Published online 8 November 2021 *Nucleic Acids Research, 2022, Vol. 50, Database issue* ***D1417–D1431***
https://doi.org/10.1093/nar/gkab1013

VARIDT 2.0: structural variability of drug transporter

Tingting Fu[1,2,†], Fengcheng Li[1,†], Yang Zhang[3,†], Jiayi Yin[1,†], Wenqi Qiu[4], Xuedong Li[3], Xingang Liu[3], Wenwen Xin[3], Chengzhao Wang[3], Lushan Yu[1], Jianqing Gao[1,5], Qingchuan Zheng[2,*], Su Zeng[1,*] and Feng Zhu[1,6,*]

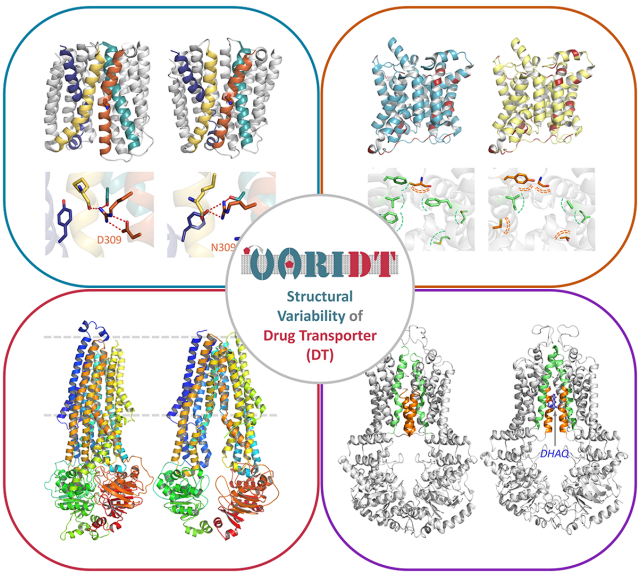
[1]College of Pharmaceutical Sciences, Zhejiang University, Hangzhou 310058, China, [2]Institute of Theoretical Chemistry, College of Chemistry, Jilin University, Changchun 130023, China, [3]Department of Pharmacology, Hebei Medical University, Shijiazhuang 050017, China, [4]Department of Surgery, HKU-SZH & Faculty of Medicine, The University of Hong Kong, Hong Kong, China, [5]Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, Zhejiang, China and [6]Innovation Institute for Artificial Intelligence in Medicine of Zhejiang University, Alibaba-Zhejiang University Joint Research Center of Future Digital Healthcare, Hangzhou 330110, China

Received September 08, 2021; Revised October 08, 2021; Editorial Decision October 11, 2021; Accepted November 04, 2021

ABSTRACT

The structural variability data of drug transporter (DT) are key for research on precision medicine and rational drug use. However, these valuable data are not sufficiently covered by the available databases. In this study, a major update of VARIDT (a database previously constructed to provide DTs' variability data) was thus described. First, the experimentally resolved structures of all DTs reported in the original VARIDT were discovered from PubMed and Protein Data Bank. Second, the structural variability data of each DT were collected by literature review, which included: (a) mutation-induced spatial variations in folded state, (b) difference among DT structures of human and model organisms, (c) outward/inward-facing DT conformations and (d) xenobiotics-driven alterations in the 3D complexes. Third, for those DTs without experimentally resolved structural variabilities, homology modeling was further applied as well-established protocol to enrich such valuable data. As a result, 145 mutation-induced spatial variations of 42 DTs, 1622 inter-species structures originating from 292 DTs, 118 outward/inward-facing conformations belonging to 59 DTs, and 822 xenobiotics-regulated structures in complex with 57 DTs were updated to VARIDT (https://idrblab.org/varidt/ and http://varidt.idrblab.net/). All in all, the newly collected structural variabilities will be indispensable for explaining drug sensitivity/selectivity, bridging preclinical research with clinical trial, revealing the mechanism underlying drug-drug interaction, and so on.

GRAPHICAL ABSTRACT



INTRODUCTION

Drug transporter (DT) is known as the key determinant of drug absorption, distribution, clearance and elimination (1,2), and its variability is reported to be critical for balancing drug efficacy with safety (3), reversing drug resistance (4) and predicting drug-drug interaction (5–8). Although the variability data of DT are essential for preclinical and

*To whom correspondence should be addressed. Tel: +86 189 8946 6518; Fax: +86 571 8820 8444; Email: zhufeng@zju.edu.cn
Correspondence may also be addressed to Su Zeng. Email: zengsu@zju.edu.cn
Correspondence may also be addressed to Qingchuan Zheng. Email: zhengqc@jlu.edu.cn
†The authors wish it to be known that, in their opinion, the first four authors should be regarded as Joint First Authors.

clinical studies (9–11), they are extensively dispersed in literatures (12). The VARIDT 1.0 database (12) was therefore developed to explicitly describe three aspects of DT variability: (i) genetic polymorphism, and epigenetic regulation of DT; (ii) species/tissue/disease-specific DT abundance and (iii) exogenous factors modulating DT activity or altering the disposition of transported drug. Meanwhile, VARIDT 1.0 also enabled the interplay analysis among multiple aspects of DT variability (12). Due to its unique data coverage, VARIDT has emerged to be the indispensable complements to other transporter-related databases (13–31) in assessing drug safety (32–35), reversing drug resistance (36) and so on (37–40).

Recently, there are increasing demands on the 3D structural variability data of DT, which are expected to provide key information for the studies on precision medicine and rational drug use (41). Such data include: (**a**) the mutation-induced spatial variations in folded state that are critical for the understanding of drug sensitivity and selectivity (42), (**b**) the difference among the DT structures of human and model organisms that are key for bridging preclinical study with clinical trials (43), (**c**) the outward-facing and inward-facing conformations that are essential for elucidating the dynamics and underlying steps of transporting cycle (44) and (**d**) the xenobiotics-driven alterations in 3D complexes of DT that are crucial for revealing the mechanism underlying drug-drug interaction (45–49). Furthermore, due to the lack of DTs' structures and their structural variability data, homology modeling is adopted as well-established protocol to construct protein structure (50), and the resulting structure together with its variability have been successfully and widely adopted to study protein-protein interaction (51), discover novel drug and epitope (52,53), promote protein engineering and design (54), and so on (55–57).

Till now, several online resources have been constructed and are still active to provide DT-related structure data (13–16,58–63). Some contain transporter structures as part of the broader collection of biological or pharmacological data (e.g. UniProt (13), TTD (58) and IUPHAR/BPS Guide to Pharmacology (16)); some others (e.g. MemProtMD (59), EncoMPASS (60), and ABCMdb (61)) provide 3D conformation collected from PDB database (14); the remaining (e.g. TCDB (15), and iMusta4SLC (62)) offer the hyperlinks redirecting to PDB entries (14). A preliminary assessment of PDB's data (14) reveals that only a small fraction (∼14%) of all DTs in VARIDT 1.0 (12) have experimentally resolved structures available. To cope with the lack of DT structures, the SWISS-MODEL (63) can thus be adopted to predict DT structure by homology modeling, and AlphaFold (64,65) can facilitate the determination of transporter's structure. However, all the existing online resources do not systematically describe the 3D structural variability data of DTs, which asks for a major update of VARIDT 1.0 to provide the comprehensive information on describing all four aspects (i–iv) of DT structural variabilities (42–45).

In this study, a major update of VARIDT (structural variability of drug transporter) was therefore conducted, which systematically described all four aspects of DT structural variability. As shown in Figure 1, these availability data included: (a) the mutation-induced spatial variations in folded state, (b) the difference among the DT structures of human and model organisms, (c) the outward-facing and inward-facing conformations of the transporting cycle and (d) the xenobiotics-driven alterations in 3D complex. With the significant advances in the field of precision medicine, these data updated to the VARIDT 2.0 version (https://idrblab.org/varidt/ and http://varidt.idrblab.net/) will be indispensable for understanding drug sensitivity/selectivity, bridging preclinical research with clinical trial, elucidating the dynamics of drug transporting cycle, revealing the mechanisms underlying a variety of drug–drug interactions, and so on so forth.

FACTUAL CONTENT AND DATA RETRIEVAL

Collection of structure data for drug transporters

The entire structures (covering the whole protein sequence) of wild-type DTs were first collected by searching the names or synonyms of all DTs in PubMed (66) and PDB (14) using the keyword combination of '[DT name] + structure', '[DT name] + conformation', '[DT name] + [substrate name] + structure', '[DT name] + [substrate name] + complex', '[DT name] + complexed with', '[DT name] + in complex with', and so on. Additionally, such entire structures of wild-type DTs could also be identified by blasting the full-length sequence of DT against all those sequences in PDB (14). Only the PubMed literature or PDB entry that explicitly described the entire structures of the wild-type DTs were collected and recorded in VARIDT 2.0. Second, for those DTs without their entire structures, the structures of their transmembrane domain or region were identified by searching or blasting against PubMed (66) and PDB (14), and the transmembrane structures were collected to represent the structures of the corresponding DTs. As a result, a total of 73 DTs were identified with experimentally resolved structure available. These structures were determined by diverse techniques (such as NMR spectroscopy, X-ray crystallography, and electron microscopy) based on various systems (such as *Escherichia coli* BL21, HEK293 cells, and SF9 cells), and the resolutions of most DT structures were within the range of 1.0 Å to 4.0 Å.

Apart from the structures of wild-type DT, many structure variants of DT were also described in PubMed (66) and PDB (14). Most of the variants (∼160 structures) were in complex with various endogenous ionic or small molecular substrates, and the remaining variants (∼50 structures) bond with different exogenous ligands (such as drugs, environmental chemicals, etc.).

Moreover, due to the lack of DTs' structures and the structural variability data, a well-established protocol of homology modeling (50) was adopted to provide the explicit data on illustrating DTs' structural variability (detailed information on homology modeling was provided in the following section entitled 'Structural variability data generated by homology modeling'). Meanwhile, a variety of DT structures predicted by AlphaFold (64,65), a popular structure prediction tool based on Artificial Intelligence (AI), were collected for referencing DTs' structural variability.

