PAPER

Computational analysis of hot spots and binding mechanism in the PD-1/PD-L1 interaction†

Dading Huang,[a] Wei Wen,[a] Xiao Liu,[a] Yang Li[a] and John Z. H. Zhang [ID] *[abcd]

Programmed cell death protein-1 (PD-1) is an important immunological checkpoint and plays a vital role in maintaining the peripheral tolerance of the human body by interacting with its ligand PD-L1. The overexpression of PD-L1 in tumor cells induces local immune suppression and helps the tumor cells to evade the endogenous anti-tumor immunity. Developing monoclonal antibodies against the PD-1/PD-L1 protein–protein interaction to block the PD-1/PD-L1 signaling pathway has demonstrated superior anti-tumor efficacy in a variety of solid tumors and has made a profound impact on the field of cancer immunotherapy in recent years. Although the X-ray crystal structure of the PD-1/PD-L1 complex has been solved, the detailed binding mechanism of the PD-1/PD-L1 interaction is not fully understood from a theoretical point of view. In this study, we performed computational alanine scanning on the PD-1/PD-L1 complex to quantitatively identify the hot spots in the PD-1/PD-L1 interaction and characterize its binding mechanisms at the atomic level. To the best of our knowledge, this is the first time that theoretical calculations have been used to systematically and quantitatively predict the hot spots in the PD-1/PD-L1 interaction. We hope that the predicted hot spots and the energy profile of the PD-1/PD-L1 interaction presented in this work can provide guidance for the design of peptide and small molecule drugs targeting PD-1 or PD-L1.

1. Introduction

Protein–protein interactions (PPIs) play vital roles in many biological processes, including gene expression control, signal transduction and immunoregulation.[1–3] Thus they are associated with many human diseases including cancer.[4–6] By studying the detailed and specific interactions of PPIs at the atomic level, we can understand their fundamental binding mechanisms and develop effective therapeutic treatments to control and cure related diseases.[7] Generally, only a fraction of PPI interface residues, called "hot spots", contribute significantly to the binding free energy of PPI while the area of the contact surface can be relatively large.[8–10] Mutation of the hot spots in PPI could dramatically decrease their binding affinity or even disrupt the protein–protein association directly. For many diseases caused by disordered PPIs (e.g. overexpression of proteins and gain-of-function mutations), pocket or part of the surface composed of hot spots could be targeted to regulate the abnormal PPIs in drug design.[11] Therefore, predicting the hot spots in PPIs becomes an important step in the many designs of drugs that block abnormal PPIs.

Experimentally, alanine scanning mutagenesis is widely used to identify hot spots in PPIs.[12] This method comprises multiple steps involving expressing the corresponding wild-type and alanine mutant proteins, measuring their binding affinities, and calculating their binding energy differences. Generally, if the binding free energy difference of the alanine mutation of a residue is greater than 2 kcal mol^{-1}, the residue is then identified as a hot spot.[10] Although this technique is well-developed, large-scale experimental alanine scanning is still costly and time-consuming.[13] Thus, the available experimental data are fairly limited and diverse efficient computational hot spot detection methods have been developed. Currently, various theoretical hot spot detection methods can be roughly classified into molecular dynamics simulation-based methods (MD-based methods),[14–21] knowledge-based methods,[22–29] empirical formula-based methods,[30–34] and network topology-based methods.[35–38] Among them, the MD-based computational alanine scanning is more quantitative and its result can be compared with experimental data directly.[14,15,17–19]

Immunologic checkpoint receptors, such as the cytotoxic T-lymphocyte-associated protein 4 (CTLA-4) and programmed cell death protein 1 (PD-1), are brakes of the human immune system that prevent the over-activation of T cells which could cause autoimmunity such as inflammation.[39–42] Some tumor

[a]State Key Laboratory for Precision Spectroscopy, Shanghai Engineering Research Center of Molecular Therapeutics & New Drug Development, School of Chemistry and Molecular Engineering, East China Normal University, Shanghai 200062, China. E-mail: john.zhang@nyu.edu

[b]NYU-ECNU Center for Computational Chemistry at NYU Shanghai, Shanghai 200062, China

[c]Department of Chemistry, New York University, NY, NY 10003, USA

[d]Collaborative Innovation Center of Extreme Optics, Shanxi University, Taiyuan, Shanxi 030006, China

† Electronic supplementary information (ESI) available: Detailed individual components of the protein–protein binding energies and some other details are provided. See DOI: 10.1039/c9ra01369e

cells utilize this characteristic of the immune system through the expression of immune checkpoint ligands to induce local immune suppression. By attenuating and exhausting the activity of T cells, tumor cell can evade the endogenous anti-tumor immunity.[43–45] For example, programmed cell death 1 ligand 1 (PD-L1), a member of the B7 family, is widely overexpressed in many types of tumors such as lung cancer, breast cancer, glioblastoma, ovarian, and bladder cancers.[46,47] The engagement of PD-1 and PD-L1 results in the activation of the inhibitory function of PD-1 in T cells.[48]

Some antibodies have been developed to block the PD-1/PD-L1 signaling pathway by targeting PD-1 (nivolumab, pembrolizumab, *etc.*) or PD-L1 (atezolizumab, avelumab, durvalumab, *etc.*) and yield unprecedented efficacy for some types of cancer (melanoma, non-small cell lung cancer, renal cell cancer, *etc.*) in clinical trials.[49–53] By inhibiting the activity of inhibitory molecules in the PD-1/PD-L1 axis, immune checkpoint inhibitors release the immune brake of T cells in the tumor microenvironment, thereby achieve anti-tumor effects.[54,55] Small-molecule inhibitors that target PD-1 or PD-L1 are also feasible and some have been applied for patents.[56–59]

Recently, the crystal structure of the extracellular part of the human PD-1/PD-L1 complex has been solved by Zak *et al.* and it provides valuable information about the binding mechanism of the PD-1/PD-L1 interaction (Fig. 1).[60] According to the crystal structure, PD-1 and PD-L1 primarily interact through a large interface formed by the β-sheets of PD-1 and PD-L1. Comparing to the crystal structures of apo-proteins (apo-PD-1 and apo-PD-L1), the CC′ loop of PD-1 is relatively flexible and shows significant conformational changes along with the PD-L1 binding, while PD-L1 is more rigid and shows less conformational differences.[61–63]

However, it is still unclear which interfacial residues on PD-1 and PD-L1 are hot spots that are critical for their binding even if the crystal structure of the complex has been revealed. In this study, we use the recently developed MM/GBSA/IE method to quantitatively identify the hot spots in the PD-1/PD-L1 interaction.[64–67] Our result shows that a hydrophobic core and a surrounding hydrogen bond network formed by the hot spots of PD-1 and PD-L1 dominate their binding. These predicted hot spots and their energy information should be helpful to the rational design of inhibitors targeting the PD-1/PD-L1 signal pathway.

2. Theoretical methods

2.1. Molecular dynamics simulation of the PD-1/PD-L1 complex

The X-ray crystal structure of the PD-1/PD-L1 complex (PDB ID: 4ZQK) was downloaded from the Protein Data Bank (PDB) database and the missing residues (residues 85–92) of PD-1 were modeled by the SWISS-MODEL server.[60,68] The tleap module in the AMBER14 suite was used to add hydrogen atoms to the complex and solvate the complex by the TIP3P water in a truncated octahedral box with a buffer distance of 12.0 Å.[69,70] Chlorine and sodium counter ions were added to the box to neutralize the charge of the system. The system was parameterized using the AMBER ff14SB force field.[71] Periodic boundary condition (PBC) was imposed on the system to eliminate the boundary effect. A cutoff distance of 10.0 Å was set for non-bonded interactions and the long-range electrostatic interactions were treated by the Particle Mesh Ewald (PME) method.[72] SHAKE algorithm was used to constrain the covalent bonds involving hydrogen atoms and the time step was set to 2 fs.[73] The temperature was controlled by the Langevin thermostat with a collision frequency of 5.0 ps^{-1} and the Berendsen barostat was used to control the pressure at 1.0 atm.[74,75] Pmemd.cuda implemented in AMBER14 was used to perform MD simulations.

In the energy minimization step, the system was firstly minimized with the PD-1/PD-L1 complex constrained, then minimized with the backbone of the PD-1/PD-L1 complex constrained and lastly without any constraints. After the energy minimization, the system was heated to 300 K with harmonic constraints on the backbone atoms of the PD-1/PD-L1 complex within 200 ps and equilibrated in NVT and NPT ensembles at 300 K for 25 ns. The RMSD of the backbone atoms of the PD-1/PD-L1 complex was monitored until it reached a steady state in the equilibration stage. Finally, we performed 5 independent 2 ns production runs (trajectories) to generate ensembles for further analysis. A total of 10 000 configurations with an interval of 0.2 ps were extracted from each production MD trajectory and used for computational alanine scanning.

