



Computational Analysis of Residue-Specific Binding Free Energies of Androgen Receptor to Ligands

Guangfeng Shao¹, Jingxiao Bao¹, Xiaolin Pan¹, Xiao He^{1,2*}, Yifei Qi^{1,2*} and John Z. H. Zhang^{1,2,3*}

¹Shanghai Engineering Research Center of Molecular Therapeutics and New Drug Development, Shanghai Key Laboratory of Green Chemistry and Chemical Process, School of Chemistry and Molecular Engineering, East China Normal University, Shanghai, China, ²NYU-ECNU Center for Computational Chemistry at NYU, Shanghai, China, ³Department of Chemistry, New York University, New York, NY, United States

OPEN ACCESS

Edited by:

Guo-Wei Wei,
Michigan State University,
United States

Reviewed by:

Tingjun Hou,
Zhejiang University, China
Kaifu Gao,
Michigan State University,
United States

*Correspondence:

John Z. H. Zhang
john.zhang@nyu.edu
Xiao He
xiaohe@phy.ecnu.edu.cn
Yifei Qi
yfqi@chem.ecnu.edu.cn

Specialty section:

This article was submitted to
Biological Modeling and Simulation,
a section of the journal
Frontiers in Molecular Biosciences

Received: 27 December 2020

Accepted: 14 January 2021

Published: 12 March 2021

Citation:

Shao G, Bao J, Pan X, He X, Qi Y and Zhang JZH (2021) Computational Analysis of Residue-Specific Binding Free Energies of Androgen Receptor to Ligands.
Front. Mol. Biosci. 8:646524.
doi: 10.3389/fmolb.2021.646524

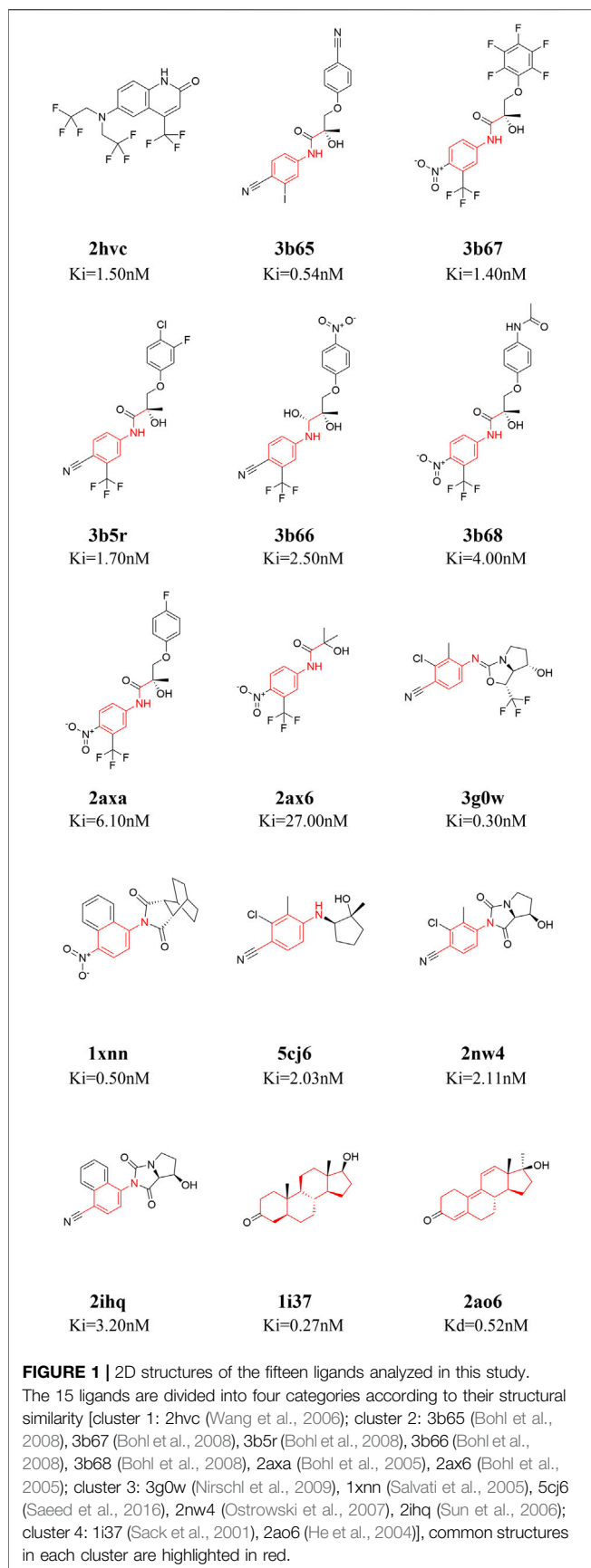
Androgen receptor (AR) is an important therapeutic target for the treatment of diseases such as prostate cancer, hypogonadism, muscle wasting, etc. In this study, the complex structures of the AR ligand-binding domain (LBD) with fifteen ligands were analyzed by molecular dynamics simulations combined with the alanine-scanning-interaction-entropy method (ASIE). The quantitative free energy contributions of the pocket residues were obtained and hotspot residues are quantitatively identified. Our calculation shows that these hotspot residues are predominantly hydrophobic and their interactions with binding ligands are mainly van der Waals interactions. The total binding free energies obtained by summing over binding contributions by individual residues are in good correlation with the experimental binding data. The current quantitative analysis of binding mechanism of AR to ligands provides important insight on the design of future inhibitors.

Keywords: androgen receptor, alanine scanning, interaction entropy, MD simulation, hotspot, GBSA

INTRODUCTION

Androgen receptor (AR) is an important target for many diseases including prostate cancer, hypogonadism, muscle wasting, osteoporosis, and benign prostate hyperplasia (Chen et al., 2004; Gao and Dalton, 2007; Dillon et al., 2010; Tan et al., 2015; Li et al., 2019). AR is expressed in many tissues, including prostate, seminal vesicle, testis, epididymis, adrenal gland, skin, skeletal muscle and central nervous system (CNS). Androgen receptor ligands can be divided into androgens (agonists) and antiandrogens (antagonists) depending on whether they activate or inhibit the transcription of AR target genes, or by ligand structure into steroidal and non-steroidal (Gao et al., 2005). Because of the rigid skeleton of steroidal compounds, the majority of recently developed ligands are non-steroidal ligands. Like other nuclear receptors, the androgen receptor is modular in structure and is composed of a N-terminal domain (NTD), a DNA binding domain (DBD), a hinge region, and a C-terminal ligand-binding domain (LBD) (Brinkmann et al., 1989). Most clinically used antiandrogens, such as flutamide (Goldspiel and Kohler, 1990), nilutamide (Kassouf et al., 2003), bicalutamide (Goa and Spencer, 1998), enzalutamide (Nadiminty et al., 2013), apalutamide (Rathkopf et al., 2013), and darolutamide (Fizazi et al., 2018), target LBD.

Understanding of protein-ligand interaction and the quantitative characterization of binding affinity are very important for the discovery, design, and development of drugs (Cheng et al., 2007;



Gilson and Zhou, 2007; Jorgensen and Thomas, 2008; Sun et al., 2013; Hu et al., 2020). Although experiments can study the thermodynamic properties of protein-ligand binding, determination of binding affinity is time-consuming, laborious and expensive. Computationally, the molecular mechanics generalized born surface area (MM/GBSA) method is often used to calculate the free energy of protein ligand binding (Massova and Kollman, 1999; Kollman et al., 2000; Moreira et al., 2007; Genheden and Ryde, 2015). Recently we developed a method called interaction entropy (IE) for practical and efficient calculation of entropy in protein-ligand and protein-protein binding (Duan et al., 2016). This method has been used in combination with alanine scanning (Massova and Kollman, 1999) (AS) and MM/GBSA to obtain the residue-specific contribution of each pocket residue (ASE method) (Yan et al., 2017; Liu et al., 2018; Qiu et al., 2018; Zhou et al., 2018; He et al., 2019; Yang et al., 2019; Wang et al., 2020).

