## **Supplement Information**

Molecular docking and MD simulations were carried out to elucidate the effect of vanillin on the stability of ligand-(P-gp) complexes, which is closely related to the absorption of actively transported drugs.

A molecular docking simulation was performed and analyzed using AutoDock 4.2.<sup>1</sup> The crystal structure of human P-gp was generated using homology modeling, which was carried out as previously reported and used for the docking study.<sup>2</sup> The structure of vanillin and other ligands (colchicine, quinidine, verapamil and berberine) was obtained from Pubchem (https://pubchem.ncbi.nlm.nih.gov/) and optimized using Chimera1.12 (Figures 1 and 2).<sup>3</sup> The lowest docked free energy structure, which was in the P-gp hydrophobic pocket (MET35, LEU299, PHE270, PHE302, PHE303, PHE638, LEU634, LEU668, GLY691, VAL884, LEU885, PHE888, THR855), was chosen. Docking images were rendered by PyMOL Molecular Graphics System (Version 1.6.0.0; Schrödinger LLC, Cambridge, MA) and ProteinsPlus.<sup>4</sup>

To observe the effect of vanillin on P-gp, homology modeling of P-gp was taken as the initial structure for CG simulation in the central region of 512 DPPC molecules, and 709 vanillin molecules (30%) were randomly placed in the simulation box. The CG force field of P-gp used Martini2.2. The whole system was solvated in water and energy minimized for 2000 steps. Then, 50 ps NVT equilibration and 1ns NPT equilibration were performed with a time step of 10 fs. After equilibration, a 1ps production simulation was run with a time step of 20 fs for analysis. The system was constructed by CHARMM-GUI and Gromacs tools. The temperature was set at 323K using the V-rescale algorithm with a coupling time of 1.0 ps. The Berendsen barostat, semi-isotropic pressure coupling at a compressibility of  $3 \times 10-4$ /bar, and a time constant of 2.0ps were used to maintain pressure (1.0 bar).

We further used all-atom MD simulations to investigate the effect of vanillin on the stability of key residues in the P-gp binding pocket and transmembrane domains (TMDs). P-gp was placed in the appropriate position of 256 DPPC membrane by FlateGro script.<sup>5</sup> Next, 348 vanillin (20%) molecules were placed into the box randomly. The entire system was solvated in SPC water and neutralized by CL-1 ions. Lipids used the Berger force field, and P-gp used the GROMOSA97 force field.<sup>6</sup> The force field of vanillin was generated from the PRODRG server.<sup>7</sup>

After 5000 steps of energy minimization, we performed 50 ps NVT equilibration and 2.0 ns NPT equilibration with a time step of 2 fs and 10000 kJ/mol·nm<sup>2</sup> position restraint. The Berendsen method was selected for temperature and pressure coupling in the equilibration. The temperature was 323K, and semi-isotropic pressure coupling maintained pressure at 1.0 bar. Particle-Mesh Ewald was set to calculate long-range electrostatic interactions, and the Van der Waals cut-off was 1.0 nm. As long as the system was well-equilibrated, we immediately released the position restraints and ran a 50 ns production simulation with V-rescale temperature coupling and Parrinello-Rahman pressure coupling.

The docking energy and conformation of four natural substrates of P-gp, colchicine, quinidine, verapamil, and berberine were compared with vanillin at the binding site of P-gp. Murine P-gp (PDB:3G5U, 3.8Å) was used as a template to model human P-gp; validation of the structure of human P-gp is shown in Supplementary Figure 1-3. As presented in Supplementary Figure 6, vanillin remained at the binding site of P-gp with a docking energy of -4.78 kcal/mol, much higher than the docking energies of colchicine, quinidine, verapamil, and berberine (Supplementary Figure 6). These results indicated that P-gp is a low-affinity receptor for vanillin.

CG MD simulations were conducted to investigate the effect of vanillin on P-gp over a long time scale. We calculated the root mean square deviation (RMSD) of P-gp embedded in the membrane to determine the stability of the protein structure and then compared the RMSD with system-added vanillin molecules (Supplementary Figure 7). All-atom MD simulations were performed to calculate the RMSD of key residues in the binding pocket and TMDs of P-gp (Supplementary Figure 7). The graph reveals that the RMSDs of all systems fluctuated slightly compared with system-added vanillin molecules, suggesting that vanillin may not affect the stability of P-gp or the P-gp binding pocket. A possible explanation for this phenomenon is that the drug-binding pocket of P-gp is composed of extremely hydrophobic and aromatic residues and is located in the center of the membrane, which is highly hydrophobic. However, the log P value of vanillin (ChemIDplus RN: 121-33-5), 1.21, was lower than that of the hydrophobic substrates of P-gp, implying that vanillin could not take up the ligand binding site and inhibit the function of P-gp.