2.2. Computational alanine scanning

There are basically two approaches to generating the alanine mutant trajectory, single- and multiple-trajectory approaches, in the MD-based computational alanine scanning.[14] In the single-trajectory approach, the alanine mutant trajectory is simply generated from the wild-type trajectory by truncating the side chain of the residue to be mutated while the C$_\beta$ atom and its linking hydrogen atoms are retained. The truncated X
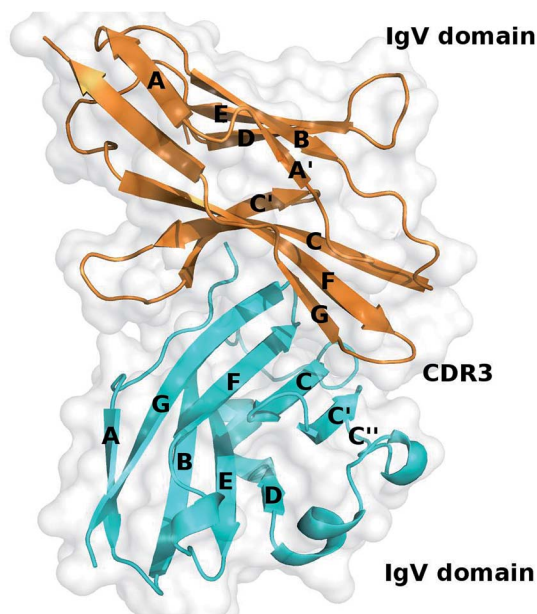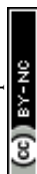

Fig. 1 Crystal structure of the PD-1/PD-L1 complex (PDB ID: 4ZQK). PD-1 and PD-L1 are shown in cartoon and colored in orange and cyan, respectively.

atom(s) is replaced by a hydrogen atom in the same direction as that of the C_β–X bond and the length of the new C_β–H bond is set to the default value according to the molecular model that used. The single-trajectory approach benefits from error cancellation and can partially offset the problem of insufficient sampling of conformational space. Previous studies have shown that the single-trajectory approach is more efficient and accurate than the multiple-trajectory approach.[14,15,17,18] In this work, we chose the single-trajectory approach to generate the alanine mutant trajectory for the above reasons.

2.3. Binding free energy difference

In order to calculate the binding free energy difference ($\Delta\Delta G_{bind}^{x \to a}$) between the wild-type and alanine mutant PPI complexes, we need to calculate the binding free energies of the wild-type (ΔG_{bind}^{x}) and alanine mutant (ΔG_{bind}^{a}), and then obtain their difference:[64,65,76,77]

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

Here, ΔG_{bind}, ΔG_{gas} and ΔG_{sol} represent the total binding free energy, the gas-phase component and the solvation free energy component of the total binding free energy, respectively. The superscript "x → a" represents that a specific residue "x" of the A^x protein mutates to alanine "a", and the superscripts "x" and "a" denote the wild-type (A^x–B) and the mutant (A^a–B), respectively.

In this work, we use the recently developed MM/GBSA/IE method which combines the MM/GBSA and IE methods to calculate the gas-phase and the solvation components in eqn (1). In the MM/GBSA/IE method, the gas-phase component is calculated as follows:[64,65]

$$\begin{aligned}\Delta\Delta G_{gas}^{x \to a} &= \Delta G_{gas}^{a} - \Delta G_{gas}^{x} \\ &\approx \Delta G_{gas}^{a-B} - \Delta G_{gas}^{x-B}\end{aligned} \quad (2)$$

where ΔG_{gas}^{a-B} and ΔG_{gas}^{x-B} are the gas-phase binding free energy of the a–B and x–B residue–protein interactions, respectively. In the MM/GBSA method, the gas-phase interaction energy is defined as the sum of molecular mechanics interaction energy and entropy contribution. Thus they can be calculated by the following equations:[14,66]

$$\begin{aligned}\Delta G_{gas}^{x-B} &= \left\langle E_{int}^{x-B}\right\rangle + KT \ln\left\langle \exp\left(\beta\Delta E_{int}^{x-B}\right)\right\rangle \\ &= \left\langle E_{int}^{x-B}\right\rangle - T\Delta S^{x-B}\end{aligned} \quad (3)$$

and

$$\begin{aligned}\Delta G_{gas}^{a-B} &= \left\langle E_{int}^{a-B}\right\rangle + KT \ln\left\langle \exp\left(\beta\Delta E_{int}^{a-B}\right)\right\rangle \\ &= \left\langle E_{int}^{a-B}\right\rangle - T\Delta S^{a-B}\end{aligned} \quad (4)$$

Here IE is defined as

$$-T\Delta S = KT \ln\langle\exp(\beta\Delta E_{int})\rangle, \quad (5)$$

where $\Delta E_{int} = E_{int} - \langle E_{int}\rangle$ is the fluctuation of interaction energy around the average. The relevant ensemble averages in the above equations can be evaluated by averaging over an MD simulation,

$$\langle E_{int}\rangle = \frac{1}{T}\int_0^T E_{int}(t)dt = \frac{1}{N}\sum_{i=1}^{N} E_{int}(t_i) \quad (6)$$

and

$$\langle\exp(\beta\Delta E_{int})\rangle = \frac{1}{N}\sum_{i=1}^{N}\exp(\beta\Delta E_{int}(t_i)) \quad (7)$$

For the solvation component ($\Delta\Delta G_{sol}^{x \to a}$), it can be expanded as follows

$$\begin{aligned}\Delta\Delta G_{sol}^{x \to a} &= \Delta G_{sol}^{a} - \Delta G_{sol}^{x} \\ &= \left(G_{sol}^{A^a B} - G_{sol}^{A^a} - G_{sol}^{B}\right) - \left(G_{sol}^{A^x B} - G_{sol}^{A^x} - G_{sol}^{B}\right) \\ &= \left(G_{sol}^{A^a B} - G_{sol}^{A^x B}\right) - \left(G_{sol}^{A^a} - G_{sol}^{A^x}\right)\end{aligned} \quad (8)$$

where $G_{sol}^{A^x}$, $G_{sol}^{A^a}$, G_{sol}^{B}, $G_{sol}^{A^x B}$ and $G_{sol}^{A^a B}$ are the solvation free energies of the A^x protein, the A^a protein, the B protein, the A^xB protein–protein complex, and the A^aB protein–protein complex, respectively. Thus the solvation free energy of B (ΔG_{sol}^{B}) is canceled and the last four solvation free energy terms need to be calculated to obtain $\Delta\Delta G_{sol}^{x \to a}$. In the MM/GBSA approach, the solvation free energy of a system is defined as the sum of two terms:

$$G_{sol} = G_{gb} + G_{np}, \quad (9)$$

where G_{gb} and G_{np} are the electrostatic solvation free energy and the nonpolar solvation free energy terms, respectively. The nonpolar component G_{np} is obtained using an empirical solvent-accessible surface area (SASA) formula:

$$G_{np} = \gamma SASA + \beta \quad (10)$$

The γ and β values we used here are the standard values of 0.00542 kcal $(mol^{-1}\ \mathring{A}^2)$ and 0.92 kcal mol^{-1}.

In this work, we run 5 trajectories for statistical averaging. For each trajectory, an interval of 10 ps was used to calculate the average of the interaction energy and the corresponding energy fluctuations. A total of 10 000 configurations with an interval of 0.2 ps were extracted from each 2 ns production trajectory for the IE calculation. This results in calculating a total of 2000 averages with 50 energy points for every average in the calculation of the fluctuation of the interaction energy for each trajectory. For MM/GBSA, a total of 100 frames equally distributed among the 10 000 configurations are used. The MM/PB(GB)SA method implemented in the AMBER14 package is applied for the MM/GBSA calculations.[69] Other detailed settings for the MM/GBSA/IE calculations are the same as in ref. 64.

3. Results and discussion

3.1. Stability of the PD-1/PD-L1complex in MD simulations

The root mean square deviation (RMSD) of the atomic positions from that of the crystal structure is usually considered to be an indicator of the stability of the system during MD simulations. Thus, we calculated the RMSDs of the heavy and C_α atoms of the complex with reference to the initial crystal structure and plotted them in Fig. 2. The RMSDs shown in Fig. 2 indicate that the complex is relatively stable with nominal fluctuations near the equilibrium state during the last 20 ns MD simulation. The average values of the RMSDs of the heavy atoms and C_α atoms are 2.11 and 1.59 Å with the standard deviations of 0.13 and 0.16 Å, respectively. This result shows the complex is similar to the crystal structure and does not have large deviations during the MD simulation.

