Journal of Medicinal Chemistry

Subscriber access provided by TULANE UNIVERSITY

Article

Structure-based discovery of a subtype-selective inhibitor targeting a transient receptor potential vanilloid channel

Hao Chai, Xi Cheng, Bin Zhou, Lifen Zhao, Xianhua Lin, Dongping Huang, Weiqiang Lu, Hao Lv, Feng Tang, Qiansen Zhang, Wei Huang, Yang Li, and Huaiyu Yang

J. Med. Chem., **Just Accepted Manuscript** • DOI: 10.1021/acs.jmedchem.8b01496 • Publication Date (Web): 08 Jan 2019

Downloaded from http://pubs.acs.org on January 9, 2019

Just Accepted

"Just Accepted" manuscripts have been peer-reviewed and accepted for publication. They are posted online prior to technical editing, formatting for publication and author proofing. The American Chemical Society provides "Just Accepted" as a service to the research community to expedite the dissemination of scientific material as soon as possible after acceptance. "Just Accepted" manuscripts appear in full in PDF format accompanied by an HTML abstract. "Just Accepted" manuscripts have been fully peer reviewed, but should not be considered the official version of record. They are citable by the Digital Object Identifier (DOI®). "Just Accepted" is an optional service offered to authors. Therefore, the "Just Accepted" Web site may not include all articles that will be published in the journal. After a manuscript is technically edited and formatted, it will be removed from the "Just Accepted" Web site and published as an ASAP article. Note that technical editing may introduce minor changes to the manuscript text and/or graphics which could affect content, and all legal disclaimers and ethical guidelines that apply to the journal pertain. ACS cannot be held responsible for errors or consequences arising from the use of information contained in these "Just Accepted" manuscripts.

ACS Publications

is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

Published by American Chemical Society. Copyright © American Chemical Society. However, no copyright claim is made to original U.S. Government works, or works produced by employees of any Commonwealth realm Crown government in the course of their duties.

Structure-based discovery of a subtype-selective inhibitor targeting a transient receptor potential vanilloid channel

Hao Chai,[†,§,‡] Xi Cheng,[†,‡] Bin Zhou,[†,§,‡] Lifen Zhao,[†] Xianhua Lin,[#] Dongping Huang,[#] Weiqiang Lu,[#] Hao Lv,[†,§] Feng Tang, [†] Qiansen Zhang,[#] Wei Huang,[†,§,] Yang Li,[†,§,*] & Huaiyu Yang[#,*]*

† State Key Laboratory of Drug Research and CAS Key Laboratory of Receptor Research, Shanghai Institute of Materia Medica, Chinese Academy of Sciences, 555 Zuchongzhi Road, Shanghai 201203, China.

Shanghai Key Laboratory of Regulatory Biology, Institute of Biomedical Sciences and School of Life Sciences, East China Normal University, 500 Dongchuan Road, Shanghai 200241, China.

§ University of Chinese Academy of Sciences, No. 19A Yuquan Road, Beijing 100049, China

‡ These authors contributed equally to this work.

* To whom correspondence should be addressed.

KEYWORDS: cancer, drug design, G-protein coupled receptor, transient receptor potential channel, lysophosphatidic acid

ABSTRACT: Discovery of potent selective inhibitors targeting a protein from a highly conserved family is challenging. Using a strategy combining structural and evolutionary information, we discovered TRP subtype-selective inhibitors (TRPV2 inhibitors). We unveiled three ligand-binding sites of TRPV2 and compounds that bind to these sites. Structural optimization of the best-hit compound provided a potent selective TRPV2 inhibitor, SET2. The molecular basis and subtype-selective inhibition mechanism were quantitatively characterized and experimentally verified. Then, as an effective chemical probe, SET2 was used to investigate the function role of TRPV2. SET2-induced inhibition of TRPV2 reduced prostate cancer migration, which indicated TRPV2 as an anti-metastasis therapeutic target. And functional assays suggested that TRPV2 was coupled to a validated metastasis mediator, LPAR1. The discovery of the potent selective inhibitor potentially leads to novel avenues for pharmacological applications and therapeutic development targeting TRPV2 channel.

INTRODUCTION

Subtype-selective compounds are of great therapeutic importance. However, given high similarity among protein members of a conserved family, development of subtype-selective compounds represents a very challenging task. As the most widely used approach in drug development, experimental high-throughput screening only has a hit rate less than 0.02% in discovering subtype-selective compounds.[1] On the other hand, the rational design approach can significantly improve the efficiency of drug development and reduces the time and cost by reducing the number of compounds that need be synthesized.[2]

Transient receptor potential (TRP) ion channels play crucial roles in nociception, taste perception, thermosensation, and mechano- and osmolarity sensing.[3] However, due to high conservation among TRP channels, it is difficult to develop selective compounds targeting a particular member of this family, which leads to a major disadvantage in the therapeutic development of TRP channels. The transient receptor potential vanilloid type 2 (TRPV2) channel is a TRP channel participating in metastasis of various cancers.[4-7] Yet, TRPV2 has no potent selective modulator. Until now, a diphenyl compound, i.e. tranilast, has been used in several reports as a TRPV2-specific antagonist. Tranilast has a half-maximal inhibition concentration (IC_{50}) of more than 10 μM and has not been fully validated as a direct TRPV2 blocker.[8] Ruthenium red can potently inhibit TRPV2 with an IC_{50} of approximately 0.6 μM, but it is a non-selective blocker of several TRP channels. For a long time, the development of TRPV2 selective compounds has been unsuccessful due to the high structural homology among transient receptor potential vanilloid (TRPV) subfamily members, especially TRPV1-TRPV4. All of these channels are homotetramers

with ion pathway located at the four-fold symmetry.[9-17] Surrounded by transmembrane helices S1-S4, the ion pathway of a TRPV channel is formed by the transmembrane helices S5-S6, pore helices and the intervening loops in the assemble tetramer.

In this study, using a protocol based on structural and evolutionary information, we identified three functional ligand-binding sites of the TRPV2 channel and discovered the inhibitors that bind to these sites. Then, we optimized the best hit and developed a highly potent selective antagonist SET2. Using SET2 as a probe, we characterized the interaction between the inhibitor and the TRPV2 channel at atomic level and revealed the mechanism of subtype selectivity among the TRPV channels. The potential role of TRPV2 as an anti-metastasis target was investigated with SET2. Furthermore, we found the functional coupling of the TRPV2 channel to the lysophosphatidic acid receptor 1 (LPAR1) might be critical in cancer progression.

RESULTS

Discovery of druggable sites and antagonists targeting these sites

The study started with the identification and characterization of drug-binding pockets. Two difficult aspects have to be considered: (i) the identification of candidate pockets that are likely to bind small subtype-selective drug-like compounds; (ii) the ranking of candidate pockets, which depends on whether the pockets might regulate the protein function. Recent tremendous advances in structural biology have provided valuable information for structure-based pocket detection methods to identify unique pockets that accept small molecular compounds. The evolutionary analysis methods, such as conservation analysis, have been widely used to unveil functionally important hotspots,

especially catalytic sites and ligand-binding sites.[18] Thus, integrating the structural-based pocket detection method with the evolutionary analysis in virtual screening may facilitate the development of subtype-selective compounds.

First, we performed multiple sequence alignment of TRPV1-TRPV6 channels. Since functionally important residues are more conserved, Jensen-Shannon divergence-based conservation analysis[18] was used to locate the top 30% ranked conserved residues of TRPV2 in multiple sequence alignment. Then, all unique residues of TRPV2 were also identified in multiple sequence alignment to reveal the hotspots, which might bind to the subtype-selective compounds. In the second stage, a structure-based pocket detection tool Fpocket[19] was used to generate a pool consisting of potential ligand binding pockets on all available wild-type (WT) full-length TRPV2 structures. Both unique residues and conserved residues ranked in the top 30% were mapped onto these ligand-binding pockets. Finally, three pockets (A, B and C) containing the largest numbers of conserved residues and at least two unique residues were detected (Figure 1A and Supporting Information Table S1).

We performed virtual screening targeting the selected pockets A, B and C. More than 200,000 commercial compounds from the Specs database were screened against each site. Fifty structure-diverse and drug-like compound candidates were selected for bioassay. In whole-cell voltage clamp experiments on HEK293T cells transiently transfected with TRPV2, six compounds (SET1, SEA6, SGE9, SGG5, SGH0 and SHB9) displayed inhibitory effects on the current of TRPV2 (Figure 1b). According to the virtual screening models, SET1 and SGH0 bound to pocket A, SEA6 and SGG5 bound to pocket B, and SGE9 and SHB9 bound to pocket C. To validate these models, we designed eight

mutations involving the following pocket residues: I165A (pocket B), Y330A (pocket B), S465A (pocket C), F467A (pocket A), L470A (pocket A), E468A (pocket C), Q525A (pocket A), and I661A (pocket C). The functionality of these mutants remained intact (Supporting Information Table S2). Compared with the WT TRPV2 channel, these mutants were less sensitive to corresponding hits (Supporting Information Figure S1). Among these six hits, SET1 is the most potent antagonist of the TRPV2 channel (Figure 1 b-d). However, because the aqueous solubility of SET1 is poor, we were not able to test this compound on TRPV2 at high concentrations (more than 10 μM).

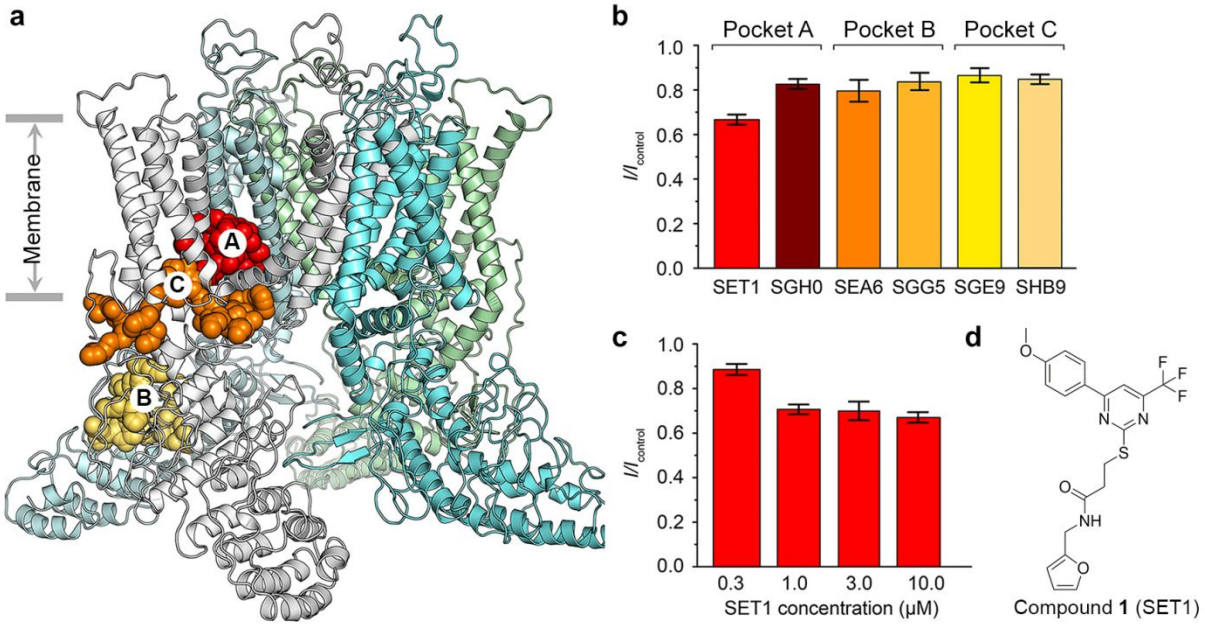


Figure 1. TRPV2 inhibitors and their binding sites. (a) Three druggable sites of the TRPV2 channel. Two TRPV2 structures (protein data bank (PDB) codes: 5AN8[11], 5HI9[10]) were employed to identify druggable sites. The protein is shown in cartoon. Ligand binding pockets A, B and C are displayed in the surface depiction. Four different subunits of TRPV2 are shown in different colors. (b) Inhibitory effects of the six compounds on

TRPV2. HEK293T cells overexpressing TRPV2 were treated with 10 μM SET1, SGH0, SEA6, SGG5, SGE9 or SHB9, respectively (n = 5). The data are shown as the mean ± standard error of mean (SEM) (c) Dose-dependent inhibition of TRPV2 by SET1 (n = 5). The data are shown as the mean ± SEM. (d) Chemical structure of SET1.