Table.S1 physicochemical information of marker drugs

ID	Compound	miLogP	TPSA	natoms	MW	nON	nOHNH	nrotb	volume
1	acyclovir	-1.61	119.06	16	225.21	8	4	4	187.75
2	hydrochlorothiazide	-0.06	118.36	17	297.75	7	4	1	202.5
3	propranolol	2.97	41.49	19	259.35	3	2	6	257.82
4	carbamazepine	2.84	48.03	18	236.27	3	2	0	215.08

Table S2. The linear regression equation for maker drugs

ange(µM)	Equation (r2)	
5-100μΜ	y=1586.6x+1330.3	
	(0.9913)	
5-100μΜ	y=2523.1x+759	
	(0.996)	
5-100μΜ	y=657.04x-3.0026	
	(0.9958)	
5-100μΜ	y=1759.7x-884.95	
	(0.9918)	
5-100µM	y=2169x-1080.2	
	(0.9907)	
	5-100µМ 5-100µМ 5-100µМ 5-100µМ	

Table S3. Binding energy and interaction results of Ligand in the binding pocket of P-gp.

Compound	Binding energy(kcal/mol)	H-bonds	Hydrophobic residues		
vanillin	-4.75	2	1		
colchicine	-7.24	1	4		
quinidine	-8.47	2	5		
verapamil	-6.49	1	6		
berberine	-8.59	0	6		