In this study, fifteen AR ligands were analyzed with the ASIE method to quantitatively characterize the detailed protein-ligand interactions, and the contribution of key binding residues on AR were identified. In addition, there is a strong correlation between the sum of the contributions of residues and the experimental binding free energy.

METHODS

Molecular Dynamics Simulations

In this study, 15 androgen receptor systems with experimental k_i/k_d values and complex structures in the Protein Data Bank (Berman et al., 2000) (PDB) were used for MD simulation (Hamann et al., 1998; He et al., 2004; Bohl et al., 2005; Salvati et al., 2005; Sun et al., 2006; van Oeveren et al., 2006; Ostrowski et al., 2007; Bohl et al., 2008; Nirschl et al., 2009; Saeed et al., 2016). The systems were simulated using pmemd.cuda (Case et al., 2005) in AMBER18 (Pearlman et al., 1995; Case et al., 2019) with the ff14SB force field (Maier et al., 2015). For each system, TIP3P (Jorgensen et al., 1983) water model and periodic boundary conditions were used to solvate the complex, and the minimum distance between solute atoms and periodic boundary was set to 12 Å. Furthermore, Sodium and chloride ions were added to neutralize the system. A two-step minimization process was carried out, where only hydrogen atoms were optimized in the first step, and all atoms were optimized in the second step. The system was then slowly heated to 300 K with Langevin dynamics temperature regulation, followed by an equilibration of 500 ps. Finally, 10-ns MD simulations were carried out in an NPT ensemble and 25,000 snapshots were saved for further analysis. Five independent replicates were simulated for each system, and the final results were obtained by averaging the calculated values on the five trajectories.

Binding Free Energy Calculation

We mutated a specific amino acid to alanine, assuming that the mutated alanine contributed little to the binding free energy, and calculated the binding free energy difference before and after the mutation. The free energy difference of a residue x mutating to ALA is defined as:

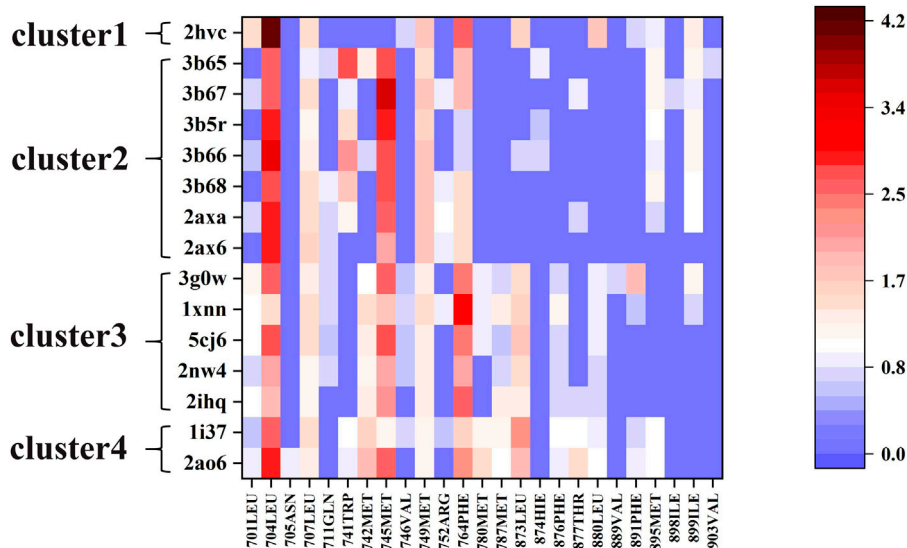


FIGURE 2 | The free energy contributions of each pocket residue in the 15 systems.

TABLE 1 | Residue-Specific Binding Free Energies (kcal/mol) of 1i37 (Sack et al., 2001) calculated by ASIE.

Residue	$\Delta\Delta E_{\text{vdw}}$	$\Delta\Delta E_{\text{ele}}$	$\Delta\Delta G_{\text{B}}$	$\Delta\Delta N_{\text{P}}$	$\Delta\Delta H$	IE	$\Delta\Delta G$
704LEU	2.39 ± 0.20	-0.01 ± 0.03	0.39 ± 0.11	0.13 ± 0.02	2.89 ± 0.29	-0.25 ± 0.10	2.64 ± 0.20
873LEU	1.81 ± 0.10	-0.05 ± 0.00	0.53 ± 0.04	0.15 ± 0.01	2.44 ± 0.15	-0.18 ± 0.02	2.25 ± 0.13
764PHE	2.11 ± 0.18	0.64 ± 0.05	-0.08 ± 0.08	0.10 ± 0.01	2.77 ± 0.16	-1.09 ± 0.16	1.68 ± 0.17
742MET	2.00 ± 0.11	0.44 ± 0.02	-0.73 ± 0.08	0.16 ± 0.01	1.87 ± 0.07	-0.25 ± 0.03	1.61 ± 0.06
707LEU	1.53 ± 0.08	0.07 ± 0.02	-0.12 ± 0.02	0.08 ± 0.01	1.57 ± 0.09	-0.15 ± 0.03	1.42 ± 0.08
745MET	1.35 ± 0.08	-0.03 ± 0.06	-0.13 ± 0.08	0.11 ± 0.01	1.29 ± 0.09	-0.09 ± 0.03	1.21 ± 0.09
780MET	1.22 ± 0.11	0.16 ± 0.05	-0.18 ± 0.11	0.12 ± 0.01	1.31 ± 0.06	-0.13 ± 0.01	1.19 ± 0.07
749MET	1.14 ± 0.10	0.26 ± 0.06	0.00 ± 0.08	0.03 ± 0.00	1.43 ± 0.10	-0.25 ± 0.06	1.17 ± 0.09
787MET	1.13 ± 0.06	0.16 ± 0.04	-0.14 ± 0.03	0.08 ± 0.00	1.23 ± 0.06	-0.10 ± 0.01	1.13 ± 0.07
877THR	0.08 ± 0.12	1.77 ± 0.05	-0.51 ± 0.03	0.05 ± 0.00	1.39 ± 0.11	-0.31 ± 0.02	1.07 ± 0.10
741TRP	1.22 ± 0.18	0.12 ± 0.01	-0.14 ± 0.02	0.07 ± 0.01	1.27 ± 0.21	-0.27 ± 0.05	1.00 ± 0.17
895MET	1.01 ± 0.09	0.03 ± 0.04	0.08 ± 0.05	0.08 ± 0.01	1.18 ± 0.08	-0.19 ± 0.02	1.00 ± 0.10
876PHE	1.04 ± 0.05	-0.03 ± 0.03	0.00 ± 0.05	0.08 ± 0.00	1.10 ± 0.03	-0.11 ± 0.01	0.99 ± 0.04
880LEU	0.68 ± 0.04	0.06 ± 0.02	0.15 ± 0.02	0.05 ± 0.00	0.94 ± 0.05	-0.10 ± 0.01	0.84 ± 0.05
746VAL	0.64 ± 0.07	-0.03 ± 0.01	0.26 ± 0.07	0.03 ± 0.01	0.90 ± 0.14	-0.06 ± 0.01	0.84 ± 0.13
891PHE	0.64 ± 0.07	0.50 ± 0.03	-0.24 ± 0.06	0.02 ± 0.00	0.92 ± 0.15	-0.09 ± 0.01	0.83 ± 0.14
701LEU	0.55 ± 0.05	0.05 ± 0.00	0.12 ± 0.02	0.04 ± 0.01	0.76 ± 0.07	-0.07 ± 0.01	0.69 ± 0.08
752ARG	0.56 ± 0.04	1.83 ± 0.07	-1.68 ± 0.05	0.05 ± 0.00	0.76 ± 0.03	-0.11 ± 0.01	0.65 ± 0.03
TOTAL	21.1 ± 0.19	5.93 ± 0.08	-2.43 ± 0.35	1.43 ± 0.02	26.02 ± 0.34	-3.80 ± 0.17	22.22 ± 0.5

$$\begin{aligned}\Delta\Delta G_{\text{bind}}^{x \rightarrow a} &= \Delta G_{\text{bind}}^a - \Delta G_{\text{bind}}^x \\ &= \Delta\Delta G_{\text{gas}}^{x \rightarrow a} + \Delta\Delta G_{\text{sol}}^{x \rightarrow a}\end{aligned}\quad (1)$$

where the gas-phase:

$$\Delta\Delta G_{\text{gas}}^{x \rightarrow a} = \Delta G_{\text{gas}}^a - \Delta G_{\text{gas}}^x \quad (2)$$

and solvation components:

$$\Delta\Delta G_{\text{sol}}^{x \rightarrow a} = \Delta G_{\text{sol}}^a - \Delta G_{\text{sol}}^x \quad (3)$$

