Computational modeling studies reveal the origin of the binding preference of 3-(3,4-di hydroisoquinolin-2(1H)-ylsulfonyl)benzoic acids for AKR1C3 over its isoforms

Xiaotian Kong | Enming Xing | Sijin Wu | Tony Zhuang | Pui-Kai Li | Chunjua Li | Xiaolin Cheng

1College of Life Science and Bioengineering, Faculty of Environment and Life, Beijing University of Technology; Beijing International Science and Technology Cooperation Base for Intelligent Physiological Measurement and Clinical Transformation, Beijing, China
2Division of Medicinal Chemistry & Pharmacognosy, College of Pharmacy, The Ohio State University, Columbus, Ohio, USA
3Dalian Institute of Chemical Physics, Chinese Academy of Sciences, Dalian, China
4J. Willis Hurst Internal Medicine Program, Department of Medicine, Emory University, Woodruff, Georgia, USA

Abstract
As a key regulator for hormone activity, human aldo-keto reductase family 1 member C3 (AKR1C3) plays crucial roles in the occurrence of various hormone-dependent or independent malignancies. It is a promising target for treating castration-resistant prostate cancer (CRPC). However, the development of AKR1C3 specific inhibitors remains challenging due to the high sequence similarity to its isoform AKR1C2. Here, we performed a combined in silico study to illuminate the inhibitory preference of 3-(3,4-dihydroisoquinolin-2(1H)-ylsulfonyl)benzoic acids for AKR1C3 over AKR1C2, of which compound 38 can achieve up to 5000-fold anti-AKR1C3 selectivity. Our umbrella sampling (US) simulations together with end-point binding free energy calculation MM/GBSA uncover that the high inhibition selectivity originates from the different binding modes, namely “Inward” and “Outward,” of this compound series in AKR1C3 and AKR1C2, respectively. In AKR1C3/38, the tetrahydroquinoline moiety of 38 is accommodated inside the SP1 pocket and interacts favorably with surrounding residues, while, in AKR1C2/38, the SP1 pocket is too small to hold the bulky tetrahydroquinoline group that instead moves out of the pocket with 38 transitioning from an “Inward” to an “Outward” state. Further 3D-QSAR and energy decomposition analyses suggest that SP1 in AKR1C3 prefers to bind with a rigid bicyclic moiety and the modification of the R3 group has important implication for the compound’s activity. This work is the first attempt to elucidate the origin of the binding preference of 3-(3,4-di hydroisoquinolin-2(1H)-ylsulfonyl)benzoic acids for AKR1C3 over its isoforms.
1 | INTRODUCTION

Aldo-keto reductase 1C3 (AKR1C3), also called type 5 17β-hydroxysteroid dehydrogenase (17β-HSD5), has been validated as an essential target for treating a variety of metabolic disorders as well as hormone-dependent and independent carcinomas.1–4 Among them, the close association of the dysregulated AKR1C3 with the progression and aggressiveness of castration-resistant prostate cancer (CRPC) attracts the greatest attention.2–9 Functionally, AKR1C3 participates in all (canonical, alternative, and backdoor) pathways of androgen biosynthesis,10,11 converting the weak androgen precursors Δ4-androstene-3,17-dione and 5α-androstan-3,17-dione, respectively, to potent androgens testosterone (T) and 5α-dihydrotestosterone (5α-DHT).12,13 As a prostaglandin F synthase, it also regulates cell proliferation and differentiation through decreasing the levels of antiproliferative PGH2 (D2) and 15-deoxy-D18,19–21 PGJ2.14–18 Multiple studies have consistently shown that inhibition of AKR1C3 could control CRPC caused by intratumoral androgen synthesis and the emergence of ligand-independent androgen receptor spliced-variants (AR-SVs).19–23 Powell et al. revealed that, in TMPRSS2-ERG fusion-positive CRPC cells, AKR1C3 was upregulated by TMPRSS2-ERG induction, whereby the ERG transcription factor could override the repressive effects of AR on AKR1C3 expression.24 AKR1C3 is also involved in the emergence of resistance to both abiraterone and enzalutamide in CRPC patients.25,26 Furthermore, Wang et al. uncovered that the AKR1C3 plays crucial roles in prostate cancer metastasis and invasion by activating ERK signaling cascade and upregulating the expression of mesenchymal markers.27 Therefore, AKR1C3-targeting drugs can potentially overcome several barriers experienced by traditional androgen receptor (AR) targeting therapies,2,5,6,8 representing a promising strategy for CRPC treatment.

Over the past decades, significant efforts have been dedicated to developing AKR1C3 inhibitors harboring distinct chemotypes4,6,9 (natural products,28–30 steroids,31,32 nonsteroidal anti-inflammatory drugs (NSAIDs) analogues,33–36 jasmonates,37 metal complexes,38,39 sulfonyleureas,40 and so on). However, so far, only two compounds had been advanced into clinical trials, ASP9521 (Phase I–IIb)41,42 and BAY-1128688 (Phase II),43 but both clinical trials failed soon afterward due to the lack of therapeutic efficacy and liver toxicity (NCT01352208 and NCT03373422).4,44 Moreover, AKR1C3 has two isoforms AKR1C1 and AKR1C2 that share high sequence identity (>84%) with AKR1C3. Both isoforms are considered as “undesired” targets for CRPC treatment, owing to their opposite roles in regulating AR activity. Unlike AKR1C3, both AKR1C1 and AKR1C2 inactivate AR by catalyzing the metabolism of potent androgens (e.g., 5α-DHT).4,44 Thus, potent and selective AKR1C3 inhibitors remain in urgent need, not only for the treatment of prostate cancer, but also as a useful chemical tool for probing the physiological and pathological functions of AKR1C3.

Partially owing to the availability of crystal structures of AKR1Cs, several kinds of selective AKR1C3 inhibitors have recently been reported, including derivatives of flufenamic acid (FLF),35,45–47 indomethacin,36 naproxen,34 cinnamic acids,28,48–50 and long-chain polyunsaturated fatty acids.51 Of them, 3-(3,4-dihydroisoquinolin-2(1H)-ylsulfonyl)benzoic acid 38 in Jamieson and colleagues’ work showed 6.1 nM inhibitory activity and 5000-fold selectivity of AKR1C3 versus AKR1C2 selectivity.46 However, the origin of the selectivity of these inhibitors has not been systematically explored and thus remains elusive. Current explanation of the agents’ binding preference is merely based on the comparison of the static crystal structures of AKR1C3 and its isoforms. For example, Adeniji et al. proposed that the SP1 pocket is lined by polar residues Ser118, Ser308, and Tyr319 in AKR1C3, while the corresponding residues in AKR1C1 and AKR1C2 are more hydrophobic, including Phe118, Leu308, and Phe319.6,34 Additionally, FLF and its derivatives were found to adopt distinct binding poses in AKR1C3 and AKR1C2 by overlaying the co-crystal structures (PDB entries: 1S2C [AKR1C3/FLF], 3R43 [AKR1C3/mefenamic acid], and 4DBU [AKR1C3/compound 1] vs. 4JQA [AKR1C2/mefenamic acid]). These differential binding modes can also be observed for other types of selective AKR1C3 inhibitors.40,41

The selectivity mechanism of inhibitors toward AKR1C3 at the atomic level, which is anticipated to propel the development of next-generation AKR1C3 inhibitors with enhanced efficacy and reduced “off-target” effect for CRPC therapy.

**KEYWORDS**

“inward” and “outward” binding mode, AKR1C3 and AKR1C2, energy decomposition, inhibitory preference, umbrella sampling
inhibitors, such as the compound series of indomethacin (PDB entries: 3UG8 [AKR1C3/indomethacin] vs. 4JQ4 [AKR1C2/indomethacin]) and naproxen (PDB entries: 3R58 [AKR1C3/S-naproxen] and 3UFY [AKR1C3/R-naproxen] vs. 4JQ1 [AKR1C2/S-naproxen]). These observations raise the questions of whether the inhibitors’ selectivity comes from their distinct binding modes in the AKR1Cs isoforms, how these different binding interactions lead to the selectivity, and what the molecular determinants for potent and selective inhibitors of AKR1C3 are.

Here, we employed an integrated computational approach to elucidate the molecular mechanisms behind the inhibitory selectivity of 3-(3,4-dihydroisoquinolin-2(1H)-ylsulfonyl)benzoic acids toward AKR1C3 over AKR1C2 (Table S1). First, our molecular docking and molecular dynamics (MD) simulations showed that this class of compounds adopts different binding poses, “Inward” and “Outward,” respectively, in the active site pockets of AKR1C3 and AKR1C2. Additionally, the endpoint free energy calculations revealed stronger interaction energies of the inhibitors with AKR1C3 than with AKR1C2. Furthermore, umbrella sampling (US) simulations were performed to explore the binding landscapes of compound 38 in AKR1C3 and AKR1C2, providing an explanation for their preferential inhibition of AKR1C3 over AKR1C2. Finally, the crucial molecular features that determine the inhibitors’ potency/selectivity against AKR1C3 were identified by 3-dimensional quantitative structure-activity relationship (3D-QSAR) and free energy decomposition analyses. We anticipate this work will provide valuable insights to facilitate the development of more effective and selective anti-AKR1C3 drug candidates.