## 1. Sequence alignment

	1 10	20	30	40	50	60
sp P08183 MDR1_HUMAN sp P21447 MDR1A_MOUSE	MDLEGDRNG <b>G</b> AKK MELEEDLK <mark>G</mark> RAD	KNFFKLNNKS KNFSKMGKKS	EKDKKEKKPT KKEKKEKKPA	/SVFSMFRYSN /SVLTMFRYAG	WLDKLYMVV WLDRLYMLV	GTLAAII GTLAAII
sp P08183 MDR1_HUMAN sp P21447 MDR1A_MOUSE	HGAGLPLMMLVFG HGVALPLMMLIFG	EMTDIFANAG DMTDSFASVG	NLEDLMSNITN NVSKNSTN	NR <mark>SDI</mark> NDTGFF M <mark>SEA</mark> DKRAMF	MNLEEDMTR AKLEEEMT	YAYYYSG YAYYYTG
sp P08183 MDR1_HUMAN sp P21447 MDR1A_MOUSE	130 IGAGVLVAAYIQV IGAGVLIVAYIQV	140 SFWCLAAGRQ SFWCLAAGRQ	150 IHKIRKQFFHZ IHKIRQKFFHZ	160 AIMRQEIGWFI AIMNQEIGWFI	170 VHDVGELNT VHDVGELNT	180 RLTDDVS RLTDDVS
sp P08183 MDR1_HUMAN sp P21447 MDR1A_MOUSE	190 KINEGIGDKIGMF KINEGIGDKIGMF	200 FQSMATFFTG FQAMATFFGG	210 FIVGFTRGWKI FI <mark>I</mark> GFTRGWKI	220 LTLVILAISPV LTLVILAISPV	230 IGLSAAVWA IGLSA <mark>GIWA</mark>	240 KILSSFT KILSSFT
sp P08183 MDR1_HUMAN sp P21447 MDR1A_MOUSE	250 DKELLAYAKAGAV DKELHAYAKAGAV	260 AEEVLAAIRT AEEVLAAIRT	270 VIAFGGOKKEI VIAFGGOKKEI	280 LERYNKNLEEA ERYNNNLEEA	290 KRIGIKKAI KRIGIKKAI	ЗОО ТАNISIG ТАNISMG
sp P08183 MDR1_HUMAN	310 AAFLLIYASYALA	320 FWYGTTLVLS	330 G <mark>eysigovltv</mark>	340 /FFSVLIGAFS	350 VGQASPSIE	360 AFANARG
sp P21447 MDRIA_MOUSE	AAFLLIYASYALA 370 AAYETFKIIDNKP	380 SIDS <mark>W</mark> SKSGH	K <u>eysigovliv</u> 390 KPDNIKGNLEE	400 TRNVHFSYPSF	410 KEVKILKGI	AFANARG 420 NLKVOSG
sp P21447 MDR1A_MOUSE	AAYE <mark>v</mark> fkiidnkp 430	SIDS <mark>F</mark> SKSGH 440	KPDNI <mark>Q</mark> GNLEE 450	F <mark>KNIHFSYPSF</mark> 460	KEV <mark>Q</mark> ILKGL	NLKV <mark>Ř</mark> SG 480
sp P08183 MDR1_HUMAN sp P21447 MDR1A_MOUSE	QTVALVGNSGCGK QTVALVGNSGCGK 490	STTVQLMQRL STTVQLMQRL 500	YDPTEGMVSVI YDPLDGMVSII 510	DGQDIRTINVF DGQDIRTINVF 520	ELREIIGVV VLREIIGVV 530	SQEPVLF SQEPVLF 540
sp P08183 MDR1_HUMAN sp P21447 MDR1A_MOUSE	ATTIAENIRYGRE ATTIAENIRYGRE 550	N <mark>VTMDEIEKA</mark> D <mark>VTMDEIEKA</mark> 560	VKEANAYDFIN VKEANAYDFIN 570	MKLPHKFDTLV MKLPHQFDTLV 580	GERGAQLSG GERGAQLSG 590	GQKQRIA GQKQRIA 600
sp P08183 MDR1_HUMAN sp P21447 MDR1A_MOUSE	IARALVRNPKILL IARALVRNPKILL 610	LDEATSALDT LDEATSALDT 620	ESEAVVQVALI ESEAVVQAALI 630	DKARKGRTTIV	IAHRLSTVR IAHRLSTVR 650	NADVIAG NADVIAG
sp P08183 MDR1_HUMAN sp P21447 MDR1A_MOUSE	FDDGVIVEKGNHD FDGGVIVEQGNHD	ELM <mark>KEKGIYF</mark> ELM <mark>R</mark> EKGIYF	KLVTMQTAGNE KLVMTQTAGNE	EVELENAADES EIELGNEACKS	KSEIDALEM KDEIDNLDM	SSNDSRS SSKDSGS
sp P08183 MDR1_HUMAN sp P21447 MDR1A_MOUSE	670 SLIRKRSTRRSVR SLIRRRSTRKSIC	680 GSQA <mark>QDRKLS</mark> GPHD <mark>QDRKLS</mark>	690 TKEALDESIPP TKEALDEDVPP	VSFWRIMKLN ASFWRILKLN	JIO LTEWPYFVVO STEWPYFVVO	720 GVFCAII GIFCAII
sp P08183 MDR1_HUMAN sp P21447 MDR1A_MOUSE	730 NGGLQPAFATIFS NGGLQPAFSVIFS	740 K <mark>IIGVFT</mark> RID K <mark>VVGVFT</mark> NGG	750 DPETKRONSNL PPETORONSNL	760 FSLLFL <mark>ALGI</mark> FSLLFL <mark>I</mark> LGI	770 ISFITFFLQC ISFITFFLQC	780 GFTFGKA GFTFGKA
sp P08183 MDR1_HUMAN sp P21447 MDR1A_MOUSE	790 GEILTKRLRYMVF GEILTKRLRYMVF	800 RSMLRODVSW KSMLRODVSW	810 FDDPKNTTGAL FDDPKNTTGAL	820 TTRLANDAAQ TTRLANDAAQ	830 VKGAIGSRLA VKGATGSRLA	840 AVITONI AVIFONI
sp P08183 MDR1_HUMAN	850 ANLGTGIIISFIYO	860 GWQLTLLLA	870 IVPIIAIAGVV	880 EMKMLSGQAL	890 KDKKELEGSO	900 GKIATEA
sp P21447 MDR1A_MOOSE sp P08183 MDR1_HUMAN	910 IENFRTVVSLTQE	920 2KFE <mark>HMYAQS</mark>	930 LQVPYRNSLRK	940 AH <mark>T</mark> FGITFSF	950 TQAMMYFSY	960 G <mark>CFRFG</mark>
sp P21447 MDR1A_MOUSE	IENFRTVVSLTREG 970 AYLVAHKIMSFED	QKFETMYAQS 980 VLLVFSAVVF	LQIPYRNAMKK 990 GAMAVGOVSSF	AHVFGITFSF 1000 APDYAKAKTS	TQAMMYFSYA 1010 Aahitmitei	AACFREG 1020 KTPLIDS
sp P21447 MDR1A_MOUSE	AYLVTQQLMTFEN 1030	VLLVFSAIVF 1040	GAMAVGQVSSF 1050 VDWDDDDDDDV10	APDYAKA TVS	ASHIIRIIEE 1070	TPEIDS 1080
sp P21447 MDR1A_MOUSE	YSTQGLKPNMLEGI 1090	NVQFSG <mark>VVFN</mark> 1100	1110	GLSLEVKKGQ 1120	TLALVGSSG TLALVGSSG 1130	CGKSTVV 1140
sp P08183 MDR1_HUMAN sp P21447 MDR1A_MOUSE	QLLERFYDPLAGK QLLERFYDPMAGS 1150	VLLDGKEIKR VFLDGKEIKQ 1160	LNVQWLRA <mark>HLG LNVQWLRA</mark> QLG 1170	IVSQEPILFD IVSQEPILFD 1180	CSIAENIAYO CSIAENIAYO 1190	GDNSRVV GDNSRVV 1200
sp P08183 MDR1_HUMAN sp P21447 MDR1A_MOUSE	S <mark>QEEIVRAAKEANI</mark> S <mark>YEEIVRAAKEANI</mark>	IHAFIESLPN IHQFIDSLPD	KY <mark>STKVGDKGT</mark> KY <mark>NTRVGDKGT</mark>	QLSGGQKQRI QLSGGQKQRI	AIARALVRQI AIARALVRQI	PHILLD
sp P08183 MDR1_HUMAN sp P21447 MDR1A_MOUSE	1210 EATSALDTESEKVY EATSALDTESEKVY	1220 VQEALDKARE VQEALDKARE	1230 GRTCIVIAHRI GRTCIVIAHRI	1240 STIQNADLIV STIQNADLIV	1250 VFQNGRVKEI VIQNGKVKEI	1260 HGTHQQL HGTHQQL
sp P08183 MDR1_HUMAN sp P21447 MDR1A_MOUSE	1270 LAQKGIYFSMVSVQ LAQKGIYFSMVSVQ	1280 QAGTKRQ QAGAKRS				