In the IE (interaction entropy) approach (Duan et al., 2016), the gas-phase component is computed by

$$\begin{aligned}\Delta G_{\text{gas}}^x &= \langle E_{\text{int}}^x \rangle - T\Delta S_{\text{int}}^x \\ &= \langle E_{\text{int}}^x \rangle + KT \ln \langle e^{\beta\Delta E_{\text{int}}^x} \rangle\end{aligned}\quad (4)$$

And

$$\Delta G_{\text{gas}}^a = \langle E_{\text{int}}^a \rangle + KT \ln \langle e^{\beta\Delta E_{\text{int}}^a} \rangle \quad (5)$$

Where E_{int}^x and E_{int}^a contain the electrostatic and van der Waals energies between the ligand and residues x and Ala. The exponential average was evaluated by discrete time averaging.

$$\langle e^{\beta\Delta E_{\text{int}}^x} \rangle = \frac{1}{N} \sum_{i=1}^N e^{\beta\Delta E_{\text{int}}^x(t_i)} \quad (6)$$

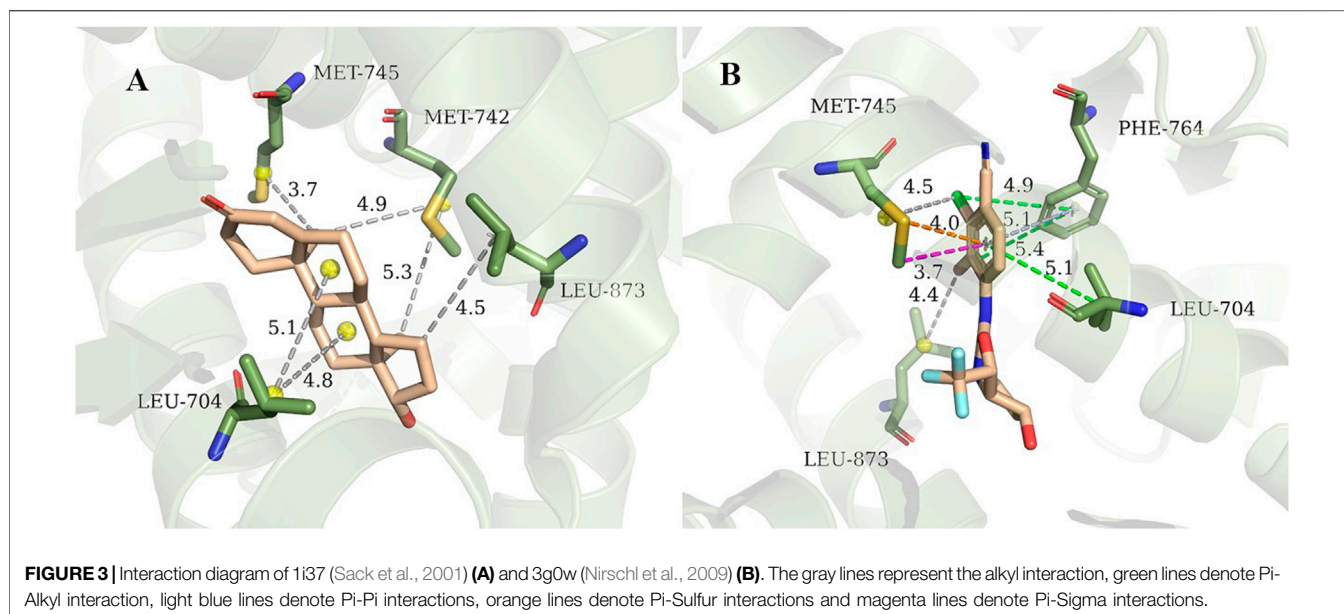


TABLE 2 | Residue-Specific Binding Free Energies (kcal/mol) of 3g0w (Nirschl et al., 2009) calculated by ASIE.

Residue	$\Delta\Delta E_{vdw}$	$\Delta\Delta E_{ele}$	$\Delta\Delta G_B$	$\Delta\Delta NP$	$\Delta\Delta H$	IE	$\Delta\Delta G$
745MET	3.58 ± 0.08	0.45 ± 0.03	-1.25 ± 0.13	0.20 ± 0.02	2.98 ± 0.11	-0.31 ± 0.05	2.67 ± 0.10
704LEU	2.39 ± 0.08	0.19 ± 0.01	0.21 ± 0.06	0.06 ± 0.00	2.85 ± 0.06	-0.24 ± 0.02	2.60 ± 0.07
764PHE	2.46 ± 0.08	0.00 ± 0.01	0.20 ± 0.03	0.08 ± 0.00	2.75 ± 0.08	-0.35 ± 0.03	2.39 ± 0.07
891PHE	1.12 ± 0.03	0.83 ± 0.02	0.13 ± 0.08	0.07 ± 0.01	2.15 ± 0.07	-0.22 ± 0.03	1.93 ± 0.07
873LEU	1.33 ± 0.04	0.03 ± 0.00	0.26 ± 0.02	0.11 ± 0.00	1.73 ± 0.05	-0.20 ± 0.02	1.54 ± 0.05
707LEU	1.32 ± 0.07	0.02 ± 0.01	0.05 ± 0.02	0.07 ± 0.00	1.47 ± 0.08	-0.15 ± 0.01	1.32 ± 0.08
749MET	1.21 ± 0.03	0.45 ± 0.02	-0.16 ± 0.04	0.02 ± 0.00	1.51 ± 0.06	-0.24 ± 0.02	1.27 ± 0.06
899ILE	0.57 ± 0.02	-0.01 ± 0.00	0.60 ± 0.05	0.07 ± 0.00	1.24 ± 0.06	-0.03 ± 0.00	1.20 ± 0.06
701LEU	0.65 ± 0.02	0.10 ± 0.01	0.49 ± 0.06	0.06 ± 0.01	1.30 ± 0.07	-0.15 ± 0.04	1.15 ± 0.03
742MET	2.18 ± 0.11	0.32 ± 0.06	-1.25 ± 0.09	0.16 ± 0.01	1.41 ± 0.17	-0.40 ± 0.07	1.01 ± 0.21
880LEU	0.69 ± 0.04	0.06 ± 0.01	0.22 ± 0.03	0.06 ± 0.00	1.03 ± 0.07	-0.08 ± 0.01	0.94 ± 0.06
780MET	1.37 ± 0.06	1.27 ± 0.11	-1.33 ± 0.09	0.12 ± 0.01	1.43 ± 0.05	-0.52 ± 0.15	0.91 ± 0.15
889VAL	0.13 ± 0.01	0.00 ± 0.00	0.72 ± 0.05	0.00 ± 0.00	0.85 ± 0.06	0.00 ± 0.00	0.84 ± 0.06
787MET	0.93 ± 0.03	-0.04 ± 0.02	-0.07 ± 0.01	0.05 ± 0.00	0.87 ± 0.03	-0.04 ± 0.00	0.83 ± 0.03
711GLN	0.91 ± 0.08	0.30 ± 0.06	-0.33 ± 0.03	0.07 ± 0.00	0.95 ± 0.05	-0.13 ± 0.02	0.82 ± 0.07
876PHE	1.00 ± 0.04	-0.30 ± 0.01	0.11 ± 0.05	0.07 ± 0.00	0.88 ± 0.06	-0.11 ± 0.01	0.77 ± 0.06
746VAL	0.59 ± 0.03	-0.07 ± 0.00	0.17 ± 0.04	0.02 ± 0.00	0.71 ± 0.05	-0.06 ± 0.02	0.65 ± 0.04
TOTAL	22.45 ± 0.15	3.59 ± 0.08	-1.22 ± 0.18	1.28 ± 0.02	26.1 ± 0.32	-3.24 ± 0.35	22.86 ± 0.37