The root mean square fluctuations (RMSFs) of the residues in the PD-1/PD-L1 complex are also calculated and plotted in Fig. 3. The RMSF curve of PD-1 shown in Fig. 3 is similar to that obtained from a 200 ns MD simulation by Shi *et al.*[78] The C'D loop (residues 85–95), which was not determined in the crystal structure and was model by the SWISS-MODEL server in this work, is the most flexible region of the liganded PD-1. Other regions of PD-1 are relatively rigid while some loops (such as BC and CC' loops) show relatively minor fluctuations. For PD-L1 in the complex, the most flexible region is the BC loop (residues 41–54). This is consistent with the recent finding of the principal component analysis (PCA) of all the crystal structures that contain the IgV domain of PD-L1.[79] The PD-1/PD-L1 complex structure obtained from the PDB database with missing residues added by the SWISS-MODEL server is stable during the MD simulations and should be reliable for further analysis.

3.2. Quantitative prediction of residue-specific binding energies

The recently developed computational alanine scanning approach based on the MM/GBSA/IE method is used to identify the hot spots in the interface of the PD-1/PD-1 interaction. The interface residues of PD-1 and PD-L1 are scanned if their
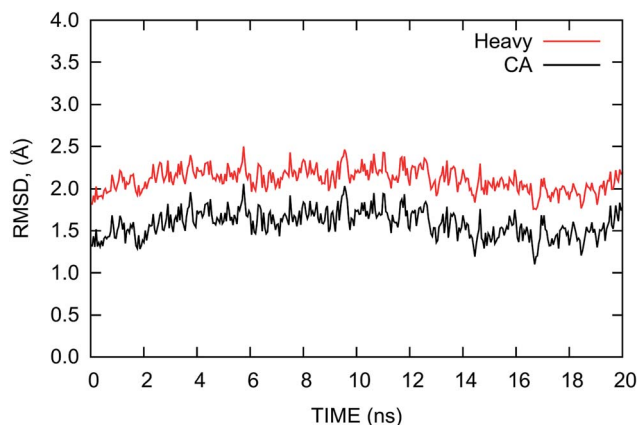

Fig. 2 RMSDs of the heavy atoms (red) and C_α atoms (black) of the PD-1/PD-L1 complex relative to the initial crystal structure as a function of time during the last 20 ns MD simulation.

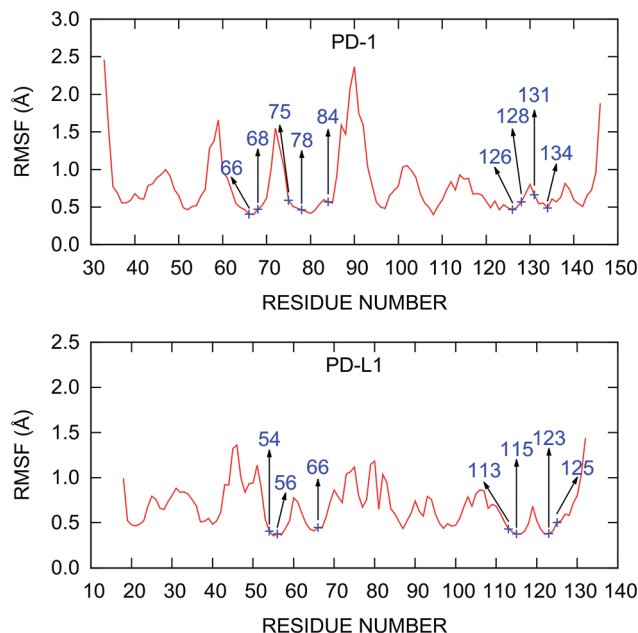Fig. 3 RMSF of the C_α atoms calculated from the last 20 ns MD trajectory. The hot and warm spots predicted by the MM/GBSA/IE method are labeled.

nearest atom-pair distance is less than 5.0 Å. A total of 22 and 20 interfacial residues of PD-1 and PD-L1 are scanned and the detailed results are ranked according to the computed $\Delta\Delta G$ values (Table S1†). Examination of the residue types of the hot spots in the PD-1/PD-L1 interaction shows that the hot spots are primarily formed by nonpolar and polar residues while some basic residues are also hot spots (Table 1). This means that both nonpolar and polar interactions are important for the PD-1/PD-L1 binding.

For PD-1, there are 8 hot spots ($_{PD-1}$Q75, $_{PD-1}$I134, $_{PD-1}$I126, $_{PD-1}$E84, $_{PD-1}$K78, $_{PD-1}$Y68, $_{PD-1}$L128, and $_{PD-1}$N66) with the $\Delta\Delta G$ values greater than 2 kcal mol^{-1} and 1 warm spot ($_{PD-1}$K131) with the $\Delta\Delta G$ between 1-2 kcal mol^{-1} (Fig. 4). For PD-L1, there are 6 hot spots ($_{PD-L1}$Y123, $_{PD-L1}$Y56, $_{PD-L1}$R125, $_{PD-L1}$M115, $_{PD-L1}$R113, and $_{PD-L1}$Q66) and 2 warm spots ($_{PD-L1}$I54 and $_{PD-L1}$K124) (Fig. 4). Hence the number of the hot and warm spots on PD-L1 is close to that on PD-1. The energy components of the hot and warm spots on PD-1 and PD-L1 are shown in Table 2

Table 1 Summary of the residue types of the hot and warm spots in the PD-1/PD-L1 interaction predicted by the MM/GBSA/IE method

Protein	Nonpolar	Polar	Basic	Acidic
Hot spots				
PD-1	$_{PD-1}$I134 $_{PD-1}$I126 $_{PD-1}$L128	$_{PD-1}$Q75 $_{PD-1}$Y68 $_{PD-1}$N66	$_{PD-1}$K78	$_{PD-1}$E84
PD-L1	$_{PD-L1}$M115	$_{PD-L1}$Y123 $_{PD-L1}$Y56 $_{PD-L1}$Q66	$_{PD-L1}$R125 $_{PD-L1}$R113	
Warm spots				
PD-1	$_{PD-L1}$I54		$_{PD-1}$K131	
PD-L1			$_{PD-L1}$K124	

Fig. 4 Hot and warm spots in the PD-1/PD-L1 interaction predicted by the MM/GBSA/IE method.
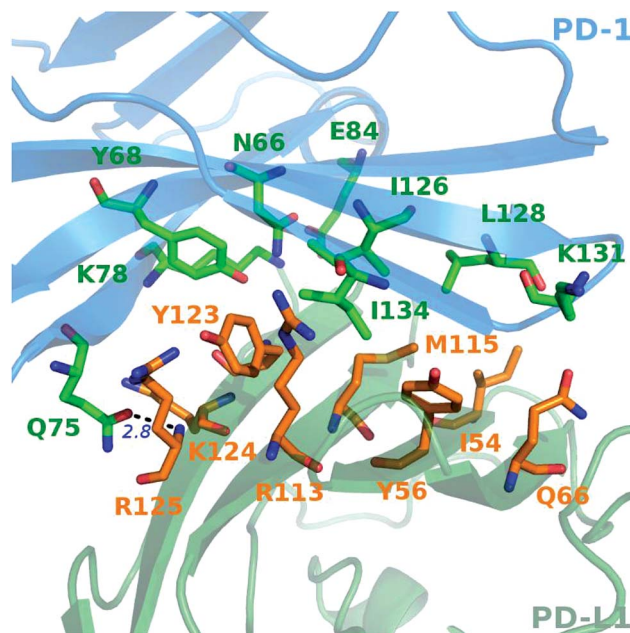

Fig. 5 Close-up view of the hot and warm spots in the PD-1/PD-L1 interaction that predicted by the MM/GBSA/IE method. The hot spots of PD-1 (blue cartoon) and PD-L1 (green cartoon) are shown in sticks with carbon atoms colored in green and orange, respectively.

and plotted in Fig. S2.† As shown in Fig. S2,† the hot spots on PD-1 and PD-L1 are mostly dominated by van der Waals (VDW) interactions except for $_{PD-1}$E84, $_{PD-1}$Y68, and $_{PD-1}$N66 which are dominated by polar and nonpolar interactions.