2 | RESULTS

2.1 | Sequence and structural comparisons of AKR1C3 with its isoforms

The sequence alignment and comparison between AKR1C3 and its isoforms (AKR1C2 and AKR1C1) were calculated using the Chimera software (Figure S1). The sequence of the full length AKR1C3 is 84.76% and 86.67% identical to AKR1C2 and AKR1C1, respectively, and the active binding site (within 10 Å of inhibitors) of AKR1C3 shares ~75.9% sequence identity with AKR1C1 and AKR1C2. The high similarity explains the emergence of many pan-AKR1Cs inhibitors implicating the challenge of developing AKR1C selective inhibitors. As illustrated in Figures S1 and S2d,e, the AKR1C enzymes all possess a highly conserved canonical αβ8 superbarrel core, capped at the N-terminus with an antiparallel hairpin. The active site is constituted by several loops (Loop A–C) located at the C-terminus of the central β-barrel, which contains a conserved catalytic tetrad consisting of Asp50, Tyr55, Lys84, and His117. The binding site of AKR1C3 can be divided into five compartments: an oxygen site (OX) comprising Tyr55, His117, and NADP+, a steroid channel (SC) formed by Trp227 and Leu54, and three subpockets including SP1, SP2, and SP3. Notably, the structural differences between the subtypes are located primarily in the active site loops, and most residues involved in substrate binding and catalysis are on the three loops A, B, and C (Figures S2a–c). Specifically, the SP1 of AKR1C3 is lined by multiple non-conserved residues (the corresponding residues in AKR1C2 are given in parenthesis), such as His304 (Arg304), Phe306 (Leu306), Asn307 (Thr307), Ser308 (Leu308), Ser310 (Ala310), His314 (Gly314), and Tyr319 (Phe319) in Loop C together with Ser118 (Phe118), and Met120 (Val120) in Loop A (Figure S2a vs. b), which gives rise to the noticeable structural difference of SP1 between two enzyme isoforms (Figure S2d). Given the extremely high sequence identity (full-length: 95.9 %, active site: 95.7 %, Figure S1) and almost completely superimposable structures (Figure S2e) of AKR1C2 and AKR1C1, we focus on the inhibitory selectivity mechanism toward AKR1C3 over AKR1C2 for a series of compounds with a common 3-(3,4-dihydroisoquinolin-2(1H)-ylsulfonyl)benzoic acid scaffold.

2.2 | Exploration of key structural features on ligand selectivity by 3D-QSAR

3D-QSAR models were built by molecular field-based QSAR tool in Maestro 11.2 (Schrodinger) to investigate the crucial structural features determining the activity and inhibition selectivity of the studied AKR1C3 inhibitors (Figure S3 and Table S2). The selected 3D-QSAR model (Table S3, PLS factor = 4) exhibited high prediction accuracy for both activity and selectivity, in view that the predicted pIC50 values agreed well with the experimental values (Activity: RMSE = 0.70, \( R^2 = 0.6534 \), \( R^2 \text{test} = 0.8767 \) and \( R^2 \text{training} = 0.9269 \); Selectivity: RMSE = 0.38, \( Q^2 = 0.8801 \), \( R^2 \text{test} = 0.9426 \) and \( R^2 \text{training} = 0.9273 \) (Figure S4).

In Table S4, the relatively high contributions of the steric (0.3108, 0.2799) and hydrophobic (0.2231, 0.2272) features indicate their major roles in governing ligand activity and selectivity, respectively. From the contour maps of both activity and selectivity 3D-QSAR models,
the bulky tetrahydroquinoline moiety of the potent compounds (26, 28, 37, and 38) was found to locate in the favorable regions of the Gaussian-steric and aromatic ring fields (Figures 1a, 3a,f, and S5a,f), suggesting this rigid bicyclic fragment may be a dominant molecular feature responsible for ligand’s activity and selectivity. This is further confirmed by the decreased activity and selectivity of compounds 89–94 with alternative aromatic groups as well as open chain compounds 56–64 (Table S1). Strikingly, we also observed favorable regions of steric and hydrophobic field distributed around the tetrahydroquinoline moiety (Figures 1a,c–f). Thus, we speculate that ligand potency can be further improved by introducing favorable groups into certain positions of this bicyclic moiety, such as the 6-position Br in 38 and 5-position NO2 in 28. However, the added group cannot be too large, as evidenced by 44 and 54 with bulky substituents but increased IC50. Moreover, the contour maps reveal positive effects on ligand selectivity of incorporating negatively charged and H-bond acceptor groups, for example, carboxylate, into the three-position of the phenyl ring (Figures 1b, S5b, 1d, and S5d), while negative effects of hydrophobic and H-bond donor groups (Figures 1c,e and S5c,e). Although the 3D-QSAR models identified some critical structural features that contribute to inhibitors’ selectivity toward AKR1C3 over AKR1C2, these models were solely based on ligand information, which may limit the scope and accuracy of the models. Thus, structure-based molecular modeling methods were subsequently carried out to provide a full understanding of the mechanism of ligand selectivity.

2.3 | High selectivity of compound 38 originates from its different binding modes in AKR1C3 and AKR1C2

2.3.1 | Compound 38 binds differently in the pockets of AKR1C2 and AKR1C3

The co-crystal structure of AKR1C3 complexed with the 3-(3,4-dihydroisoquinolin-2(1H)-ylsulfonyl)benzoic acid derivative 38 has been resolved (PDB entry: 4FAM), where the bulky tetrahydroquinoline moiety is buried deeply into the SP1 pocket (Figure 2a,c), referred to as an “Inward” binding mode below. The structure of the AKR1C2/38 complex is not yet available. Nevertheless, given the high sequence and structural similarity between AKR1C3 and

FIGURE 1 Contour maps for Gaussian-based 3D-QSAR model (PLS = 4) on AKR1C3 selectivity with the alignment of compounds 28, 37, 38, 44, 54, and 61. (a) Steric favorable and unfavorable regions are shown in green and yellow, respectively; (b) Electrostatic positive and negative regions are shown in green and yellow, respectively; (c) Hydrophobic favorable and unfavorable regions are shown in blue and red, respectively; (d) Hydrogen bond acceptor favorable and unfavorable regions are shown in red and magenta, respectively; (e) Hydrogen bond donor favorable and unfavorable regions are shown in purple and cyan, respectively; and (f) Aromatic ring favorable and unfavorable regions are shown in orange and gray, respectively
AKR1C2, the compound could likely adopt a similar “Inward” binding mode in ARK1C2 as in AKR1C3. On the other hand, considering that compound 38 was derived from flufenamic acid (FLF), we predicted the AKR1C2/38 complex structure by docking 38 into the binding pocket of AKR1C2 bound with a FLF analogue, mefenamic acid (PDB entry: 4JQA), in view that Glide docking method exhibits strong ability to reproduce the crystal poses of the known AKR1C2/3 complexes (Figure S6, RMSD: 0.19, 0.32, 0.46, and 0.32 Å for 4JQA, 4FAM, 4FAL, and 4FA3, respectively). Surprisingly, 38 was found to bind AKR1C2 in an outward posture (referred to as “Outward1” below) with the tetrahydroquinoline completely outside of SP1 (Figure 2b).

To discern which conformation, “Inward” or “Outward1,” is the preferred binding pose of 38 in ARK1C2, MD simulations were carried out on the AKR1C2/38 complex in both binding conformations. Figure S7 shows that the system starting from the “Outward1” binding conformation reaches equilibrium at ~220 ns. However, the “Inward” conformation does not reach equilibrium within 500 ns, and the MD trajectory shows a sudden RMSD increase at ~450 ns. We thus extended the “Inward” simulation for another 500 ns. Visual inspection of the MD trajectory reveals that the sudden RMSD increase in the “Inward” AKR1C2/38 simulation corresponds to an “Inward” to “Outward” binding mode transition of 38 in AKR1C2 (Figure 3a–c). After

---

**Figure 2**  
(a) Co-crystal structure of AKR1C3 in complex with 38 that adopts an “Inward” binding conformation (red), and computationally docked structures of AKR1C2 bound with 38 in (b) an “Outward1” (olive) and (c) an “Inward” binding mode (yellow).

**Figure 3**  
(a) The root-mean-square deviations (RMSDs) of backbone heavy atoms as a function of time in the MD simulation of the AKR1C2/38-Inward complex. The “Inward” binding mode (b) remains until 400 ns, and then transitions to an “Outward2” binding mode (c–e) for the rest of the simulation time; (f) Structural overlay of the “Outward1” and “Outward2” binding poses of 38 in the active site of AKR1C2.
the transition, the system reaches equilibrium for the rest of the 600 ns simulation, where 38 adopts a new outward binding conformation (referred to as “Outward2” below) with the tetrahydroquinoline moiety partially out of SP1 (Figure 3c–e). The “Outward2” binding pose derived from the MD simulation exhibits an RMSD of ~1.5 Å from the “Outward1” conformation obtained in the docking calculation (Figure 3f). In the “Outward1” binding pose, 38 interacts with Ile 310 and Phe311 in SP1, Trp86, Val128, Ile129, and Trp227 in SP2, Tyr24, His222, and Glu225 in SP3, and the OX residue Val54 (Figure 4a), while the surrounding residues of 38 in the “Outward2” binding mode include His118, Val120, Ile137, Ile310, Phe311, Tyr317 in SP1, and Trp86, Ile129, and Trp227 in SP2 (Figure 4b).