After that, Eq. 2 becomes:

$$\begin{aligned} \Delta\Delta G_{gas}^{x \rightarrow a} &= \Delta\Delta E_{gas}^{x \rightarrow a} - T\Delta\Delta S_{gas}^{x \rightarrow a} \\ &= \langle E_{int}^a \rangle - \langle E_{int}^x \rangle + KT [\ln \langle e^{\beta\Delta E_{int}^a} \rangle - \ln \langle e^{\beta\Delta E_{int}^x} \rangle] \end{aligned} \quad (7)$$

The solvation free energy was calculated by the MM/GBSA method:

$$\Delta G_{sol} = \Delta G_{gb} + \Delta G_{np} \quad (8)$$

The polarization part ΔG_{gb} is obtained by the generalized Born (GB) model. The non-polar term ΔG_{np} can be obtained by using the solvent accessible surface area (SASA) formula:

$$\Delta G_{np} = \gamma SASA + \beta \quad (9)$$

Finally, the free binding energy of protein ligands can be expressed as (Liu et al., 2018; Zhou et al., 2018)

$$\Delta G_{bind} = -\sum_x \Delta\Delta G_{bind}^{x \rightarrow a} \quad (10)$$

For each system, 25,000 frames were extracted from the entire 10-ns trajectory at an interval of 400 fs for IE calculation. Hundred frames were uniformly extracted from the 25,000 frames for MM/GBSA calculation with “igb” set to 8 (Ryckaert et al., 1977; Nguyen et al., 2013) because igb = 8 is the latest GB model, and Nguyen et al. proved that the GB-Neck2 model (igb = 8) shows significant improvement in solvation energy and effective radii calculation as compared to GB-OBC (igb = 2, igb = 5) and GB-Neck (igb = 7) (Jorgensen et al., 1983). Following

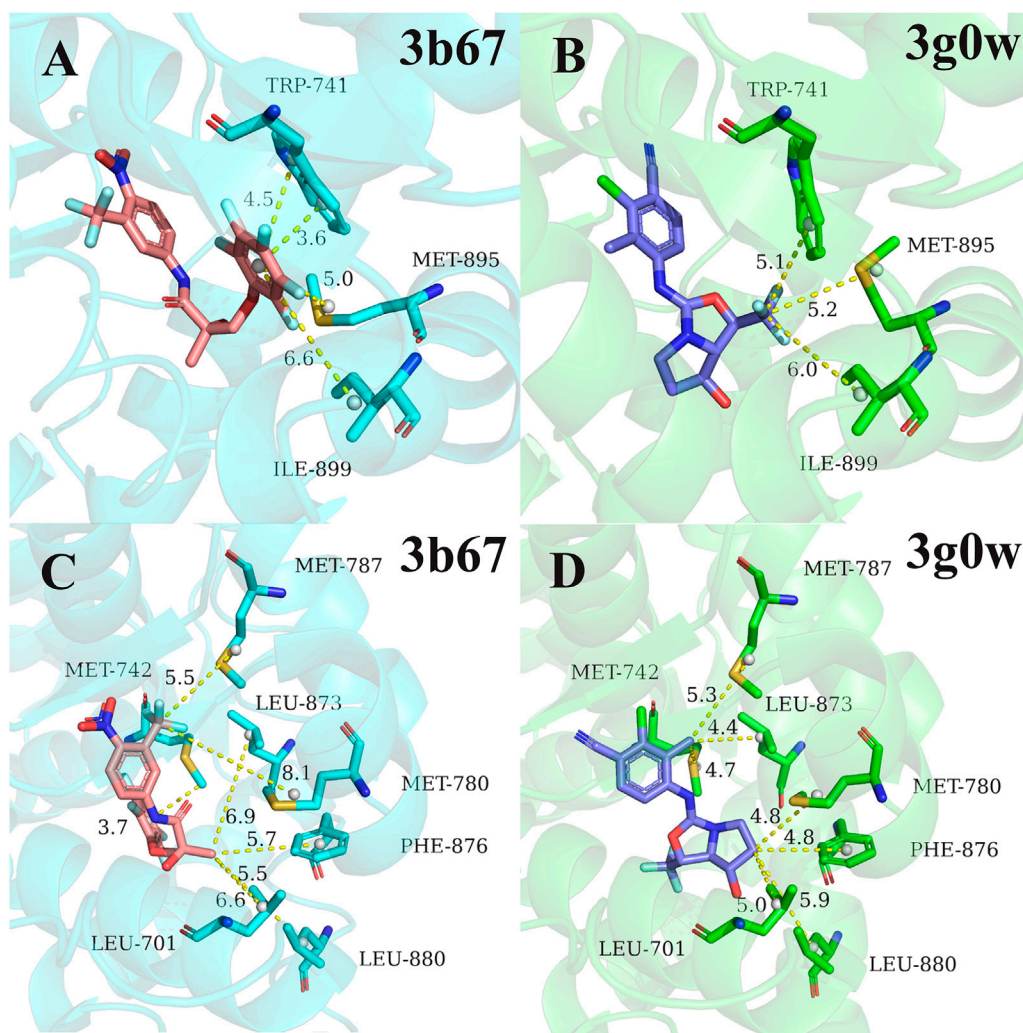


FIGURE 4 | Comparison of pocket residues in 3b67 (Bohl et al., 2008) (cluster2) and 3g0w (Nirschl et al., 2009) (cluster3). The relative position of residues 741TRP, 895MET, 899ILE in 3b67 **(A)** and 3g0w **(B)** with the ligand in the crystal structure. The relative position of the residues 701LEU, 742MET, 780MET, 787MET, 873LEU, 876PHE, 880LEU in the crystal structure to the ligand in 3b67 **(C)** and 3g0w **(D)**. The dashed lines and numbers indicated distance (in Å) between the specific groups of the ligands and residues.

our previous protocol, the dielectric constant for nonpolar, polar, and charged residues are 1, 3, 5, respectively (Hou et al., 2011; Petukh et al., 2015). We also calculated binding energy with the conventional MM/GBSA method for comparison, where the dielectric constant was set to one for all residues. The free energy and its standard deviation were obtained from five free energy values calculated from five independent trajectories.