A close-up analysis of the predicted hot spots in the interface of the complex indicates a hydrophobic core formed by $_{PD-1}$I134, $_{PD-1}$I126, and $_{PD-1}$L128 of PD-1 and $_{PD-L1}$Y123, $_{PD-L1}$Y56, $_{PD-L1}$M115, $_{PD-L1}$R113, and $_{PD-L1}$I54 of PD-L1 (Fig. 5 and 6). The hot spot $_{PD-1}$I134 is an important residue of PD-1, its side chain interacts directly with the side chains of $_{PD-L1}$Y123, $_{PD-L1}$R113,

$_{PD-L1}$M115, and $_{PD-L1}$Y56 by van der Waals interactions (Fig. S2†). The $_{PD-1}$I134A mutation leads to the loss of about 4 kcal mol^{-1} of the binding free energy (Table 2A), highlighting the importance of these hydrophobic interactions. The hot spots $_{PD-1}$I126, $_{PD-1}$L128, and $_{PD-1}$I134, together with $_{PD-1}$A132, form a hydrophobic pocket to accommodate $_{PD-L1}$M115 (named M115 pocket) (Fig. 6A). The $_{PD-1}$I126A and $_{PD-1}$L128A mutations lead to a decrease of about 3 kcal mol^{-1} and 2 kcal mol^{-1} of the binding free energy, respectively. Indicating that $_{PD-1}$I126 and

Table 2 Hot and warm spots in the PD-1/PD-L1 interaction predicted by the MM/GBSA/IE method

Mutation	$\Delta\Delta E_{vdW}$	$\Delta\Delta E_{ele}$	$\Delta\Delta G_{gb}$	$\Delta\Delta G_{np}$	$\Delta\Delta H$	SD1[a]	$\Delta\Delta IE$	SD2[b]	$\Delta\Delta G$	SD[c]
(A) PD-1										
Q75A	4.14	10.77	−8.48	0.34	6.78	0.27	−1.48	0.28	5.30	0.54
I134A	5.91	−0.29	−0.80	0.12	4.95	0.22	−0.87	0.26	4.09	0.46
I126A	4.53	−0.01	−0.41	−0.08	4.03	0.12	−0.60	0.03	3.42	0.10
E84A	2.03	11.58	−8.62	0.09	5.08	0.14	−2.08	0.18	3.01	0.24
K78A	2.98	4.33	−2.93	0.07	4.46	0.17	−1.58	0.07	2.88	0.21
Y68A	1.80	5.23	−3.74	−0.08	3.22	0.23	−0.84	0.07	2.38	0.27
L128A	3.56	−0.38	−0.67	0.14	2.65	0.10	−0.38	0.03	2.27	0.12
N66A	1.39	3.15	−2.06	−0.06	2.42	0.04	−0.38	0.03	2.04	0.07
K131A	1.69	7.31	−7.43	0.19	1.76	0.08	−0.58	0.06	1.18	0.09
(B) PD-L1										
Y123A	7.14	3.96	−3.98	0.49	7.60	0.50	−1.69	0.11	5.92	0.46
Y56A	5.24	0.50	−0.68	0.04	5.10	0.28	−0.52	0.04	4.59	0.27
R125A	4.64	−2.21	1.50	0.55	4.48	0.38	−1.39	0.10	3.09	0.41
M115A	3.93	−0.96	−0.03	0.06	2.99	0.12	−0.47	0.04	2.52	0.10
R113A	2.59	0.48	0.62	0.30	3.99	0.72	−1.84	0.10	2.15	0.65
Q66A	2.25	3.16	−2.92	0.04	2.53	0.03	−0.43	0.04	2.11	0.04
I54A	1.88	−0.08	−0.38	0.13	1.55	0.05	−0.14	0.00	1.41	0.04
K124A	3.02	3.21	−3.56	0.34	3.01	0.11	−1.78	0.52	1.23	0.59

[a] Standard deviation of $\Delta\Delta H$. [b] Standard deviation of $\Delta\Delta IE$. [c] Standard deviation of $\Delta\Delta G$.
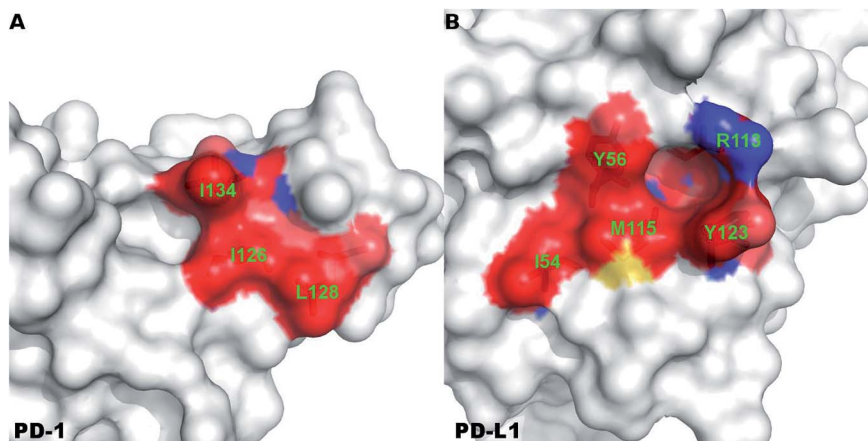
Fig. 6 Surface representation of the hydrophobic core that formed by the hot and warm spots in the PD-1/PD-L1 interaction. The hot and warm spots that compose the hydrophobic core are colored surface and labeled in green font.

$_{PD-1}$L128 contribute less than $_{PD-1}$I134 for the PD-1/PD-L1 binding. For PD-L1, the hot spot $_{PD-L1}$M115 binds to the center of the hydrophobic pocket that formed by the hot spots on PD-1 (Fig. 5). Our calculation shows that the $_{PD-L1}$M115A mutation leads to a decrease of over 2 kcal mol^{-1} of the binding free energy (Table 2B). This's less than the binding free energy differences of the $_{PD-L1}$Y123A, $_{PD-L1}$Y56A, and $_{PD-L1}$R125A mutations. The $_{PD-L1}$Y123A mutation results in a loss of $\Delta\Delta G$

greater than 5 kcal mol^{-1}. The side chain of $_{PD-L1}$Y123 interacts with the side chain of $_{PD-1}$I134 and forms a $\pi-\pi$ interaction with $_{PD-1}$Y68 (Fig. 5). Moreover, the hydroxyl group of the side chain of $_{PD-L1}$Y123 forms a stable hydrogen bond with the side chain of $_{PD-1}$E136 (see below). These various interactions formed by the side chain of $_{PD-L1}$Y123 and the high $\Delta\Delta G$ value of the $_{PD-L1}$Y123A mutation demonstrate that $_{PD-L1}$Y123 is a significant residue in the PD-1/PD-L1 interaction.
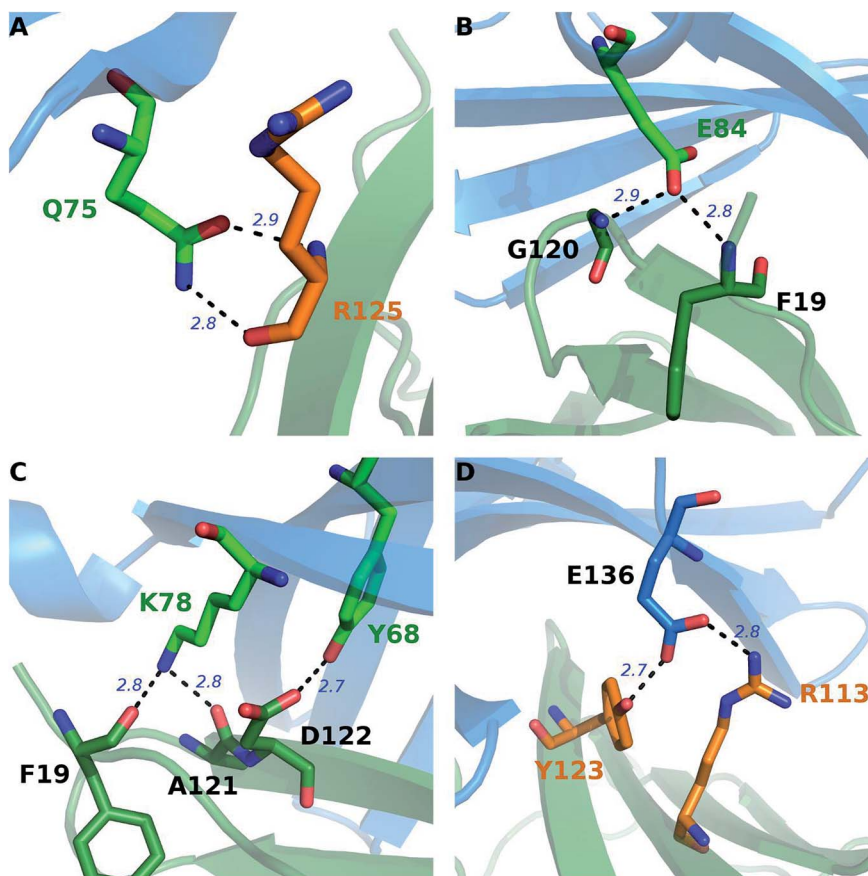

Fig. 7 Stable hydrogen bonds in the interface of the PD-1/PD-L1 complex during the MD simulations. The hot spots of PD-1 (blue cartoon) and PD-L1 (green cartoon) are shown in sticks with carbon atoms colored in green and orange, respectively.