2.3.2 "Inward" binding mode leads to higher binding affinity in AKR1C3.

As shown in Table 1, the predicted binding energy of 38 with AKR1C3 is much lower than that with AKR1C2, regardless of which “Outward” binding pose in AKR1C2 (−44.4, −34.1, and −33.4 kcal/mol for AKR1C3/38-
Inward, AKR1C2/38-Outward1, and AKR1C2/38-Outward2, respectively). The energy component analysis indicates that van der Waals and electrostatic interaction are both dominant factor for the stronger protein-ligand interaction in AKR1C3 than AKR1C2 (Table 1). In the “Inward” bound AKR1C3-38 structure (Figure 4c), the tetrahydroquinoline moiety is surrounded by multiple residues in SP1 as well as Asn167 and Trp227 in SP2, and thus forms more extensive interaction with the protein in AKR1C3 than in AKR1C2 for both “Outward1” and “Outward2” poses (−23.5 vs. −17.3 and −15.4 kcal/mol, Table S5). This may offer an explanation why the replacement of the phenyl ring in FLF by a tetrahydroquinoline fragment can dramatically improve the compounds’ inhibitory selectivity toward AKR1C3 over AKR1C2. However, given the small difference in the calculated binding free energies between the two different 38-bound AKR1C2 structures (“Outward1” and “Outward2”), it is impossible to determine which of the two outward binding poses is the preferred binding mode of 38 in AKR1C2. To this end, we performed US simulations to investigate how 38 is stabilized in the active site of AKR1C2.

### Table 1

| System          | ΔE_{ele} | ΔE_{vdW} | ΔG_{SA} | ΔG_{GB} | ΔG_{bind} | IC50 (μM) |
|-----------------|----------|----------|---------|---------|-----------|-----------|
| C3/38 Inward    | −44.9 ± 1.23 | −46.1 ± 0.73 | −5.9 ± 0.04 | 52.4 ± 1.15 | −44.4 ± 0.46 | 0.0061     |
| C2/38 (Outward1)| −23.0 ± 3.30 | −35.5 ± 0.48 | −4.8 ± 0.11 | 27.3 ± 2.43 | −34.1 ± 1.16 | >30        |
| C2/38 (Outward2)| −23.5 ± 1.03 | −34.6 ± 1.37 | −4.9 ± 0.20 | 29.6 ± 1.37 | −33.4 ± 1.20 |            |

*Electrostatic interaction.

| van der Waals interaction.  |
| Nonpolar contribution of the solvation effect.  |
| Polar contribution of the solvation effect.  |
| Binding free energy.  |
| The standard deviation were estimated based on 20 blocks.  |

![Figure 5](image-url) (a) The volume value of the tetrahydroquinoline moiety (colored in green) of compound 38. The representation of space volume of SP1 in (b) AKR1C3 and (c) AKR1C2 crystal structures (PDB entry: 4FAM and 4JQA, respectively).
AKR1C3 and AKR1C2. Clustering method was employed to generate 20 representative structures from the trajectory of two MD simulations, respectively. The value comparisons of the SP1's volume suggest that the binding pocket of AKR1C3 is ~20% larger than that of AKR1C2 (226.3 ± 21.73 vs. 185.3 ± 18.59, Table S6). Moreover, the SP1 volume of all AKR1C3 structures are larger than the size of the tetrahydroquinoline fragment, while 20% of AKR1C2’s SP1 cannot hold the tetrahydroquinoline group because of their smaller size. The different size of SP1 can be explained by the non-conserved residues Ser118 and Ser308 in AKR1C3, which, in AKR1C2, were replaced by the bulky residues Phe118 and Leu308 (Figure S8). Therefore, the tetrahydroquinoline moiety can snug into the more spacious SP1 of AKR1C3 without insurmountable steric hindrance, while the SP1 in AKR1C2 is too crowded to bind the bulky tetrahydroquinoline group.

2.4.2 | High flexibility of SP1 in AKR1C3 enables the “Inward” binding of 38

The root-mean-square fluctuations (RMSFs) of Cα atoms were calculated for the six simulated systems using the cpptraj module. As shown in Figure S9, the apo-AKR1C3 is more flexible than the apo-AKR1C2 in several regions, especially Loops A–C that constitute SP1, reflected by the larger RMSF values of the apo-AKR1C3 (Figure S9a). The GNM-driven protein mobility analysis (Figures S9b,c) confirms the higher inherent flexibility of Loops B and C in AKR1C3 than AKR1C2 (Figure S2). Additionally, the binding of 38 to AKR1C3 leads to strong interaction of the tetrahydroquinoline with and thus stabilizes the residues in SP1 of AKR1C3, particularly those in Loops B and C (Table S5, Figures 4 and S2). By comparison, docking of the tetrahydroquinoline into SP1 of AKR1C2 leads to unfavorable interaction with SP1 and dramatically amplifies the dynamics of Loop C (Figure 6c), while the

![FIGURE 6](image-url)  
*Root mean square fluctuations (RMSFs) of heavy atoms in MD simulations of (a) AKR1C3/38-Inward and apo-AKR1C3; (b) AKR1C2/38-Outward1 and apo-AKR1C2; (c) AKR1C2/38-Inward and apo-AKR1C2; and (d) AKR1C2/38-Outward2 and apo-AKR1C2. Loop A: 117–143, Loop B: 217–238, H2: 290–297, Loop C: 298–320; and α2: 53–71, α8: 274–280 are labeled.*
“Outward” binding mode of 38 turns out more favorable in AKR1C2, supported by the decreased RMSFs of Loop B in the “Outward” AKR1C2-38 simulations (Figure 6b,d).

2.4.3 | The “Inward” to “Outward” conformational transition in ARK1C3 and AKR1C2

The MD simulation on the AKR1C2/Inward-38 system showed a spontaneous transition from an “Inward” to an “Outward” conformation (Figure 3a–e). However, considering the limited conformation sampling in conventional MD simulations, the obtained “Outward” conformation is unlikely the final binding mode of 38 in AKR1C2. Thus, US simulations were performed to quantitatively probe the thermodynamically stable binding conformations of compound 38 in ARK1C2 and ARK1C3. As shown in Figure S10, the computed PMF curves reached convergence after ~6 ns US simulations, which were used for further analyses below (Figures 7 and 8).

Figure 7f shows the PMF profile for the “Inward” to “Outward” transition in ARK1C2 by rotating the C7–S1 bond in 38. Different binding poses of 38 are captured along the PMF with the representative ones illustrated in Figures 7a–e and 7a–e, indicating the torsional angle C3-C7-S1-N1 is a sound reaction coordinate for the US calculation. As shown in Figure 7f, the “Inward” conformations corresponding to points d’ and e’ on the PMF curve (Figures 7d’,e’) are located at a metastable energy well with relatively high PMF values of ~4.5 and

---

**FIGURE 7** (a–e, a’–e’) Representative conformations sampled in the umbrella sampling (US) simulations of AKR1C2/38; and (f) Potential mean force (PMF) profile as a function of the C3-C7-S1-N1 dihedral angle using AKR1C2/38-Inward (d’) as the input structure. (g) PMF profile for the transition between AKR1C2/38-Outward2 (g’) and AKR1C2/38-Outward1 (c’).
~11.3 kcal/mol, respectively. The tetrahydroquinoline moiety of the “Inward” bound 38 can rotate out of the SP1 pocket by overcoming an energetic barrier of ~6 kcal/mol (the local PMF maximum between point c’ and d’). Multiple “Outward” binding conformations (states a’, b’, c’, a, and b) of 38 are sampled, shown in Figure 7a’–c’ and Figure 7a,b. Among them, the outward state b’ with the lowest PMF value of ~3.0 kcal/mol exhibits characteristics of the “Outward2” conformation generated by the MD simulations (Figures 2e vs. Figure 7b’). In both complexes, the tetrahydroquinoline fragment of 38 is partially out of the SP1 cavity. The decreased PMF value of state b’ relative to the “Inward” conformation d’ provides an explanation for the spontaneous “Inward” to “Outward2” conformational transition of 38 in the MD simulation of AKR1C2/38-Inward. After flipping out of the SP1 pocket, the AKR1C2/38 system samples three representative states c, d, and e with the tetrahydroquinoline group completely out of SP1 (Figure 7c–e). Structural comparison reveals that state c that corresponds to the global minimum (~0 kcal/mol) of the PMF curve (Figure 7f) represents the “Outward1” binding pose predicted by the docking (Figures 2b vs. Figure 7c). It is evident from the PMF curve that the “Outward1” conformations of 38 in AKR1C2 are more stable than the “Inward” and “Outward2” bound states (Figure 7c vs. 7c’–e’ and 7b’), suggesting that the “Outward1” state is likely the preferred binding pose of 38 in AKR1C2. To further corroborate this, we performed an additional set of US simulations using AKR1C2/38-
Outward2 as the starting structure to explore the local transition between the two outward states. The computed PMF profile (Figure 7g) is broadly consistent with the corresponding region in the full PMF for the Outward2 to Outward1 transition, with “Outward2” corresponding to a local minimum g (3.7 kcal/mol), which is less stable than “Outward1” corresponding to the global minimum c (0 kcal/mol). Therefore, we conclude that 38 adopts the “Outward1” binding mode in the active site of AKR1C2.