Identifying SET2 as a novel selective antagonist of TRPV2

To obtain antagonists with better solubility and potency, we designed and synthesized ten analogues of SET1 (Figure 2a). The most potent and soluble compound SET2, which has a tertiary amine instead of the phenyl group in SET1 (Figure 2b-d), inhibited TRPV2 channels with an IC_{50} of 0.46 ± 0.01 μM (Figure 2d). We then evaluated the subtype selectivity of SET2. TRPV1, TRPV3 and TRPV4 channels are the closet homologs of TRPV2. Similar to the TRPV2 channel, both TRPV1 and TRPV3 can be activated by 2-aminoethoxydiphenylborate (2-APB). TRPV4 can be activated by GSK-1016790A. Using HEK293T cells expressing TRPV1, TRPV3 or TRPV4, we found that none of TRPV1, TRPV3 or TRPV4 was sensitive to SET2 (Figure 2d and Supporting Information Figure S2). These results suggested SET2 as a potent selective TRPV2 antagonist. Based on the activities of SET1-SET11, it seems that the 4-position substituent of the pyrimidine prefers alkylamino groups rather than rigid ring structures for TRPV2 binding.

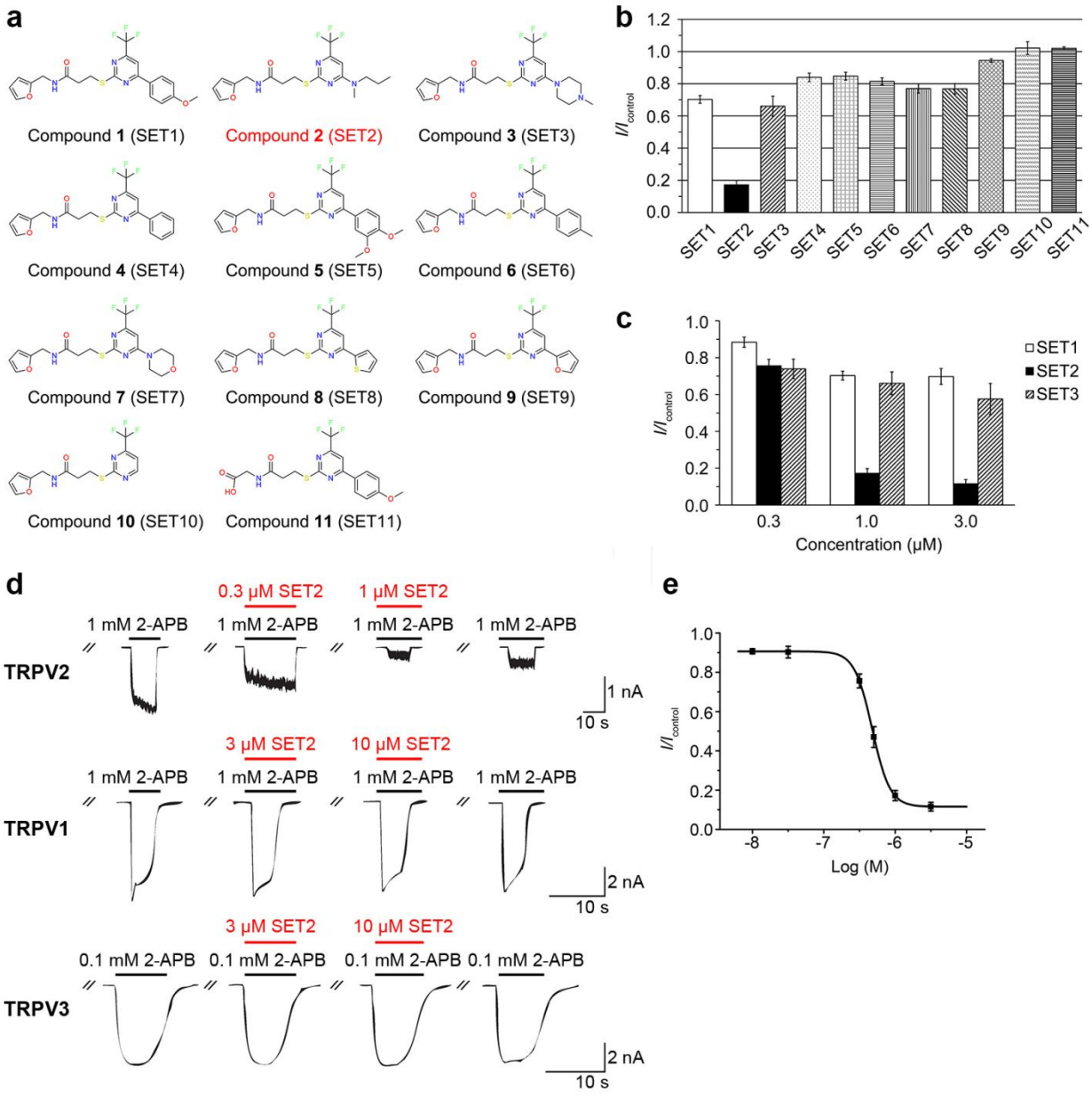


Figure 2. SET2 as a potent selective TRPV2 antagonist. (a) Chemical structures of SET1 and its analogues. (b) Inhibitory effects of 1 μM SET1 or its analogues on TRPV2 channel (n = 4). (c) Dose-dependent inhibition of TRPV2 channel by SET1, SET2 and SET3 (n = 4). The data are shown as the mean ± SEM. (d) Representative whole-cell current traces recorded from HEK293T cells overexpressing TRPV1, TRPV2 or TRPV3

under SET2 treatments. Currents were evoked by 2-APB. (e) Dose-dependent inhibition of TRPV2 by SET2 (n = 5).

Characterization of the TRPV2/SET2 binding mode

After discovering a selective potent TRPV2 inhibitor, we paid attentions to characterization of the TRPV2/SET2 binding mode. We combined several advanced modeling methods with a mutagenesis approach. First, Glide standard molecular docking analysis[20] indicated the insertion of SET2 in pocket A, which is the SET1-binding site. Then, considering the receptor flexibility and environmental effects, the RosettaLigand program and RosettaMembrane energy function[21] were used to build accurate ligand-binding models (Figure 3a and Supporting Information Figure S3). In these models, the furan ring of SET2 pointed downward to a cavity between the S4-S5 linker and S5 of an adjacent subunit, while the pyrimidine group extended upwards between S3-S4 of one subunit and between S5-S6 of the adjacent subunit (Figure 3a). To identify the key protein residues interacting with SET2, we decomposed the binding energies of the models on a per-residue basis. Tyr466, Phe467, Gln525 and Leu470 had low van der Waals (VDW) energy scores (less than -1.0 in Rosetta energy units (REU)), suggesting that they might be the main contributors to the binding of SET2 (Figure 3b). We mutated these four residues into alanine, respectively. Compared with the WT TRPV2, the TRPV2 mutants F467A, Q525A and L470A were less sensitive to SET2 (Figure 3c). We could not observe the effects of SET2 on Y466A, because it was not activated by 2-APB. These results supported the predicted SET2-binding model. In this model, Phe467, Gln525 and Leu470 stabilized the orientation of the SET2 pyrimidine group in the

transmembrane pocket (Figure 3a). These mutagenesis studies of TRPV2 validated that SET2 bound to the site used for virtual screening.

We also used explicit (all-atom) molecular dynamics simulations to simulate the process of SET2 antagonizing the TRPV2 channel. In the simulations, the dominant pose of SET2 was similar to the one observed in the docking model (Figure 3a and Supporting Information Figure S4). We found that SET2 also frequently interacted with the S5 and S6 helices of the adjacent subunit (Val538, Phe542, Leu627 and Leu631) (Supporting Information Figure S4). Since the central ion-conductive pore region includes the S5 and S6 helices, SET2 might stabilize the channel in the resting state and might prevent channel activation upon binding to the transmembrane pocket (Supporting Information Table S3).

Mechanism of the subtype selectivity

Elucidating the subtype selectivity mechanism of homologous proteins is always challenging. After validating the binding site of SET2 on TRPV2, it is possible to elucidate the subtype selectivity by comparing TRPV2 with other TRPV channels. None of TRPV1, TRPV3 or TRPV4 is sensitive to SET2 (Figure 2 and Supporting Information Figure S2). For TRPV3 or TRPV4, it is most likely that its low similarity to TRPV2 in the transmembrane domains leads to the distinct channel sensitivity towards SET2. In contrast, the TRPV1 channel is structurally similar to the TRPV2 channel, with 43% amino acid sequence identity in the transmembrane domains. We performed sequence alignment between TRPV1 and TRPV2, and found three unique residues Phe467, Leu505 and Gln525 within the SET2-binding TRPV2 pocket (Figure 3d).

We first made mutagenesis study on TRPV2. We substituted these residues (Phe467, Leu505 and Gln525) of TRPV2 with the corresponding residues (Ser513, Thr551 and Glu571) in TRPV1. Compared with the WT TRPV2, the TRPV2 mutants F467S and L505T were less sensitive to SET2 (Figure 3c). SET2 had decreased activities on both F467A and F467S mutants, which suggested the importance of Phe467 in the interaction with the SET2 molecule. In addition, a hydrophobic leucine was more favorable than a polar threonine at the residue position 505 for SET2-binding possibly because the hydrophobic side chain of Leu505 stabilized the pyrimidine group of SET2 (Figure 3a). We could not observe the effects of SET2 on Q525E, since this mutant was rapidly desensitized to 2-APB. SET2 did not inhibit the mutant Q525A, as shown above.

TRPV1 is not sensitive to SET2. We tried to identify TRPV1 mutants that are sensitive to SET2. Along with the mutations of TRPV2, such mutations of TRPV1 could provide very important information of the subtype selectivity mechanism. Because Phe467 and Leu505 are unique and crucial residues for SET2 binding in the TRPV2 channel, we substituted the corresponding residues in TRPV1 (Ser513 and Thr551) to design a double mutation of TRPV1, i.e., S513F/T551L. As shown in Figure 3g, the S513F/T551L mutant was inhibited by SET2, indicating that this ligand bound to the transmembrane pocket of the TRPV1 mutant. We also tested S513F and T551L single-point mutants in bioassays. While S513F was not sensitive to SET2, T551L was inhibited by SET2 at high concentrations (no less than 1 μM) (Supporting Information Figure S5). To further characterize the structural basis for this gain-of-function mutation, we built models of the S513F/T551L mutant and performed docking simulations of SET2 to these models. In the docking models, both Phe513 and Leu551 had low VDW energy scores (less than -1.0

REU), suggesting that they made significant hydrophobic contributions to SET2 binding. In the WT TRPV1, the corresponding residues Ser513 and Thr551 have polar side chains, which might interfere in the interactions with the hydrophobic pyrimidine groups and furan ring of the SET2 molecule. These identified gain-of-function mutations of TRPV1 further convincingly demonstrate the binding site and the subtype selectivity mechanism.