Table 3 Stable hydrogen bonds (and salt bridges) between PD-1 and PD-L1 in the PD-1/PD-L1 interaction. Hot-spot acceptor and hot-spot donor (A); hot-spot acceptor and null-spot donor (B); null-spot acceptor and hot-spot donor (C); null-spot acceptor and null-spot donor (D)

Acceptor[a]	Donor	Occupancy[b]
(A) H–H		
$_{PD-L1}$R125@O	$_{PD-1}$Q75@NE2	0.90
$_{PD-1}$Q75@OE1	$_{PD-L1}$R125@N	0.51
(B) H–NH		
$_{PD-1}$E84@OE	$_{PD-L1}$F19@N	0.69
$_{PD-1}$E84@OE	$_{PD-L1}$G120@N	0.64
(C) NH–H		
$_{PD-L1}$D122@OD	$_{PD-1}$Y68@OH	0.98
*$_{PD-1}$E136@OE	$_{PD-L1}$R113@NH	0.94
$_{PD-L1}$F19@O	$_{PD-1}$K78@NZ	0.87
$_{PD-L1}$A121@O	$_{PD-1}$K78@NZ	0.80
$_{PD-1}$E136@OE	$_{PD-L1}$Y123@OH	0.79
$_{PD-1}$D26@OD	$_{PD-1}$Q75@NE2	0.78
$_{PD-L1}$A121@O	$_{PD-1}$N66@ND2	0.68
(D) NH–NH		
$_{PD-1}$D26@OD	$_{PD-1}$S73@OG	0.57

[a] The salt bridge is labeled with *. [b] The hydrogen bonds are determined by the default criteria (angle of acceptor hydrogen donor atoms greater than 135° and distance of acceptor donor atoms less than 3.0 Å) in the CPPTRAJ module of the AMBER14 package. Only hydrogen bonds with occupancy ≥ 0.50 are listed.

The $_{PD-L1}$Y56A mutation also leads to a large decrease of the binding free energy (over 4 kcal mol^{-1}) since $_{PD-L1}$Y56 interacts with $_{PD-1}$I134 and $_{PD-1}$A132 through the van der Waals interactions of its side chain (Fig. S2†). Therefore, $_{PD-L1}$Y123, $_{PD-L1}$Y56, and $_{PD-L1}$M115 are the most important hot spots that constitute the hydrophobic pocket on PD-L1 to accommodate $_{PD-1}$I134 (named I134 pocket) (Fig. 6B). In addition, the hot spot $_{PD-L1}$R113 and warm spot $_{PD-L1}$I54 interact directly with $_{PD-1}$I134 and $_{PD-1}$L128, respectively (Fig. 5). The computational alanine scanning shows the $_{PD-L1}$R113A and $_{PD-L1}$I54A mutations lead to smaller losses of the binding free energy (2.15 kcal mol^{-1} and 1.41 kcal mol^{-1}). The warm spot $_{PD-L1}$I54 is located at the edge of the hydrophobic region of PD-L1 and interacts with the side chain of $_{PD-1}$L128 only (Fig. 5 and 6B). This can explain why the $_{PD-L1}$I54A mutation causes a less loss of the binding free energy.

Besides the hydrophobic core hot spots, $_{PD-1}$Q75 and $_{PD-L1}$R125 interact through hydrogen bonds and side-chain van der Waals interactions (Fig. 7A). The energy components of the $_{PD-1}$Q75A mutation indicate that both polar and nonpolar interactions are significant for its interaction with PD-L1 while the energy components of the $_{PD-L1}$R125A mutation indicate that $_{PD-L1}$R125 is dominated by nonpolar interactions (Fig. S2†). Hydrogen bond analysis shows that the side chain amine of $_{PD-1}$Q75 forms a strong hydrogen bond with the backbone oxygen atom of $_{PD-L1}$R125 with an occupancy of 90% and the side-chain carbonyl group of $_{PD-1}$Q75 forms a relatively weaker hydrogen bond with the backbone nitrogen atom of $_{PD-L1}$R125 with an occupancy of 51% (Table 3A). This can explain why the polar energy component is significant in the $_{PD-1}$Q75A

mutation. These two hydrogen bonds that formed by the two hot spots are stable during the MD simulations and the energy components of their alanine mutations show that the van der Waals interactions between their side chains are also significant. These results highlight the importance of $_{PD-1}$Q75/$_{PD-L1}$R125 interaction for the PD-1/PD-L1 binding.

Recently, Liu *et al.* studied the binding mechanism of the PD-1/PD-L1 interaction based on MD simulations.[80] They found that both the open and closed conformations exist in apo-PD-1, and PD-L1 stabilizes the CC′ loop of PD-1 in the closed conformation and induces the structural rearrangement of the CC′ loop after binding. These results suggest a binding mechanism that involves both conformational selection and induced fit. Here, we can see that this binding mechanism is mainly controlled by the interaction between $_{PD-1}$Q75 and $_{PD-L1}$R125 because $_{PD-1}$Q75 is located in the CC′ loop of PD-1 and is the most important hot spot in PD-1 (Fig. 4). In addition, it's evident that hot spots usually have smaller RMSF values (Fig. 3), meaning they are more stable residues in the PD-1/PD-L1 interaction.

3.3. Hydrogen bond analysis

Except for the hot spots (hydrophobic core hot spots, $_{PD-1}$Q75, and $_{PD-L1}$R125) that have been mentioned in the above section, our alanine scanning calculation predicts that some other polar and charged residues are also hot spots ($_{PD-1}$E84, $_{PD-1}$K78, $_{PD-1}$N66, and $_{PD-L1}$Q66) (Fig. 4 and 5). In order to obtain more information about their interactions, we performed an exhaustive hydrogen bond analysis between PD-1 and PD-L1 (Table 3). As shown in Table 3 and Fig. 5, the most important hydrogen bonds between PD-1 and PD-L1 are mainly formed by the polar and charged hot spots that around the hydrophobic core and some backbone atoms of null spots ($\Delta\Delta G$ less than 1.00 kcal mol^{-1}) are also involved in the hydrogen bonds formation.

For PD-1, the side-chain carboxyl group of the hot spot $_{PD-1}$E84 forms two hydrogen bonds with the backbone nitrogen atoms of $_{PD-L1}$F19 and $_{PD-L1}$G120 ($_{PD-1}$E84@OE···$_{PD-L1}$F19@N; $_{PD-1}$E84@OE···$_{PD-L1}$G120@N) (Fig. 7B), and the side-chain amine group of the hot spot $_{PD-1}$K78 forms two hydrogen bonds with the backbone carbonyl oxygen atom of $_{PD-L1}$F19 and $_{PD-L1}$A121 ($_{PD-L1}$F19@O···$_{PD-1}$K78@NZ; $_{PD-L1}$A121@O···$_{PD-1}$K78@NZ) (Fig. 7C). Additionally, the side-chain amine group of the hot spot $_{PD-1}$N66 forms a hydrogen bond with the backbone carbonyl oxygen of $_{PD-L1}$A121 ($_{PD-L1}$A121@O···$_{PD-1}$N66@ND2) (Table 3C). These five hydrogen bonds that formed by the side chains of the hot spots of PD-1 and the backbones of the null spots of PD-L1 are stable during the MD simulations (Table 3B and C). Indicating the side chain of these hot spots ($_{PD-1}$E84, $_{PD-1}$K78, and $_{PD-1}$N66) are important for the stabilization of the PD-1/PD-L1 interaction. Moreover, the side-chain hydroxyl group of the hot spot $_{PD-1}$Y68 forms a hydrogen bond with the side-chain carboxyl group of $_{PD-L1}$D122 ($_{PD-L1}$D122@OD···$_{PD-1}$Y68@OH) (Fig. 7C and Table 3C). This hydrogen bond can stabilize the orientation of the side chain of $_{PD-1}$Y68 while $_{PD-1}$Y68 additionally takes part in a π–π interaction with $_{PD-L1}$Y123 (Fig. 5). Although the side chain of $_{PD-L1}$D122

Table 4 Hot spots in the PD-1/PD-L1 interaction predicted by four different methods. The x denotes that the residue is not calculated by the method, 0 denotes the predicted null spot, and 1 denotes the predicted hot spot

Residue	Robetta server[a]	KFC server	HotRegion database	MM/GBSA/IE	Percentage[b]
(A) PD-1					
Y68	1	1	1	1	100
K78	1	1	1	1	100
I126	1	1	1	1	100
A132	x	1	1	x	100
I134	1	1	1	1	100
Q75	1	1	0	1	75
L128	1	0	1	1	75
N66	1	0	0	1	50
T76	0	1	1	0	50
G124	x	0	1	x	50
E84	0	0	x	1	33
V64	0	0	1	0	25
(B) PD-L1					
M115	1	1	1	1	100
A121	x	1	1	x	100
Y123	1	1	1	1	100
Y56	1	1	0	1	75
R113	1	0	1	1	75
Q66	1	0	0	1	50
D122	0	1	1	0	50
K124	1	1	0	0	50
R125	0	0	1	1	50
E58	1	0	x	0	33
N63	1	0	x	0	33
F19	0	0	1	0	25
V23	0	0	1	0	25
I54	0	0	1	0	25

[a] For the Robetta server, the residue is a hot spot if $\Delta\Delta G \geq 1.0$ kcal mol^{-1}. [b] Percentage of hot spot predictions (%). Percentage = HP/NP, where HP is the number of hot spot predictions and NP is the number of predictions. Residues with the percentage greater than 0 are shown.

participates in this hydrogen bond, it is a null spot and the $\Delta\Delta G$ value of its alanine mutation is 0.53 kcal mol^{-1} (Table S1†).