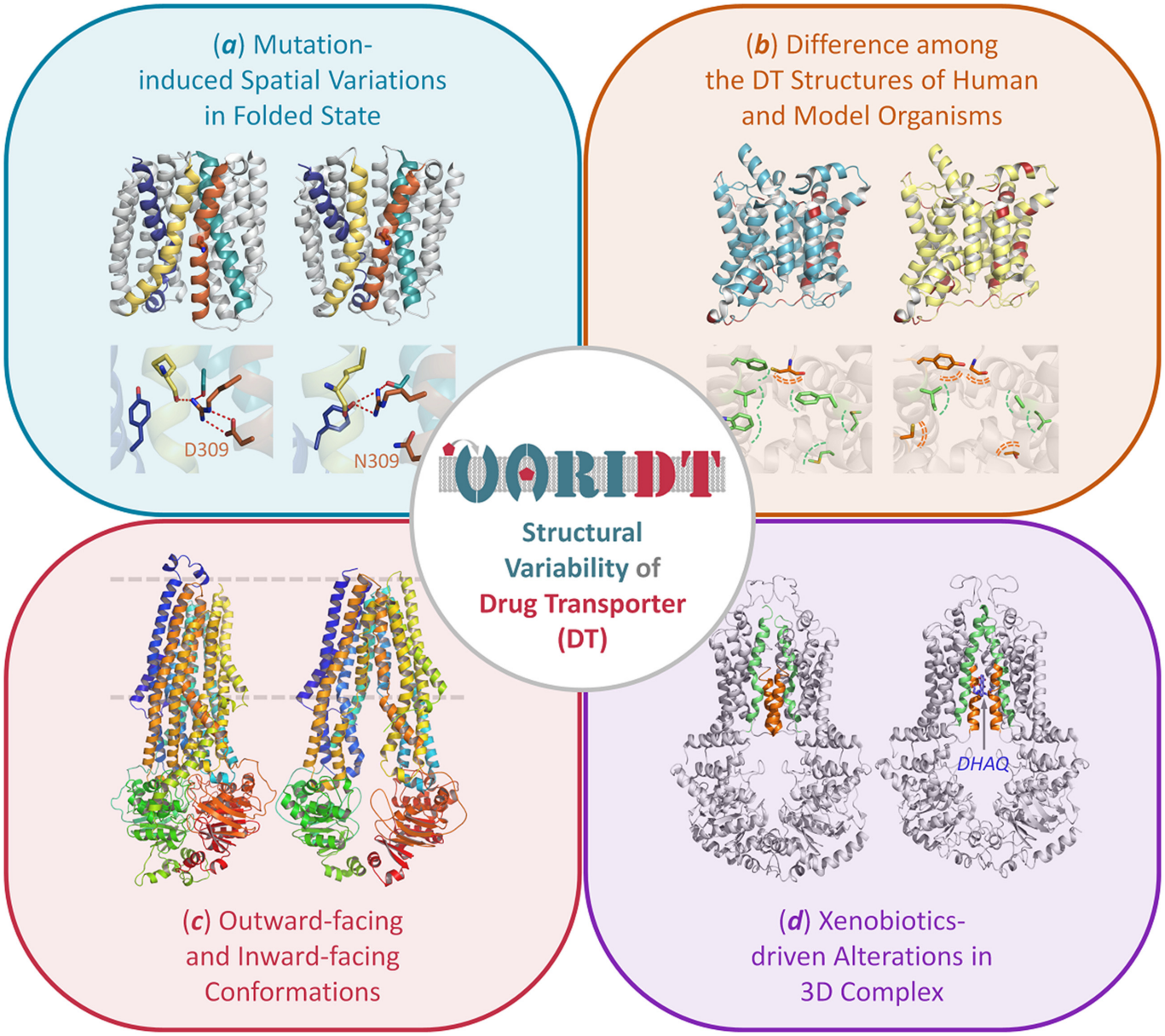


Figure 1. The structural variability data of drug transporter (DT) updated to the VARIDT 2.0. (**a**) the mutation-induced spatial variations in folded state; (**b**) the difference among the DT structures of human and model organisms; (**c**) the outward/inward-facing conformations of the transporting cycle; (**d**) the xenobiotics-driven alterations in 3D complex.

Structural variability data provided in VARIDT 2.0

Mutation-induced spatial variations in folded state. Mutation-induced spatial variations in the folded state of DT are essential for understanding drug sensitivity & selectivity (42,67), because of the resulting variations of interaction pattern, charge environment, hydrophobic property, and volume of each residue in drug binding pocket (68–74). Figure 2 showed the spatial variations in the folded state of human monocarboxylate transporter 1 (hMCT1) induced by a point mutation (D309N), which were explicitly recorded and described in VARIDT 2.0. hMCT1 was reported to drive the transmembrane transportation of several drugs that treated metabolic diseases, such as atorvastatin (75), and was known to be a promising target for immunosuppression (76). As described in the online VARIDT 2.0 and Figure 2A, the carboxyl group of D309 in wild-type hMCT1 electron microscopy structure interacted with the guanidine group of R313 through hydrogen bonds (68). In contrast, the hydrogen bonds between the above residues vanished, as shown in Figure 2B, due to the introduction of N309′s amino group in the mutated (D309N) structure of hMCT1. In other words, the point mutation of D309N released the key R313 residue, which broke its hydrogen bond interaction with the carbonyl oxygen of M151 in transmembrane domain 5 (TM5) and formed a new critical interaction with Y34 in TM1.

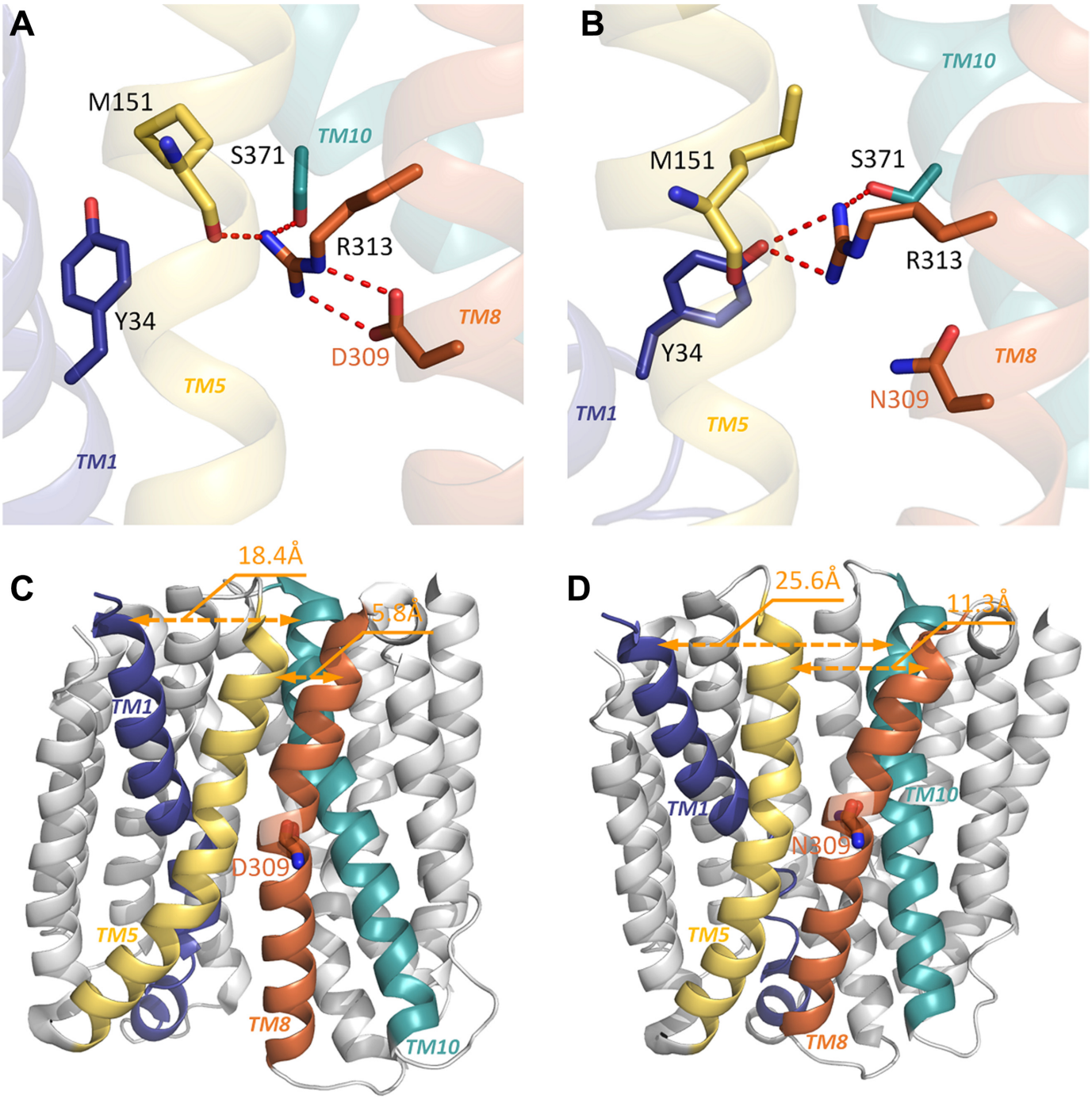


Figure 2. Spatial variations in the folded state of human monocarboxylate transporter 1 (hMCT1) induced by a point mutation (D309N), which were described in VARIDT. (**A**) the carboxyl group of D309 in wild-type hMCT1 electron microscopy structure interacted with the guanidine group of R313 through hydrogen bonds (key residues were shown as sticks, and hydrogen bonds were highlighted as red dashed lines); (**B**) the point mutation of D309N released the key R313 residue, which broke its hydrogen bond interaction with the carbonyl oxygen of M151 in transmembrane domain 5 (TM5) and formed new critical interaction with Y34 in TM1 (key residues were provided as stick, and hydrogen bonds were highlighted as red dashed line); (**C**) the structure of wild-type hMCT1 folded to inward-facing state (D309 was shown in stick, and the distances between key TMs were labeled); (**D**) the structure of hMCT1 mutant folded to outward-facing state (N309 was shown in stick, and the distances between key TMs were labeled).

Moreover, as described in Figure 2C (68), the structure of wild-type hMCT1 folded to an inward-facing state. Because of the variations between interaction patterns in Figure 2A and B, the restrictions between TM8 and TM5 were loosened, while a new constraint between TM8 and TM1 was established, making the mutated structure of hMCT1 an outward-facing conformation (Figure 2D). Particularly, the distance between G19 in TM1 and M379 in TM10 was significantly enlarged from 18.4 Å in wild-type hMCT1 (Figure 2C) to 25.6 Å in the mutated structure (Figure 2D), and the distance between P144 in TM5 and N321 in TM8 was also expanded from 5.8 Å in wild-type hMCT1 to 11.3 Å in the mutated one (Figure 2C and D). All in all, the spatial variations between the folded states of wild-type and mutated DTs that were provided in VARIDT 2.0 could alter the affinity between drug and DT, and thus affect drug sensitivity/selectivity (68,77,78).

Difference among the DT structures of human and model organisms

Various model organisms were adopted in biomedical research with their distinct advantages and limitations, and the information obtained from these model organisms may not be able to reflect human physiology (79–81). The differences among DT structures of human and model organisms were therefore essential for bridging preclinical study with clinical trial (43,82). Figure 3 showed the difference between the crystal structures of human urea transporter 1 (hUT1) and bovine urea transporter 1 (UT1b), which were systematically recorded and described in VARIDT. As reported, the sequence identity between hUT1 and UT1b equaled to 80.99%, which guaranteed a similarly rapid urea diffusion following concentration gradient in both species (83). As shown in VARIDT 2.0 website and Figure 3A and B, although over 40 residues were non-conserved between hUT1 and UT1b (highlighted in red on the ribbon plot), their conformations looked highly similar, and it was reasonable to deduce that their functions in drug transportation might be similar.

However, an in-depth investigation of those non-conserved residues in UT1′s transporting tunnel may tell you otherwise (Figure 3C and D). As shown, compared with the non-conserved residues (C87, A278, I292, T295 and G333) in UT1b's drug transporting tunnel, the corresponding amino acids (W92, F283, M297, M300 and C338) in hUT1 presented larger side chains, which resulted in strong steric hindrance and hindered drug transports. Moreover, three non-polar residues (W92, F136 and M300) in hUT1 were replaced by the polar ones (C87, Y131 and T295) in UT1b, which completely re-defined the physicochemical profiles of the drug transporting tunnel (shifting from a non-polar tunnel primarily colored in green to a polar one mainly colored in orange). All in all, those above structural difference shown in VARIDT were essential for bridging preclinical study with clinical trial and understanding the difference among species in drug transportation.