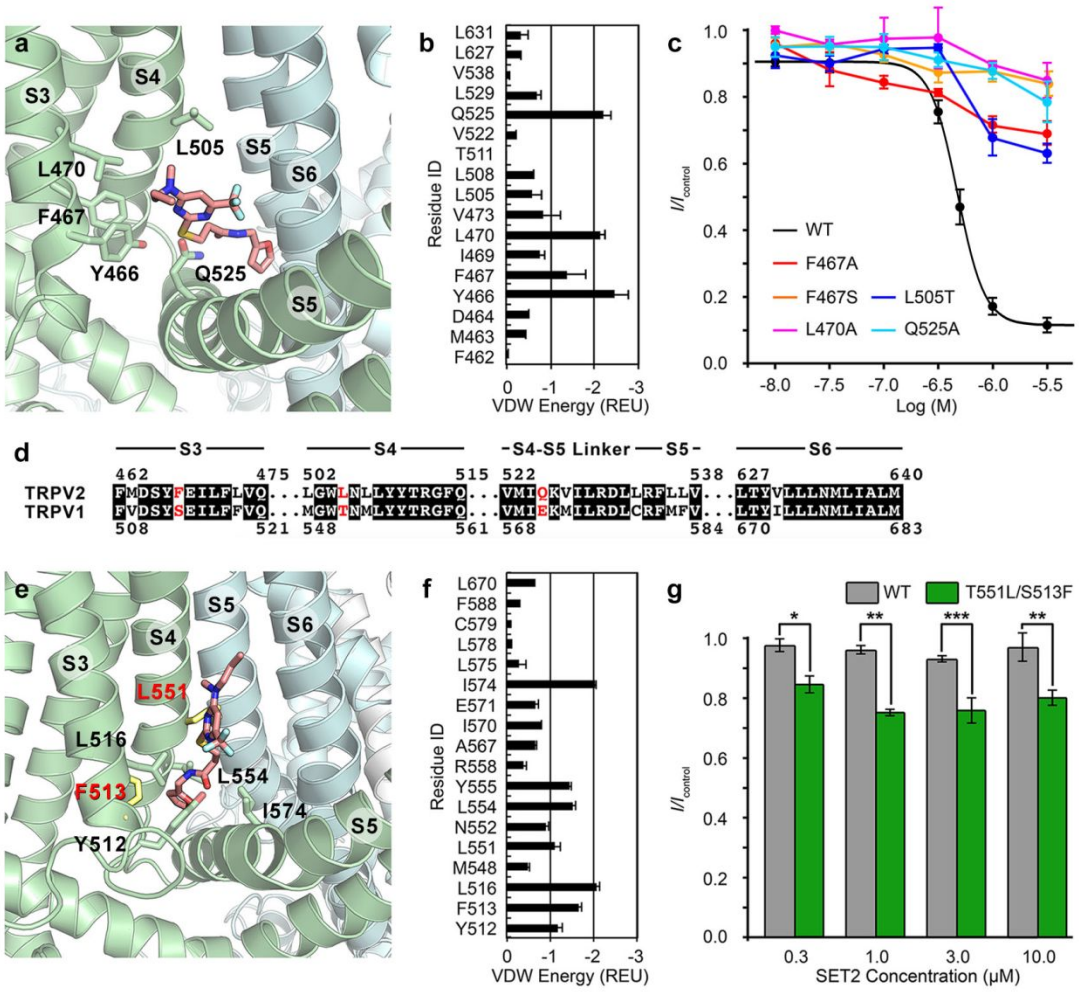


Figure 3. Inhibition of WT TRPV2, mutant TRPV2 and mutant TRPV1 channels by SET2. (a) Binding of SET2 to the TRPV2 channel (PDB code: 5AN8[11]). The protein is shown in cartoon. SET2 and the residues involved in binding are displayed as sticks. (b) Distributions of the average VDW interactions between SET2 and TRPV2 (top 10

docking models). (c) Dose-dependent inhibition of WT and mutant TRPV2 channels by SET2 (n = 4-6). (d) Sequence alignment of the transmembrane pockets of the TRPV1 and TRPV2 channels. Unique residues are colored in red. (e) Binding of SET2 to the TRPV1 S513F/T551L mutant channel (PDB code: 5AN8[11]). Residues Phe513 and Leu551 are highlighted in yellow. (f) Distributions of average VDW interactions between SET2 and the TRPV1 S513F/T551L mutant channel (top 10 docking models). (g) Dose-dependent inhibition of the WT and S513F/T551L mutant TRPV1 channels by SET2 (n = 4-5). Unpaired t-test was used for statistical analysis [t (8) = 3.572 for 0.3 μM SET2-treatment, t (7) = 8.933 for 0.1 μM SET2-treatment, t (6) = 3.780 for 3.0 μM SET2-treatment, t (8) = 3.174 for 10.0 μM SET2-treatment]; * indicates $P < 0.05$, ** indicates $P < 0.01$, *** indicates $P < 0.005$. The data are shown as the mean ± SEM.

Functional role of TRPV2 in metastasis revealed by SET2

TRPV2 has high expression levels in metastatic prostate cancers.[4] It has been reported that small-interfering RNA (siRNA) silencing of the TRPV2 reduces cancer cell migration.[5] Therefore, TRPV2 is a potential anti-metastasis target. In accordance with previous reports,[5] our semi-quantitative reverse transcription polymerase chain reaction (RT-PCR) and immunoblotting analysis of the human prostate cancer cell line PC-3M revealed the expression of TRPV2 (Supporting Information Figure S6a, b). And our transwell-based cell migration assays showed that the siRNA silencing of TRPV2 reduced PC-3M migration (Supporting Information Figure S6c, d). SET2 treatment had no significant influence on the migration of the PC-3M that were transfected with siRNA (siTRPV2) (Supporting Information Figure S6c, d). In contrast, SET2 treatment

abrogated the migration of PC-3M cells expressing TRPV2 (Figure 4a, b). Remarkably, SET2 had no significant effect on cell survival at all tested concentrations (Figure 4c). These results supported that the negative modulation of TRPV2 by small molecules inhibited the migration of prostate cancer cells, confirming TRPV2 as an anti-metastasis target.

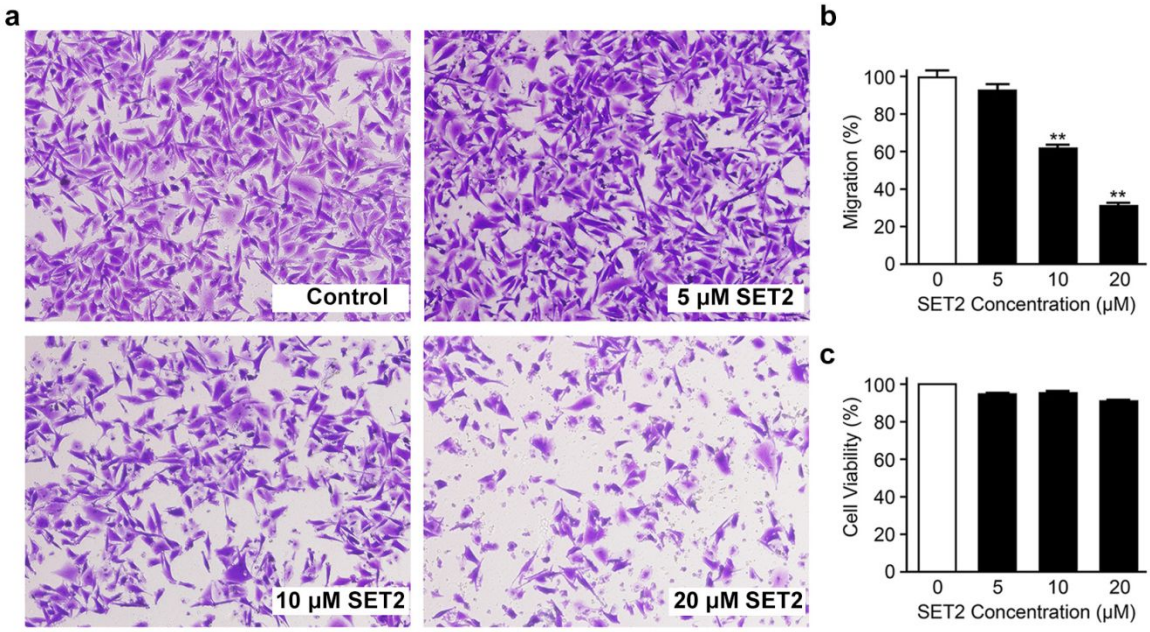


Figure 4. Effects of SET2 treatment on the migration of prostate cancer cells. (a) Representative images showing SET2 inhibition of prostate cancer cell (PC-3M) migration. For PC-3M migration assessment, PC-3M cells (8×10^4) cells were allowed to migrate for 24 hours with the indicated treatments. (b) Migration ratios of PC-3M after treatment with SET2 at different concentrations (n = 3). One-way ANOVA with Dunnett's multiple comparison test was used for statistical analysis [F (4, 15) = 141.2]. The data are shown as the mean ± SEM; ** indicates $P < 0.01$. (c) Cell viability of PC-3M after treatments with SET2 for 48 hours (n = 3).

As an effective chemical tool, SET2 facilitated the mechanism study of TRPV2 in simulating cancer cell migration. The lysophosphatidic acid (LPA) signaling increases cancer cell migration in numerous ways.[22] It has been reported that LPA can indirectly or directly up-regulate TRPV1.[23, 24] Considering high homology between TRPV1 and TRPV2, LPA might also regulate TRPV2. To validate this assumption, we used SET2 and LPA to treat PC-3M in migration assays. Consistent with previous reports,[25, 26] after treating 0.1 μM LPA, a significant LPA-induced increase in cell migration was observed (Figure 5a, b). Pretreatment with SET2 reduced the LPA-induced migration of PC-3M, suggesting that TRPV2 participated in the LPA-induced signal process. Next we studied whether LPA directly regulate TRPV2. The effects of LPA on TRPV2 were investigated using patch-clamp recordings. LPA failed to directly activate or potentiate TRPV2 (Supporting Information Figure S7). These observations indicated that LPA indirectly regulated TRPV2 through an upstream effector expressed in PC-3M cells. In the RT-PCR and immunoblotting analysis of PC-3M, we found that a G-protein coupled receptor of LPA, i.e. LPAR1, was highly expressed (Supporting Information Figure S6e, f). We treated PC-3M with a selective LPAR1 antagonist AM966 in a cell migration assay and observed that LPA-induced cell migration was abolished by AM966 (Figure 5a, b). These findings indicated that LPA activated LPAR1 to induce PC-3M cell migration. Altogether, migration assays suggested that both TRPV2 and LPAR1 participated in LPA-induced PC-3M cell migration.

We further studied the relationship between LPAR1 and TRPV2 using patch-clamp recordings. We assumed that LPA potentiated TRPV2 via regulating LPAR1. In the

recording experiments, we used HEK293T cells transiently co-transfected with TRPV2 and LPAR1. The TRPV2 agonist 2-APB was used to invoke currents. In co-transfected cells, 2-APB-evoked currents were inhibited by SET2 (Figure 5c and Supporting Information Figure S8a), indicating that these currents were primarily carried by TRPV2. LPA potentiated (2-APB-evoked) TRPV2 currents in the TRPV2/LPAR1 co-transfected cells (Figure 5c and Supporting Information Figure S8b), but it did not potentiate TRPV2 currents in the TRPV2 transfected cells (Supporting Information Figure S7b, c). These results supported our assumption that LPA potentiated TRPV2 via modulating LPAR1. In co-transfected cells, the LPA-induced potentiation of TRPV2 currents was blocked by a LPAR1 antagonist AM966 (Figure 5c and Supporting Information Figure S8c), indicating that LPA potentiated TRPV2 by activating LPAR1. We also studied the relationship between LPAR1 and TRPV2 in calcium responses, which are essential in cancer cell migration. Application of LPA induced a cytoplasmic calcium increase in PC-3M (Supporting Information Figure S9). The LPA-induced cytoplasmic calcium increases were inhibited by the TRPV2 antagonist SET2 (Supporting Information Figure S9).