For PD-L1, the side chain guanidino group of the hot spot $_{PD-L1}$R113 forms a stable salt bridge with the side-chain carboxyl group of $_{PD-1}$E136 (*$_{PD-1}$E136@OE⋯$_{PD-L1}$R113@NH) (Fig. 7D and Table 3C). Besides this salt bridge, the hot spot $_{PD-L1}$Y123 forms a relatively weak hydrogen bond with $_{PD-1}$E136 ($_{PD-1}$E136@OE⋯$_{PD-L1}$Y123@OH) (Fig. 7D and Table 3C). Interestingly, $_{PD-1}$E136 is a null spot although its side-chain atoms form a salt bridge and a hydrogen bond with the side chains of $_{PD-L1}$R113 and $_{PD-L1}$Y123, respectively (Table S1†). This is validated by some other hot spot detection methods (see below). The computational alanine scanning result shows the $_{PD-1}$E136A mutation is favorable for the PD-1/PD-L1 interaction due to the benefits of van der Waals interactions and entropy contribution (Table S1†). Except these stable salt bridge and hydrogen bond interactions, many weaker hydrogen bonds and salt bridges (such as $_{PD-1}$E84@OE⋯$_{PD-L1}$A18@N; *$_{PD-1}$E136@OE⋯$_{PD-L1}$R113@NE; $_{PD-L1}$Q66@OE1⋯$_{PD-1}$A132@N; $_{PD-1}$T76@O⋯$_{PD-L1}$K124@NZ and $_{PD-L1}$Y123@O⋯$_{PD-1}$T76@OG1) (data not shown) that formed by the interface residues of PD-1 and PD-L1 should also contribute to the stabilization of the PD-1/PD-L1 interaction.

The hydrogen bond analysis shows that the polar and charged hot spots play an important role in the PD-1/PD-L1 interaction. The hydrogen bonds and salt bridges between the side chains of the hot spots and the backbones of the null spots form a stable hydrogen bond network around the hydrophobic core between PD-1 and PD-L1. These interactions, together with the hydrophobic interactions of the hydrophobic core, are responsible for the stable PD-1/PD-L1 interaction.

3.4. Hot spot detection using different methods

We have identified the hot spots in the PD-1/PD-L1 complex by using the computational alanine scanning and obtained some valuable information about their interactions. However, in silico alanine scanning approaches for hot spot detections are not perfect and false predictions may exist. To the best of our knowledge, the experimental alanine scanning of the PD-1/PD-L1 interaction has not been studied so far. Thus, there are no systemic experimental results that can be used to compare with our calculated results.

In order to validate our hot spot predictions, we compared our predictions to those from some other hot spot detection methods. Up to now, a number of in silico methods have been developed and widely used to detect the hot spots in PPI
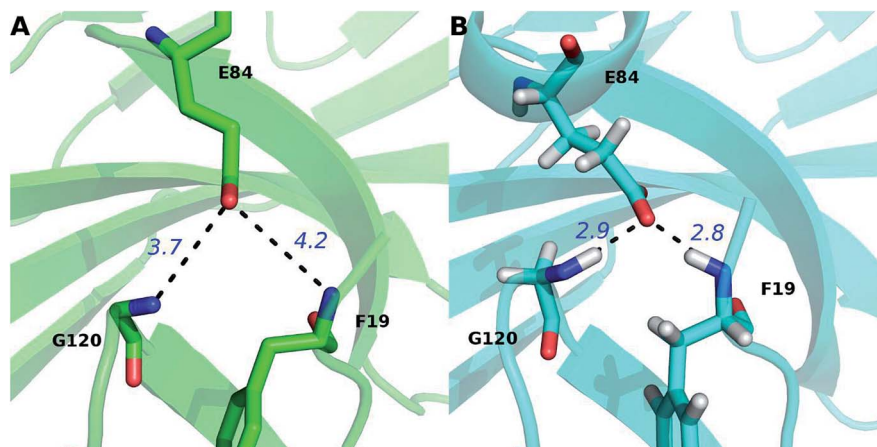
Fig. 8 Comparison of the hydrogen bonds of $_{PD-1}$E84 in the PDB (A) and MD equilibrated (B) structures. The hydrogen bonds between $_{PD-1}$E84 and the two residues of PD-L1 ($_{PD-L1}$G120 and $_{PD-L1}$F19) are shown in dotted lines and their distances are labeled in blue font.

interface.[16–18,21–28,31–35] In this work, we used three knowledge-based methods (Robetta server, KFC server, and HotRegion database) for comparison.[16,23,28,29] These methods are generally accurate for hot spot identification and relatively fast because they use simple scoring functions and a single protein–protein complex structure for hot spot prediction. However, they usually do not predict the binding free energy difference quantitatively like the MD-based computational alanine scanning methods.

The hot spots predicted by the four different methods are summarized in Table 4. As shown in Table 4, there are 7 ($_{PD-1}$Y68, $_{PD-1}$K78, $_{PD-1}$I126, $_{PD-1}$A132, $_{PD-1}$I134, $_{PD-1}$Q75, and $_{PD-1}$L128) and 5 ($_{PD-L1}$M115, $_{PD-1}$A121, $_{PD-L1}$Y56, $_{PD-L1}$Y123, and
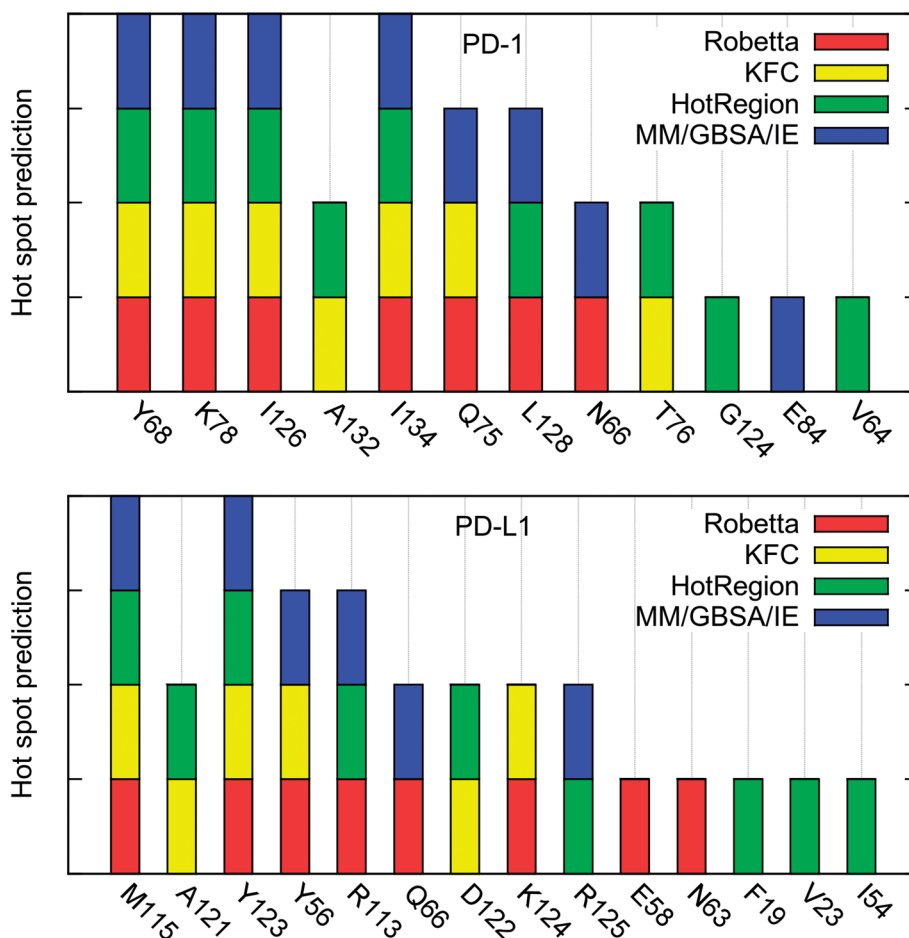

Fig. 9 Comparison of the hot spots on PD-1 and PD-L1 that predicted by the four different methods. The hot spots that predicted by the Robetta server, KFC server, HotRegion database, and MM/GBSA/IE methods are shown in red, yellow, green, and blue bars, respectively.