In contrast, as shown in Figure 8a–e and a–e’, the true “Outward” states were not sampled in the US simulations of the AKR1C3/38-inward system in which the tetrahydroquinoline moiety of 38 stays strongly bound and does not escape from the SP1 cavity. Hence, states b–e can also be defined as “Inward” conformations, as the tetrahydroquinoline fragment rotates only inside the SP1 pocket (Figure 8b–e), which is accompanied by the greatly elevated PMF values (Figures 8f). The inward bound state b is the most stable binding pose in AKR1C3 (Figure 8b'), as evidenced by the lowest PMF value (0 kcal/mol) in Figure 8f, which is also consistent with the strong binding interaction of 38's tetrahydroquinoline fragment with surrounding residues in SP1 of AKR1C3 (Figure 4 and Table S5).

To sum up, the US simulations quantitatively confirm that compound 38 binds AKR1C2 and AKR1C3 in distinct poses, an “Inward” conformation in AKR1C3 versus an “Outward1” conformation in AKR1C2, consistent with our above results that the tetrahydroquinoline group of 38 can fit well in the SP1 pocket of AKR1C3, while AKR1C2's SP1 pocket is too crowded and rigid to accommodate this bulky group.

3 DISCUSSION

The high selectivity of compound 38 toward AKR1C3 over AKR1C2 was shown to originate from its distinct binding modes, “Inward” and “Outward1,” respectively, in the two isoforms. Specifically, in AKR1C3, the tetrahydroquinoline moiety of 38 can fit in SP1 and interact favorably with the surrounding residues. In AKR1C2, however, the tetrahydroquinoline moiety cannot be accommodated in SP1 and thus binds outside of the SP1 pocket, partially exposed to the solvent, which leads to weaker binding with AKR1C2 (−23.5 vs. −17.3 kcal/mol, Table S5). Therefore, modifications harboring a bulky group, such as tetrahydroquinoline or similar moieties, tend to enhance the compound’s inhibition activity toward AKR1C3, while maintaining its low activity toward AKR1C2 (Table S1), as evidenced by the contour maps for the activity and selectivity of these compounds from the 3D-QSAR models (Figures 1 and S5).

In view of the dominant role of tetrahydroquinoline in compound potency and selectivity, we selected several analogues (28, 38, 61, and 85) with various modifications on the tetrahydroquinoline core to probe the binding mechanism with AKR1C3. Of them, analogues 28 and 38 with a NO2 and Br group added at the 5- and 6-position of the tetrahydroquinoline core, respectively, are the most potent (IC50: 0.0061 and 0.0089 μM, respectively, Table S1) AKR1C3 inhibitors. In both 28 and 38 bound AKR1C3, the inhibitor's tetrahydroquinoline forms favorable van der Waals contact with SP1 residues Ser118, Met120, Phe306, Phe311, Tyr317, Pro318, and Tyr319, and SP2 residue Asn167 (Figure 9 and Table S5 and S7). Among them, the ΔGbind values of Phe306 and Phe311 are particularly strong, −5.64 and −5.26 kcal/mol for 28, and −5.74 and −5.78 kcal/mol for 38, respectively, due to favorable π–π stacking interactions, in line with the contour maps of the 3D-QSAR model in which the tetrahydroquinoline is located in the region favorable for aromatic ring interaction. Additionally, 28 and 38 form van de Waals interactions with Asn167, −4.08 and −4.02 kcal/mol, respectively. Furthermore, the 3D-QSAR analysis revealed an important role of electrostatic interaction, especially the negative charge carried by the tetrahydroquinoline, in governing both inhibitory activity and selectivity (Table S4, Figures 1 and S5). The energy decomposition calculations confirmed that the electrostatic component is a crucial part of the total protein–ligand interaction energy with ΔEele of −18.1 and −16.1 kcal/mol between SP1 residues and the tetrahydroquinoline moiety in 28 and 38, respectively. It can also be noted that minor changes on the tetrahydroquinoline core cause large changes in binding energy and potency (Figure 9 and Tables S5 and S7). For example, in 28, the electro-withdrawing NO2 group in tetrahydroquinoline enhances its interactions with Ser118, Pro119, and Met120, while the 6-bromo analogue 38 shows stronger interaction with Tyr319 than 28, which can be attributed to the formation of a halogen bond between 38 and Tyr319. Therefore, introducing specific substituents at certain positions of tetrahydroquinoline can be a viable strategy to enhance the compounds’ anti-AKR1C3 activity. However, despite the size-tolerance and plasticity of SP1 in AKR1C3, the added substituents still have a size limit to avoid steric clash, as exemplified by a 215-fold potency decrease of 44 that contains a large 4-trifluoromethoxyphenylacetylene group (Table S1).

In analogue 85, the two groups connected by the sulfonamide were partially reversed, with the tetrahydroquinoline and phenyl ring replaced by naphthyl and piperidine, respectively. Despite alternative orientations of naphthyl and tetrahydroquinoline, the interaction energies of 85's naphthyl group with residues in SP1 do not
change much, and those of its piperidine with Trp86 and Trp227 in SP2, and Tyr24 in SP3 increase slightly (Table S7 and Figure 9), which accounts for a moderate anti-AKR1C3 activity decrease of 85 (IC\textsubscript{50}: 0.032 \(\mu\)M). By contrast, for the open chain analogue 61, its anti-AKR1C3 potency is markedly reduced (IC\textsubscript{50} = 1.44 \(\mu\)M, Table S1), which can be explained by its far attenuated interactions with several critical residues Ser118, Pro119, Met120, Phe306, and Phe311 than its enhanced interactions with Asn307 and Ser308 (Figure 9 and Table S7), indicating that AKR1C3’s SP1 prefers the binding of a rigid bicyclic moiety.

Furthermore, a wide range of IC\textsubscript{50} values due to the changes of the R\textsubscript{3} substituent in the phenyl ring (Table S1) highlights its importance for the inhibitor’s activity. The contour maps of the Gaussian-based 3D-QSAR fields clearly show the negative effect of incorporating a hydrophobic group into the three-position of the phenyl ring (Figures 1 and S5c), which instead favors a positively charged and H-bond acceptor group (Figures 1b, S5b, 1d, and S5d). Consistently, as shown in Figure 10 and Table S8, unlike 38, both 67 and 68 undergo binding pose changes with the substituted NH\textsubscript{2} and CF\textsubscript{3} pointing toward the entrance of the active site pocket, resulting in reduced electrostatic interaction of the R\textsubscript{3} moiety with His117 in the OX site (−6.70, −3.00, and −2.82 kcal/mol in 38, 67, and 68, respectively). These structural changes also lead to decreased interactions of the compound with several other residues, such as Lys84, Tyr24, Leu54, and Phe311. A strong H-bond between His117 and the benzoic moiety seems to anchor the ligand in a proper binding pose to strengthen its interactions with AKR1C3-specific residues, supported by the high anti-AKR1C3 activity of analogues 28, 38, and 74 that all contains a H-bond acceptor, acid or its isostere tetrazole at the 3-position of the phenyl ring. Therefore, the property and position of the substituents on the phenyl ring are the molecular determinants dictating the compounds' binding conformation and affinity, and thus its anti-AKR1C3 potency and selectivity.

**CONCLUSIONS**

In this study, an integrated computational approach was employed to offer molecular-level insights into the high selectivity mechanism of 3-(3,4-dihydroisoquinolin-2 (1H)-ylsulfonoyl)benzoic acid derivatives toward
AKR1C3 over its close isofrom AKR1C2. Of them, the most AKR1C3-selective inhibitor 38, up to 5000-fold, was chosen for in-depth analysis. The calculations showed that 38 interacts strongly with AKR1C3 in an “Inward” binding mode with its tetrahydroquinoline moiety stabilized in the sub pocket1 (SP1), while the crowded and rigid SP1 of AKR1C2 cannot accommodate the bulky tetrahydroquinoline moiety of 38, leading to an “Outward1” binding mode and accordingly a lower binding affinity of 38 in AKR1C2. The US MD simulations confirmed that the AKR1C3/38-Inward is the most stable binding pose in AKR1C3 while the AKR1C2/38-Outward1 corresponds to the lowest free energy (PMF) state in AKR1C2. Hence, 38’s inhibitory preference toward AKR1C3 can be explained by its distinct binding modes in AKR1C2 and AKR1C3, which is in turn owing to the inherent size and flexibility difference between the binding pockets of the two isoforms. Moreover, 3D-QSAR combined with residue-based energy decomposition analysis revealed that SP1 in AKR1C3 prefers to bind with a rigid bicyclic moiety and the interaction of the R3 substituent with His117 in the OX site is key for stabilizing the ligand in AKR1C3. Overall, this study represents a first attempt to elucidate the inhibitor selectivity mechanism between two AKR1C isoforms, which will pave the way for developing novel AKR1C3 inhibitors with improved efficacy while avoiding “off-target” effects for CRPC therapy.