Figure. S1 Sequence alignment of 3G5U and human P-gp obtained from ClustalW

## 2. 3D-structure of Human P-glycoprotein



Figure. S2 The 3D structure of homology modeling of human P-gp.

3. Ramachandran plot for homology model



Figure. S3 Ramachandran plot for the Chain A of P-gp homology mode

4. Simulation snapshots of maker drugs located at various sites of membrane



Figure. S4 Maker drugs with vanillin located at various sites of membrane: A.ACV B.HTZ C.PRO

D.CBZ. In each snapshot, Maker drugs were shown as purple, the vanillin in green and the phosphorus atoms in brown. The remaining membrane atoms are shown as grey lines and water as

red balls.

5. Simulation snapshots of protein-ligand system



Figure. S5

Protein-lipid

system for 10ns equilibration. The protein is shown as gray and the head of lipid as brown. Water have been removed for clarity.

6. Drug docking models in the human P-gp binding pocket.



**Figure. S6** Drug docking models in the human P-gp binding pocket. A. vanillin; B. colchicine; C. quinidine; D. verapamil; E. berberine. The energy of the binding pose and the number of runs were shown below. The information of H-bonds and hydrophibic interactions between drugs and P-gp are highlighted in the right. Color code: drug=yellow; residues=green; O=red; S= orange; N= blue;

H=white.

7. Plots of RMSD of P-gp versus time(ns) obtained after 50ns and 1ps of production run.



Figure. S7 Plots of RMSD of P-gp versus time(ns) obtained after 50ns and 1ps of production run. A.Coarse-grained MD simulations. B-C: All-atom MD simulations. A. Protein; B. Drug binding residues;C. TMDs of P-gp. The red plots are the system with 20% vanillin molecules and the black plots are the control group.

8. Typical chromatogram



chromatogram of (A) HBSS buffer, (B) HBSS buffer spiked with ACV, HTZ, PRO, VIN, CBZ (50μM), (C) HBSS buffer obtained from Caco-2 cell bi-directional transport experiment. (peak 1: ACV; peak 2: HTZ; peak 3: PRO; peak 4: VIN; peak5: CBZ)

## References

 G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson. Autodock4 and AutoDockTools4: automated docking with selective receptor flexiblity. J Comput Chem, 2009, 16, 2785-2791.

- J. R. Daddam, M. R. Dowlathabad, S. Panthangi and P. Jasti. Molecular docking and Pglycoprotein inhibitory activity of flavonoids. *Interdiscip Sci*, 2014, 6, 167-175.
- E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng and T. E. Ferrin. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*, 2004, 25, 1605-1612.
- R. Fährrolfes, S. Bietz, F. Flachsenberg, A. Meyder, E. Nittinger, T. Otto, A. Volkamer and M. Rarey. ProteinsPlus: a web portal for structure analysis of macromolecules. *Nucleic Acids Res*, 2017, 45, W337-W343.
- C. Kandt, W. L. Ash, D. P. Tieleman. Setting up and running molecular dynamics simulations of membrane proteins. *Methods*, 2007, 41, 475-488.
- O. Berger, O. Edholm and F. Jähnig. Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature. *Biophys J*, 1997, 72, 2002-2013.
- A. W. Schüttelkopf and D. M. van Aalten. PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr D Biol Crystallogr*, 2004, 60, 1355-1363.