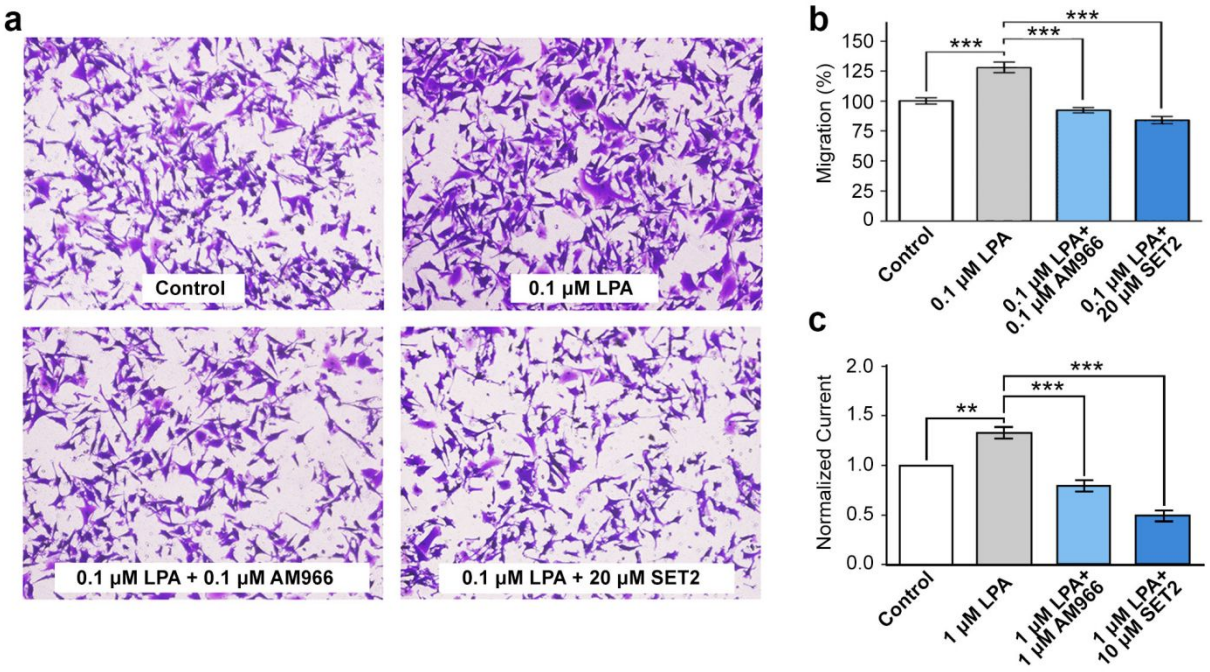


Figure 5. Effects of SET2 and AM966 on the LPA-induced migration of PC-3M. (a) Representative images showing that SET2 and AM966 inhibited LPA-induced PC-3M migration. For the LPA-induced PC-3M migration assay, PC-3M cells (4×10^4) were allowed to migrate for 24 hours with the indicated treatments. (b) Migration ratios of PC-3M after treatment with dimethyl sulfoxide (DMSO), LPA, AM966 and SET2 (n = 3). One-way ANOVA with Dunnett's multiple comparison test was used for statistical analysis [F (3, 12) = 30.63]. (c) Effects of LPA, AM966 and SET2 on currents of HEK293T cells, which were co-transfected with LPAR1 and TRPV2. The currents were evoked by 2-APB. 3 min after 1 μM LPA perfusion on the co-transfected cells, the currents were significantly increased (n = 10). Both AM966 (n = 7) and SET2 (n = 7) eliminated the LPA-induced potentiation of currents. On-way ANOVA with Dunnett's T3 comparison test was used for statistical analysis [F (3, 30) = 59.49]. The data are shown as the mean ± SEM. ** indicates $P < 0.01$, *** indicates $P < 0.005$.

Discussion

A solid drug target validation requires both genetic evidence and specific chemical probe. It is more challenging to discover chemical probes than obtaining genetic evidence. As a result, there might be a massive genetic evidence of a particular protein suggesting its association with certain diseases, while its target validation could not be fully performed due to a lack of chemical probes. Despite decades of studies, the development of potent selective TRPV2 antagonists has been unsuccessful. New strategies and information for drug development are urgently needed. In this study, we introduced Jensen-Shannon divergence-based conservation analysis in the pocket prediction and successfully identified three pockets containing multiple functionally important residues. New TRPV2 antagonists with different structure types were discovered via screening against these pockets, which demonstrates the efficiency of the strategy in drug development. From the molecular mechanistic point of view, it is interesting to reveal three ligand-binding sites regulating the gating of TRPV2. Pocket A includes the S4-S5 linker, a domain known to transduce activation conformational changes in TRP channels.[10, 11, 14, 15, 27, 28] Close to pocket A, pocket C is formed by the transmembrane helices S1-S4 and the TRP helix following the S6 helix. Pocket B is an intracellular site involving the N-terminal ankyrin repeat domain (ARD), which is connected to the S1 helix. The twisting motion of the ARD might trigger the TRP domain, which pulls the S6 helix and causes the lower gate to open.[29] Interestingly, these three proposed druggable pockets comprise three pivotal functional elements the S4-S5 linker, the TRP helix and the ARD, which are interdependent in TRPV2 gating. These findings

not only provide new hot spots for drug design but also help us understand the TRP channel machinery.

Many evidence demonstrate that the TRPV2 expression correlates with tumor grades in various cancers.[4, 30] Previous studies have shown that siRNA mediated silencing of TRPV2 reduces the growth and migration of metastatic cancer cells *in vitro* and *in vivo*.[5] Thus, TRPV2 is a potential drug target for cancers. Consistent with results of the gene silencing assays, we found that SET2 effectively blocked the TRPV2 channel and reduced basal migration of prostate cancer cells. This result supports TRPV2 as an anti-metastasis target. Considering the efficiency of SET2, it could also be used in preclinical validation of TRPV2. And new chemical probes may be developed based on SET2.

LPA has gained immense importance due to its excess production in carcinoma, glioma, ovarian, prostate, and breast cancers, promoting their invasive properties.[31-35] Several studies have reported that LPA and its receptor LPAR1 regulates the motion dynamics in cancer cells via mediating cell spreading and lamellipodia formation.[36] Moreover, LPA modulates adherent junction disassembly, cytoskeletal disorganization and formation of focal adhesion in ovarian and colon cancers.[36, 37] LPA-mediated signaling is critical in mediating cancer cell migration, but the downstream effectors in this signaling pathway remain largely unknown. In this work, we showed that LPA potentiated TRPV2 via activating LPAR1 and identified the LPA/LPAR1/TRPV2 signaling pathway in the migratory process of prostate cancer cells. Since the autotaxin/LPA axis is essential in cancer cell migration, identification of TRPV2 as a downstream mediator in this signaling pathway further validates its importance and potential as a therapeutic target for cancers. Considering that cancer cell migration

remains one of the less-understood processes leading to cancer progression, further study of the LPA/LPAR1/TRPV2 pathway may lead to new therapeutic treatments for cancers.

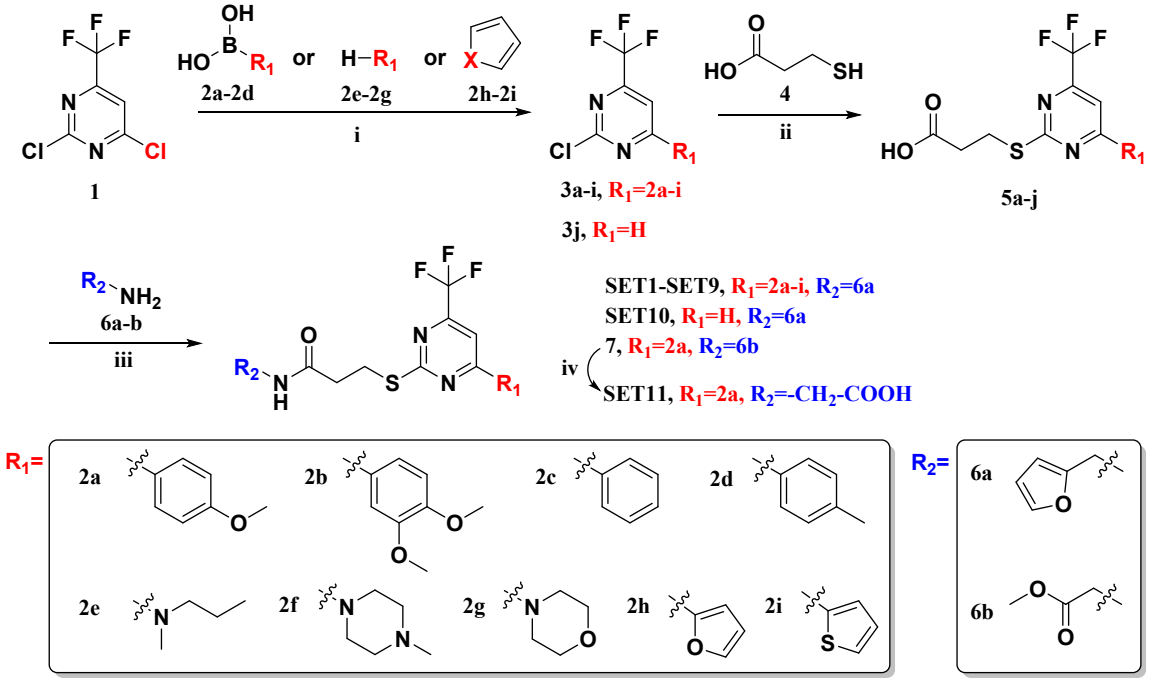
In summary, by applying an advanced strategy of drug development, we designed a novel compound SET2, which potently selectively inhibited the TRPV2 channel at submicromolar concentrations. Combining electrophysiology, mutagenesis and computational modeling, we revealed the molecular basis of the subtype selectivity of SET2 at the atomic level. Using SET2 as a chemical probe, we unveiled the functional coupling between LPAR1 and TRPV2 in the physiological processes of cancer metastasis. These findings not only demonstrate the feasibility of exploiting the TRPV2 channel as a therapeutic target for metastatic prostate cancer, but also suggest SET2 as a promising drug lead and an efficient tool for further target validation studies in the future.

EXPERIMENTAL SECTION

Material and Instrumentation. All reagents and solvents were commercially purchased from Sinopharm, Bide Pharmatech Ltd. and Shanghai Titan Ltd. (Shanghai, China) and used without further purification. Analytical thin layer chromatography (TLC) was performed on TLC plates precoated with silica gel $HSGF_{254}$ (200 □□30 μm thickness). Analytic RP-HPLC analysis was performed on a Beijing Chuang Xin Tong Heng LC-3000 (analytic model) instrument with a C-18 column (5 μm, 4.6 x 150 mm) at 40 ^{o}C. The column was eluted with a gradient of 2-90% acetonitrile containing 0.1% TFA in 30 min at a flow rate of 1 mL/min. HPLC analysis showed that the purity of all the final products were more than 95%. ^{1}H-NMR and ^{13}C-NMR spectra were recorded on a BRUKER Ascend^{TM} 400 MHz or 600 MHz instrument. Chemical shifts were assigned in

ppm and coupling constants were assigned in Hz. Constant-time ^{1}H-NMR and HSQC were recorded on AVANCE III HD BRUKER Ascend™ 600 MHz instrument. ESI-MS spectra were measured with an Agilent 6230 LC-TOF MS spectrometer.

