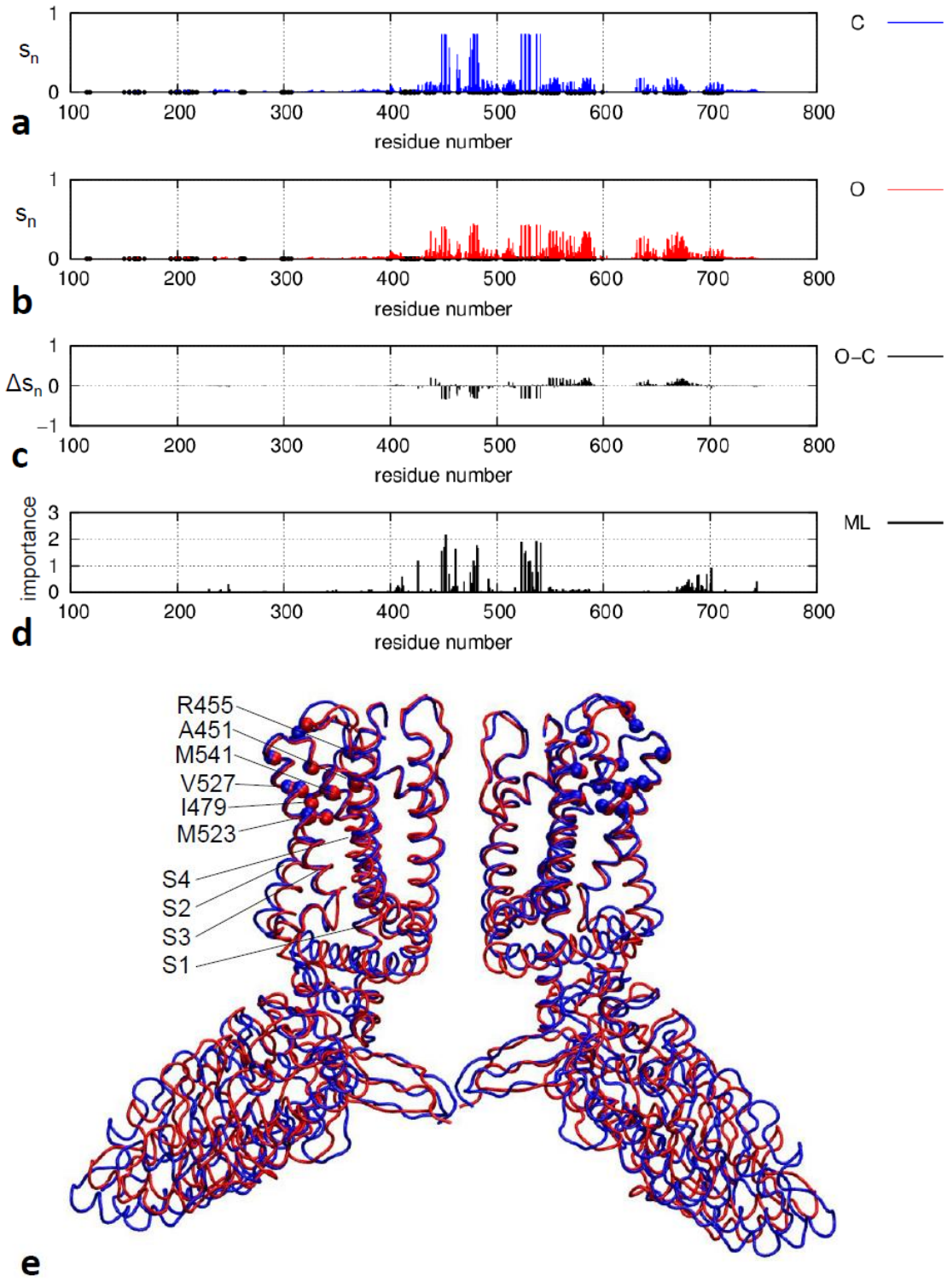
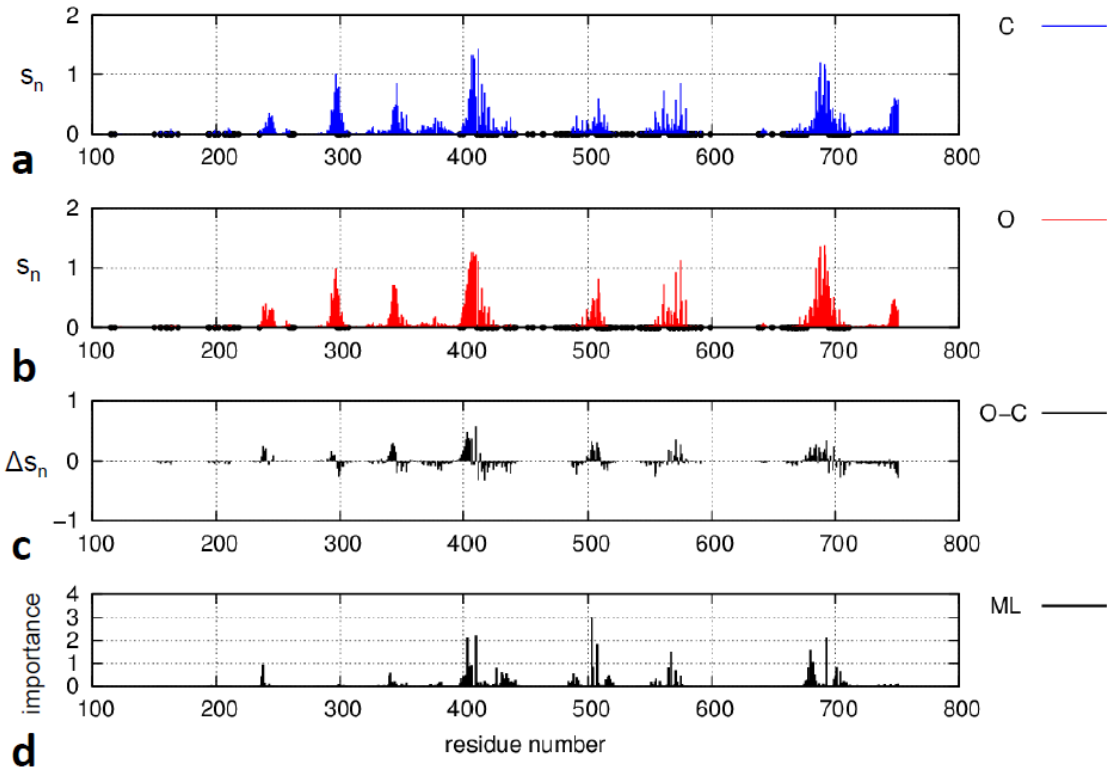