RESULTS AND DISCUSSION

Fifteen ligands binding to AR-LBD systems with experimental k_i/k_d data were used for the binding energy calculations [PDB ID: 1i37 (Sack et al., 2001), 1xnn (Salvati et al., 2005), 2a06 (He et al., 2004), 2ax6 (Bohl et al., 2005), 2axa (Bohl et al., 2005), 2hvc

(Wang et al., 2006), 2ihq (Sun et al., 2006), 2nw4 (Ostrowski et al., 2007), 3b5r (Bohl et al., 2008), 3b65 (Bohl et al., 2008), 3b66 (Bohl et al., 2008), 3b67 (Bohl et al., 2008), 3b68 (Bohl et al., 2008), 3g0w (Nirschl et al., 2009), 5cj6 (Saeed et al., 2016)]. The 2D structures of these 15 small molecules are shown in **Figure 1**. The 15 molecules are divided into four categories according to their core structures (cluster 1: 2hvc; cluster 2: 3b65, 3b67, 3b5r, 3b66, 3b68, 2axa, 2ax6, these compounds share a N-ethylaniline; cluster 3: 3g0w, 1xnn, 5cj6, 2nw4, 2ihq, these compounds share a p-toluidine; cluster 4: 1i37, 2a06, both of these are steroidal compounds, **Figure 1**). There are two steroidal compounds, where the ligand in 1i37 is the natural AR agonist dihydrotestosterone and ligand in 2a06 is the synthetic steroid agonist. These two compounds have strong binding affinities with AR in the experiment (He et al., 2004; Bohl et al., 2008).

TABLE 3 | Free energy components of the binding energies (kcal/mol).

PDB-ID	ΔE_{vdw}	ΔE_{ele}	ΔG_B	ΔNP	ΔH	IE	ΔG	ΔG (exp) ^a
2hvc	12.08 ± 0.19	0.66 ± 0.23	6.88 ± 0.76	0.75 ± 0.01	20.37 ± 0.81	-1.78 ± 0.05	18.59 ± 0.78	12.05
3b65	23.23 ± 0.60	3.75 ± 0.31	-5.16 ± 0.18	1.33 ± 0.02	23.15 ± 0.31	-4.58 ± 0.49	18.57 ± 0.60	12.65
3b67	19.77 ± 0.17	6.38 ± 0.34	-5.91 ± 0.43	1.15 ± 0.04	21.39 ± 0.34	-3.75 ± 0.25	17.64 ± 0.51	12.09
3b5r	17.83 ± 0.52	3.25 ± 0.14	-4.46 ± 0.21	0.92 ± 0.03	17.55 ± 0.34	-3.71 ± 0.25	13.84 ± 0.32	11.97
3b66	23.26 ± 0.27	4.37 ± 0.15	-6.30 ± 0.29	1.48 ± 0.03	22.81 ± 0.39	-5.27 ± 0.34	17.55 ± 0.46	11.75
3b68	19.27 ± 0.32	5.70 ± 0.27	-6.57 ± 0.35	0.99 ± 0.03	19.39 ± 0.40	-3.48 ± 0.33	15.91 ± 0.33	11.47
2axa	20.03 ± 0.28	5.62 ± 0.15	-5.96 ± 0.13	1.17 ± 0.03	20.87 ± 0.24	-4.37 ± 0.27	16.49 ± 0.09	11.22
2ax6	11.83 ± 0.32	3.14 ± 0.15	-2.48 ± 0.28	0.54 ± 0.04	13.03 ± 0.20	-1.44 ± 0.15	11.60 ± 0.24	10.33
3g0w	22.45 ± 0.15	3.59 ± 0.08	-1.22 ± 0.18	1.28 ± 0.02	26.10 ± 0.32	-3.24 ± 0.35	22.86 ± 0.37	13.00
1xnn	24.32 ± 0.28	4.01 ± 1.02	-4.19 ± 0.64	1.28 ± 0.02	25.42 ± 0.69	-3.64 ± 0.18	21.78 ± 0.77	12.70
5cj6	18.78 ± 0.11	1.68 ± 0.12	0.28 ± 0.18	1.10 ± 0.01	21.84 ± 0.21	-2.31 ± 0.10	19.52 ± 0.25	11.87
2nw4	19.55 ± 0.25	1.73 ± 0.14	-3.75 ± 0.14	0.98 ± 0.03	18.50 ± 0.31	-2.83 ± 0.18	15.68 ± 0.42	11.85
2ihq	19.19 ± 0.26	2.71 ± 0.14	-3.85 ± 0.12	0.97 ± 0.01	19.03 ± 0.12	-2.50 ± 0.16	16.52 ± 0.20	11.60
1i37	21.10 ± 0.19	5.93 ± 0.08	-2.43 ± 0.35	1.43 ± 0.02	26.02 ± 0.34	-3.80 ± 0.17	22.22 ± 0.50	13.07
2ao6	21.10 ± 0.27	6.59 ± 0.10	-1.16 ± 0.10	1.22 ± 0.01	27.75 ± 0.33	-2.94 ± 0.09	24.81 ± 0.41	12.68

^aThe experimental value is obtained by using the relation ΔG (exp) = $-RT \ln Ki/Kd$ at $T = 298K$ and multiplied by a minus sign.

TABLE 4 | Binding Free Energies (kcal/mol) calculated by ASIE method compared to the experimental data (kcal/mol).

PDB-ID	ΔH	ASIE		MM/GBSA ^a		EXP ^b
		- ΔS	ΔG	ΔH	ΔG (exp)	
2hvc	20.37 ± 0.81	-1.78 ± 0.05	18.59 ± 0.78	8.95 ± 0.83	12.05	
3b65	23.15 ± 0.31	-4.58 ± 0.49	18.57 ± 0.60	48.65 ± 0.19	12.65	
3b67	21.39 ± 0.34	-3.75 ± 0.25	17.64 ± 0.51	38.58 ± 0.70	12.09	
3b5r	17.55 ± 0.34	-3.71 ± 0.25	13.84 ± 0.32	40.28 ± 0.80	11.97	
3b66	22.81 ± 0.39	-5.27 ± 0.34	17.55 ± 0.46	40.23 ± 0.57	11.75	
3b68	19.39 ± 0.40	-3.48 ± 0.33	15.91 ± 0.33	42.37 ± 1.18	11.47	
2axa	20.87 ± 0.24	-4.37 ± 0.27	16.49 ± 0.09	38.14 ± 0.51	11.22	
2ax6	13.03 ± 0.20	-1.44 ± 0.15	11.60 ± 0.24	29.56 ± 0.40	10.33	
3g0w	26.10 ± 0.32	-3.24 ± 0.35	22.86 ± 0.37	35.02 ± 0.89	13.00	
1xnn	25.42 ± 0.69	-3.64 ± 0.18	21.78 ± 0.77	44.45 ± 0.57	12.70	
5cj6	21.84 ± 0.21	-2.31 ± 0.10	19.52 ± 0.25	37.77 ± 0.55	11.87	
2nw4	18.50 ± 0.31	-2.83 ± 0.18	15.68 ± 0.42	40.59 ± 0.67	11.85	
2ihq	19.03 ± 0.12	-2.50 ± 0.16	16.52 ± 0.20	41.39 ± 0.48	11.60	
1i37	26.02 ± 0.34	-3.80 ± 0.17	22.22 ± 0.50	47.24 ± 0.52	13.07	
2ao6	27.75 ± 0.33	-2.94 ± 0.09	24.81 ± 0.41	48.79 ± 0.27	12.68	
R	0.87		0.85	0.34		
MAE	9.53		6.22	27.19		
RMSE	10.03		6.85	28.31		

^aMM/GBSA results added a minus sign.

^bThe experimental value is obtained by using the relation ΔG (exp) = $-RT \ln Ki/Kd$ at $T = 298 K$ and multiplied by a minus sign.