Outward-facing and inward-facing conformations. The transportation of a drug through DTs was a multi-step process, and the DT structure of high-resolution was the basis for understanding the structure dynamics of this complex process (68,84). Among these structures, the outward- and inward-facing conformations of DTs were reported as essential for elucidating the underlying steps of a transporting cycle (44,84–86). Figure 4 showed the outward-facing and inward-facing conformations of two typical DTs recoded in VARIDT 2.0: human monocarboxylate transporter 1 (hMCT1) and human P-glycoprotein (hMDR1). Both DTs were represented using ribbon plots and cylindrical helices on the left and right side, respectively. As described in Figure 4A, the outward-facing and inward-facing conformations of hMCT1 were viewed from three different perspectives: (i) parallel to the membrane, (ii) from the intra-cellular space and (iii) from the extra-cellular space. Clearly, there were distinct conformation variations between these two critical snaps of a drug transporting process, which could provide the valuable information to reconstruct the dynamic process of drug absorption, distribution or exclusion, and to design novel ligands that interfered with the hMCT1-affiliated signaling pathways.

Similarly, those outward-facing and inward-facing conformations of hMDR1 were illustrated in Figure 4B (viewed parallel to cell membrane). hMDR1 was one of the most famous DTs in drug exclusion and resistance, which was able to exclude both drug and toxic molecule from the cells, thereby protecting tissues from toxic substances (87). Compared with hMDR1′s outward-facing conformation, some transmembrane domains (TMs), in the inward-facing conformation, moved as a rigid body, while the remaining TMs underwent major rearrangement (85,88). All in all, the outward/inward-facing conformations of all DTs shown in VARIDT laid the solid foundation for depicting the mechanisms and dynamic processes underlying drug transportation.

Xenobiotics-driven alterations in 3D complexes. The affinities of DTs in transporting drugs could be extensively modulated by xenobiotics, which, in turn, affected the pharmacokinetics, efficacy and safety of the transported drugs (89). In other words, the xenobiotics-driven alterations in DT-included complexes were essential for revealing mechanism underlying drug-drug interactions (45,90–93). Figure 5 provided xenobiotics-driven alterations in the 3D complexes of two representative DTs recoded in VARIDT 2.0: human breast cancer resistance protein (hABCG2) and human serotonin transporter (hSERT). hABCG2 could exclude a variety of chemotherapeutic drugs from cancer cells, which resulted in a resistance to these anti-cancer drugs (84,94,95). As shown in Figure 5A, two typical conformational states of hABCG2 were explicitly described in VARIDT 2.0: (i) apo-closed conformation on the left side and (ii) inhibitor trapped inward-facing conformation on the right side (84). Particularly, in apo-closed state, TM2 and TM5 helixes form a tightly packed helical bundle near the cytosolic region, sealing off the crevice at the dimer interface. Meanwhile, the xenobiotics mitoxantrone (DHAQ) drove the conformation alterations in TM2 and TM5, which resulted in the shifts of hABCG2 to its inward facing conformation. In other words, DHAQ stabilized hABCG2 in its inward-

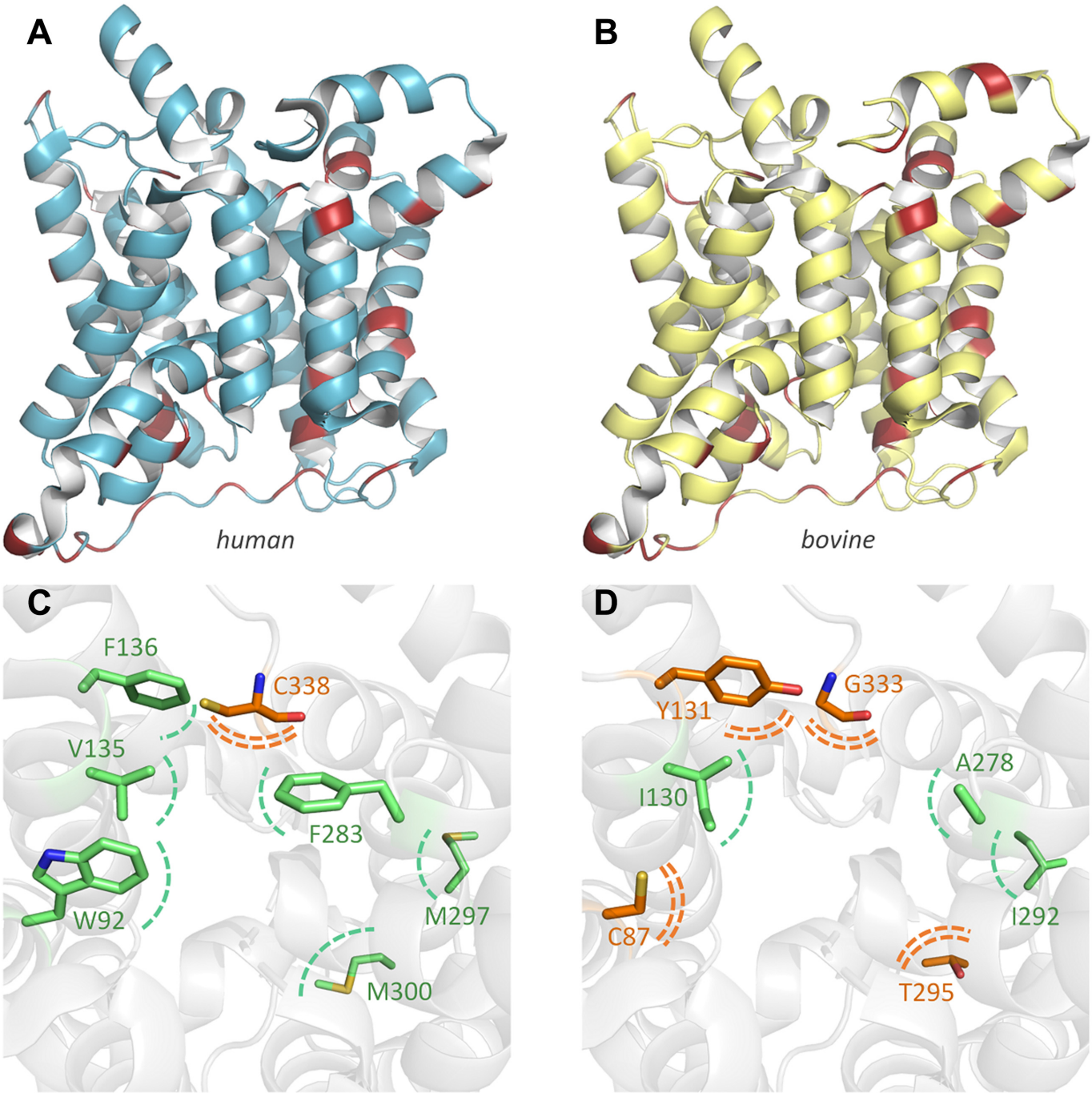


Figure 3. The differences between the crystal structures of *human* urea transporter 1 (hUT1) and *bovine* urea transporter 1 (UT1b), which were provided in VARIDT. Crystal structures (described in ribbon diagrams) of and non-conserved residues (highlighted in red) between hUT1 and UT1b were shown in (**A**) and (**B**), respectively. Drug transporting tunnel (grey ribbon diagram) and key residues (colored in green and orange for the non-polar and polar residues, respectively) of hUT1 and UT1b were illustrated in (**C**) and (**D**), respectively. Key residues were represented in sticks.

facing conformation, and could thus affect not only the transporting cycle but also the drug efficacy.

Moreover, the structural information of DT complexes collected in VARIDT 2.0 varied according to their complexed ligands (xenobiotics). Particularly, hSERT was popular target of the marketed antidepressants, and Figure 5B showed its experimentally resolved structures in complex with two drugs (*paroxetine* & *fluvoxamine*, both were the selective serotonin reuptake inhibitors). The binding of different drugs could induce alteration in the S1 binding pocket (shown in violet). In VARIDT, the alteration data driven by different xenobiotics in the 3D complexes of DTs could thus give key information for revealing the mechanism underlying drug–drug interaction.

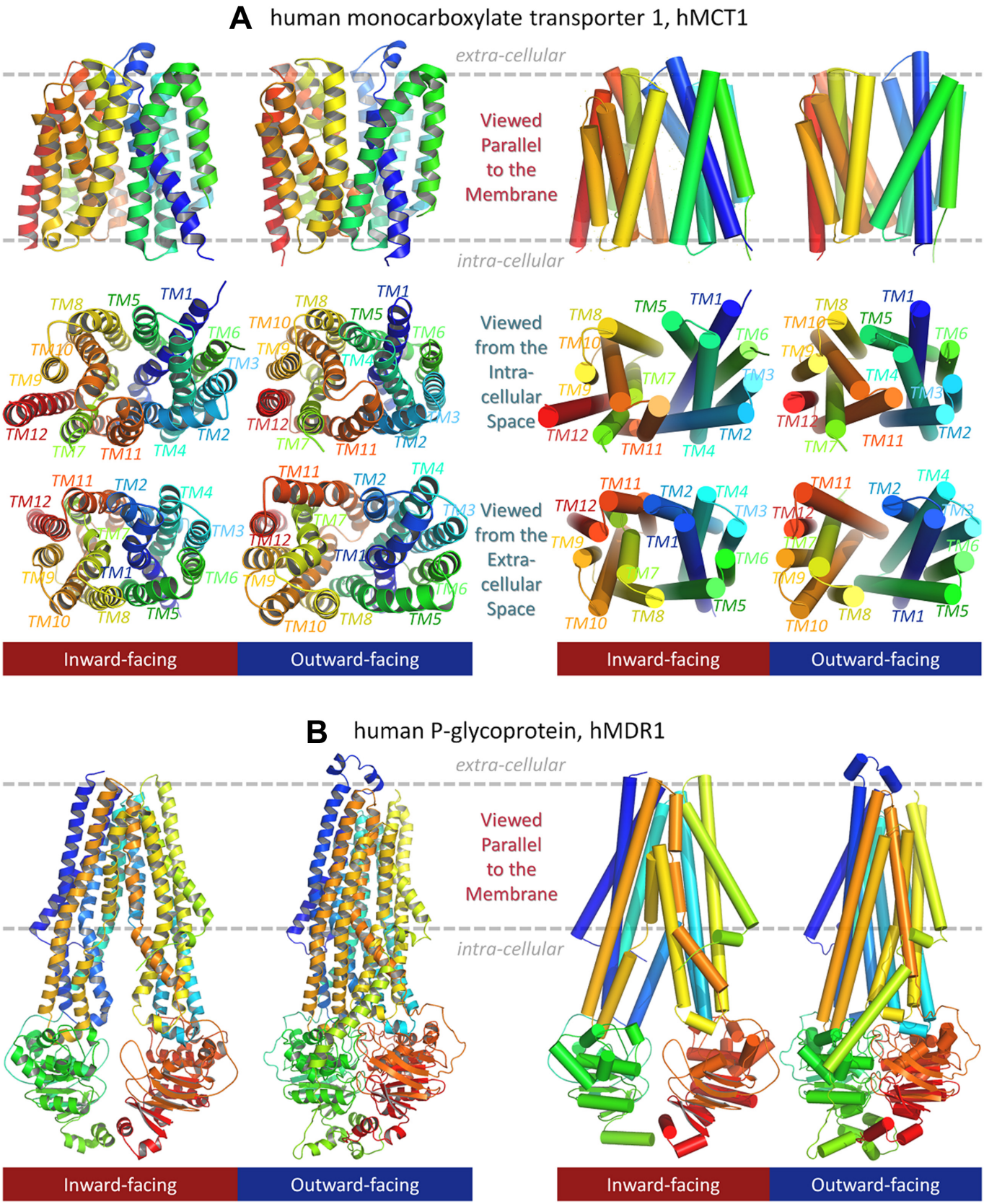


Figure 4. Outward-facing and inward-facing conformations of two DTs recoded in VARIDT 2.0, human monocarboxylate transporter 1 (hMCT1) & human P-glycoprotein (hMDR1). (**A**) hMCT1 conformations were viewed from different perspectives: parallel to the membrane (top), from the intra-cellular space (middle), and from extra-cellular space (bottom); (**B**) hMDR1 conformations were viewed parallel to the membrane. These structural conformations were shown using ribbon (on the left side) and cylindrical helices (on the right side).