N-(furan-2-ylmethyl)-3-((4-(4-methoxyphenyl)-6-(trifluoromethyl)pyrimidin-2-yl)thio)propanamide (SET1) and other compounds applied in virtual screening were purchased from Specs. We synthesized 10 analogues of **SET1**, including **N-(furan-2-ylmethyl)-3-((4-(N'-methyl-N'-propylamino)-6-(trifluoromethyl)-pyrimidin-2-yl)thio)-propanamide (SET2)** (Scheme 1).



Scheme 1. Synthesis of derivatives **SET1-SET11**[a].

[a]**Reagents and conditions:** i) for synthesis of **3a-3d** and **3h-3i**: **2a-2d** or **2h-2i**, Trans-Dichlorobis-(triphenyl-phosphine)Palladium(II), 1,4-dioxane, sat. Na_2CO_3, 75 °C; for synthesis of **3e-3g**: **2e-2g**, TEA, DCM, 0 °C; ii) 1,8-diazabicyclo[5.4.0]undec-7-ene, 1,4-dioxane, 95 °C; iii) EDC, HOBt, TEA, DCM, rt; iv) NaOH, THF/MeOH/H_2O, rt.

General Procedures for synthesis of intermediates 3a-3d and 3h-3i. To a solution of 2,4-dichloro-6-(trifluoromethyl)-pyrimidine (**1**, 1g, 1 eq) in 1,4-dioxane (10 mL), **2a-2d** and **2h-2i** (1.1 eq), saturated solution of sodium carbonate (5 mL) and Trans-Dichlorobis-(triphenyl-phosphine)Palladium(II) (342 mg, 0.1 eq) were added and the mixture was stirred at 75℃ for 5 h. After cooled to room temperature, the solution was diluted with dichloromethane (20 mL) and washed with brine and water successively. The organic layer was dried over anhydrous sodium sulfate, filtered and evaporated in vacuo. The residue was subject to chromatography purification through a silica gel column and eluted with petroleum ether/ethyl acetate (20/1) to give the desired product.

General procedures for synthesis of intermediates 3e-3g. To a solution of 2,4-dichloro-6-(trifluoro-methyl)pyrimidine (**1**, 1 g, 1 eq) in dichloromethane (30 mL), **2e-2g** (1.1 eq) and triethylamine (1.4 g, 3 eq) were added and the mixture was stirred under ice-cooling for 5 h. Then, the solution was diluted with dichloromethane (50 mL) and washed with brine and water successively. The organic layer was dried over anhydrous sodium sulfate, filtered and evaporated in vacuo. The residue was subject to chromatography purification through a silica gel column and eluted with petroleum ether/ethyl acetate (20/1) to give the desired product.

General procedures for synthesis of SET1-SET10 and intermediate 7. To synthesize **SET1-SET10** and intermediate **7**, we prepared the intermediates **5a-j** via chlor-substitution of prior intermediates **3a-i** and 3-((4-(trifluoromethyl)-pyrimidin-2-yl)thio)propanoic acid (**3j**). First, **3a-j** (1 g, 1 eq) and 3-mercaptopropanoic acid (**4**, 1.1 g, 3 eq) were dissolved in 1,4-dioxane (10 mL) and 1,8-diazabicyclo[5.4.0]undec-7-ene

(1.58 g, 3 eq) was added. The solution was stirred at 95 °C for 10 h. After cooled to room temperature, the solution was diluted with dichloromethane (50 mL) and washed with brine and water successively. The organic layer was dried over anhydrous sodium sulfate, filtered and evaporated in vacuo. The residue was subject to chromatography purification through a silica gel column and eluted with petroleum ether/ethyl acetate (2/1) to give desired product **5a-j**. Then, A mixture of **5a-j** (1 eq), furfurylamine or glycinemethylester hydrochloride (**6a-b**, 2 eq), 1-ethyl-3-(3-dimethylaminopropyl)-carbodimide hydrochloride (550 mg, 2 eq), 1-hydroxybenzotriazole hydrate (377 mg, 2 eq), triethylamine (423 mg, 3 eq), and dichloromethane (5 mL) was stirred at room temperature for 10 h. After complete conversion, the solution was diluted with dichloromethane (20 mL) and washed with brine and water successively. The organic layer was dried over anhydrous sodium sulfate, filtered and evaporated in vacuo. The residue subject to chromatography purification through a silica gel column and eluted with petroleum ether/ethyl acetate (3/1) to give desired products **SET1-SET10** and intermediate **7**.

N-(furan-2-ylmethyl)-3-((4-(4-methoxyphenyl)-6-(trifluoromethyl)pyrimidin-2-yl)thio)propanamide (SET1). Yield (final step) 85%. ^{1}H NMR (400 MHz, Chloroform-d) δ 8.14 (d, J = 2.2 Hz, 1H), 8.12 (s, 1H), 7.61 (s, 1H), 7.35 (m, 1H), 7.05 (s, 1H), 7.03 (s, 1H), 6.33 (dd, J = 3.2, 1.8 Hz, 1H), 6.25 (d, J = 3.2 Hz, 1H), 5.91 (s, 1H), 4.48 (d, J = 5.2 Hz, 2H), 3.92 (s, 3H), 3.55 (t, J = 7.1 Hz, 2H), 2.76 (t, J = 7.1 Hz, 2H). ^{13}C NMR (101 MHz, CDCl3) δ 173.1, 170.6, 165.9, 163.1, 156.1, 151.1, 142.2, 129.2, 129.23, 127.4, 120.6, 114.5, 114.5, 110.4, 107.4, 106.6, 55.5, 36.5, 36.1, 26.9. ESI-HRMS: calcd. for $C_{20}H_{18}F_3N_3O_3S$, 438.1099 $(M+H)^+$, found, 438.1098.

N-(furan-2-ylmethyl)-3-((4-(N'-methyl-N'-propylamino)-6-(trifluoromethyl)-pyrimidin-2-yl)thio)-propanamide (SET2). Yield (final step) 80%. ^{1}H NMR (400 MHz, Chloroform-d) δ 7.35 (m, 1H), 6.41 (s, 1H), 6.32 (m, 1H), 6.24 (d, *J* = 3.2 Hz, 1H), 6.03 (s, 1H), 4.46 (d, *J* = 6.8 Hz, 2H), 3.65 (m, 1H), 3.39 (t, *J* = 7.2 Hz, 2H), 3.33 (m, 1H), 3.18-3.06 (m, 3H), 2.69 (t, *J* = 7.2 Hz, 2H), 1.64 (m, 2H), 0.95 (m, 3H). 13C NMR (101 MHz, Chloroform-d) δ 171.2, 161.5, 151.1, 142.2, 122.3, 120.5, 110.4, 107.4, 94.7, 51.5, 36.8, 36.5, 35.9, 26.6, 20.4, 11.4. ESI-HRMS: calcd. for $C_{17}H_{21}F_3N_4O_2S$, 403.1416 $[M+H]^+$, found, 403.1477.

N-(furan-2-ylmethyl)-3-((4-(4-methyl-piperazin-1-yl)-6-(trifluoromethyl)-pyrimidin-2-yl)thio)-propanamide (SET3). Yield (final step) 37%. ^{1}H NMR (400 MHz, Chloroform-d) δ 7.36 (m, 1H), 6.52 (s, 1H), 6.33 (dd, J = 3.2, 1.9 Hz, 1H), 6.24 (m, 1H), 6.01 (s, 1H), 4.46 (d, J = 5.6 Hz, 2H), 3.74 (s, 4H), 3.39 (t, J = 7.3 Hz, 2H), 2.68 (t, J = 7.2 Hz, 2H), 2.50 (t, J = 5.2 Hz, 4H), 2.36 (s, 3H). ^{13}C NMR (101 MHz, CDCl3) δ 172.0, 170.9, 161.3, 154.6, 151.1, 142.1, 120.8, 110.4, 107.4, 94.9, 54.4, 54.4, 45.9, 43.9, 36.7, 36.5, 29.6, 26.8. ESI-HRMS: calcd. for $C_{18}H_{22}F_3N_5O_2S$, 430.1525 $[M+H]^+$, found, 430.1542.

N-(furan-2-ylmethyl)-3-((4-(phenyl)-6-(trifluoromethyl)-pyrimidin-2-yl)thio)-propanamide (SET4). Yield (final step) 65%. ^{1}H NMR (400 MHz, Chloroform-d) δ 8.14 (m, 2H), 7.68 (s, 1H), 7.62 – 7.57 (m, 1H), 7.57 – 7.51 (m, 2H), 7.35 (m, 1H), 6.33 (dd, J = 3.3, 1.9 Hz, 1H), 6.24 (m, 1H), 5.95 (s, 1H), 4.48 (d, J = 5.5 Hz, 2H), 3.56 (t, J = 7.1 Hz, 2H), 2.76 (t, J = 7.1 Hz, 2H). ^{13}C NMR (101 MHz, CDCl3) δ 173.5, 170.7, 166.6, 156.4, 151.0, 142.2, 135.1, 132.2, 129.2, 129.2, 127.5, 127.5, 120.5, 110.5, 107.6, 107.5,

36.6, 36.1 27.0. ESI-HRMS: calcd. for $C_{19}H_{16}F_3N_3O_2S$, 408.0994 $[M+H]^+$, found, 408.1065.

N-(furan-2-ylmethyl)-3-((4-(3,4-dimethoxyl-phenyl)-6-(trifluoromethyl)-pyrimidin-2-yl)thio)-propanamide (SET5). Yield (final step) 88%. ^{1}H NMR (400 MHz, Chloroform-d) δ 7.76 (d, J = 2.1 Hz, 1H), 7.71 (dd, J = 8.4, 2.1 Hz, 1H), 7.61 (s, 1H), 7.35 (m, 1H), 6.99 (d, J = 8.5 Hz, 1H), 6.32 (dd, J = 3.3, 1.9 Hz, 1H), 6.24 (m, 1H), 5.90 (s, 1H), 4.47 (d, J = 5.5 Hz, 2H), 4.02 (s, 3H), 3.99 (s, 3H), 3.55 (t, J = 7.2 Hz, 2H), 2.76 (t, J = 7.2 Hz, 2H). ^{13}C NMR (101 MHz, CDCl3) δ 173.1, 170.6, 165.9, 156.0, 152.7, 151.0, 149.6, 142.2, 127.7, 121.1, 120.6, 111.1, 110.4, 109.8, 107.5, 106.8, 56.1, 56.1, 36.5, 36.1, 26.9. ESI-HRMS: calcd. for $C_{21}H_{20}F_3N_4O_4S$, 468.1205 $[M+H]^+$, found, 468.1211.

N-(furan-2-ylmethyl)-3-((4-(4-methyl-phenyl)-6-(trifluoromethyl)-pyrimidin-2-yl)thio)-propanamide (SET6). Yield (final step) 56%. ^{1}H NMR (400 MHz, Chloroform-d) δ 8.03 (s, 1H), 8.01 (s, 1H), 7.63 (s, 1H), 7.34 (d, J = 1.8 Hz, 2H), 7.32 (s, 1H), 6.32 (dd, J = 3.2, 1.9 Hz, 1H), 6.23 (d, J = 3.2 Hz, 1H), 6.07 (t, J = 5.6 Hz, 1H), 4.46 (d, J = 5.5 Hz, 2H), 3.53 (t, J = 7.1 Hz, 2H), 2.75 (t, J = 7.1 Hz, 2H), 2.45 (s, 3H). ^{13}C NMR (101 MHz, CDCl3) δ 173.3, 170.6, 166.4, 156.23, 151.1, 143.0, 142.2, 132.3, 129.9, 129.94, 127.4, 127.4, 120.5, 110.4, 107.5, 107.21, 36.58, 36.1, 26.9, 21.6. ESI-HRMS: calcd. for $C_{20}H_{18}F_3N_3O_2S$, 422.1150 $[M+H]^+$, found, 422.1157.