Hot-Spots Residues in the Complex Structures

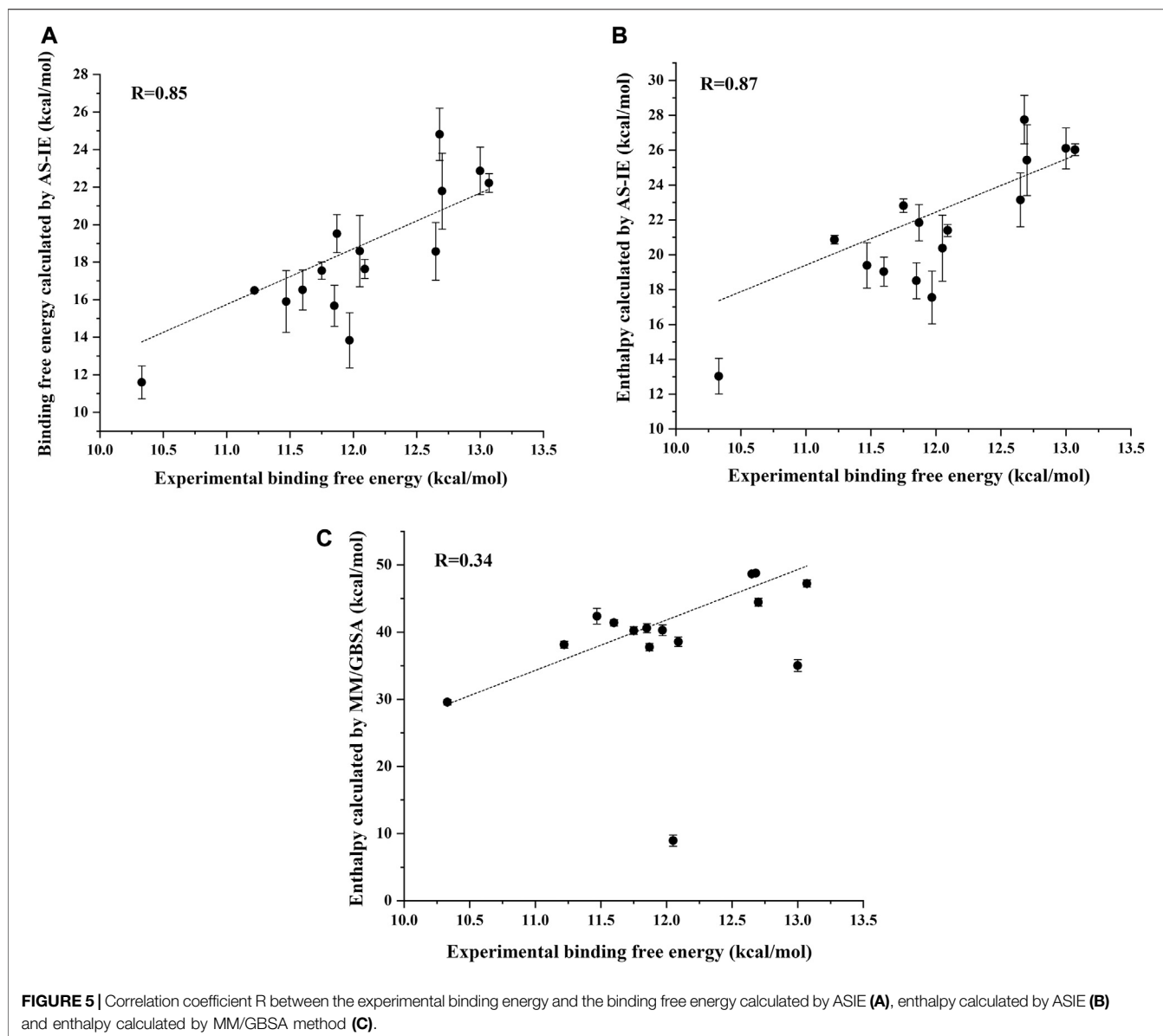
We used the ASIE method to quantitatively analyze the contribution of each pocket residue on AR when binds to different ligands. **Figure 2** shows the specific binding free energy values of the pocket residues in the 15 systems. First of all, 704LEU, 707LEU, 745MET, 749MET and 764PHE contributed much more to the binding free energy than other residues in these 15 complexes, which are identified as hotspot residues. Secondly, there are clear differences in the contribution of residues to binding energy in these four types of ligands. For example, in cluster 2, residues 741TRP, 895MET, and 899ILE generally have higher binding energy contribution compared to

those in cluster 3. In cluster 3, residues 701LEU, 742MET, 780MET, 787MET, 873LEU, 876PHE, and 880LEU have a more prominent binding energy contribution compared to cluster 2.

3g0w (Nirschl et al., 2009) and 1i37 (Sack et al., 2001) from cluster3 and cluster4 are selected for detailed analysis of hot spot residues, because these two systems have the strongest experimental value of binding free energy (Bohl et al., 2008; Nirschl et al., 2009), and the ligands belong to non-steroidal and steroid respectively. In 1i37, 704LEU, 873LEU, 764PHE, 742MET, 745MET, 780MET, 749MET, 787MET, 877THR, 741TRP, and 895MET contribute more than 1 kcal/mol to the total binding energy and are identified as hot residues (**Table 1**). The crystal structure of 1i37 shows that, LEU704, which is the major contributing residue, forms two Alkyl interactions with the ligand dihydrotestosterone at a distance of 4.8 Å and 5.1 Å (**Figure 3A**).

For 3g0w, ASIE calculation shows that 745MET, 704LEU, 764PHE, 891PHE, 873LEU, 707LEU, 749MET, 899ILE, 701LEU, and 742MET can be identified as hot residues with a binding free energy contribution of 2.67, 2.60, 2.39, 1.93, 1.54, 1.32, 1.27, 1.20, 1.15, and 1.01 kcal/mol (**Table 2**). The crystal structure of 3g0w suggests that the interaction between the ligand and the surrounding residues is more abundant than that of 1i37, including the interaction of alkyl, Pi-Alkyl, Pi-Pi, Pi-Sulfur and Pi-Sigma (**Figure 3B**). The strong contribution of 745MET in calculation is well rationalized in the crystal structure, which shows Pi-Sulfur (4.0 Å), Pi-sigma (3.7 Å) and alkyl (4.5 Å) interactions with the ligand.

Next, for the different mechanism of those four types of ligands binding to androgen receptor, we analyzed cluster2 and cluster3, among which the most obvious difference was found. As can be seen from **Figure 2**, residues 741TRP, 895MET and 899ILE in cluster2 has more binding free energy contribution than those in cluster3. **Figure 4A,B** respectively show the relative positions of these residues and ligands in the crystal structure of 3b67 (Cluster2) and 3g0w (Cluster3). It can be seen that the position of ligands in 3b67 is closer to these three



residues than that in 3g0w. In cluster3, 701LEU, 742MET, 780MET, 787MET, 873LEU, 876PHE and 880LEU generally have a stronger binding free energy contribution than those in cluster2. In addition, **Figure 4C,D** suggest that these residues are generally closer to the ligand in 3g0w (cluster3), which is in good agreement with our calculation results.

The quantitative analysis of these residues specific binding energies provides important clues to design high affinity AR LBD ligands. As shown in **Figure 2**, the substituents of 2ax6 are shorter than other ligands in Cluster2, AR LBD has fewer residues to interact with it, and its binding free energy is also the lowest in Cluster2. 704LEU has a strong binding free energy contribution in all 15 systems, and in 2hvc, 704LEU has a very high binding energy contribution. This is due to its interaction with the benzene ring on the 2hvc ligand, and the trifluoromethyl on the ligand also interacts with it. Due to its polycyclic structure, cluster4's two steroidal

compounds can interact with more residues in AR LBD than the other three types of molecules, which is also of certain reference significance for the design of AR LBD ligand.

The Total AR-Ligand Binding Energy

We next calculate the total binding energy by summing up the free energy contribution of each residue and compare the results with those obtained using the conventional MM/GBSA method. **Table 3** shows the contributions of each energy to the binding free energy of the 15 systems, including enthalpy and entropy. Furthermore, enthalpy is decomposed into van der Waals, electrostatic energy, GB, and nonpolar solvation energy, and it is clear that van der Waals interaction provided most of the enthalpy of each systems. **Table 4** shows the computational details of the fifteen systems, including the enthalpy, entropy and binding free energy calculated by ASIE method, the

enthalpy calculated by MM/GBSA method and their correlation coefficient, mean absolute error (MAE) and root mean square error (RMSE) with the experimental binding free energy. It can be seen that the binding free energy calculated by ASIE method has a good correlation with the experimental values, and there are also acceptable MAE values and RMSE values.