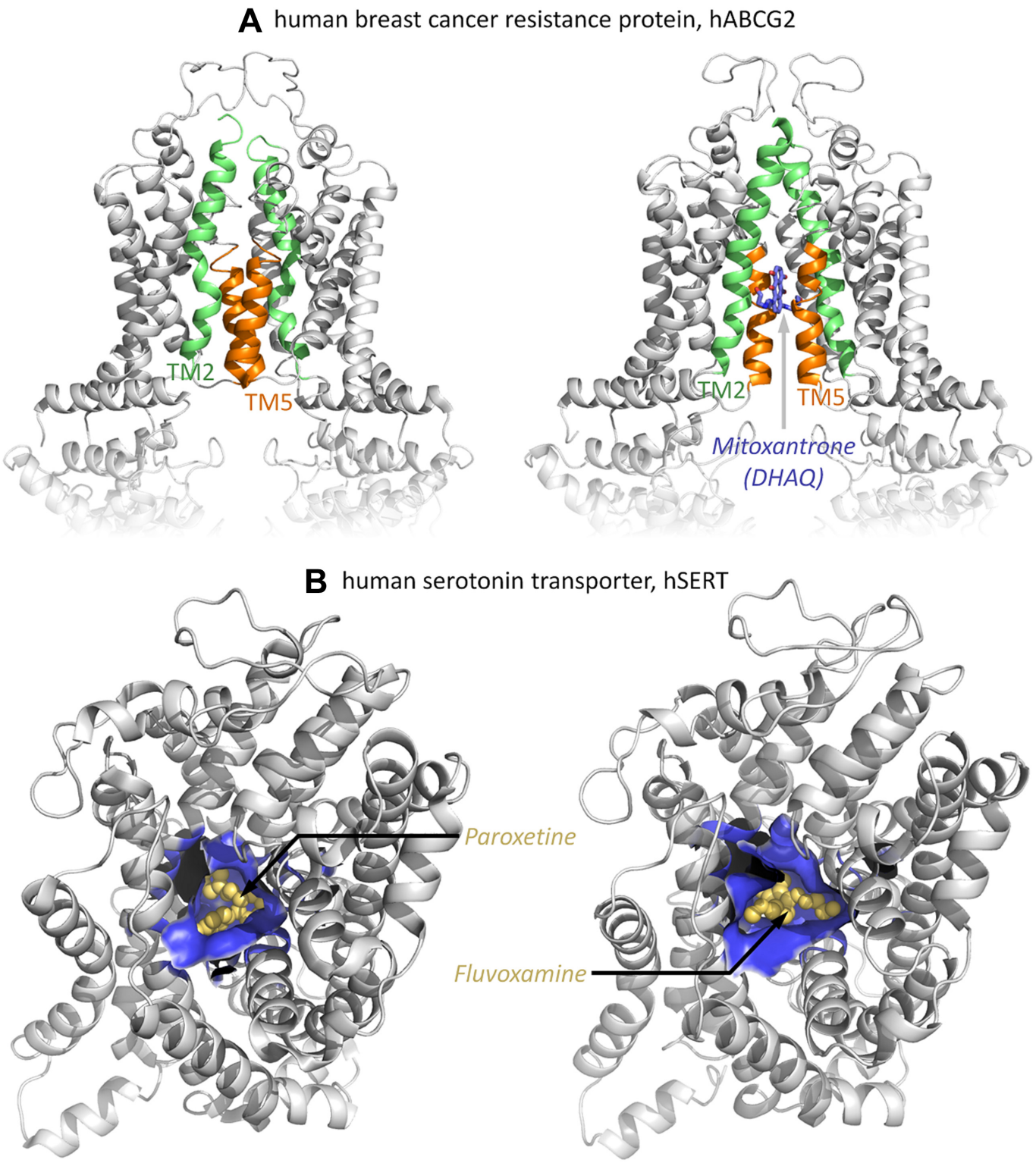


Figure 5. Xenobiotics-driven alterations in the 3D complexes of two representative DTs recoded in VARIDT: human breast cancer resistance protein (hABCG2) and human serotonin transporter (hSERT). (**A**) two typical conformational states of hABCG2 were described in VARIDT 2.0: the apo-closed conformation on the left side & the inhibitor trapped inward-facing conformation by mitoxantrone (DHAQ, colored in violet) on the right side (two key domains TM2 and TM5 were colored in green and orange, respectively); (**B**) the experimental hSERT structures in complexed with two antidepressants (*paroxetine* & *fluvoxamine*) were illustrated, and the S1 site was shown by violet surface and both drugs were represented by yellow spheres.

Structural variability data generated by homology modeling

A preliminary assessment of PDB (14) revealed that only a small fraction of the DTs in VARIDT 1.0 (12) have experimentally resolved structure available. These results were consistent with the previous reports (50,96), which highlighted the challenges in the expression or crystallization of membrane proteins (97). Therefore, homology modeling had emerged as a valuable approach for generating transporter structures (98). So far, it had been applied as well-established protocol to build protein structure (50), and the resulting structures and their variability data had been widely and successfully adopted to study protein-protein interactions (51), discover new drug or epitope (52,53), promote protein engineering or design (54,99), identify thermostabilizing mutations (55), and predict the molecular mechanisms underlying drug transportation (100).

In this study, the well-established protocol *RosettaCM* was adopted to generate structural models and variabilities for each DT (50). As shown in Figure 6, the adopted protocol consisted of four key steps. (S1) identifying suitable structure templates through sequence alignment against PDB (14) & threading DT sequence over single/multiple top-ranked templates; (S2) parsing sequence to 3mer and 9mer fragments to generate fragment file by *Rosetta* server (101); (S3) forming five model candidates for comparatively modeling each DT using *RosettaCM Hybridize* module; (S4) optimizing the generated structures by energy minimization, and evaluating the structure quality by assessing from two different perspectives. To describe the reliability of the generated structure, the resulting energy scores of each computationally modelled structure and its quality evaluation outcomes were fully provided in VARIDT, which were further described in the following section entitled 'Evaluation of the modelled structural variability data'.

For modeling the mutation-induced spatial variations, the nucleotide polymorphism data within the coding region of DTs affecting drugs' transportation were first identified from VARIDT (12). Then, these polymorphisms were mapped to DT sequence for template discovery. For modeling, the species-specific structure differences, DT sequences of various model organisms (such as rat, mouse, zebrafish, bovine and rabbit) and human were first collected. These sequences were *then* used to identify their structural templates. For modeling outward & inward-facing conformations, the special attentions were paid to the conformation orientation of the modelled DTs during their identifications of structural template. For modeling xenobiotics-driven alteration, the xenobiotics capable of altering DT activity were first identified from VARIDT (12). *Then*, the widely applied flexible docking strategy, induced fit docking (102), was employed to generate the complex.

Evaluation of the modelled structural variability data

Two popular criteria were used to evaluate the structure generated by homology modeling, which included the (a) percentage of the modelled residues within the favored region of *Ramachandran* plot (103) and (b) QMEANBrane score specifically designed to measure modeling qualities for membrane protein (104). Particularly, the *Ramachandran* plot described the GLOBAL geometric attributes of the modelled structure by visualizing the dihedral angles of all residues in a structure, and the QMEANBrane score illustrated the LOCAL qualities of an alpha-helical transmembrane protein model by combining statistical potential with a per-residue weighting scheme (103,104). Due to the innate independence between these two criteria, they were collectively considered in VARIDT 2.0 to evaluate those modelled structural variabilities. For criterion (a), the percentage of modelled residues within the favored *Ramachandran* region ($\geq$98%, 95–98% and $<$95%) were accepted in VARIDT 2.0 database as: *Excellent*, *Medium* and *Poor*, respectively (Figure 7A). For criterion (b), the QMEANBrane score of $\geq$0.8, 0.6–0.8 and $<$0.6 was considered in VARIDT 2.0 as *High*, *Medium* and *Poor*, respectively (Figure 7A). When the evaluation result of any criterion was classified to '*Poor*', the corresponding modelled structure would be considered as unreliable. As a result, the evaluation outcome based on both criteria for each computationally modelled DT structure were fully described in VARIDT 2.0 to indicate the level of precision and reliability of the generated structural variability data.

Taking serotonin transporter as an example, three experimentally resolved structures determined by X-ray crystallography or electron microscopy were collected in the first place, which included *human* serotonin transporter (hSERT; PDB ID: 6VRH), *Drosophila* dopamine transporter (dDAT; PDB ID: 4XPT), *Aquifex aeolicus* leucine transporter (LeuT; PDB ID: 3F3A). Second, the crystal structures of dDAT and LeuT were adopted to model two hSERT structures, and the experimental structure of hSERT (6VRH) was utilized to assess the quality of the two modelled outcomes. As provided in Figure 7B, the structure in grey represented the experimentally resolved structure of hSERT (6VRH), and the structures in blue and red indicated those structures modelled using the templates of dDAT (4XPT) and LeuT (3F3A), respectively. Compared with the red structure, the blue one presented the much better consistency with the experimental structure in grey, since its RMSD of the backbone atoms of all helix residues (1.19 Å, blue structure) was much lower than that (3.51 Å) of the red one, and its RMSD of the backbone atoms of all S1 pocket residues (0.84 Å, blue structure) was much lower than that (1.81 Å) of the red. Particularly, only two residues (F341 & V501) in the S1 pocket of blue structure showed a substantial shift from their original positions in the experimentally resolved structure, while six resides (I172, F341, Y95, S336, I168 & F334) in red structure presented major shift from their original positions in the experimentally resolved structure. These variations above in the qualities of modelled structures were consistent with the evaluation outcomes, as the percentage of the residues of modelled blue structure within favored *Ramachandran* regions was much higher than that of the red one (96.3% and 90.8% for blue and red structures in Figure 7B, respectively), and the QMEANBrane score of the blue structure was also much higher than that of the red one (0.82 and 0.75 for blue and red structures in Figure 7B, respectively). All in all, this example provided clear illustration on the effectiveness of those two independent criteria in evaluating the computationally modelled structural variability data.