N-(furan-2-ylmethyl)-3-((4-(morpholin-4-yl)-6-(trifluoromethyl)-pyrimidin-2-yl)thio)-propanamide (SET7). Yield (final step) 40%. ^{1}H NMR (400 MHz, Chloroform-d) δ 7.36 (m, 1H), 6.51 (s, 1H), 6.33 (dd, J = 3.2, 1.9 Hz, 1H), 6.24 (m, 1H), 5.98 (s, 1H), 4.46 (d, J = 5.5 Hz, 2H), 3.80 (m, 4H), 3.70 (s, 4H), 3.39 (t, J = 7.3 Hz, 2H), 2.67 (t, J =

7.3 Hz, 2H). ^{13}C NMR (101 MHz, CDCl3) δ 172.1, 170.8, 161.6, 154.7, 151.1, 142.2, 120.7, 110.4, 107.4, 94.8, 94.7, 66.3, 66.3, 44.3, 36.6, 36.6, 26.7. ESI-HRMS: calcd. for $C_{17}H_{19}F_3N_4O_3S$, 417.1208 $[M+H]^+$, found, 417.1241.

N-(furan-2-ylmethyl)-3-((4-(thiophen-2-yl)-6-(trifluoromethyl)-pyrimidin-2-yl)thio)-propanamide (SET8). Yield (final step) 46%. ^{1}H NMR (400 MHz, Chloroform-d) δ 7.86 (dd, J = 3.8, 1.2 Hz, 1H), 7.61 (dd, J = 5.0, 1.1 Hz, 1H), 7.50 (s, 1H), 7.36 (m, 1H), 7.21 (dd, J = 5.0, 3.8 Hz, 1H), 6.34 (dd, J = 3.2, 1.9 Hz, 1H), 6.26 (d, J = 3.2 Hz, 1H), 5.93 (s, 1H), 4.49 (d, J = 5.5 Hz, 2H), 3.52 (t, J = 7.2 Hz, 2H), 2.76 (t, J = 7.2 Hz, 2H). ^{13}C NMR (101 MHz, CDCl3) δ 173.4, 170.6, 161.2, 156.1, 151.1, 142.2, 140.8, 132.1, 129.3, 128.8, 120.3, 110.4, 107.5, 105.9, 36.6, 36.1, 27.0. ESI-HRMS: calcd. for $C_{17}H_{14}F_3N_3O_2S_2$, 414.0558 $[M+H]^+$, found, 414.0613.

N-(furan-2-ylmethyl)-3-((4-(furan-2-yl)-6-(trifluoromethyl)-pyrimidin-2-yl)thio)-propanamide (SET9). Yield (final step) 78%. ^{1}H NMR (400 MHz, Chloroform-d) δ 7.64 (m, 1H), 7.55 (s, 1H), 7.40 (m, 1H), 7.33 (m, 1H), 6.62 (dd, J = 3.6, 1.7 Hz, 1H), 6.31 (dd, J = 3.3, 1.9 Hz, 1H), 6.23 (m, 1H), 5.95 (s, 1H), 4.46 (d, J = 5.5 Hz, 2H), 3.49 (t, J = 7.2 Hz, 2H), 2.71 (t, J = 7.2 Hz, 2H). ^{13}C NMR (101 MHz, CDCl3) δ 173.2, 170.6, 157.4, 156.2, 151.0, 150.6, 146.3, 142.2, 120.3, 114.8, 113.1, 110.4, 107.5, 105.7, 36.5, 36.1, 27.0. ESI-HRMS: calcd. for $C_{17}H_{14}F_3N_3O_3S$, 398.0786 $[M+H]^+$, found, 398.0891.

N-(furan-2-ylmethyl)-3-((6-(trifluoromethyl)-pyrimidin-2-yl)thio)-propanamide (SET10). Yield (final step) 85%. ^{1}H NMR (400 MHz, Chloroform-d) δ 8.76 (d, J = 4.9 Hz, 1H), 7.36 (m, 1H), 7.30 (d, J = 5.0 Hz, 1H), 6.34 (dd, J = 3.2, 1.9 Hz, 1H), 6.28 – 6.14 (m, 1H), 5.92 (s, 1H), 4.48 (d, J = 5.5 Hz, 2H), 3.49 (t, J = 7.1 Hz, 2H), 2.71 (t, J = 7.1 Hz, 2H). ^{13}C NMR (101 MHz, CDCl3) δ 173.8, 170.4, 159.5, 155.8, 151.0, 142.3,

120.2, 111.9, 110.5, 107.7, 36.6, 35.9, 26.9. ESI-HRMS: calcd. for $C_{13}H_{12}F_3N_3O_2S$, 332.0681 $[M+H]^+$, found, 332.0813.

N-(hydroxycarbonylmethyl)-3-((4-(4-methoxyl-phenyl)-6-(trifluoromethyl)-pyrimidin-2-yl)thio)-propanamide (SET11). To a solution of **SET11** precursor **7** (300 mg, 1 eq) in the mixture of tetrahydrofuran, methanol and water (3:1:1, 5 mL), a proper amount of 1 M sodium hydroxide solution was added to adjust pH to 12 and the mixture was stirred for 4 h at room temperature. Then, the solution was diluted with dichloromethane (10 mL) and washed with brine and water successively which was furtherly concentrated under reduced pressure. After that, the residue was subject to chromatography purification through a silica gel column and eluted with petroleum ether/ethyl acetate (2/1) to give **SET11** (275 mg, final-step yeild 95%). ^{1}H NMR (400 MHz, DMSO-d6) δ 8.35 (s, 1H), 8.33 (d, J = 2.1 Hz, 1H), 8.28 (t, J = 5.8 Hz, 1H), 8.17 (s, 1H), 7.14 (d, J = 2.0 Hz, 1H), 7.12 (d, J = 2.1 Hz, 1H), 3.88 (s, 3H), 3.78 (d, J = 5.8 Hz, 2H), 3.42 (t, J = 6.9 Hz, 2H), 2.70 (t, J = 7.0 Hz, 2H). ^{13}C NMR (101 MHz, DMSO) δ 172.6, 171.7, 171.1, 166.1, 163.3, 155.5, 130.2, 130.2, 127.4, 121.0, 115.1, 115.1, 107.9, 56.0,41.1, 34.7, 26.8. ESI-HRMS: calcd. for $C_{17}H_{16}F_3N_3O_4S$, 416.0892 $(M+H)^+$, found, 416.0834.

Molecular biology and cell transfection. Complementary DNA (cDNA) of the mouse TRPV1, TRPV2, TRPV3 and TRPV4 channels were subcloned into the pEGFPN1 expression vector (Invitrogen), respectively. cDNA of TRPV2 and cDNA of LPAR1 were co-transfected at a 1:5 ratio. cDNA encoding green fluorescent proteins were used to facilitate identification of transfected cells. Mutations were introduced into TRPV1 or

TRPV2 by PCR using the QuickChange XL sitedirected mutagenesis kit (Agilent Technologies) and were subsequently confirmed by DNA sequencing.

Human embryonic kidney 293T (HEK293T) cells were provided by the Cell Bank of Chinese Academy of Sciences (Shanghai, China). HEK293T cells were maintained in Dulbecco's modified Eagle's medium/high glucose (Gibco) supplemented with 10% fetal bovine serum (FBS) (Gibco) and 100 $\mu g \cdot ml^{-1}$ penicillin-streptomycin (Cellgro) in a humidified incubator at 37 °C (5% CO_2). For electrophysiology experiments, cells were plated in 35 mm tissue culture dishes (Corning Incorporated) were transiently transfected with the plasmids using Lipofectamine 2000 (Life Technologies) following the manufacturer's protocol. Primers used in this study are listed in Supporting Information Table S4.

The human prostate cancer cell lines PC-3M were obtained from American Type Cell Culture (Manassas, Virginia, USA). PC-3M cell line was a metastasis-derived variant of PC3. The cells were maintained in 1640 medium (Gibco) supplemented with 10% FBS (Gibco). PC-3M was cultured in a humidified incubator (5% CO_2) at 37 °C.

Small interfering RNA transfection. PC-3M, the human prostate cancer line, was used in the experiment. These cells were grown in Roswell Park Memorial Institute (RPMI) 1640 Medium (GIBCO, USA) supplemented with 10% FBS in humidified room air containing 5% CO_2 at 37℃. The day before transfection, PC-3M cells were seeded in 6cm cell culture dishes (Corning, USA) at a density of 106 cells/dish. Cells were transfected with siRNA (negative control, positive control, TRPV2) and Lipofectamine RNAi MAX (Invitrogen, USA) according to the manufacturer's protocol. The siRNA

sequence were list: siRNA-TRPV2: 5'-UAAGAGUCAACCUCAACUTT-3', negative control siRNA: 5'-UUCUCCGAACGUGUCACGUTT-3', positive control siRNA (siRNA-β-actin): 5'-UGAAGAUCAAGAUCAUUGCTT-3'. 36 h post−transfection, Cells were used.

Electrophysiology. All electrophysiological recordings were obtained using patch-clamp recordings for macroscopic current dose-response curves. All patch-clamp recordings were conducted 36-96 h after transfection. Patch-clamp recordings were performed with Axopatch-200B amplifier and Axon Digidata 1550A driven by Clampex10 software (Molecular Devices). Membrane potential was held at -60 mV. The microelectrodes fashioned from 1.5 mm thin-walled borosilicate glass with filament were pulled from Flaming/Brown type micropipette puller (P-97; SUTTER INSTRUMENT) with the resistances of 3-7 MΩ for whole-cell recordings. The bath solutions contained 165 mM NaCl, 10 mM HEPES, pH 7.3. The pipette solutions contained 140 mM KCl, 10 mM HEPES, pH 7.3. All recordings were performed at room temperature (approximately 25 °C). Temperature variation was less than 1 °C. The current signals were filtered at 1 kHz and digitized at a 10 kHz sampling frequency. The bath level was kept as low as possible to reduce noise.

To apply 2-APB or other drugs during patch-clamp recording, a rapid solution changer with a gravity-driven perfusion system was used (RSC-200, Bio-Logic). Each solution was delivered through a separate tube so there was no mixing of solutions. Pipette tip with a membrane patch or a whole-cell patch was placed right in front of the perfusion outlet during recording to ensure the solution exchange was complete.

Transwell migration assay. 24-well transwell chamber with 8.0 μm pore polycarbonate membrane inserted (Corning) was used for migration assay. For PC-3M basal migration assessing, cells (8×10^4) were resuspended in 200 μl of serum-free 1640 medium with indicated concentrations of SET2 and seeded on the upper chamber and 500 μl 1640 medium with 10% FBS were added in the lower chamber. For LPA induced PC-3M migration assay, cells (4×10^4) were resuspended in 200 μl of serum-free 1640 medium with indicated concentrations of SET2 or AM966 and seeded on the upper chamber and 500 μl 1640 medium with 0.1 μM LPA and 10% FBS were added in the lower chamber. After incubating at 37 °C with 5% CO_2 for 24 h, the upper surface cells were scraped off with swabs and the invaded cells were fixed by preheated 4% paraformaldehyde for 30 min. Then, cells were stained with 0.2% crystal violet for 20 min and photographed by inverted phase contrast microscope (Olympus, DP71, Japan). The results were presented as counted cells per field at 160× magnification.

Cell viability assay. PC-3M cells were seeded into 96-well microplates at a density of 8000 cells per well and cultured overnight. Then, cells were treated with indicated concentrations of SET2 for 48 h. Cell viability was determined by sulpharhodamine B (SRB) assay. Cells were fixed with 10% trichloroacetic acid for 1 h at 4°C, washed five times with flowing water, and then air-dried. Cells were stained with 50 μl 0.4% (weight/volume) SRB for 20 min at room temperature, washed five times with 1% acetic acid, and then air-dried. Then 100 μl 10 mM Tris was added per well, and absorbance was measured at 515 nm.