5 | MATERIALS AND METHODS

5.1 | System setup and MD simulations

The co-crystal structures of AKR1C3 in complex with 3-(3,4-dihydroisoquinolin-2(1H)-ylsulfonyl)benzoic acids 17, 80, and 85 (PDB entries: 4FAM [17], 3FAL [80], and 4FA3 [86]) were utilized as templates to predict the binding modes of the selected analogues (28, 38, 61, 67, 68, and 85) using Glide in Schrodinger 2018. In AKR1C3, this class of compounds was found to adopt a conformation with the bulky tetrahydroquinoline moiety deeply buried into the sub pocket1 (SP1), which was defined as the “Inward” conformation (Figure 2a). However, how this class of compounds interacts with AKR1C2 is unknown. Hence, we employed two approaches to probe compound 38’s binding modes in
AKR1C2. Compound 38 was chosen because it is the most potent AKR1C3-selective inhibitor in the series, which was obtained by optimizing an AKR1Cs pan-inhibitor, Flufenamic acid (FLF). First, given the high sequence identity of the binding pockets between AKR1C3 and AKR1C2, compound 38 may bind AKR1C2 in a similar manner as AKR1C3. Thus, we constructed an AKR1C2/38-Inward complex through the structural alignment of AKR1C3/38 and AKR1C2/mefenamic acid (Figure 2c). Our second approach was to dock 38 into the pocket of the AKR1C2/mefenamic acid structure (PDB entry: 4IQA), in which mefenamic acid is an analogue of FLF.\textsuperscript{45,59} 38 was found to bind AKR1C2 in an “Outward” pose with the tetrahydroyquinoline group located outside of the pocket (Figure 2b). The Loops/Refine Structure module in UCSF Chimera package\textsuperscript{56} was employed to model the missing side-chains and loop structures of the proteins. Proteins and ligands were, respectively, minimized by Protein Preparation Wizard\textsuperscript{60} and ligprep\textsuperscript{60} module using OPLS3e force field.\textsuperscript{51} Collectively, two binary complexes bound with cofactor only (apo-AKR1C3/NADP\textsuperscript{+} and -AKR1C2/NADP\textsuperscript{+}) and eight ternary complexes bound with both cofactor and ligand, including AKR1C3/28/NADP\textsuperscript{+}, AKR1C3/38/NADP\textsuperscript{+}, AKR1C3/61/NADP\textsuperscript{+}, AKR1C3/67/NADP\textsuperscript{+}, AKR1C3/68/NADP\textsuperscript{+}, AKR1C3/85/NADP\textsuperscript{+}, AKR1C2/38-Outward/NADP\textsuperscript{+}, and AKR1C2/Inward/NADP\textsuperscript{+}, were subjected to MD simulations. For convenience, AKR1C3/38/NADP\textsuperscript{+} and AKR1C2/38-Outward/NADP\textsuperscript{+} and AKR1C2/38-Inward/NADP\textsuperscript{+} were, respectively, referred to as AKR1C3/38, AKR1C2/Outward, and AKR1C2/38-Inward, and the same naming rule was applied to all other complexes.

Prior to MD simulations, all bound ligands, including selected 3-(3,4-dihydroisoquinolin-2(1H)-ylsulfonyl)benzoic acids and the coenzyme NADP\textsuperscript{+}, were optimized by the Hartree–Fock (HF) method\textsuperscript{62} at the level of B3LYP/6-31G\* theory\textsuperscript{63} with Gaussian 09.\textsuperscript{64} Partial atomic charges for the optimized geometries of the ligands were calculated by the restrained electrostatic potential (RESP) algorithm.\textsuperscript{65} The AMBER14SB force field\textsuperscript{66} was used for the proteins, while the atom types and force field parameters of ligands were assigned using the general AMBER force field (gaff).\textsuperscript{67} All the complexes were immersed into a periodic TIP3P water box\textsuperscript{68} with at least 12 Å water padding for each side of the rectangular box, and appropriate numbers of counter-ions (Na\textsuperscript{+} or Cl\textsuperscript{−}) were added to neutralize each system. The missing hydrogen atoms were added to the heavy atoms using the leap module.\textsuperscript{69} A stepwise minimization was then followed using the pmemd program in Amber 18\textsuperscript{70}: (1) 1000 cycles of steepest descent and 4000 cycles of conjugate gradient minimization with the non-hydrogen atoms restrained by 5 kcal/mol/Å\textsuperscript{2} force constant; (2) 5000 cycles of minimization (1000 cycles of steepest descent and 4000 cycles of conjugate gradient minimization), where a 5 kcal/mol/(mol Å\textsuperscript{2}) force was used to fix the non-hydrogen of the complex; (3) 1000 cycles of steepest descent and 4000 cycles of conjugate gradient minimization with the heavy atoms of the complex restrained (2 kcal/mol/Å\textsuperscript{2}); and (4) the whole system was subject to 10,000 cycles of optimization without any restraints (5000 cycles of steepest descent and 5000 cycles of conjugate gradient minimization).

The minimized systems were gradually heated up to 300 K in the NVT ensemble with a restraint of 2 kcal/mol Å\textsuperscript{2} on the backbone of each complex (100 ps). Each system was then equilibrated in the NPT ensemble (T = 300 K and P = 1 atm) with and without the 2 kcal/mol Å\textsuperscript{2} restraints for 500 ps, respectively. Afterward, 1 ns additional equilibration in the NPT ensemble was carried out for each unrestrained system using the algorithms of Berendsen barostat\textsuperscript{71} and Langevin thermostat.\textsuperscript{72} Finally, all AKR1C2 and AKR1C3 systems were subjected to 500 ns NPT MD simulations. In particular, the MD simulation of the AKR1C2/38-Inward complex was extended to 1 μs. The Particle Mesh Ewald (PME) technique\textsuperscript{73} was utilized to treat the long-range electrostatic interactions with a cutoff distance of 10 Å, and the same threshold value was used for the truncation of Lennard–Jones potentials. The SHAKE algorithm\textsuperscript{74} was applied to constrain all bonds involving hydrogen atoms to their equilibrium length. The time step was set to 2 fs and the snapshots were saved every 1 ps for subsequent structural and energetic analysis.

5.2 | Structural motion analysis and pocket size calculations of AKR1C2 and AKR1C3

MD simulation trajectories were analyzed using the cpptraj module\textsuperscript{75} in AmberTools18. The root mean square deviations (RMSDs) as a function of the simulation time (t) were computed to assess the stability of the protein structure in the simulation. The flexibility and dynamics of individual residues were evaluated by the root mean square fluctuations (RMSFs). Besides, the residue mobility were also quantified by the slowest GNM modes using the DynOomics ENM server.\textsuperscript{76}

For each system, 5000 conformations were evenly extracted from the 500 ns MD trajectory. Then, these conformations were clustered using the k-means clustering algorithm through iterative minimization of the sum of the pairwise root-mean-square displacements (RMSDs) between each conformation and its cluster centroid over
all clusters. Twenty representative structures were generated for each trajectory of the apo AKR1C3 and apo AKR1C2 simulations. The pocket size was calculated as the average over the 20 obtained structures using the Computed Atlas of Surface Topography of proteins (CASTp).

5.3 Protein-ligand interaction energy calculations

The binding affinity of a drug with its target was estimated by the Molecular Mechanics/Generalized Born Solvent Area (MM/GBSA) methodology. According to Equation (1), the total binding free energy ($\Delta G_{\text{bind}}$) can be subdivided into several terms, including the van der Waals interaction ($\Delta E_{\text{vdW}}$), the electrostatic interaction ($\Delta E_{\text{ele}}$), the polar ($\Delta G_{\text{polar}}$), and non-polar ($\Delta G_{\text{SA}}$) parts of the solvation free energy ($\Delta G_{\text{solvolution}}$), and the conformational entropy upon ligand binding ($\pm T\Delta S$).

$$\Delta G_{\text{bind}} = G_{\text{com}} - (G_{\text{rec}} + G_{\text{lig}}) = \Delta E_{\text{MM}} + \Delta G_{\text{solvolution}} - T\Delta S$$

$$= \Delta E_{\text{int}} + \Delta E_{\text{ele}} + \Delta E_{\text{vdW}} + \Delta G_{\text{GB}} + \Delta G_{\text{SA}} - T\Delta S$$

Herein, the pmemd module in Amber 18 was used to compute the molecular mechanics gas-phase energy ($\Delta E_{\text{MM}}$), ($\Delta E_{\text{vdW}}$, and $\Delta E_{\text{ele}}$), where $\Delta E_{\text{int}}$ change of the intramolecular energy upon ligand binding, can be neglected due to the use of the single trajectory strategy. $\Delta G_{\text{GB}}$ was calculated using the modified GB model developed by Onufriev et al. ($\text{GB}^{\text{OBC1}}$, igb=2) with the solute ($\varepsilon_{\text{in}}$) and solvent ($\varepsilon_{\text{out}}$) dielectric constants set to 2.0 and 80, respectively. $\Delta G_{\text{SA}}$ was determined by the change of the solvent-accessible surface areas (DSASA) with the LCPO algorithm: $\Delta G_{\text{SA}} = \gamma \times \Delta \text{SASA} + \beta$, where $\gamma$ and $\beta$ were set to 0.0072 kcal/(mol Å$^2$) and 0 kcal/mol, respectively. $-T\Delta S$ is not considered here since the studied compounds are analogues with high structural similarity, and the inclusion of the entropy term demands expensive computational cost but does not necessarily improve the prediction accuracy.