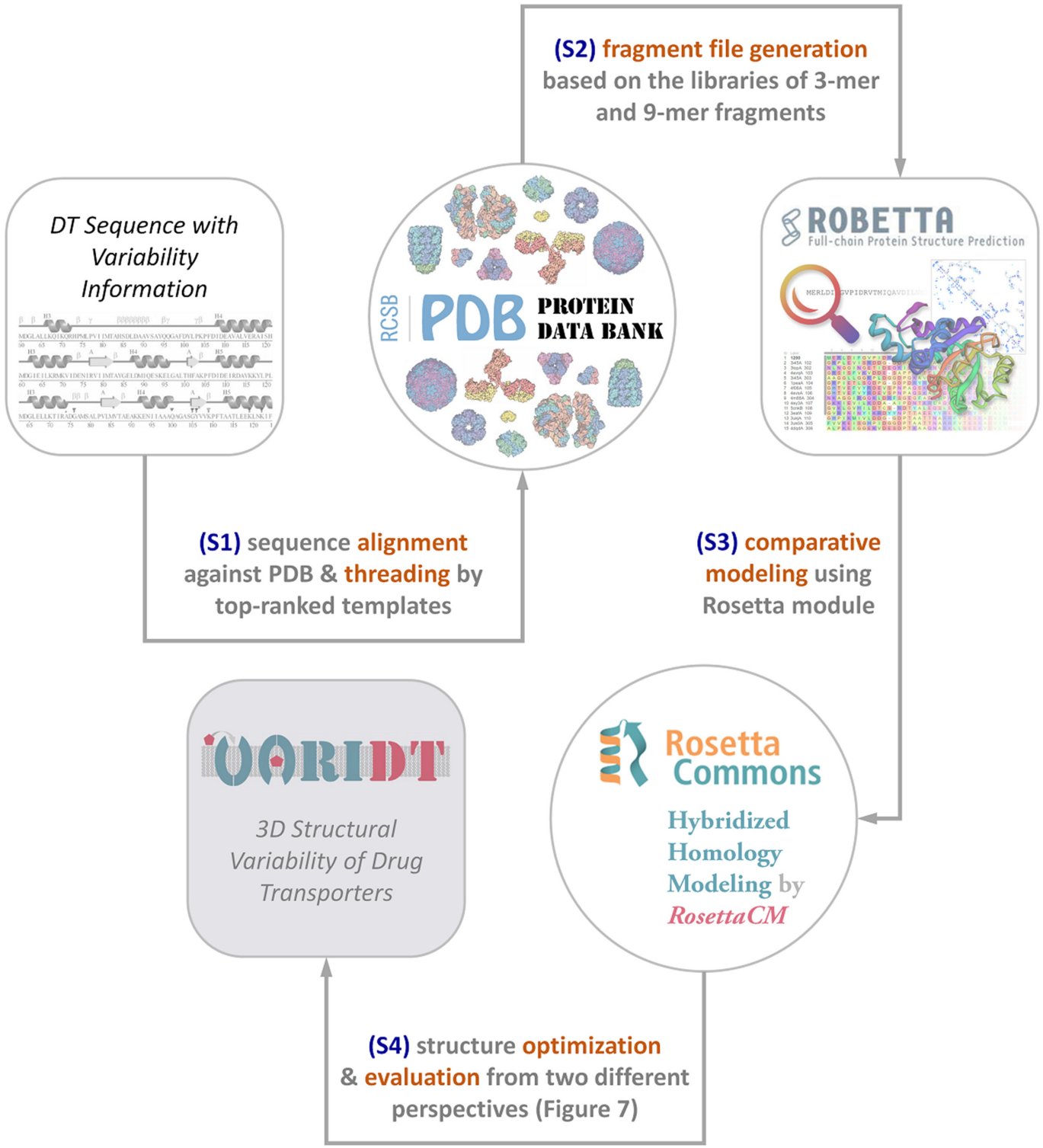


Figure 6. The protocol of comparative modeling used in this study to generate structural models of DTs together with their structural variability data. (**S1**) identifying suitable structure templates through sequence alignment against PDB and threading DT sequence over top-ranked templates; (**S2**) parsing sequence into 3mer and 9mer fragments to generate fragment file; (**S3**) forming five model candidates for comparatively modeling each DT; (**S4**) optimizing the generated structures by energy minimization, and evaluating the quality of the generated structures.

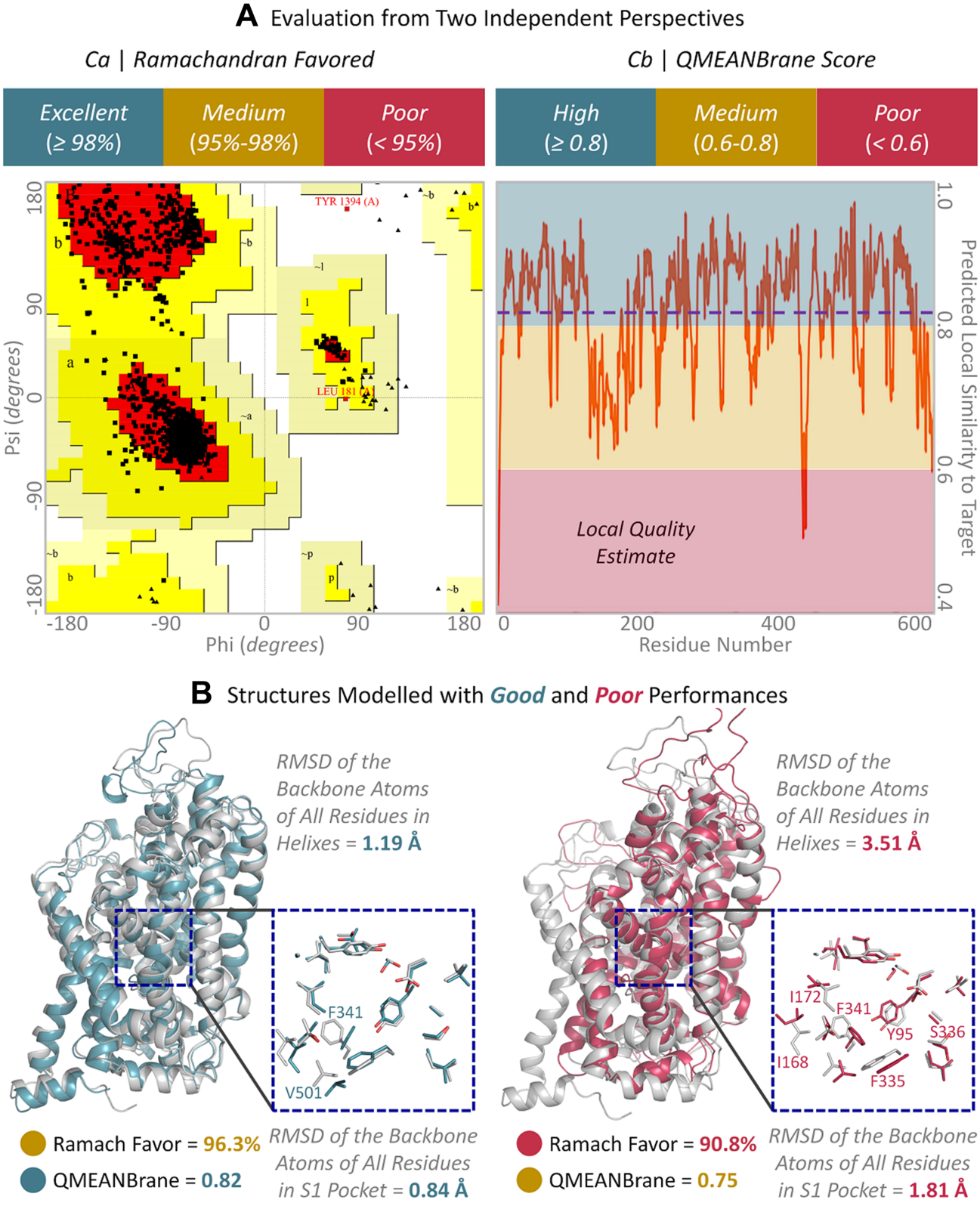


Figure 7. Multiple criteria for model assessment. (**A**) criteria used for assessing generated models: *Ramachandran plot* (*Ca*, left) and *QMEANBrane score* (*Cb*, right). The percentages of modelled residues in the favored regions of *Ramachandran* plot (≥98%, 95–98% and <95%) were adopted in VARIDT as *Excellent* (green), *Medium* (yellow), and *Poor* (red), respectively. QMEANBrane scores of ≥0.8, 0.6–0.8 and <0.6 were considered in VARIDT as *High* (green), *Medium* (yellow) and *Poor* (red), respectively. (**B**) superposition between hSERT structure (grey) and the structures modelled based on *Drosophila* dopamine transporter (dDAT; blue), and *Aquifex aeolicus* leucine transporter (LeuT; red). The percentage of the residues of modelled blue structure within favored *Ramachandran* regions was much higher than that of the red one (96.3% and 90.8% for blue and red structures, respectively), and QMEANBrane score of the blue structure was also much higher than that of the red one (0.82 and 0.75 for blue and red structures, respectively). S1 residues were shown in lines.

Table 1. The statistics of all data updated to VARIDT 2.0. DT: drug transporter; TCDB: Transporter Classification Database; ICD11: the latest WHO International Classification of Diseases. TCDB-defined DT families were directly adopted from the third-level of TCDB families

		No. of DTs (no. of structures)	No. of TCDB-defined DT families	No. of drugs transported by DTs (approved/in clinical trial)	No. of ICD11-defined disease classes	Detailed description on DT structural variability provided in VARIDT 2.0
All DTs and their various structure data available in VARIDT 2.0		416 (2498)	54	885 (585/132)	414	Four types of DT structural variability data were updated as illustrated in Figure 1
Structural variabilities in VARIDT 2.0	(a) Mutation-induced spatial variations in folded state of DT	42 (145)	16	655 (456/92)	233	A total of 266 pharmacokinetic variations measured by drug clearance, response, etc.
	(b) Difference among the DT structures of multiple species	292 (1622)	44	841 (569/119)	267	A total of 17 DT's species origins such as human, rat, mouse, dog, zebrafish, *etc.*
	(c) Outward-facing and inward-facing DT conformations	59 (118)	6	653 (450/98)	235	A total of 118 paired outward- and inward-facing conformations of 59 human DTs
	(d) Xenobiotics-driven alterations in 3D complex of DT	57 (822)	15	709 (495/104)	239	A total of 409 regulating activity data from 214 xenobiotics against 18 human DTs

Statistics, data standardization, access and retrieval

As described in Table 1, four types of DTs' structural variabilities were collected, which included both the experimentally resolved and the computationally modelled structures. (a) for mutation-induced spatial variations in the folded state of DT, a total of 145 structures originating from 42 DTs were collected, (b) for differences among the DT structures of human and model organisms, a total of 1622 inter-species structures covering 292 DTs of 17 species (human, rat, mouse, dog, zebrafish, rabbit, bovine, pig, chicken, sheep, fruit fly, frog, orangutan, monkey, hamster, guinea pig and horse) were described, (c) for outward- and inward-facing conformations, a total of 118 conformations belonging to 59 DTs were provided and (d) for xenobiotics-driven alterations in 3D DT complexes, a total of 822 xenobiotics-regulated structures of 506 xenobiotics in complex with 57 DTs were collected and described. As a result, 292 DTs with at least one type of structure variability data were shown in VARIDT, which were closely related to the transportation of 844 drugs (including 570 approved & 274 clinical/preclinical drugs) for the treatment of 271 disease classes (such as coronavirus infections, lung cancer, diabetes mellitus, depression, parkinsonism, hypertension and asthma) as defined by the latest WHO International Classification of Diseases (105). Moreover, as one popular structure prediction software tool using AI technique, AlphaFold (64,65) has been used to predict the structure of protein, and a hyperlink to the structure predicted by AlphaFold was also provided in the corresponding DT page of VARIDT 2.0, which could be used as a valuable reference when describing DTs' structural variability.