Western blotting. PC-3M cells cultured in 60 cm dish were lysed in 100 μL of CelLytic™ M solution (Sigma, USA) for 30 min at 4 °C and the lysates were centrifuged (12000 rpm, 10 min, 4 °C). The Supernatant was diluted in 6X SDS loading buffer, resolved by SDS-polyacrylamide gel electrophoresis, and transferred to polyvinylidene fluoride membranes. Blocking was performed using 5% non-fat dry milk in Tris-buffered saline containing 0.1% Tween 20 for 2 h at room temperature. After having been incubated overnight at 4 °C with rabbit anti-TRPV2 primary antibody (1:500, Thermofisher, USA), rabbit anti-LPAR1 primary antibody (1:200, Abnova, Taiwan, China) or rabbit anti-β-actin primary antibody (1:3000, CST, USA), the membranes were incubated with horseradish peroxidase-conjugated anti-rabbit secondary antibody (1:5000, CST, USA) for 1 h at room temperature. Proteins were detected using enhanced chemiluminescent kit (Thermofisher, USA).

Reverse transcription PCR. Total RNAs were isolated from PC-3M cells followed by PCR using One-Step gDNA Removal and cDNA Synthesis SuperMix (TransGen Biotech, China) according to the instruction from the manufacturer. The cDNAs were amplified with primer sets for TRPV2, LPAR1 and β-actin: TRPV2 forward, 5'-AAAGGGAACAGGTGCCAGTCA-3'; TRPV2 reverse, 5'-TCCCACTGCTTGGTGCAAGCG-3'; LPAR1 forward, 5'-TAGTGGTGGTCATTGTGGTCA-3'; LPAR1 reverse, 5'-GGCAGAGTTGAATTCAGCAAG-3'; β-actin forward, 5'-CAGAGCAAGAGAGGCATCCT-3'; β-actin reverse, 5'-

GTTGAAGGTCTCAAACATGATC-3'. PCR products were visualized using Gel Image System (Tanon, China).

Ca^{2+} measurements using Fluo-4 acetoxymethyl ester. Before fluorescence measurements, PC-3M cells were cultured in 96-well plates for 24 h. The cells were incubated with 2 μM Fluo-4 AM diluted in HBSS solution (0.4 $g{\cdot}L^{-1}$ KCl, 0.12 $g{\cdot}L^{-1}$ $Na_2HPO_4{\cdot}12H_2O$, 0.06 $g{\cdot}L^{-1}$ KH_2PO_4, 0.35 $g{\cdot}L^{-1}$ $NaHCO_3$, 0.14 $g{\cdot}L^{-1}$ $CaCl_2$, 0.10 $g{\cdot}L^{-1}$ $MgCl_2{\cdot}6H_2O$, 0.05 $g{\cdot}L^{-1}$ $MgSO_4$ and 8.0 $g{\cdot}L^{-1}$ NaCl) at 37 ℃ for 45 min. After dye loading, the cells were treated with the compounds of interest. The relative cytoplasmic calcium increase (relative fluorescence unit) was measured using Flexstation 3 (Molecular Device, USA) with fluorescence excition made at 485 nm and emission at 525 nm.

Pocket prediction and virtual screening. The protein gene name from UniProt database (www.uniprot.org) was used to collect relevant structures of TRPV2 channels from the RCSB Protein Data Bank (www.rcsb.org). Based on these structures, Modeller[38] was used to create homology models of mouse TRPV2 channels. Models with the lowest root mean square deviation from their template structures were selected for further analysis. Fpocket 2.0[19] was used to detect potential binding pockets for these selected models with all parameters set at recommended default. Pockets with pocket score of more than 10 were predicted as ligand-binding pockets. Then, we used Clustal Omega program[39] to align protein sequences of the mouse TRPV subfamily (TRPV1-TRPV6) and produced a multiple sequence alignment file. Jensen-Shannon divergence score[18] was

used to assess sequence conservation of the multiple sequence alignment. The conservation scores for all columns of multiple sequence alignment were computed. In addition, the unique residues of TRPV2 was also identified as hotspots binding to subtype-selective compounds. The top 30% ranked conserved residues and unique residues of TRPV2 were mapped onto the predicted ligand-binding pockets. We selected three pockets containing the largest number of conserved residues and at least two unique residues for virtual screening. We used schrodinger Glide software[20] to perform structure-based virtual screening against these selected pockets. For each pocket, approximately 200,000 commercially available compounds from the Specs database (www.specs.net) were screened in the first run using the high-throughput virtual screening module HVS in Glide. Glide score was used to rank the docking result and the top-ranked 20,000 candidates were selected for each pocket. Since the high-throughput virtual screening module applies a simplified scoring function, these selected candidates were screened again against each pocket using a standard docking module SP in Glide. The top-ranked 10,000 candidates were rescored by the consensus score module CSCORE of SYBYL 6.8 (Trips Inc.). We selected compounds by two criteria: (1) compounds having consensus scores of four or five; (2) compounds ranked in top 10% by at least four scoring functions. Using the Cluster Modules module in Pipeline Pilot 7.5 (Scitegic, Inc.), these selected compounds were clustered based on their two-dimensional structures. To ensure the chemical-structural diversity, two to three compounds with good drug-like properties (molecular weight is less than 500, log P is less than 5 and polar surface area is less than 140 $Å^2$) were selected from each cluster. Fifty compounds were selected for bioassay.

Molecular docking. Modeller[38] was used to create homology models of mouse TRPV2 channel based on the cryo-EM structure of the apo rabbit TRPV2 (PDB code: 5AN8). The model with the lowest root mean square deviation (RMSD) from the template was applied to ligand docking using RossetaLigand and RosettaMembrane applications as previously described.[28] Top 10 models with the lowest binding energy (*interface_delta_X*) were identified as the candidates. In the same way mentioned above, the docking models of SET2 and mutant TRPV1 S513F/T551L were generated. To quantitatively analyze the docking results, binding energy was decomposed as mainly van der Waals (VDW) energy.[28] The VDW energies were further mapped on a per residue basis to the channel by Rosetta's residue_energy_breakdown utility. Average values of VDW energy were calculated based on the top 10 ranked models with the lowest binding energy.

Molecular dynamics simulation. To build a TRPV2/SET2 simulation system, we docked four SET2 molecules into four equivalent transmembrane pockets of TRPV2 channel (one for each), and then placed this TRPV2/SET2 complex into a 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine lipid bilayer. The docking model was generated using the Rosetta protocol described above. In the molecular dynamics simulation system, the docking model and lipid bilayer were solvated in a periodic boundary condition box (16.0 nm × 16.0 nm × 16.6 nm) filled with TIP3P water molecules[40] and 0.15 M NaCl. Based on the CHARMM all-atom force field.[41, 42], two molecular dynamics simulations were performed using GROMACS version 4.52.[43, 44] After 100-ns equilibration, a 500-ns production run was carried out for each simulation. All

productions were performed in the NPT ensemble at a temperature of 303.15 K and a pressure of 1 atm. Temperature and pressure were controlled using the velocity-rescale thermostat[45] and the Parrinello-Rahman barostat with isotropic coupling,[46] respectively. Equations of motion were integrated with a 2 fs time step, and the LINCS algorithm was used to constrain bonds lengths.[47] Nonbonded pair lists were generated every 10 steps using a distance cutoff of 1.4 nm. A cutoff of 1.2 nm was used for Lennard-Jones (excluding scaled 1-4) interactions, which were smoothly switched off between 1 and 1.2 nm. Electrostatic interactions were computed using the Particle-Mesh-Ewald algorithm with a real-space cutoff of 1.2 nm. Analysis of the simulations was performed using GROMACS. The last 200-ns trajectory of each simulation was used for analysis.

Data analysis and statistics. Whole-cell voltage clamp experimental data of the inhibition of TRP channels by SET2 molecule were fitted using dose-response function. The half inhibition concentrations of SET2 were derived from fits of the dose–response curves to the function:

$$\frac{I}{I_0} = A_1 + \frac{A_2 - A_1}{1 + 10^{(logx_0 - x)P}} \quad (1)$$

Where I_0 and I are current amplitudes before and after application of inhibitors; A_1 and A_2 are constants between 0 and 1; x is the concentration of inhibitor; x_o is the concentration when 50% inhibition was reached (IC_{50}), P is the Hill constant.

In statistical analyses, the experimental data were expressed as the mean ± SEM. The averaged IC_{50} was collected from the fitting results of each patch (lasting for five inhibitor concentrations), and then was used to compare between different groups. Unless specified, an unpaired *t*-test was used when two means are compared; a one-way analysis

of variance (ANOVA) was used when more than two means are compared. $P < 0.05$ was considered statistically significant.

ASSOCIATED CONTENT

Supporting Information.

Additional figures showing the effects of different antagonists and agonists on the different channels and receptors, figures showing molecular basis of the antagonist binding, a figure showing cellular expression of TRPV2 and LPAR1, a table showing information of ligand-binding sites, a table showing the effects of 2-APB on different channels, a table showing helix orientations, a table listing primers used in the study, and a table listing Molecular Formula Strings of compounds. This material is available free of charge via the Internet at http://pubs.acs.org.

Accession Codes

PDB code for TRPV2 channel model with bound SET2 is 5AN8.

PDB code for mutant TRPV1 channel (S513F/T551L) model with bound SET2 is 5AN8.

Authors will release the atomic coordinates upon article publication.

AUTHOR INFORMATION

Corresponding Authors

Huaiyu Yang, E-mail: hyyang@bio.ecnu.edu.cn, Phone: (+86)-21-24206524

Yang Li, E-mail: liyang@simm.ac.cn, Phone: (+86)-21-50803107

Wei Huang, E-mail: huangwei@simm.ac.cn, Phone: (+86)-21-20231000-2517

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. ‡These authors contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENT

The research was supported in part by National Science & Technology Major Project "Key New Drug Creation and Manufacturing Program" of China (2018ZX09711002), the National Natural Science Foundation of China (21422208, 31671049, and 31371066), the E-Institutes of Shanghai Municipal Education Commission (E09013) and the Special Program for Applied Research on Super Computation of the NSFC-Guangdong Joint Fund (the second phase) under Grant No. U1501501, and the State Key Laboratory of Bioorganic and Natural Products Chemistry. We also thank the computer center of East China Normal University for computational resources.

ABBREVIATIONS USED

2-APB, 2-aminoethoxydiphenylborate; ARD, N-terminal ankyrin repeat domain; cDNA, complementary DNA; DMSO, dimethyl sulfoxide; FBS, fetal bovine serum; HEK293T, human embryonic kidney 293T; IC_{50}, half-maximal inhibition concentration; LPA, lysophosphatidic acid; LPAR1, lysophosphatidic acid receptor 1; PDB, protein data bank; REU, Rosetta energy units; RMSD, root mean square deviation; RT-PCR, reverse transcription polymerase chain reaction; TRP, transient receptor potential channel; TRPV, transient receptor potential vanilloid channel; TRPV2, transient receptor potential vanilloid type 2 channel; SEM, standard error of mean; siRNA, small-interfering RNA; SRB, sulpharhodamine B, VDW, van der Waals; WT, wild-type.

REFERENCES

1. Wagner, J. R.; Lee, C. T.; Durrant, J. D.; Malmstrom, R. D.; Feher, V. A.; Amaro, R. E. Emerging Computational Methods for the Rational Discovery of Allosteric Drugs. *Chem Rev* **2016,** 116, 6370-6390.
2. Jorgensen, W. L. The Many Roles of Computation in Drug Discovery. *Science* **2004,** 303, 1813-1818.
3. Bandell, M.; Macpherson, L. J.; Patapoutian, A. From Chills to Chilis: Mechanisms for Thermosensation and Chemesthesis Via Thermotrps. *Curr Opin Neurobiol* **2007,** 17, 490-497.
4. Peralvarez-Marin, A.; Donate-Macian, P.; Gaudet, R. What Do We Know About the Transient Receptor Potential Vanilloid 2 (Trpv2) Ion Channel? *FEBS J* **2013,** 280, 5471-5487.