Figure 5A shows the correlation between the binding free energy obtained from experiments and the binding free energy calculated by ASIE method. The calculated binding energy shows good correlation with the experimental values ($R = 0.85$), although the experimental values range is very narrow. **Figure 5B** shows the correlation between the enthalpy calculated by the MM/GBSA method and the experimental binding free energy ($R = 0.34$). Furthermore, even if the system with the largest error (2HVC) was removed from the MM/GBSA results, the correlation (0.66) between the calculated value and the experimental value was still lower than that of ASIE.

CONCLUSION

Androgen receptor (AR) is an important target for many diseases. In our study, we used the ASIE method to quantitatively analyze the contribution of residues when AR binds to different ligands. The residues that contribute most to each ligand are determined. Furthermore, the sum of the contributions of each residue was relatively consistent with the experimental binding energy values. The results of these calculations will be useful for the design and analysis of AR ligands.

REFERENCES

- Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., et al. (2000). The protein data bank. *Nucleic Acids Res.* 28, 235–242. doi:10.1093/nar/28.1.235
- Bohl, C. E., Miller, D. D., Chen, J., Bell, C. E., and Dalton, J. T. (2005). Structural basis for accommodation of nonsteroidal ligands in the androgen receptor. *J. Biol. Chem.* 280, 37747–37754. doi:10.1074/jbc.M507464200
- Bohl, C. E., Wu, Z., Chen, J., Mohler, M. L., Yang, J., Hwang, D. J., et al. (2008). Effect of B-ring substitution pattern on binding mode of propionamide selective androgen receptor modulators. *Bioorg. Med. Chem. Lett.* 18, 5567–5570. doi:10.1016/j.bmcl.2008.09.002
- Brinkmann, A. O., Klaasen, P., Kuiper, G. G., van Der Korput, J. A., Bolt, J., de Boer, W., et al. (1989). Structure and function of the androgen receptor. *Urol. Res.* 17, 87–93. doi:10.1007/bf00262026
- Case, D. A., Ben-Shalom, I. Y., Brozell, S. R., Cerutti, D. S., Cheatham, T. E. I., Cruzeiro, V. W. D., et al. (2019). *AMBER 2019*, San Francisco, CA: University of California.
- Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., et al. (2005). The Amber biomolecular simulation programs. *J. Comput. Chem.* 26, 1668–1688. doi:10.1002/jcc.20290
- Chen, C. D., Welsbie, D. S., Tran, C., Baek, S. H., Chen, R., Vessella, R., et al. (2004). Molecular determinants of resistance to antiandrogen therapy. *Nat. Med.* 10, 33–39. doi:10.1038/nm972
- Cheng, A. C., Coleman, R. G., Smyth, K. T., Cao, Q., Soulard, P., Caffrey, D. R., et al. (2007). Structure-based maximal affinity model predicts small-molecule druggability. *Nat. Biotechnol.* 25, 71–75. doi:10.1038/nbt1273
- Dillon, E. L., Durham, W. J., Urban, R. J., and Sheffield-Moore, M. (2010). Hormone treatment and muscle anabolism during aging: androgens. *Clin. Nutr.* 29, 697–700. doi:10.1016/j.clnu.2010.03.010

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

GS performed MD simulation JB helped with ASIE coding XP helped with system analysis XH participated in design of the project and discussion YQ help wrote the paper JZ supervised the overall project.

FUNDING

This work was supported by National Key R&D Program of China (Grant Nos. 2016YFA0501700 and 2019YFA0905201), the National Natural Science Foundation of China (Grant Nos. 21922301, 21761132022, 21673074, 91753103, and 21933010), and the Natural Science Foundation of Shanghai Municipality (Grant No. 18ZR1412600).

ACKNOWLEDGMENTS

We also thank NYU Shanghai and the Supercomputer Center of East China Normal University (ECNU Multifunctional Platform for Innovation 001) for providing computer resources.

- Duan, L., Liu, X., and Zhang, J. Z. (2016). Interaction entropy: a new paradigm for highly efficient and reliable computation of protein-ligand binding free energy. *J. Am. Chem. Soc.* 138, 5722–5728. doi:10.1021/jacs.6b02682
- Fizazi, K., Smith, M. R., and Tombal, B. (2018). Clinical development of darolutamide: a novel androgen receptor antagonist for the treatment of prostate cancer. *Clin. Genitourin. Cancer* 16, 332–340. doi:10.1016/j.clgc.2018.07.017
- Gao, W., Bohl, C. E., and Dalton, J. T. (2005). Chemistry and structural biology of androgen receptor. *Chem. Rev.* 105, 3352–3370. doi:10.1021/cr020456u
- Gao, W., and Dalton, J. T. (2007). Expanding the therapeutic use of androgens via selective androgen receptor modulators (SARMs). *Drug Discov. Today* 12, 241–248. doi:10.1016/j.drudis.2007.01.003
- Genheden, S., and Ryde, U. (2015). The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. *Expert Opin. Drug Discov.* 10, 449–461. doi:10.1517/17460441.2015.1032936
- Gilson, M. K., and Zhou, H. X. (2007). Calculation of protein-ligand binding affinities. *Annu. Rev. Biophys. Biomol. Struct.* 36, 21–42. doi:10.1146/annurev.biophys.36.040306.132550
- Goa, K. L., and Spencer, C. M. (1998). Bicalutamide in advanced prostate cancer. A review. *Drugs Aging* 12, 401–422. doi:10.2165/00002512-199812050-00006
- Goldspiel, B. R., and Kohler, D. R. (1990). Flutamide: an antiandrogen for advanced prostate cancer. *DIAP* 24, 616–623. doi:10.1177/106002809002400612
- Hamann, L. G., Higuchi, R. I., Zhi, L., Edwards, J. P., Wang, X. N., Marschke, K. B., et al. (1998). Synthesis and biological activity of a novel series of nonsteroidal, peripherally selective androgen receptor antagonists derived from 1,2-dihydropyridono[5,6-g]quinolines. *J. Med. Chem.* 41, 623–639. doi:10.1021/jm970699s
- He, B., Gampe, R. T., Kole, A. J., Hnat, A. T., Stanley, T. B., An, G., et al. (2004). Structural basis for androgen receptor interdomain and coactivator interactions suggests a transition in nuclear receptor activation function dominance. *Mol. Cell* 16, 425–438. doi:10.1016/j.molcel.2004.09.036