Binding free energy decomposition as implemented in MMPBSA.py was used to identify and compare the residue-ligand interaction differences between AKR1C2 and AKR1C3. The interaction energy of individual residues with the ligand can be formulated as: $\Delta G_{\text{Bind}} = \Delta E_{\text{vdW}} + \Delta E_{\text{ele}} + \Delta G_{\text{GB}} + \Delta G_{\text{SA}}$. Except for $\Delta G_{\text{SA}}$, which was calculated by the ICOSA algorithm, all other terms were calculated based on the same parameters used in the above MM/GBSA calculations.

5.4 Gaussian-based 3D-QSAR study

5.4.1 Dataset collection and 3D-QSAR model construction

A total of 77 derivatives of 3-(3,4-dihydroisoquinolin-2(1H)-yl)sulfonyl)benzoic acids (compounds 17–94, Table S1) that covered a reasonable range of chemical diversity and inhibitory activity were collected from Jamieson and colleagues’ work. The biological activities of these compounds against AKR1C2 and AKR1C3 (IC50,C2 and IC50,C3, μM), together with their structures, were summarized in Table S1. The 3D structures of all molecules were sketched using Maestro and were then optimized using Ligprep module with the OPLS3e force field in Schrödinger 2018. The flexible ligand alignment method was employed to align all the compounds using 38 as a template given its highest IC50 value (Figure S3). Herein, the IC50 values of compounds were converted into pIC50, pIC50,C2, and pIC50,C3 (pIC50,C3-pIC50,C2) were then used as the dependent variables in 3D-QSAR models for AKR1C3 activity and selectivity, respectively (Table S2). The dataset was divided randomly into the training set and test set with a ratio of 4:1. The extended Gaussian-based potential function of Schrodinger 2018 Suite was employed to construct the 3D-QSAR models using Partial Least Squares (PLS) regression, where six Gaussian fields, including steric, electrostatic, hydrophobic, hydrogen bond donor (HBD), hydrogen bond acceptor (HBA), and aromatic ring potential fields, were calculated to evaluate each type of interactions using Gaussian equations. Parameters involved in 3D-QSAR model building were set using the default values.

5.4.2 The validation and analysis of 3D-QSAR models

The accuracy and usefulness of the Gaussian-based 3D-QSAR models were evaluated based on the internal and external validation parameters, especially test set (RMSE, $R^2$, $Q^2$), and Pearson-r test) and training set ($R^2$) statistics. RMSE indicates root-mean-square error in the test set predictions. $Q^2$ is directly analogous to $R^2$ for the regression, correlation coefficient, 0–1), but based on the test set predictions. Pearson-r indicates the correlation between the predicted and observed activity for the test set (0–1). In a validated 3D-QSAR model, the “field fractions” provide a general idea of the relative contribution of each field to the activity and selectivity of the molecules. For example, if the steric and hydrophobic Gaussian field fractions are much larger than all other types, then most of the binding energy...
comes from then hydrophobic interactions. Besides, the alignment of the studied compounds with the 3D-QSAR contour maps of each field can help identify the favorable (positive) and unfavorable (negative) structural features for ligand selectivity/activity.

5.5 US simulations

US simulation, a common enhanced sampling approach, was used to investigate the interconversion between the “Inward” and “Outward” (Outward1 and Outward2) conformations of 38 in the binding cavity of AKR1C3 and AKR1C2, in which AKR1C3/38-Inward (Figure 2a), AKR1C2/38-Outward2 (Figure 2e) and AKR1C2/38-Inward (Figure 2c) structures were respectively used as the input structure. In US simulations, biasing potentials are imposed on the reaction coordinates (RCs) to drive the system from one conformational state to the other. The transition of 38 between the “Inward” and “Outward” conformations can be realized through rotating 38’s tetrahydroquinoline fragment around the C7–S1 bond by 360°, so the torsional angle C3-C7-S1-N1 was chosen as the RC. The full torsional angle range was divided into 36 windows (every 10°), and each window was subjected to 10 ns US simulations. The harmonic potentials applied to each window for biased sampling in that window are shown in Equation (2).

\[ u_{i}^{bias} = \frac{1}{2} k_i (\theta - \theta_{i}^{ref}) ^2 \] (2)

where \( k_i \) is the spring constant in window \( i \) (100 kcal/mol rad²), and \( \theta_{i}^{ref} \) and \( \theta \) denote the reference and instantaneous angle of window \( i \), respectively. All US simulations were performed using Amber 18. The weighted histogram analysis method (WHAM) was employed to estimate the potential mean force (PMF) curve along the RC. WHAM is an effective method to reconstruct unbiased probability distribution from biased probability distribution by umbrella sampling, which contains three steps: (1) calculate the unbiased probability density \( P_i(x) \) of each window, (2) obtain the overall \( P(x) \) of each RC by linearly combining the \( P_i(x) \) of each window, (3) compute the free energy of each RC point and finally draw the PMF curve.

AUTHOR CONTRIBUTIONS

Xiaotian Kong: Conceptualization (lead); data curation (lead); formal analysis (lead); funding acquisition (lead); investigation (lead); methodology (lead); project administration (lead); resources (equal); software (equal); validation (lead); visualization (lead); writing – original draft (lead); writing – review and editing (equal). Enming Xing: Data curation (supporting); validation (supporting); writing – review and editing (supporting). Sijin Wu: Validation (supporting); writing – review and editing (supporting). Tony Zhuang: Writing – review and editing (supporting). Pui-Kai Li: Writing – review and editing (supporting). Chunhua Li: Writing – review and editing (supporting). Xiaolin Cheng: Conceptualization (lead); data curation (lead); formal analysis (lead); investigation (lead); methodology (lead); project administration (lead); software (lead); supervision (lead); validation (lead); visualization (lead); writing – review and editing (lead).

ACKNOWLEDGMENTS

This work was supported by the startup fund from the Faculty of Environment and Life at Beijing University of Technology. Computation was performed using the computing resources at the Ohio Supercomputer Center and the Faculty of Environment and Life at Beijing University of Technology.

CONFLICT OF INTEREST

The authors declare no competing financial interest.

DATA AND SOFTWARE AVAILABILITY

Data and software used in this manuscript are all available.

The co-crystal structure of 3-(3,4-dihydroisoquinolin-2(1H)-ylsulfonyl)benzoic acids 17, 80, and 85 in complex with AKR1C3 (PDB entries: 4FAM [17], 3FAL [80], and 4FA3 [86]) and the AKR1C2/mefenamic acid structure (PDB entry: 4JQA) were downloaded from the RCSB PDB database.

The software packages utilized in this manuscript include ChemDraw Ultra 8.0, Glide, Gaussian 09, Amber18, UCSF Chimera 1.14, and OriginPro 2016. All input and output data files are made available upon request. The DynOomics ENM portal employed in this manuscript can be accessed at http://enm.pitt.edu/ or http://dyn.life.nthu.edu.tw/oENM/.

ORCID

Xiaotian Kong https://orcid.org/0000-0001-7557-1674

REFERENCES

1. Chang TS, Lin H-K, Rogers KA, et al. Expression of aldo-keto reductase family 1 member C3 (AKR1C3) in neuroendocrine tumors & adenocarcinomas of pancreas, gastrointestinal tract, and lung. Int J Clin Exp Pathol. 2013;6:2419.
2. Zeng C-M, Chang L-L, Ying M-D, et al. Aldo–keto reductase AKR1C1–AKR1C4: Functions, regulation, and intervention for anti-cancer therapy. Front Pharmacol. 2017;8:119.
3. Penning TM. AKR1C3 (type 5 17β-hydroxysteroid dehydrogenase/prostaglandin F synthase): Roles in malignancy and endocrine disorders. Mol Cell Endocrinol. 2019;489:82–91.

4. Liu Y, He S, Chen Y, et al. Overview of AKR1C3: Inhibitor achievements and disease insights. J Med Chem. 2020;63:11305–11329.

5. Hamid ARA, Pfeiffer MJ, Verhaegh GW, et al. Aldo-keto reductase (AKR) 1C3 inhibitors: A patent review. Expert Opin Ther Pat. 2017;27:1329–1340.

6. Pippione AC, Carnovale IM, Bonanni D, et al. Potent and selective aldo-keto reductase 1C3 inhibitors: A mini-review. Curr Cancer Drug Tar. 2017;17:603–616.

7. Sun S-Q, Gu X, Gao X-S, et al. Overexpression of AKR1C3 significantly enhances human prostate cancer cells resistance to radiation. Oncotarget. 2016;7:48050.

8. Karunasinghe N, Masters J, Flanagan JU, Ferguson LR. Influence of Aldo-keto reductase IC3 in prostate cancer—A mini-review. Curr Cancer Drug Tar. 2017;17:603–616.

9. Penning TM. Aldo-keto reductase (AKR) 1C3 inhibitors: A patent review. Expert Opin Ther Pat. 2017;27:1329–1340.

10. Pippione AC, Boschi D, Fores K, Oliaro-Bosso S, Lolli ML. Androgen-AR axis in primary and metastatic prostate cancer: Chasing steroidogenic enzymes for therapeutic intervention. J Cancer Metastasis Treat. 2017a;3:328–361.