$_{PD-L1}$R113) hot spots on PD-1 and PD-L1, respectively, if we use the percentage value greater than 50% as the criterion of a hot spot. All these hot spots are correctly predicted by the MM/GBSA/IE method except the two alanine hot spots ($_{PD-1}$A132 and $_{PD-L1}$A121) which are not calculated by the MM/GBSA/IE method.

The percentage value of $_{PD-1}$N66 is 50%, both the Robetta server and the MM/GBSA/IE method predict that it is a hot spot, but the KFC server and the HotRegion database predict that it is a null spot (Table 4). Thus, the status of $_{PD-1}$N66 can't be determined currently. The percentage value of $_{PD-1}$E84 is 33%, only the MM/GBSA/IE method predicts that it is a hot spot and the HotRegion database doesn't give a prediction for this residue. Actually, the distance of the two hydrogen bonds that formed by the side chain of $_{PD-1}$E84 and the residues of PD-L1 ($_{PD-L1}$G120 and $_{PD-L1}$F19) is smaller than that in the crystal structure during the MD simulations (Fig. 8). This can explain the different results between the MM/GBSA/IE method and the other three methods because the MM/GBSA/IE method calculates $\Delta\Delta G$ according to the snapshots of MD trajectories and the other three methods identify hot spots based on a single X-ray crystal structure. For PD-L1, the result of $_{PD-L1}$Q66 is the same as that of $_{PD-1}$N66 (Table 4). Both the Robetta server and the MM/GBSA/IE method predict that it is a hot spot while the KFC server and the HotRegion database predict that it is a null spot. For $_{PD-L1}$R125, both the Hotregion database and the MM/GBSA/IE method predict that it is a hot spot. But the Robetta server and the KFC server predict that it is a null spot. However, the guanidino group of the side chain of $_{PD-L1}$R125 is missing in the X-ray crystal structure. Thus, the predictions of this residue that based on the crystal structure should be less reliable.

Hydrogen bond analysis in the above section identified multiple hydrogen bonds stably formed by the side chain of $_{PD-1}$E136 and the residues on PD-L1, but the computational alanine scanning based on the MM/GBSA/IE method identified $_{PD-1}$E136 as a null spot. Here, the other three methods also predict that it is a null spot (it is not listed in Table 4 since its percentage value equals to zero). The hot spots predicted by the MM/GBSA/IE method are basically consistent with the other three methods. Combining with the predicted results of the other methods, we identified two additional hot spots ($_{PD-1}$A132 and $_{PD-L1}$A121) which could not be calculated by computational alanine scanning approaches (Fig. S3†). Fig. 9 is added for a direct comparison of the hot spots that predicted by these four methods.

4. Conclusions

By using the MM/GBSA/IE method-based computational alanine scanning, we have quantitatively identified the hot spots in the PD-1/PD-L1 interaction. Briefly, the results of the computational alanine scanning revealed a shallow hydrophobic pocket on the surfaces of PD-1 and PD-L1, respectively. The hydrophobic pocket on PD-1 (M115 pocket) is formed by the hot spots $_{PD-1}$I134, $_{PD-1}$I126, and $_{PD-1}$L128 which accommodates the hot spot $_{PD-L1}$M115. And the hydrophobic pocket on PD-L1 (I134 pocket) is composed of the hot spots $_{PD-L1}$Y123,

$_{PD-L1}$Y56, $_{PD-L1}$M115, and $_{PD-L1}$R113 which accommodates the hot spot $_{PD-1}$I134. The residues of the two pockets, together with $_{PD-1}$A132, $_{PD-L1}$I54, and $_{PD-L1}$A121, are tightly compacted to form a hydrophobic core in the interface of the PD-1/PD-L1 complex. The energy component analysis of the computational alanine scanning shows the PD-1/PD-L1 interaction is dominated by the van der Waals interactions of these hydrophobic hot spots. Thus, these hydrophobic residues are potential targets for designing inhibitors that disrupt the PD-1/PD-L1 interaction. Moreover, our result shows $_{PD-1}$Q75 and $_{PD-L1}$R125 are hot spots that have direct and stable hydrogen bonds and van der Waals interactions, which means their interactions are important for the PD-1/PD-L1 recognition and binding. This finding is consistent with some recent works.[78,80]

In addition, the hydrogen bond analysis revealed a stable hydrogen bond network between PD-1 and PD-L1. This unique hydrogen bond network around the hydrophobic core is formed by the side chains of polar and charged hot spots and the backbone atoms of null spots. Rational design of inhibitors that targeting both the hydrophobic and hydrophilic hot spots could maximize the specificity and affinity to PD-1 or PD-L1. One interesting thing that we have found is $_{PD-1}$E136 is a null spot even if its side chain forms multiple hydrogen bonds with PD-L1.

Moreover, we compared the hot spots identified by the MM/GBSA/IE method with other three hot spot detection methods. The hot spots predicted by the four methods are basically in agreement while some minor differences are discovered. Specifically, the different predictions of $_{PD-1}$E84 between the MM/GBSA/IE method and the other three methods reveal a gap between the MD-based and crystal structure-based methods. One should keep in mind that the MD-based methods are not always more reliable. Because there are still many defects in the present mainstream force fields, especially for polar and charged residues which usually have significant polarization effects. By combing the results of different methods, we predicted the hot spots in the PD-1/PD-L1 interaction with more confidence and detected the alanine hot spots in the PD-1/PD-L1 interaction.

In a word, the hot spots determined in this study help us understand the complex binding mechanism of the PD-1/PD-L1 interaction. We hope that the current study provides useful guidance for the rational design of novel drugs for immune regulation.

Conflicts of interest

There are no conflicts of interest to declare.

Acknowledgements

References

1 G. Sowmya, E. J. Breen and S. Ranganathan, Linking structural features of protein complexes and biological function, *Protein Sci.*, 2015, **24**(9), 1486–1494.

2 S. H. Yook, Z. N. Oltvai and A. L. Barabasi, Functional and topological characterization of protein interaction networks, *Proteomics*, 2004, **4**(4), 928–942.

3 D. Eisenberg, E. M. Marcotte, I. Xenarios and T. O. Yeates, Protein function in the post-genomic era, *Nature*, 2000, **405**(6788), 823–826.

4 A. A. Ivanov, F. R. Khuri and H. A. Fu, Targeting protein-protein interactions as an anticancer strategy, *Trends Pharmacol. Sci.*, 2013, **34**(7), 393–400.

5 D. P. Ryan and J. M. Matthews, Protein-protein interactions in human disease, *Curr. Opin. Struct. Biol.*, 2005, **15**(4), 441–446.

6 M. W. Gonzalez and M. G. Kann, Chapter 4: Protein Interactions and Disease, *PLoS Comput. Biol.*, 2012, **8**(12).

7 R. R. Halehalli and H. A. Nagarajaram, Molecular principles of human virus protein-protein interactions, *Bioinformatics*, 2015, **31**(7), 1025–1033.

8 O. Keskin, B. Ma and R. Nussinov, Hot regions in protein–protein interactions: the organization and contribution of structurally conserved hot spot residues, *J. Mol. Biol.*, 2005, **345**(5), 1281–1294.

9 T. Clackson and J. A. Wells, A Hot-Spot of Binding-Energy in a Hormone-Receptor Interface, *Science*, 1995, **267**(5196), 383–386.

10 A. A. Bogan and K. S. Thorn, Anatomy of hot spots in protein interfaces, *J. Mol. Biol.*, 1998, **280**(1), 1–9.

11 X. Li, O. Keskin, B. Y. Ma, R. Nussinov and J. Liang, Protein-protein interactions: hot spots and structurally conserved residues often locate in complemented pockets that pre-organized in the unbound states: implications for docking, *J. Mol. Biol.*, 2004, **344**(3), 781–795.

12 K. L. Morrison and G. A. Weiss, Combinatorial alanine-scanning, *Curr. Opin. Chem. Biol.*, 2001, **5**(3), 302–307.

13 I. S. Moreira, P. A. Fernandes and M. J. Ramos, Hot spots-a review of the protein-protein interface determinant amino-acid residues, *Proteins*, 2007, **68**(4), 803–812.

14 I. Massova and P. A. Kollman, Computational alanine scanning to probe protein-protein interactions: a novel approach to evaluate binding free energies, *J. Am. Chem. Soc.*, 1999, **121**(36), 8133–8143.

15 S. Huo, I. Massova and P. A. Kollman, Computational alanine scanning of the 1 : 1 human growth hormone-receptor complex, *J. Comput. Chem.*, 2002, **23**(1), 15–27.

16 T. Kortemme, D. E. Kim and D. Baker, Computational alanine scanning of protein-protein interfaces, *Sci. STKE*, 2004, **2004**(219), pl2.