To make the access and analysis of VARIDT data convenient for all users, the collected raw data were carefully cleaned up and then systematically standardized. These standardizations included: (i) all DTs were standardized by and crosslinked to several available databases, and the extended data of each DT could be identified by the diverse hyperlinks to UniProt (13), ClinicalTrials.gov (106), TTD (58), PDB (14), INTEDE (107), PubChem (108), TCDB (15), ICD-11 (105), ChEBI (109), NCBI Gene (66), dbSNP (110), Drugs@FDA (111) and CAS Registry Number (112). (ii) The web page of each DT was re-organized by categorizing all drugs and endogenous metabolites into single superclass entitled 'Molecular Transporting Profiles of This DT', and all transported molecules were grouped according to their clinical statuses (such as: approved, clinical trial and preclinical). (iii) The web page that described the epigenetic regulations data for each DT was also re-organized by categorizing the epigenetic regulations according to their corresponding diseases. All structure variability data can be viewed, assessed, and downloaded from VARIDT 2.0, which is freely assessable without login requirement by all users at: https://idrblab.org/varidt/.

FUNDING

Natural Science Foundation of Zhejiang Province [LR21H300001]; National Natural Science Foundation of China [81973394, 81872798, U1909208]; Leading Talent of the 'Ten Thousand Plan'—National High-Level Talents Special Support Plan of China; Fundamental Research Fund for Central Universities [2018QNA7023]; 'Double Top-Class' University Project [181201*194232101]; Key R&D Program of Zhejiang Province [2020C03010]; Westlake Laboratory (Westlake Laboratory of Life Sciences and Biomedicine); Alibaba-Zhejiang University Joint Research Center of Future Digital Healthcare; Alibaba Cloud; Information Technology Center of Zhejiang University.

Funding for open access charge: National Natural Science Foundation of China [81872798].
Conflict of interest statement. None declared.

REFERENCES

1. DeGorter,M.K., Xia,C.Q., Yang,J.J. and Kim,R.B. (2012) Drug transporters in drug efficacy and toxicity. *Annu. Rev. Pharmacol. Toxicol.*, **52**, 249–273.
2. Yi,W., Tu,M.J., Liu,Z., Zhang,C., Batra,N., Yu,A.X. and Yu,A.M. (2020) Bioengineered miR-328-3p modulates GLUT1-mediated glucose uptake and metabolism to exert synergistic antiproliferative effects with chemotherapeutics. *Acta Pharm Sin B*, **10**, 159–170.
3. Nixon,M., Mackenzie,S.D., Taylor,A.I., Homer,N.Z., Livingstone,D.E., Mouras,R., Morgan,R.A., Mole,D.J., Stimson,R.H., Reynolds,R.M. *et al.* (2016) ABCC1 confers tissue-specific sensitivity to cortisol versus corticosterone: a rationale for safer glucocorticoid replacement therapy. *Sci. Transl. Med.*, **8**, 352ra109.
4. Liu,Y., Zheng,X., Yu,Q., Wang,H., Tan,F., Zhu,Q., Yuan,L., Jiang,H., Yu,L. and Zeng,S. (2016) Epigenetic activation of the drug transporter OCT2 sensitizes renal cell carcinoma to oxaliplatin. *Sci. Transl. Med.*, **8**, 348ra397.
5. Montanari,F. and Ecker,G.F. (2015) Prediction of drug-ABC-transporter interaction–recent advances and future challenges. *Adv. Drug. Deliv. Rev.*, **86**, 17–26.
6. Zhang,J., Bi,R., Meng,Q., Wang,C., Huo,X., Liu,Z., Wang,C., Sun,P., Sun,H., Ma,X. *et al.* (2019) Catalpol alleviates adriamycin-induced nephropathy by activating the SIRT1 signalling pathway in vivo and in vitro. *Br. J. Pharmacol.*, **176**, 4558–4573.
7. Ali,Y., Shams,T., Wang,K., Cheng,Z., Li,Y., Shu,W., Bao,X., Zhu,L., Murray,M. and Zhou,F. (2020) The involvement of human organic anion transporting polypeptides (OATPs) in drug-herb/food interactions. *Chin Med*, **15**, 71.
8. Yang,Q., Wang,Y., Zhang,Y., Li,F., Xia,W., Zhou,Y., Qiu,Y., Li,H. and Zhu,F. (2020) NOREVA: enhanced normalization and evaluation of time-course and multi-class metabolomic data. *Nucleic Acids Res.*, **48**, W436–W448.
9. Hahn,D., Emoto,C., Euteneuer,J.C., Mizuno,T., Vinks,A.A. and Fukuda,T. (2019) Influence of OCT1 ontogeny and genetic variation on morphine disposition in critically Ill neonates: lessons from PBPK modeling and clinical study. *Clin. Pharmacol. Ther.*, **105**, 761–768.
10. Li,P.C., Tu,M.J., Ho,P.Y., Batra,N., Tran,M.M.L., Qiu,J.X., Wun,T., Lara,P.N., Hu,X., Yu,A.X. *et al.* (2021) In vivo fermentation production of humanized noncoding RNAs carrying payload miRNAs for targeted anticancer therapy. *Theranostics*, **11**, 4858–4871.
11. Li,J., Wang,X., Ning,C., Wang,Z., Wang,Y., Zheng,M., Zhang,S., Lu,Y., Zhang,Y., Li,N. *et al.* (2020) Influences of ABC transporter and CYP3A4/5 genetic polymorphisms on the pharmacokinetics of lenvatinib in chinese healthy subjects. *Eur. J. Clin. Pharmacol.*, **76**, 1125–1133.
12. Yin,J., Sun,W., Li,F., Hong,J., Li,X., Zhou,Y., Lu,Y., Liu,M., Zhang,X., Chen,N. *et al.* (2020) VARIDT 1.0: variability of drug transporter database. *Nucleic Acids Res.*, **48**, D1042–D1050.
13. UniProt,C. (2021) UniProt: the universal protein knowledgebase in 2021. *Nucleic Acids Res.*, **49**, D480–D489.
14. Burley,S.K., Bhikadiya,C., Bi,C., Bittrich,S., Chen,L., Crichlow,G.V., Christie,C.H., Dalenberg,K., Di Costanzo,L., Duarte,J.M. *et al.* (2021) RCSB Protein Data Bank: powerful new tools for exploring 3D structures of biological macromolecules for basic and applied research and education in fundamental biology, biomedicine, biotechnology, bioengineering and energy sciences. *Nucleic Acids Res.*, **49**, D437–D451.
15. Saier,M.H., Reddy,V.S., Moreno-Hagelsieb,G., Hendargo,K.J., Zhang,Y., Iddamsetty,V., Lam,K.J.K., Tian,N., Russum,S., Wang,J. *et al.* (2021) The transporter classification database (TCDB): 2021 update. *Nucleic Acids Res.*, **49**, D461–D467.
16. Armstrong,J.F., Faccenda,E., Harding,S.D., Pawson,A.J., Southan,C., Sharman,J.L., Campo,B., Cavanagh,D.R., Alexander,S.P.H., Davenport,A.P. *et al.* (2020) The IUPHAR/BPS guide to pharmacology in 2020: extending immunopharmacology content and introducing the IUPHAR/MMV guide to malaria pharmacology. *Nucleic Acids Res.*, **48**, D1006–D1021.
17. Kanehisa,M., Furumichi,M., Sato,Y., Ishiguro-Watanabe,M. and Tanabe,M. (2021) KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res.*, **49**, D545–D551.
18. Wang,Y., Zhang,S., Li,F., Zhou,Y., Zhang,Y., Wang,Z., Zhang,R., Zhu,J., Ren,Y., Tan,Y. *et al.* (2020) Therapeutic target database 2020: enriched resource for facilitating research and early development of targeted therapeutics. *Nucleic Acids Res.*, **48**, D1031–D1041.
19. Mendez,D., Gaulton,A., Bento,A.P., Chambers,J., De Veij,M., Felix,E., Magarinos,M.P., Mosquera,J.F., Mutowo,P., Nowotka,M. *et al.* (2019) ChEMBL: towards direct deposition of bioassay data. *Nucleic Acids Res.*, **47**, D930–D940.
20. Wishart,D.S., Feunang,Y.D., Guo,A.C., Lo,E.J., Marcu,A., Grant,J.R., Sajed,T., Johnson,D., Li,C., Sayeeda,Z. *et al.* (2018) DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.*, **46**, D1074–D1082.
21. Elbourne,L.D., Tetu,S.G., Hassan,K.A. and Paulsen,I.T. (2017) TransportDB 2.0: a database for exploring membrane transporters in sequenced genomes from all domains of life. *Nucleic Acids Res.*, **45**, D320–D324.
22. Kim,M.S. and Yi,G.S. (2013) HMPAS: human membrane protein analysis system. *Proteome Sci.*, **11**, S7.
23. Geffers,L., Tetzlaff,B., Cui,X., Yan,J. and Eichele,G. (2013) METscout: a pathfinder exploring the landscape of metabolites, enzymes and transporters. *Nucleic Acids Res.*, **41**, D1047–D1054.
24. Trezza,A., Bernini,A., Langella,A., Ascher,D.B., Pires,D.E.V., Sodi,A., Passerini,I., Pelo,E., Rizzo,S., Niccolai,N. *et al.* (2017) A computational approach from gene to structure analysis of the human ABCA4 transporter involved in genetic retinal diseases. *Invest. Ophthalmol. Vis. Sci.*, **58**, 5320–5328.
25. Meng,H., Xu,H.Q., Yu,L., Lin,G.W., He,N., Su,T., Shi,Y.W., Li,B., Wang,J., Liu,X.R. *et al.* (2015) The SCN1A mutation database: updating information and analysis of the relationships among genotype, functional alteration, and phenotype. *Hum. Mutat.*, **36**, 573–580.
26. Barbarino,J.M., Whirl-Carrillo,M., Altman,R.B. and Klein,T.E. (2018) PharmGKB: a worldwide resource for pharmacogenomic information. *Wiley Interdiscip. Rev. Syst. Biol. Med.*, **10**, e1417.
27. Amberger,J.S., Bocchini,C.A., Scott,A.F. and Hamosh,A. (2019) OMIM.org: leveraging knowledge across phenotype-gene relationships. *Nucleic Acids Res.*, **47**, D1038–D1043.
28. Hediger,M.A., Clemencon,B., Burrier,R.E. and Bruford,E.A. (2013) The ABCs of membrane transporters in health and disease (SLC series): introduction. *Mol. Aspects Med.*, **34**, 95–107.
29. Morrissey,K.M., Wen,C.C., Johns,S.J., Zhang,L., Huang,S.M. and Giacomini,K.M. (2012) The UCSF-FDA transportal: a public drug transporter database. *Clin. Pharmacol. Ther.*, **92**, 545–546.
30. Hoffmann,M.F., Preissner,S.C., Nickel,J., Dunkel,M., Preissner,R. and Preissner,S. (2014) The transformer database: biotransformation of xenobiotics. *Nucleic Acids Res.*, **42**, D1113–D1117.
31. Mak,L., Marcus,D., Howlett,A., Yarova,G., Duchateau,G., Klaffke,W., Bender,A. and Glen,R.C. (2015) Metrabase: a cheminformatics and bioinformatics database for small molecule transporter data analysis and (Q)SAR modeling. *J. Cheminform*, **7**, 31.
32. Zhou,Y., Ye,C., Lou,Y., Liu,J., Ye,S., Chen,L., Lei,J., Guo,S., Zeng,S. and Yu,L. (2020) Epigenetic mechanisms underlying organic solute transporter beta repression in colorectal cancer. *Mol. Pharmacol.*, **97**, 259–266.
33. Xu,L., Zhang,L., Wang,T., Wu,Y., Pu,X., Li,M. and Guo,Y. (2020) ExoceRNA atlas: a database of cancer ceRNAs in human blood exosomes. *Life Sci.*, **257**, 118092.
34. Yu,A.M., Choi,Y.H. and Tu,M.J. (2020) RNA drugs and RNA targets for small molecules: principles, progress, and challenges. *Pharmacol. Rev.*, **72**, 862–898.
35. Li,Y.H., Li,X.X., Hong,J.J., Wang,Y.X., Fu,J.B., Yang,H., Yu,C.Y., Li,F.C., Hu,J., Xue,W.W. *et al.* (2020) Clinical trials, progression-speed differentiating features and swiftness rule of the innovative targets of first-in-class drugs. *Brief. Bioinform.*, **21**, 649–662.
36. Wang,Y., Zhu,Q., Hu,H., Zhu,H., Yang,B., He,Q., Yu,L. and Zeng,S. (2021) Upregulation of histone acetylation reverses organic anion transporter 2 repression and enhances 5-fluorouracil