11. Sumńska M, Bogusz-Górna K, Wegner D, Fichna M. Nongenomic disorder of adrenal steroidogenesis and clinical dilemmas in 21-hydroxylase deficiency combined with backdoor androgen pathway. Mini-review and case report. Int J Mol Sci. 2020;21:4622.

12. Lin H-K, Jez JM, Schlegel BP, Peehl DM, Pachter JA, Penning TM. Expression and characterization of recombinant type 2 3α-hydroxysteroid dehydrogenase (HSD) from human prostate: Demonstration of bifunctional 3α/17β-HSD activity and cellular distribution. Mol Endocrinol. 1997;11:1971–1984.

13. Penning TM, Burczynski ME, Jez JM, et al. Human 3α-hydroxysteroid dehydrogenase isofoms (AKR1C1–AKR1C4) of the aldro-keto reductase superfamily: Functional plasticity and tissue distribution reveals roles in the inactivation and formation of male and female sex hormones. Biochem J. 2000;351:67–77.

14. Forman BM, Tontonoz P, Chen J, Brun RP, Spiegelman BM, Evans RM. 15-deoxy-Δ12, 14-prostaglandin J2 is a ligand for the adipocyte determination factor PPARγ. Cell. 1995;83:803–812.

15. Matsuura K, Shiraishi H, Harai A, et al. Identification of a principal mRNA species for human 3α-hydroxysteroid dehydrogenase isoform (AKR1C3) that exhibits high prostaglandin D2 11-ketoreductase activity. J Biochem. 1998;124:940–946.

16. Suzuki-Yamamoto T, Nishizawa M, Fukui M, et al. cDNA cloning, expression and characterization of human prostaglandin F synthase. FEBS Lett. 1999;462:335–340.

17. Desmond JC, Mountford JC, Drayson MT, et al. The aldo-keto reductase AKR1C3 is a novel suppressor of cell differentiation that provides a plausible target for the non-cyclooxygenase-dependent antineoplastic actions of nonsteroidal anti-inflammatory drugs. Cancer Res. 2003;63:505–512.

18. Wang T, Xu J, Yu X, Yang R, Han ZC. Peroxisome proliferator-activated receptor γ in malignant diseases. Crit Rev Oncol Hemat. 2006;58:1–14.

19. Cai C, Balk SP. Intratumoral androgen biosynthesis in prostate cancer pathogenesis and response to therapy. Endocr Relat Cancer. 2011;18:R175–R182.

20. Knuttila M, Yatkin E, Kallio J, et al. Castration induces up-regulation of intratumoral androgen biosynthesis and androgen receptor expression in an orthotopic VCaP human prostate xenograft model. Am J Pathol. 2014;184:2163–2173.

21. Liu C, Yang JC, Armstrong CM, et al. AKR1C3 promotes AR-V7 protein stabilization and confers resistance to AR-targeted therapies in advanced prostate cancer. Mol Cancer Ther. 2019; 18:1875–1886.

22. Wang B, Wu S, Fang Y, et al. The AKR1C3/AR-V7 complex maintains CRPC tumour growth by repressing B4GALT1 expression. J Cell Mol Med. 2020;24:12032–12043.

23. Zhao J, Ning S, Lou W, et al. Cross-resistance among next-generation androgen drugs through the AKR1C3/AR-V7 axis in advanced prostate cancer. Mol Cancer Ther. 2020;19:1708–1718.

24. Powell K, Semaan L, Conley-LaComb MK, et al. ERG/AKR1C3/AR constitutes a feed-forward loop for AR signaling in prostate cancer cells. Clin Cancer Res. 2015;21:2569–2579.

25. Liu C, Lou W, Zhu Y, et al. Intracrine androgens and AKR1C3 activation confer resistance to enzalutamide in prostate cancer. Cancer Res. 2015;75:1413–1422.

26. Liu C, Armstrong CM, Lou W, Lombard A, Evans CP, Gao AC. Inhibition of AKR1C3 activation overcomes resistance to abiraterone in advanced prostate cancer. Mol Cancer Ther. 2017;16:35–44.

27. Wang B, Gu Y, Hui K, et al. AKR1C3, a crucial androgenic enzyme in prostate cancer, promotes epithelial-mesenchymal transition and metastasis through activating ERK signaling. Urol Oncol. 2018;36:472.e11–472.e20.

28. Zang T, Verma K, Chen M, Jin Y, Trippier FC, Penning TM. Screening baccharin analogs as selective inhibitors against type 5 17β-hydroxysteroid dehydrogenase (AKR1C3). Chem Biol Interact. 2015;234:339–348.

29. Tian Y, Zhao L, Wang Y, et al. Berberine inhibits androgen synthesis by interaction with aldo-keto reductase 1C3 in 22Rv1 prostate cancer cells. Asian J Androl. 2016;18:607.

30. Hulcová D, Breiterová K, Zemanová L, et al. AKR1C3 inhibitory potency of naturally-occurring amaryllidaceae alkaloids of different structural types. Nat Prod Commun. 2017;12:1934578X1701200226.

31. Byrns MC, Jin Y, Penning TM. Inhibitors of type 5 17β-hydroxysteroid dehydrogenase (AKR1C3): Overview and structural insights. J Steroid Biochem. 2011;125:95–104.

32. Savić MP, Ajduković JJ, Plavša JJ, et al. Evaluation of A-ring fused pyridine d-modified androstane derivatives for anti-proliferative and aldo-keto reductase 1C3 inhibitory activity. MedChemComm. 2018;9:969–981.

33. Hendriks CM, Penning TM, Zang T, et al. Pentafluorosulfanyl-containing flufenamic acid analogs: Syntheses, properties and biological activities. Bioorg Med Chem Lett. 2015;25:4437–4440.

34. Adeniji A, Uddin MJ, Zang T, et al. Discovery of (R)-2-(6-methoxynaphthalen-2-yl) butanoic acid as a potent and selective aldo-keto reductase 1C3 inhibitor. J Med Chem. 2016;59:7431–7444.

35. Pippione AC, Carnovale IM, Bonanni D, et al. Potent and selective aldo-keto reductase 1C3 (AKR1C3) inhibitors based on the
benzoisoxazole moiety: Application of a bioisosteric scaffold hopping approach to flufenamic acid. Eur J Med Chem. 2018;150:930–945.

36. Lollì ML, Carnovale IM, Pippione AC, et al. Bioisosteres of indomethacin as inhibitors of aldo-keto reductase 1C3. ACS Med Chem Lett. 2019;10:437–443.

37. Davies NJ, Hayden RE, Simpson PJ, et al. AKR1C isoforms represent a novel cellular target for jasmonates alongside their mitochondrial-mediated effects. Cancer Res. 2009;69:4769–4775.

38. Traven K, Sinreich M, Stojan J, et al. Ruthenium complexes as inhibitors of the aldo–keto reductases AKR1C1–1C3. Chem Biol Interact. 2015;234:349–359.

39. Kljun J, Anko M, Traven K, et al. Pyrithione-based ruthenium complexes as inhibitors of aldo–keto reductase 1C enzymes and anticancer agents. Dalton Trans. 2016;45:11791–11800.

40. Zhao Y, Zheng X, Zhang H, et al. In vitro inhibition of AKR1Cs by sulphonylureas and the structural basis. Chem Biol Interact. 2015;240:310–315.

41. Kikuchi A, Furutani T, Azami H, et al. In vitro and in vivo characterisation of ASP9521: A novel, selective, orally bioavailable inhibitor of 17β-hydroxysteroid dehydrogenase type 5 (17βHSD5; AKR1C3). Invest New Drugs. 2014;32:860–870.

42. Loriot Y, Fizazi K, Jones RJ, et al. Safety, tolerability and anti-tumour activity of the androgen biosynthesis inhibitor ASP9521 in patients with metastatic castration-resistant prostate cancer: Multi-centre phase I/II study. Invest New Drug. 2014;32:995–1004.

43. Rižner TL, Penning TM. Aldo-keto reductase 1C3—Assessment as a new target for the treatment of endometriosis. Pharmacol Res. 2020;152:104446.

44. Endo S, Oguri H, Segawa J, et al. Development of novel AKR1C3 inhibitors as new potential treatment for castration-resistant prostate cancer. J Med Chem. 2020;63(18):10396–10411.

45. Adeniji AO, Twenter BM, Byrns MC, et al. Development of potent and selective inhibitors of aldo–keto reductase IC3 (type 5 17β-hydroxysteroid dehydrogenase) based on N-phenylamino-benzoates and their structure–activity relationships. J Med Chem. 2012;55:2311–2323.

46. Jamieson SM, Brooke DG, Heinrich D, et al. 3-(3,4-dihydroisocoumarin-2-(1 H)-ylsulfonyl) benzoic acids: Highly potent and selective inhibitors of the type 5 17β-hydroxysteroid dehydrogenase AKR1C3. J Med Chem. 2012;55:7746–7758.

47. Pippione AC, Giraudo A, Bonanni D, et al. Hydroxysteriazole derivatives as potent and selective aldo-keto reductase 1C3 (AKR1C3) inhibitors discovered by bioisosteric scaffold hopping approach. Eur J Med Chem. 2017b;139:936–946.