a
b
c
d
e
C
O
O–C
ML
s_n
Δs_n
importance
residue number
R455
A451
M541
V527
I479
M523
S4
S2
S3
S1

Figure 5


a
b
c
d
C
O
O–C
ML
s_n
Δs_n
importance
residue number

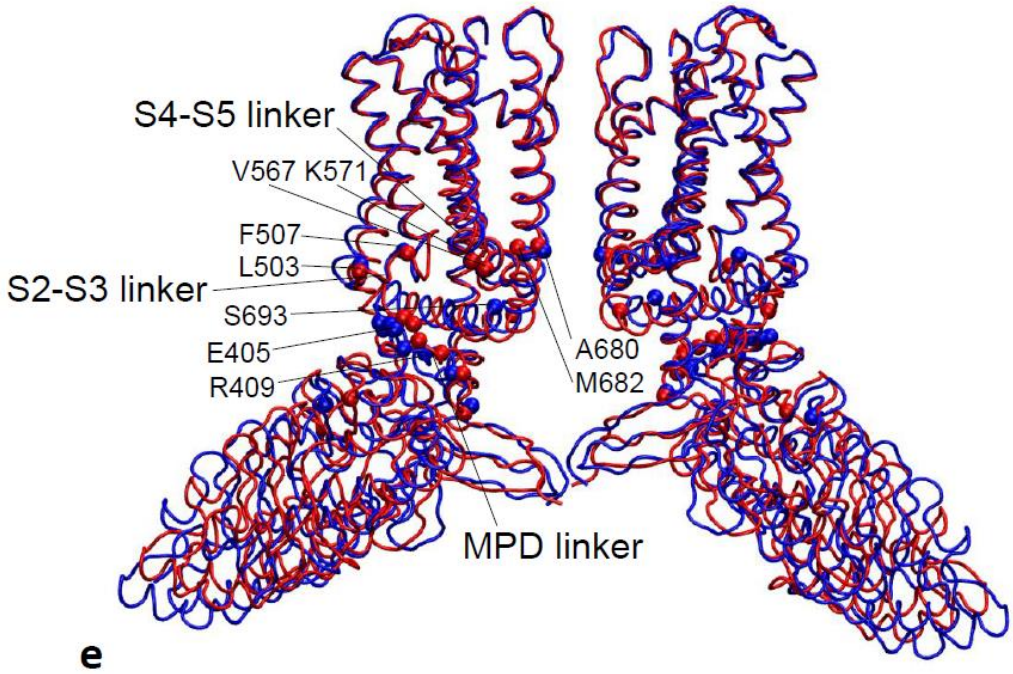

S4-S5 linker
V567 K571
F507
L503
S2-S3 linker
S693
E405
R409
A680
M682
MPD linker
e

Figure 6