sensitivity in hepatocellular carcinoma. *Biochem. Pharmacol.*, **188**, 114546.
37. Tao,Z., Li,Y., Teng,Z. and Zhao,Y. (2020) A method for identifying vesicle transport proteins based on LibSVM and MRMD. *Comput. Math. Methods Med.*, **2020**, 8926750.
38. Engelhart,D.C., Azad,P., Ali,S., Granados,J.C., Haddad,G.G. and Nigam,S.K. (2020) Drosophila SLC22 orthologs related to OATs, OCTs, and OCTNs regulate development and responsiveness to oxidative stress. *Int. J. Mol. Sci.*, **21**, 2002.
39. Li,B., Tang,J., Yang,Q., Li,S., Cui,X., Li,Y., Chen,Y., Xue,W., Li,X. and Zhu,F. (2017) NOREVA: normalization and evaluation of MS-based metabolomics data. *Nucleic Acids Res.*, **45**, W162–W170.
40. Xu,X., Liu,Y., Du,G., Ledesma-Amaro,R. and Liu,L. (2020) Microbial chassis development for natural product biosynthesis. *Trends Biotechnol.*, **38**, 779–796.
41. Kim,J., Tan,Y.Z., Wicht,K.J., Erramilli,S.K., Dhingra,S.K., Okombo,J., Vendome,J., Hagenah,L.M., Giacometti,S.I., Warren,A.L. *et al.* (2019) Structure and drug resistance of the *Plasmodium falciparum* transporter PfCRT. *Nature*, **576**, 315–320.
42. Hicks,M., Bartha,I., di Iulio,J., Venter,J.C. and Telenti,A. (2019) Functional characterization of 3D protein structures informed by human genetic diversity. *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 8960–8965.
43. Alam,A., Kowal,J., Broude,E., Roninson,I. and Locher,K.P. (2019) Structural insight into substrate and inhibitor discrimination by human P-glycoprotein. *Science*, **363**, 753–756.
44. Zakrzewska,S., Mehdipour,A.R., Malviya,V.N., Nonaka,T., Koepke,J., Muenke,C., Hausner,W., Hummer,G., Safarian,S. and Michel,H. (2019) Inward-facing conformation of a multidrug resistance MATE family transporter. *Proc. Natl. Acad. Sci. U.S.A.*, **116**, 12275–12284.
45. Coleman,J.A., Navratna,V., Antermite,D., Yang,D., Bull,J.A. and Gouaux,E. (2020) Chemical and structural investigation of the paroxetine-human serotonin transporter complex. *Elife*, **9**, e56427.
46. Zhang,J., Zhao,T., Wang,C., Meng,Q., Huo,X., Wang,C., Sun,P., Sun,H., Ma,X., Wu,J. *et al.* (2021) Catalpol-induced AMPK activation alleviates cisplatin-induced nephrotoxicity through the mitochondrial-dependent pathway without compromising its anticancer properties. *Oxid. Med. Cell Longev.*, **2021**, 7467156.
47. Yin,J., Li,X., Li,F., Lu,Y., Zeng,S. and Zhu,F. (2021) Identification of the key target profiles underlying the drugs of narrow therapeutic index for treating cancer and cardiovascular disease. *Comput. Struct. Biotechnol. J.*, **19**, 2318–2328.
48. Chen,J.Y., Brockmoller,J., Tzvetkov,M.V., Wang,L.J. and Chen,X.J. (2019) An in vitro study on interaction of anisodine and monocrotaline with organic cation transporters of the SLC22 and SLC47 families. *Chin. J. Nat. Med.*, **17**, 490–497.
49. Fu,T.T., Tu,G., Ping,M., Zheng,G.X., Yang,F.Y., Yang,J.Y., Zhang,Y., Yao,X.J., Xue,W.W. and Zhu,F. (2021) Subtype-selective mechanisms of negative allosteric modulators binding to group I metabotropic glutamate receptors. *Acta Pharmacol. Sin.*, **42**, 1354–1367.
50. Shen,Y. and Bax,A. (2015) Homology modeling of larger proteins guided by chemical shifts. *Nat. Methods*, **12**, 747–750.
51. Roel-Touris,J., Jimenez-Garcia,B. and Bonvin,A. (2020) Integrative modeling of membrane-associated protein assemblies. *Nat. Commun.*, **11**, 6210.
52. Feng,X., Ambia,J., Chen,K.M., Young,M. and Barth,P. (2017) Computational design of ligand-binding membrane receptors with high selectivity. *Nat. Chem. Biol.*, **13**, 715–723.
53. Shahsavani,N., Sheikhha,M.H., Yousefi,H. and Sefid,F. (2018) In silico homology modeling and epitope prediction of NadA as a potential vaccine candidate in neisseria meningitidis. *Int. J. Mol. Cell Med.*, **7**, 53–68.
54. Kuhlman,B. and Bradley,P. (2019) Advances in protein structure prediction and design. *Nat. Rev. Mol. Cell Biol.*, **20**, 681–697.
55. Jana,S., Ghosh,S., Muk,S., Levy,B. and Vaidehi,N. (2019) Prediction of conformation specific thermostabilizing mutations for class A G protein-coupled receptors. *J. Chem. Inf. Model.*, **59**, 3744–3754.
56. Zhang,Y., Ying,J.B., Hong,J.J., Li,F.C., Fu,T.T., Yang,F.Y., Zheng,G.X., Yao,X.J., Lou,Y., Qiu,Y. *et al.* (2019) How does chirality determine the selective inhibition of histone deacetylase 6? A lesson from trichostatin a enantiomers based on molecular dynamics. *ACS Chem. Neurosci.*, **10**, 2467–2480.
57. Xue,W., Yang,F., Wang,P., Zheng,G., Chen,Y., Yao,X. and Zhu,F. (2018) What contributes to serotonin-norepinephrine reuptake inhibitors' dual-targeting mechanism? The key role of transmembrane domain 6 in human serotonin and norepinephrine transporters revealed by molecular dynamics simulation. *ACS Chem. Neurosci.*, **9**, 1128–1140.
58. Li,Y.H., Yu,C.Y., Li,X.X., Zhang,P., Tang,J., Yang,Q., Fu,T., Zhang,X., Cui,X., Tu,G. *et al.* (2018) Therapeutic target database update 2018: enriched resource for facilitating bench-to-clinic research of targeted therapeutics. *Nucleic Acids Res.*, **46**, D1121–D1127.
59. Newport,T.D., Sansom,M.S.P. and Stansfeld,P.J. (2019) The MemProtMD database: a resource for membrane-embedded protein structures and their lipid interactions. *Nucleic Acids Res.*, **47**, D390–D397.
60. Sarti,E., Aleksandrova,A.A., Ganta,S.K., Yavatkar,A.S. and Forrest,L.R. (2019) EncoMPASS: an online database for analyzing structure and symmetry in membrane proteins. *Nucleic Acids Res.*, **47**, D315–D321.
61. Tordai,H., Jakab,K., Gyimesi,G., Andras,K., Brozik,A., Sarkadi,B. and Hegedus,T. (2017) ABCMdb reloaded: updates on mutations in ATP binding cassette proteins. *Database*, **2017**, bax023.
62. Higuchi,A., Nonaka,N. and Yura,K. (2018) iMusta4SLC: database for the structural property and variations of solute carrier transporters. *Biophys. Physicobiol.*, **15**, 94–103.
63. Waterhouse,A., Bertoni,M., Bienert,S., Studer,G., Tauriello,G., Gumienny,R., Heer,F.T., de Beer,T.A.P., Rempfer,C., Bordoli,L. *et al.* (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.*, **46**, W296–W303.
64. Jumper,J., Evans,R., Pritzel,A., Green,T., Figurnov,M., Ronneberger,O., Tunyasuvunakool,K., Bates,R., Zidek,A., Potapenko,A. *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature*, **596**, 583–589.
65. Tunyasuvunakool,K., Adler,J., Wu,Z., Green,T., Zielinski,M., Zidek,A., Bridgland,A., Cowie,A., Meyer,C., Laydon,A. *et al.* (2021) Highly accurate protein structure prediction for the human proteome. *Nature*, **596**, 590–596.
66. Sayers,E.W., Beck,J., Brister,J.R., Bolton,E.E., Canese,K., Comeau,D.C., Funk,K., Ketter,A., Kim,S., Kimchi,A. *et al.* (2020) Database resources of the national center for biotechnology information. *Nucleic Acids Res.*, **48**, D9–D16.
67. Zhou,F., Zhu,L., Wang,K. and Murray,M. (2017) Recent advance in the pharmacogenomics of human solute carrier transporters (SLCs) in drug disposition. *Adv. Drug. Deliv. Rev.*, **116**, 21–36.
68. Wang,N., Jiang,X., Zhang,S., Zhu,A., Yuan,Y., Xu,H., Lei,J. and Yan,C. (2021) Structural basis of human monocarboxylate transporter 1 inhibition by anti-cancer drug candidates. *Cell*, **184**, 370–383.
69. Li,F., Leier,A., Liu,Q., Wang,Y., Xiang,D., Akutsu,T., Webb,G.I., Smith,A.I., Marquez-Lago,T., Li,J. *et al.* (2020) Procleave: predicting protease-specific substrate cleavage sites by combining sequence and structural information. *Genomics Proteomics Bioinformatics*, **18**, 52–64.
70. Yang,Q.X., Wang,Y.X., Li,F.C., Zhang,S., Luo,Y.C., Li,Y., Tang,J., Li,B., Chen,Y.Z., Xue,W.W. *et al.* (2019) Identification of the gene signature reflecting schizophrenia's etiology by constructing artificial intelligence-based method of enhanced reproducibility. *CNS Neurosci. Ther.*, **25**, 1054–1063.
71. Khan,S. and Vihinen,M. (2007) Spectrum of disease-causing mutations in protein secondary structures. *BMC Struct. Biol.*, **7**, 56.
72. Ferrer-Costa,C., Orozco,M. and de la Cruz,X. (2007) Characterization of compensated mutations in terms of structural and physico-chemical properties. *J. Mol. Biol.*, **365**, 249–256.
73. Han,Z.J., Xue,W.W., Tao,L. and Zhu,F. (2018) Identification of novel immune-relevant drug target genes for Alzheimer's Disease by combining ontology inference with network analysis. *CNS Neurosci. Ther.*, **24**, 1253–1263.
74. Song,J., Wang,Y., Li,F., Akutsu,T., Rawlings,N.D., Webb,G.I. and Chou,K.C. (2019) iProt-Sub: a comprehensive package for accurately mapping and predicting protease-specific substrates and cleavage sites. *Brief. Bioinform.*, **20**, 638–658.
75. Thwaites,D.T. and Anderson,C.M. (2007) H+-coupled nutrient, micronutrient and drug transporters in the mammalian small intestine. *Exp. Physiol.*, **92**, 603–619.