48. Endo S, Matsunaga T, Kanamori A, et al. Selective inhibition of human type-5 17β-hydroxysteroid dehydrogenase (AKR1C3) by baccharin, a component of Brazilian propolis. J Nat Prod. 2012;75:716–721.

49. Endo S, Hu D, Matsunaga T, et al. Synthesis of non-prenyl analogues of baccharin as selective and potent inhibitors for aldo-keto reductase 1C3. Bioorga Med Chem. 2014;22:5220–5233.

50. Verma K, Gupta N, Zang T, et al. AKR1C3 inhibitor KV-37 exhibits antineoplastic effects and potentiates enalaprilamide in combination therapy in prostate adenocarcinoma cells. Mol Cancer Ther. 2018;17:1833–1845.

51. Hara A, Endo S, Matsunaga T, Soda M, Yashiro K, El-Kabbani O. Long-chain fatty acids inhibit human members of the aldo-keto reductase 1C subfamily. J Biochem. 2017;162:371–379.

52. Chen M, Adeniji AO, Twenter BM, Winkler JD, Christianson DW, Penning TM. Crystal structures of AKR1C3 containing an N-(aryl) amino-benzoate inhibitor and a bifunctional AKR1C3 inhibitor and androgen receptor antagonist. Therapeutic leads for castrate resistant prostate cancer. Bioorg Med Chem Lett. 2012;22:3492–3497.

53. Meng EC, Pettersen EF, Couch GS, Huang CC, Ferrin TE. Tools for integrated sequence-structure analysis with UCSF Chimera. BMC Bioinform. 2006;7:1–10.

54. Penning TM. The aldo-keto reductases (AKRs): Overview. Chem Biol Interact. 2015;234:236–246.

55. Yadav DK, Sharma P, Misra S, et al. Studies of the benzopyanor class of selective COX-2 inhibitors using 3D-QSAR and molecular docking. Arch Pharm Res. 2018;41:1178–1189.

56. Yang Z, Lasker K, Schneidman-Duhovny D, et al. UCSF Chimera, MODELLER, and IMP: An integrated modeling system. J Struct Biol. 2012;179:269–278.

57. Friesner RA, Banks JL, Murphy RB, et al. Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem. 2004;47:1739–1749.

58. Schroedinger L. Schroedinger software suite. New York, NY: Schrödinger, LLC; 2011:670.

59. Flanagan JU, Yosaatmadja Y, Teague RM, Chai MZ, Turnbull AP, Squire CJ. Crystal structures of three classes of non-steroidal anti-inflammatory drugs in complex with aldo-keto reductase 1C3. PLoS One. 2012;7:e43965.

60. Schroedinger L. Release suite 2017-4 protein preparation wizard. New York: Epik, Schrödinger, LLC; 2017.

61. Roos K, Wu C, Damm W, et al. OPLS3e: Extending force field coverage for drug-like small molecules. J Chem Theory Comput. 2019;15:1863–1874.

62. Gill PM. Density functional theory (DFT), Hartree-Fock (HF), and the self-consistent field. J Chem Phys. 1994;100:5066–5075.

63. Bouzine SM, Bouzakraoui S, Bouachrine M, Hamidi M. Density functional theory (B3LYP/6-31G*) study of oligothiophenes in their aromatic and polaronic states. J Mol Struct-Theochem. 2005;726:271–276.

64. Frisch M, Trucks G, Schlegel HB, et al. Gaussian 09, Revision D.01. Wallingford, CT: Gaussian Inc, 2009.

65. Wang J, Cieplak P, Kollman PA. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? J Comput Chem. 2000;21:1049–1074.

66. Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. J Chem Theory Comput. 2015;11:3696–3713.

67. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. J Comput Chem. 2004;25:1157–1174.

68. Paschek D, Day R, García AE. Influence of water–protein hydrogen bonding on the stability of Trp-cage miniprotein. A comparison between the TIP3P and TIP4P-Ew water models. J Chem Theory Comput. 2015;11:19840–19847.

69. Pearlman DA, Case DA, Caldwell JW, et al. AMBER, a package of computer programs for applying molecular mechanics.
normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. Comput Phys Commun. 1995;91:1–41.

70. Salomon-Ferrer R, Case DA, Walker RC. An overview of the Amber biomolecular simulation package. WIRES Comput Mol Sci. 2013;3:198–210.

71. Lin Y, Pan D, Li J, Zhang L, Shao X. Application of berendsen barostat in dissipative particle dynamics for nonequilibrium dynamic simulation. J Chem Phys. 2017;146:124108.

72. Liu J, Li D, Liu X. A simple and accurate algorithm for path integral molecular dynamics with the Langevin thermostat. J Chem Phys. 2016;145:024103.

73. Harvey M, De Fabritiis G. An implementation of the smooth particle mesh Ewald method on GPU hardware. J Chem Theory Comput. 2009;5:2371–2377.

74. Elber R, Ruymgaart AP, Hess B. SHAKE parallelization. Eur Phys J Spec Top. 2011;200:211–223.

75. Roe DR, Cheatham TE III. Parallelization of CPPTRAJ enables large scale analysis of molecular dynamics trajectory data. J Comput Chem. 2018;39:2110–2117.

76. Li H, Chang Y-Y, Lee JY, Bahar I, Yang L-W. DynOmines: Dynamics of structural proteome and beyond. Nucleic Acids Res. 2017;45:E374–E380.

77. Tian W, Chen C, Lei X, Zhao J, Liang J. CASTp 3.0: Computed atlas of surface topography of proteins. Nucleic Acids Res. 2018;46:W363–W367.

78. Rastelli G, Rio AD, Degliesposti G, Sgobba M. Fast and accurate predictions of binding free energies using MM-PBSA and MM-GBSA. J Comput Chem. 2010;31:797–810.

79. Sun H, Li Y, Tian S, Xu L, Hou T. Assessing the performance of MM/PBSA and MM/GBSA methods. 4. Accuracies of MM/PBSA and MM/GBSA methodologies evaluated by various simulation protocols using PDBbind data set. Phys Chem Chem Phys. 2014;16:16719–16729.

80. Kong X, Pan P, Li D, Tian S, Li Y, Hou T. Importance of protein flexibility in ranking inhibitor affinities: Modeling the binding mechanisms of piperidine carboxamides as type I/2 ALK inhibitors. Phys Chem Chem Phys. 2015;17:6098–6113.

81. Wang E, Sun H, Wang J, et al. End-point binding free energy calculation with MM/PBSA and MM/GBSA: Strategies and applications in drug design. Chem Rev. 2019;119:9478–9508.

82. Onufriev A, Bashford D, Case DA. Exploring protein native states and large-scale conformational changes with a modified generalized born model. Proteins. 2004;55:383–394.

83. Weiser J, Shenkin PS, Still WC. Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO). J Comput Chem. 1999;20:217–230.

84. Frederick KK, Marlow MS, Valentine KG, Wand AJ. Conformational entropy in molecular recognition by proteins. Nature. 2007;448:325–329.

85. Tzeng S-R, Kalodimos CG. Protein activity regulation by conformational entropy. Nature. 2012;488:236–240.

86. Miller BR III, Mcgee TD Jr, Swails JM, Homeyer N, Gohlke H, Roitberg AE. MM-PBSA. py: An efficient program for end-state free energy calculations. J Chem Theory Comput. 2012;8:3314–3321.

87. Elhefnawy W, Chen L, Han Y, Li Y. ICOASA: A distance-dependent, orientation-specific coarse-grained contact potential for protein structure modeling. J Mol Biol. 2015;427:2562–2576.

88. Release S. 2: LigPrep. New York, NY: Schrödinger, LLC; 2017.

89. Kalva S, Vinod D, Saleema LM. Field-and Gaussian-based 3D-QSAR studies on barbiturate analogs as MMP-9 inhibitors. Med Chem Res. 2013;22:5303–5313.

90. Lindgren F, Rännar S. Alternative partial least-squares (PLS) algorithms. Perspect Drug Discov. 1998;12:14:105–113.

91. Warmflash A, Bhimalapuram P, Dinner AR. Umbrella sampling for nonequilibrium processes. J Chem Phys. 2007;127:114109.

92. Kästner J. Umbrella sampling. WIREs Comput Mol Sci. 2011;1:932–942.

93. Kong X, Sun H, Pan P, et al. Molecular principle of the cyclin-dependent kinase selectivity of 4-(thiazol-5-yl)-2-(phenylamino) pyrimidine-5-carbonitrile derivatives revealed by molecular modeling studies. Phys Chem Chem Phys. 2016;18:2034–2046.

94. Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. J Comput Chem. 1992;13:1011–1021.

95. Souaille M, Roux B. Extension to the weighted histogram analysis method: Combining umbrella sampling with free energy calculations. Comput Phys Commun. 2001;135:40–57.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Kong X, Xing E, Wu S, Zhuang T, Li P-K, Li C, et al. Computational modeling studies reveal the origin of the binding preference of 3-(3,4-di hydrosoquinolin-2 (1H)-ylsulfonyl)benzoic acids for AKR1C3 over its isoforms. Protein Science. 2022;31(12):e4499. https://doi.org/10.1002/pro.4499