- He, L., Bao, J., Yang, Y., Dong, S., Zhang, L., Qi, Y., et al. (2019). Study of SHMT2 inhibitors and their binding mechanism by computational alanine scanning. *J. Chem. Inf. Model.* 59, 3871–3878. doi:10.1021/acs.jcim.9b00370
- Hou, T., Wang, J., Li, Y., and Wang, W. (2011). Assessing the performance of the MM/PBSA and MM/GBSA methods. 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. *J. Chem. Inf. Model.* 51, 69–82. doi:10.1021/ci100275a
- Hu, X., Chai, X., Wang, X., Duan, M., Pang, J., Fu, W., et al. (2020). Advances in the computational development of androgen receptor antagonists. *Drug Discov. Today* 25, 1453–1461. doi:10.1016/j.drudis.2020.04.004
- Jorgensen, W. L., and Thomas, L. L. (2008). Perspective on free-energy perturbation calculations for chemical equilibria. *J. Chem. Theor. Comput.* 4, 869–876. doi:10.1021/ct800011m
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79, 926–935. doi:10.1063/1.445869
- Kassouf, W., Tanguay, S., and Aprikian, A. G. (2003). Nilutamide as second line hormone therapy for prostate cancer after androgen ablation fails. *J. Urol.* 169, 1742–1744. doi:10.1097/01.ju.0000057795.97626.66
- Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., et al. (2000). Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. *Acc. Chem. Res.* 33, 889–897. doi:10.1021/ar000033j
- Li, D., Zhou, W., Pang, J., Tang, Q., Zhong, B., Shen, C., et al. (2019). A magic drug target: androgen receptor. *Med. Res. Rev.* 39, 1485–1514. doi:10.1002/med.21558
- Liu, X., Peng, L., Zhou, Y., Zhang, Y., and Zhang, J. Z. H. (2018). Computational alanine scanning with interaction entropy for protein-ligand binding free energies. *J. Chem. Theor. Comput.* 14, 1772–1780. doi:10.1021/acs.jctc.7b01295
- Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015). ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theor. Comput.* 11, 3696–3713. doi:10.1021/acs.jctc.5b00255
- Massova, I., and Kollman, P. A. (1999). Computational alanine scanning to probe Protein-Protein interactions: a novel approach to evaluate binding free energies. *J. Am. Chem. Soc.* 121, 8133–8143. doi:10.1021/ja990935j
- Moreira, I. S., Fernandes, P. A., and Ramos, M. J. (2007). Computational alanine scanning mutagenesis—an improved methodological approach. *J. Comput. Chem.* 28, 644–654. doi:10.1002/jcc.20566
- Nadiminty, N., Tummala, R., Liu, C., Yang, J., Lou, W., Evans, C. P., et al. (2013). NF- κ B2/p52 induces resistance to enzalutamide in prostate cancer: role of androgen receptor and its variants. *Mol. Cancer Ther.* 12, 1629–1637. doi:10.1158/1535-7163.Mct-13-0027
- Nguyen, H., Roe, D. R., and Simmerling, C. (2013). Improved generalized born solvent model parameters for protein simulations. *J. Chem. Theor. Comput.* 9, 2020–2034. doi:10.1021/ct3010485
- Nirschl, A. A., Zou, Y., Krystek, S. R., Jr., Sutton, J. C., Simpkins, L. M., Lupisella, J. A., et al. (2009). N-Aryl-oxazolidin-2-imine muscle selective androgen receptor modulators enhance potency through pharmacophore reorientation. *J. Med. Chem.* 52, 2794–2798. doi:10.1021/jm801583j
- Ostrowski, J., Kuhns, J. E., Lupisella, J. A., Manfredi, M. C., Beehler, B. C., Krystek, S. R., Jr., et al. (2007). Pharmacological and x-ray structural characterization of a novel selective androgen receptor modulator: potent hyperandrogenic stimulation of skeletal muscle with hypostimulation of prostate in rats. *Endocrinology* 148, 4–12. doi:10.1210/en.2006-0843
- Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., III, Debolt, S., et al. (1995). 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.* 91, 1–41. doi:10.1016/0010-4655(95)00041-d
- Petukh, M., Li, M., and Alexov, E. (2015). Predicting binding free energy change caused by point mutations with knowledge-modified MM/PBSA method. *PLoS Comput. Biol.* 11, e1004276. doi:10.1371/journal.pcbi.1004276
- Qiu, L., Yan, Y., Sun, Z., Song, J., and Zhang, J. Z. H. (2018). Interaction entropy for computational alanine scanning in protein-protein binding. *Wires Comput. Mol. Sci.* 8, e1342. doi:10.1002/wcms.1342
- Rathkopf, D. E., Morris, M. J., Fox, J. J., Danila, D. C., Slovin, S. F., Hager, J. H., et al. (2013). Phase I study of ARN-509, a novel antiandrogen, in the treatment of castration-resistant prostate cancer. *J. Clin. Oncol.* 31, 3525–3530. doi:10.1200/jco.2013.50.1684
- Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. C. (1977). Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* 23, 327–341. doi:10.1016/0021-9991(77)90098-5
- Sack, J. S., Kish, K. F., Wang, C., Attar, R. M., Kiefer, S. E., An, Y., et al. (2001). Crystallographic structures of the ligand-binding domains of the androgen receptor and its T877A mutant complexed with the natural agonist dihydrotestosterone. *Proc. Natl. Acad. Sci. USA* 98, 4904–4909. doi:10.1073/pnas.081565498
- Saeed, A., Vaught, G. M., Gavardinas, K., Matthews, D., Green, J. E., Losada, P. G., et al. (2016). 2-Chloro-4-[[[(1R,2R)-2-hydroxy-2-methyl-cyclopentyl]amino]-3-methyl-benzonitrile: a transdermal selective androgen receptor modulator (sarm) for muscle atrophy. *J. Med. Chem.* 59, 750–755. doi:10.1021/acs.jmedchem.5b01168
- Salvati, M. E., Balog, A., Shan, W., Wei, D. D., Pickering, D., Attar, R. M., et al. (2005). Structure based approach to the design of bicyclic-1H-isoindole-1,3(2H)-dione based androgen receptor antagonists. *Bioorg. Med. Chem. Lett.* 15, 271–276. doi:10.1016/j.bmcl.2004.10.085
- Sun, C., Robl, J. A., Wang, T. C., Huang, Y., Kuhns, J. E., Lupisella, J. A., et al. (2006). Discovery of potent, orally-active, and muscle-selective androgen receptor modulators based on an N-aryl-hydroxybicyclohydantoin scaffold. *J. Med. Chem.* 49, 7596–7599. doi:10.1021/jm061101w
- Sun, H., Li, Y., Li, D., and Hou, T. (2013). Insight into crizotinib resistance mechanisms caused by three mutations in ALK tyrosine kinase using free energy calculation approaches. *J. Chem. Inf. Model.* 53, 2376–2389. doi:10.1021/ci400188q
- Tan, M. H., Li, J., Xu, H. E., Melcher, K., and Yong, E. L. (2015). Androgen receptor: structure, role in prostate cancer and drug discovery. *Acta Pharmacol. Sin.* 36, 3–23. doi:10.1038/aps.2014.18
- van Oeveren, A., Motamedi, M., Mani, N. S., Marschke, K. B., López, F. J., Schrader, W. T., et al. (2006). Discovery of 6-N,N-bis(2,2,2-trifluoroethyl)amino-4-trifluoromethylquinolin-2(1H)-one as a novel selective androgen receptor modulator. *J. Med. Chem.* 49, 6143–6146. doi:10.1021/jm060792t
- Wang, F., Liu, X. Q., Li, H., Liang, K. N., Miner, J. N., Hong, M., et al. (2006). Structure of the ligand-binding domain (LBD) of human androgen receptor in complex with a selective modulator LGD2226. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 62, 1067–1071. doi:10.1107/s1744309106039340
- Wang, R., Cong, Y., Li, M., Bao, J., Qi, Y., and Zhang, J. Z. H. (2020). Molecular mechanism of selective binding of NMS-P118 to PARP-1 and PARP-2: a computational perspective. *Front. Mol. Biosci.* 7, doi:10.3389/fmolb.2020.00050
- Yan, Y., Yang, M., Ji, C. G., and Zhang, J. Z. H. (2017). Interaction entropy for computational alanine scanning. *J. Chem. Inf. Model.* 57, 1112–1122. doi:10.1021/acs.jcim.6b00734
- Yang, Y.-P., He, L.-P., Bao, J.-X., Qi, Y.-F., and Zhang, J. Z. H. (2019). Computational analysis for residue-specific CDK2-inhibitor bindings. *Chin. J. Chem. Phys.* 32, 134–142. doi:10.1063/1674-0068/cjcp1901012
- Zhou, Y., Liu, X., Zhang, Y., Peng, L., and Zhang, J. Z. H. (2018). Residue-specific free energy analysis in ligand bindings to JAK2. *Mol. Phys.* 116, 2633–2641. doi:10.1080/00268976.2018.1442596

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Shao, Bao, Pan, He, Qi and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.