76. Murray,C.M., Hutchinson,R., Bantick,J.R., Belfield,G.P., Benjamin,A.D., Brazma,D., Bundick,R.V., Cook,I.D., Craggs,R.I., Edwards,S. *et al.* (2005) Monocarboxylate transporter MCT1 is a target for immunosuppression. *Nat. Chem. Biol.*, **1**, 371–376.
77. Zhang,B., Jin,Q., Xu,L., Li,N., Meng,Y., Chang,S., Zheng,X., Wang,J., Chen,Y., Neculai,D. *et al.* (2020) Cooperative transport mechanism of human monocarboxylate transporter 2. *Nat. Commun.*, **11**, 2429.
78. Yang,Q., Li,B., Tang,J., Cui,X., Wang,Y., Li,X., Hu,J., Chen,Y., Xue,W., Lou,Y. *et al.* (2020) Consistent gene signature of schizophrenia identified by a novel feature selection strategy from comprehensive sets of transcriptomic data. *Brief. Bioinform.*, **21**, 1058–1068.
79. Kim,J., Koo,B.K. and Knoblich,J.A. (2020) Human organoids: model systems for human biology and medicine. *Nat. Rev. Mol. Cell Biol.*, **21**, 571–584.
80. Emoto,C., Johnson,T.N., Hahn,D., Christians,U., Alloway,R.R., Vinks,A.A. and Fukuda,T. (2019) A theoretical physiologically-based pharmacokinetic approach to ascertain covariates explaining the large interpatient variability in tacrolimus disposition. *CPT Pharmacometrics Syst. Pharmacol.*, **8**, 273–284.
81. Li,F., Zhou,Y., Zhang,X., Tang,J., Yang,Q., Zhang,Y., Luo,Y., Hu,J., Xue,W., Qiu,Y. *et al.* (2020) SSizer: determining the sample sufficiency for comparative biological study. *J. Mol. Biol.*, **432**, 3411–3421.
82. Hahn,D., Emoto,C., Vinks,A.A. and Fukuda,T. (2017) Developmental changes in hepatic organic cation transporter OCT1 protein expression from neonates to children. *Drug Metab. Dispos.*, **45**, 23–26.
83. Levin,E.J., Cao,Y., Enkavi,G., Quick,M., Pan,Y., Tajkhorshid,E. and Zhou,M. (2012) Structure and permeation mechanism of a mammalian urea transporter. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 11194–11199.
84. Orlando,B.J. and Liao,M. (2020) ABCG2 transports anticancer drugs via a closed-to-open switch. *Nat. Commun.*, **11**, 2264.
85. Nosol,K., Romane,K., Irobalieva,R.N., Alam,A., Kowal,J., Fujita,N. and Locher,K.P. (2020) Cryo-EM structures reveal distinct mechanisms of inhibition of the human multidrug transporter ABCB1. *Proc. Natl. Acad. Sci. U.S.A.*, **117**, 26245–26253.
86. Tang,J., Fu,J., Wang,Y., Luo,Y., Yang,Q., Li,B., Tu,G., Hong,J., Cui,X., Chen,Y. *et al.* (2019) Simultaneous improvement in the precision, accuracy, and robustness of label-free proteome quantification by optimizing data manipulation chains. *Mol. Cell. Proteomics*, **18**, 1683–1699.
87. Fromm,M.F. (2004) Importance of P-glycoprotein at blood-tissue barriers. *Trends Pharmacol. Sci.*, **25**, 423–429.
88. Kim,Y. and Chen,J. (2018) Molecular structure of human P-glycoprotein in the ATP-bound, outward-facing conformation. *Science*, **359**, 915–919.
89. Giacomini,K.M., Huang,S.M., Tweedie,D.J., Benet,L.Z., Brouwer,K.L., Chu,X., Dahlin,A., Evers,R., Fischer,V., Hillgren,K.M. *et al.* (2010) Membrane transporters in drug development. *Nat. Rev. Drug Discov.*, **9**, 215–236.
90. Huo,X., Meng,Q., Wang,C., Wu,J., Zhu,Y., Sun,P., Ma,X., Sun,H. and Liu,K. (2020) Targeting renal OATs to develop renal protective agent from traditional chinese medicines: protective effect of apigenin against Imipenem-induced nephrotoxicity. *Phytother. Res.*, **34**, 2998–3010.
91. Ali,Y., Shams,T., Cheng,Z., Li,Y., Chun,C.S., Shu,W., Bao,X., Zhu,L., Murray,M. and Zhou,F. (2021) Impaired transport activity of human organic anion transporters (OATs) and organic anion transporting polypeptides (OATPs) by wnt inhibitors. *J. Pharm. Sci.*, **110**, 914–924.
92. Hong,J., Luo,Y., Mou,M., Fu,J., Zhang,Y., Xue,W., Xie,T., Tao,L., Lou,Y. and Zhu,F. (2020) Convolutional neural network-based annotation of bacterial type IV secretion system effectors with enhanced accuracy and reduced false discovery. *Brief. Bioinform.*, **21**, 1825–1836.
93. Hong,J., Luo,Y., Zhang,Y., Ying,J., Xue,W., Xie,T., Tao,L. and Zhu,F. (2020) Protein functional annotation of simultaneously improved stability, accuracy and false discovery rate achieved by a sequence-based deep learning. *Brief. Bioinform.*, **21**, 1437–1447.
94. Yeboah,D., Sun,M., Kingdom,J., Baczyk,D., Lye,S.J., Matthews,S.G. and Gibb,W. (2006) Expression of breast cancer resistance protein (BCRP/ABCG2) in human placenta throughout gestation and at term before and after labor. *Can. J. Physiol. Pharmacol.*, **84**, 1251–1258.
95. Ren,G., Qin,Z., Yang,N., Chen,H., Fu,K., Lu,C., Lu,Y., Li,N., Zhang,Y., Chen,X. *et al.* (2020) Interactions between oroxylin A with the solute carrier transporters and ATP-binding cassette transporters: drug transporters profile for this flavonoid. *Chem. Biol. Interact.*, **324**, 109097.
96. Song,Y., DiMaio,F., Wang,R.Y., Kim,D., Miles,C., Brunette,T., Thompson,J. and Baker,D. (2013) High-resolution comparative modeling with RosettaCM. *Structure*, **21**, 1735–1742.
97. Carpenter,E.P., Beis,K., Cameron,A.D. and Iwata,S. (2008) Overcoming the challenges of membrane protein crystallography. *Curr. Opin. Struct. Biol.*, **18**, 581–586.
98. Ravna,A.W. and Sylte,I. (2012) Homology modeling of transporter proteins (carriers and ion channels). *Methods Mol. Biol.*, **857**, 281–299.
99. Zhang,Y., Yu,S., Xie,R., Li,J., Leier,A., Marquez-Lago,T.T., Akutsu,T., Smith,A.I., Ge,Z., Wang,J. *et al.* (2020) PeNGaRoo, a combined gradient boosting and ensemble learning framework for predicting non-classical secreted proteins. *Bioinformatics*, **36**, 704–712.
100. Yao,S.Y.M. and Young,J.D. (2018) Inward- and outward-facing homology modeling of human concentrative nucleoside transporter 3 (hCNT3) predicts an elevator-type transport mechanism. *Channels*, **12**, 291–298.
101. Kim,D.E., Chivian,D. and Baker,D. (2004) Protein structure prediction and analysis using the Robetta server. *Nucleic Acids Res.*, **32**, W526–W531.
102. Miao,Y., Goldfeld,D.A., Moo,E.V., Sexton,P.M., Christopoulos,A., McCammon,J.A. and Valant,C. (2016) Accelerated structure-based design of chemically diverse allosteric modulators of a muscarinic G protein-coupled receptor. *Proc. Natl. Acad. Sci. U.S.A.*, **113**, E5675–E5684.
103. Lovell,S.C., Davis,I.W., Arendall,W.B. 3rd, de Bakker,P.I., Word,J.M., Prisant,M.G., Richardson,J.S. and Richardson,D.C. (2003) Structure validation by Calpha geometry: phi,psi and Cbeta deviation. *Proteins*, **50**, 437–450.
104. Studer,G., Biasini,M. and Schwede,T. (2014) Assessing the local structural quality of transmembrane protein models using statistical potentials (QMEANBrane). *Bioinformatics*, **30**, 505–511.
105. Lancet,T. (2019) ICD-11. *Lancet*, **393**, 2275.
106. Zarin,D.A., Fain,K.M., Dobbins,H.D., Tse,T. and Williams,R.J. (2019) 10-year update on study results submitted to ClinicalTrials.gov. *N. Engl. J. Med.*, **381**, 1966–1974.
107. Yin,J., Li,F., Zhou,Y., Mou,M., Lu,Y., Chen,K., Xue,J., Luo,Y., Fu,J., He,X. *et al.* (2021) INTEDE: interactome of drug-metabolizing enzymes. *Nucleic Acids Res.*, **49**, D1233–D1243.
108. Kim,S., Chen,J., Cheng,T., Gindulyte,A., He,J., He,S., Li,Q., Shoemaker,B.A., Thiessen,P.A., Yu,B. *et al.* (2021) PubChem in 2021: new data content and improved web interfaces. *Nucleic Acids Res.*, **49**, D1388–D1395.
109. Yang,Q., Hong,J., Li,Y., Xue,W., Li,S., Yang,H. and Zhu,F. (2020) A novel bioinformatics approach to identify the consistently well-performing normalization strategy for current metabolomic studies. *Brief. Bioinform.*, **21**, 2142–2152.
110. Sherry,S.T., Ward,M.H., Kholodov,M., Baker,J., Phan,L., Smigielski,E.M. and Sirotkin,K. (2001) dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.*, **29**, 308–311.
111. Tang,J., Fu,J., Wang,Y., Li,B., Li,Y., Yang,Q., Cui,X., Hong,J., Li,X., Chen,Y. *et al.* (2020) ANPELA: analysis and performance assessment of the label-free quantification workflow for metaproteomic studies. *Brief. Bioinform.*, **21**, 621–636.
112. Stobaugh,R.E. (1988) Chemical abstracts service chemical registry system. 11. substance-related statistics: update and additions. *J. Chem. Inf. Comput. Sci.*, **28**, 180–187.