5. Monet, M.; Lehen'kyi, V.; Gackiere, F.; Firlej, V.; Vandenberghe, M.; Roudbaraki, M.; Gkika, D.; Pourtier, A.; Bidaux, G.; Slomianny, C.; Delcourt, P.; Rassendren, F.; Bergerat, J. P.; Ceraline, J.; Cabon, F.; Humez, S.; Prevarskaya, N. Role of Cationic Channel Trpv2 in Promoting Prostate Cancer Migration and Progression to Androgen Resistance. *Cancer Res* **2010,** 70, 1225-1235.
6. Monet, M.; Gkika, D.; Lehen'kyi, V.; Pourtier, A.; Vanden Abeele, F.; Bidaux, G.; Juvin, V.; Rassendren, F.; Humez, S.; Prevarsakaya, N. Lysophospholipids Stimulate Prostate Cancer Cell Migration Via Trpv2 Channel Activation. *Biochim Biophys Acta* **2009,** 1793, 528-539.
7. Moolenaar, W. H.; Perrakis, A. Insights into Autotaxin: How to Produce and Present a Lipid Mediator. *Nat Rev Mol Cell Biol* **2011,** 12, 674-679.
8. Shiozaki, A.; Kudou, M.; Ichikawa, D.; Fujiwara, H.; Shimizu, H.; Ishimoto, T.; Arita, T.; Kosuga, T.; Konishi, H.; Komatsu, S.; Okamoto, K.; Marunaka, Y.; Otsuji, E. Esophageal Cancer Stem Cells Are Suppressed by Tranilast, a Trpv2 Channel Inhibitor. *J Gastroenterol* **2017,** 2, 197-207.
9. Zubcevic, L.; Le, S.; Yang, H.; Lee, S. Y. Conformational Plasticity in the Selectivity Filter of the Trpv2 Ion Channel. *Nat Struct Mol Biol* **2018,** 25, 405-415.
10. Huynh, K. W.; Cohen, M. R.; Jiang, J.; Samanta, A.; Lodowski, D. T.; Zhou, Z. H.; Moiseenkova-Bell, V. Y. Structure of the Full-Length Trpv2 Channel by Cryo-Em. *Nat Commun* **2016,** 7, 11130.
11. Zubcevic, L.; Herzik, M. A., Jr.; Chung, B. C.; Liu, Z.; Lander, G. C.; Lee, S. Y. Cryo-Electron Microscopy Structure of the Trpv2 Ion Channel. *Nat Struct Mol Biol* **2016,** 23, 180-186.

12. Gao, Y.; Cao, E.; Julius, D.; Cheng, Y. Trpv1 Structures in Nanodiscs Reveal Mechanisms of Ligand and Lipid Action. *Nature* **2016,** 534, 347-351.
13. Barad, B. A.; Echols, N.; Wang, R. Y.; Cheng, Y.; DiMaio, F.; Adams, P. D.; Fraser, J. S. Emringer: Side Chain-Directed Model and Map Validation for 3d Cryo-Electron Microscopy. *Nat Methods* **2015,** 12, 943-946.
14. Liao, M.; Cao, E.; Julius, D.; Cheng, Y. Structure of the Trpv1 Ion Channel Determined by Electron Cryo-Microscopy. *Nature* **2013,** 504, 107-112.
15. Cao, E.; Liao, M.; Cheng, Y.; Julius, D. Trpv1 Structures in Distinct Conformations Reveal Activation Mechanisms. *Nature* **2013,** 504, 113-118.
16. Singh, A. K.; McGoldrick, L. L.; Sobolevsky, A. I. Structure and Gating Mechanism of the Transient Receptor Potential Channel Trpv3. *Nat Struct Mol Biol* **2018,** 25, 805-813.
17. Deng, Z.; Paknejad, N.; Maksaev, G.; Sala-Rabanal, M.; Nichols, C. G.; Hite, R. K.; Yuan, P. Cryo-Em and X-Ray Structures of Trpv4 Reveal Insight into Ion Permeation and Gating Mechanisms. *Nat Struct Mol Biol* **2018,** 25, 252-260.
18. Capra, J. A.; Singh, M. Predicting Functionally Important Residues from Sequence Conservation. *Bioinformatics* **2007,** 23, 1875-1882.
19. Le Guilloux, V.; Schmidtke, P.; Tuffery, P. Fpocket: An Open Source Platform for Ligand Pocket Detection. *BMC Bioinformatics* **2009,** 10, 168.
20. Friesner, R. A.; Banks, J. L.; Murphy, R. B.; Halgren, T. A.; Klicic, J. J.; Mainz, D. T.; Repasky, M. P.; Knoll, E. H.; Shelley, M.; Perry, J. K.; Shaw, D. E.; Francis, P.; Shenkin, P. S. Glide: A New Approach for Rapid, Accurate Docking and Scoring. 1. Method and Assessment of Docking Accuracy. *J Med Chem* **2004,** 47, 1739-1749.

21. Lemmon, G.; Meiler, J. Rosetta Ligand Docking with Flexible Xml Protocols. *Methods Mol Biol* **2012,** 819, 143-155.

22. Mills, G. B.; Moolenaar, W. H. The Emerging Role of Lysophosphatidic Acid in Cancer. *Nat Rev Cancer* **2003,** 3, 582-591.

23. Pan, H. L.; Zhang, Y. Q.; Zhao, Z. Q. Involvement of Lysophosphatidic Acid in Bone Cancer Pain by Potentiation of Trpv1 Via Pkcepsilon Pathway in Dorsal Root Ganglion Neurons. *Mol Pain* **2010,** 6, 85.

24. Nieto-Posadas, A.; Picazo-Juarez, G.; Llorente, I.; Jara-Oseguera, A.; Morales-Lazaro, S.; Escalante-Alcalde, D.; Islas, L. D.; Rosenbaum, T. Lysophosphatidic Acid Directly Activates Trpv1 through a C-Terminal Binding Site. *Nat Chem Biol* **2011,** 8, 78-85.

25. Hao, F.; Tan, M.; Xu, X.; Han, J.; Miller, D. D.; Tigyi, G.; Cui, M. Z. Lysophosphatidic Acid Induces Prostate Cancer Pc3 Cell Migration Via Activation of Lpa(1), P42 and P38alpha. *Biochim Biophys Acta* **2007,** 1771, 883-892.

26. Shin, S. H.; Kwon, Y. W.; Heo, S. C.; Jeong, G. O.; Kim, B. R.; Seo, E. J.; Kim, J. H. Kruppel-Like Factor 4 Mediates Lysophosphatidic Acid-Stimulated Migration and Proliferation of Pc3m Prostate Cancer Cells. *Exp Mol Med* **2014,** 46, e104.

27. Salazar, H.; Jara-Oseguera, A.; Hernandez-Garcia, E.; Llorente, I.; Arias, O., II; Soriano-Garcia, M.; Islas, L. D.; Rosenbaum, T. Structural Determinants of Gating in the Trpv1 Channel. *Nat Struct Mol Biol* **2009,** 16, 704-710.

28. Yang, F.; Xiao, X.; Cheng, W.; Yang, W.; Yu, P.; Song, Z.; Yarov-Yarovoy, V.; Zheng, J. Structural Mechanism Underlying Capsaicin Binding and Activation of the Trpv1 Ion Channel. *Nat Chem Biol* **2015,** 11, 518-524.

29. Zheng, W.; Qin, F. A Combined Coarse-Grained and All-Atom Simulation of Trpv1 Channel Gating and Heat Activation. *J Gen Physiol* **2015,** 145, 443-456.

30. Liberati, S.; Morelli, M. B.; Amantini, C.; Farfariello, V.; Santoni, M.; Conti, A.; Nabissi, M.; Cascinu, S.; Santoni, G. Loss of Trpv2 Homeostatic Control of Cell Proliferation Drives Tumor Progression. *Cells* **2014,** 3, 112-128.

31. Nakanaga, K.; Hama, K.; Aoki, J. Autotaxin--an Lpa Producing Enzyme with Diverse Functions. *J Biochem* **2010,** 148, 13-24.

32. Houben, A. J.; Moolenaar, W. H. Autotaxin and Lpa Receptor Signaling in Cancer. *Cancer Metastasis Rev* **2011,** 30, 557-565.

33. Wang, J.; Sun, Y.; Qu, J.; Yan, Y.; Yang, Y.; Cai, H. Roles of Lpa Receptor Signaling in Breast Cancer. *Expert Rev Mol Diagn* **2016,** 16, 1103-1111.

34. Lopane, C.; Agosti, P.; Gigante, I.; Sabba, C.; Mazzocca, A. Implications of the Lysophosphatidic Acid Signaling Axis in Liver Cancer. *Biochim Biophys Acta* **2017,** 1868, 277-282.

35. Ray, U.; Roy, S. S.; Chowdhury, S. R. Lysophosphatidic Acid Promotes Epithelial to Mesenchymal Transition in Ovarian Cancer Cells by Repressing Sirt1. *Cell Physiol Biochem* **2017,** 41, 795-805.

36. Van Leeuwen, F. N.; Olivo, C.; Grivell, S.; Giepmans, B. N.; Collard, J. G.; Moolenaar, W. H. Rac Activation by Lysophosphatidic Acid Lpa1 Receptors through the Guanine Nucleotide Exchange Factor Tiam1. *J Biol Chem* **2003,** 278, 400-406.

37. Bian, D.; Mahanivong, C.; Yu, J.; Frisch, S. M.; Pan, Z. K.; Ye, R. D.; Huang, S. The G12/13-Rhoa Signaling Pathway Contributes to Efficient Lysophosphatidic Acid-Stimulated Cell Migration. *Oncogene* **2006,** 25, 2234-2244.

38. Sali, A.; Blundell, T. L. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *J Mol Biol* **1993,** 234, 779-815.

39. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Soding, J.; Thompson, J. D.; Higgins, D. G. Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega. *Mol Syst Biol* **2011,** 7, 539.

40. Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J of Chem Phys* **1983,** 79, 926-935.

41. Guvench, O.; Mallajosyula, S. S.; Raman, E. P.; Hatcher, E.; Vanommeslaeghe, K.; Foster, T. J.; Jamison, F. W., 2nd; Mackerell, A. D., Jr. Charmm Additive All-Atom Force Field for Carbohydrate Derivatives and Its Utility in Polysaccharide and Carbohydrate-Protein Modeling. *J Chem Theory Comput* **2011,** 7, 3162-3180.

42. MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J Phys Chem B* **1998,** 102, 3586-3616.

43. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. Gromacs 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J Chem Theory Comput* **2008,** 4, 435-447.

44. Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. Gromacs: Fast, Flexible, and Free. *J Comput Chem* **2005,** 26, 1701-1718.

45. Bussi, G.; Donadio, D.; Parrinello, M. Canonical Sampling through Velocity Rescaling. *J Chem Phys* **2007,** 126, 014101.

46. Aoki, K. M.; Yonezawa, F. Constant-Pressure Molecular-Dynamics Simulations of the Crystal-Smectic Transition in Systems of Soft Parallel Spherocylinders. *Phys Rev A* **1992,** 46, 6541-6549.

47. Hess, B. P-Lincs: A Parallel Linear Constraint Solver for Molecular Simulation. *J Chem Theory Comput* **2008,** 4, 116-122.

TABLE OF CONTENTS GRAPHIC