17 I. S. Moreira, P. A. Fernandes and M. J. Ramos, Computational alanine scanning mutagenesis - an improved methodological approach, *J. Comput. Chem.*, 2007, **28**(3), 644–654.

18 S. A. Martins, M. A. S. Perez, I. S. Moreira, S. F. Sousa, M. J. Ramos and P. A. Fernandes, Computational Alanine Scanning Mutagenesis: MM-PBSA vs. TI, *J. Chem. Theory Comput.*, 2013, **9**(3), 1311–1319.

19 R. M. Ramos and I. S. Moreira, Computational Alanine Scanning Mutagenesis-An Improved Methodological Approach for Protein-DNA Complexes, *J. Chem. Theory Comput.*, 2013, **9**(9), 4243–4256.

20 M. Petukh, M. H. Li and E. Alexov, Predicting Binding Free Energy Change Caused by Point Mutations with Knowledge-Modified MM/PBSA Method, *PLoS Comput. Biol.*, 2015, **11**(7), e1004276.

21 I. C. M. Simoes, I. P. D. Costa, J. T. S. Coimbra, M. J. Ramos and P. A. Fernandes, New Parameters for Higher Accuracy in the Computation of Binding Free Energy Differences upon Alanine Scanning Mutagenesis on Protein-Protein Interfaces, *J. Chem. Inf. Model.*, 2017, **57**(1), 60–72.

22 S. J. Darnell, D. Page and J. C. Mitchell, An automated decision-tree approach to predicting protein interaction hot spots, *Proteins*, 2007, **68**(4), 813–823.

23 S. J. Darnell, L. LeGault, J. C. Mitchell and K. F. C. Server, Interactive forecasting of protein interaction hot spots, *Nucleic Acids Res.*, 2008, **36**, W265–W269.

24 K. I. Cho, D. Kim and D. Lee, A feature-based approach to modeling protein-protein interaction hot spots, *Nucleic Acids Res.*, 2009, **37**(8), 2672–2687.

25 N. Tuncbag, A. Gursoy and O. Keskin, Identification of computational hot spots in protein interfaces: combining solvent accessibility and inter-residue potentials improves the accuracy, *Bioinformatics*, 2009, **25**(12), 1513–1520.

26 N. Tuncbag, O. Keskin and A. Gursoy, HotPoint: hot spot prediction server for protein interfaces, *Nucleic Acids Res.*, 2010, **38**, W402–W406.

27 J. F. Xia, X. M. Zhao, J. Song and D. S. Huang, APIS: accurate prediction of hot spots in protein interfaces by combining protrusion index with solvent accessibility, *BMC Bioinf.*, 2010, **11**, 174.

28 X. Zhu and J. C. Mitchell, KFC2: a knowledge-based hot spot prediction method based on interface solvation, atomic density, and plasticity features, *Proteins*, 2011, **79**(9), 2671–2683.

29 E. Cukuroglu, A. Gursoy and O. Keskin, HotRegion: a database of predicted hot spot clusters, *Nucleic Acids Res.*, 2012, **40**, D829–D833.

30 A. Shulman-Peleg, M. Shatsky, R. Nussinov and H. J. Wolfson, Spatial chemical conservation of hot spot interactions in protein-protein complexes, *BMC Biol.*, 2007, **5**, 43.

31 E. Guney, N. Tuncbag, O. Keskin and A. Gursoy, HotSprint: database of computational hot spots in protein interfaces, *Nucleic Acids Res.*, 2008, **36**, D662–D666.

32 A. Pavelka, E. Chovancova and J. Damborsky, HotSpot Wizard: a web server for identification of hot spots in
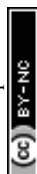
protein engineering, *Nucleic Acids Res.*, 2009, **37**, W376–W383.

33 D. M. Kruger and H. Gohlke, DrugScorePPI webserver: fast and accurate in silico alanine scanning for scoring protein-protein interactions, *Nucleic Acids Res.*, 2010, **38**, W480–W486.

34 T. Geppert, B. Hoy, S. Wessler and G. Schneider, Context-Based Identification of Protein-Protein Interfaces and "Hot-Spot" Residues, *Chem. Biol.*, 2011, **18**(3), 344–353.

35 A. del Sol and P. O'Meara, Small-world network approach to identify key residues in protein-protein interaction, *Proteins: Struct., Funct., Bioinf.*, 2005, **58**(3), 672–682.

36 J. Y. Li and Q. Liu, Double water exclusion: a hypothesis refining the O-ring theory for the hot spots at protein interfaces, *Bioinformatics*, 2009, **25**(6), 743–750.

37 N. Tuncbag, F. S. Salman, O. Keskin and A. Gursoy, Analysis and network representation of hotspots in protein interfaces using minimum cut trees, *Proteins: Struct., Funct., Bioinf.*, 2010, **78**(10), 2283–2294.

38 C. Pons, F. Glaser and J. Fernandez-Recio, Prediction of protein-binding areas by small-world residue networks and application to docking, *BMC Bioinf.*, 2011, **12**, 378.

39 T. L. Walunas, C. Y. Bakker and J. A. Bluestone, CTLA-4 ligation blocks CD28-dependent T cell activation, *J. Exp. Med.*, 1996, **183**(6), 2541–2550.

40 T. L. Walunas, D. J. Lenschow, C. Y. Bakker, P. S. Linsley, G. J. Freeman, J. M. Green, C. B. Thompson and J. A. Bluestone, CTLA-4 can function as a negative regulator of T cell activation, *Immunity*, 1994, **1**(5), 405–413.

41 L. L. Carter, L. A. Fouser, J. Jussif, L. Fitz, B. Deng, C. R. Wood, M. Collins, T. Honjo, G. J. Freeman and B. M. Carreno, PD-1 : PD-L inhibitory pathway affects both CD4(+) and CD8(+) T cells and is overcome by IL-2, *Eur. J. Immunol.*, 2002, **32**(3), 634–643.

42 A. D. Salama, T. Chitnis, J. Imitola, H. Akiba, F. Tushima, M. Azuma, H. Yagita, M. H. Sayegh and S. J. Khoury, Critical role of the programmed death-1 (PD-1) pathway in regulation of experimental autoimmune encephalomyelitis, *J. Exp. Med.*, 2003, **198**(1), 71–78.

43 E. J. Wherry, T cell exhaustion, *Nat. Immunol.*, 2011, **12**(6), 492–499.

44 W. Zou and L. Chen, Inhibitory B7-family molecules in the tumour microenvironment, *Nat. Rev. Immunol.*, 2008, **8**(6), 467–477.

45 D. S. Vinay, E. P. Ryan, G. Pawelec, W. H. Talib, J. Stagg, E. Elkord, T. Lichtor, W. K. Decker, R. L. Whelan, H. Kumara, E. Signori, K. Honoki, A. G. Georgakilas, A. Amin, W. G. Helferich, C. S. Boosani, G. Guha, M. R. Ciriolo, S. Chen, S. I. Mohammed, A. S. Azmi, W. N. Keith, A. Bilsland, D. Bhakta, D. Halicka, H. Fujii, K. Aquilano, S. S. Ashraf, S. Nowsheen, X. Yang, B. K. Choi and B. S. Kwon, Immune evasion in cancer: mechanistic basis and therapeutic strategies, *Semin. Canc. Biol.*, 2015, **35**(suppl.), S185–S198.

46 M. J. Butte, M. E. Keir, T. B. Phamduy, A. H. Sharpe and G. J. Freeman, Programmed death-1 ligand 1 interacts specifically with the B7-1 costimulatory molecule to inhibit T cell responses, *Immunity*, 2007, **27**(1), 111–122.

47 H. Dong, S. E. Strome, D. R. Salomao, H. Tamura, F. Hirano, D. B. Flies, P. C. Roche, J. Lu, G. Zhu, K. Tamada, V. A. Lennon, E. Celis and L. Chen, Tumor-associated B7-H1 promotes T-cell apoptosis: a potential mechanism of immune evasion, *Nat. Med.*, 2002, **8**(8), 793–800.

48 C. Sun, R. Mezzadra and T. N. Schumacher, Regulation and Function of the PD-L1 Checkpoint, *Immunity*, 2018, **48**(3), 434–452.

49 M. M. Zhan, X. Q. Hu, X. X. Liu, B. F. Ruan, J. Xu and C. Liao, From monoclonal antibodies to small molecules: the development of inhibitors targeting the PD-1/PD-L1 pathway, *Drug Discov. Today*, 2016, **21**(6), 1027–1036.

50 T. Powles, J. P. Eder, G. D. Fine, F. S. Braiteh, Y. Loriot, C. Cruz, J. Bellmunt, H. A. Burris, D. P. Petrylak, S. L. Teng, X. Shen, Z. Boyd, P. S. Hegde, D. S. Chen and N. J. Vogelzang, MPDL3280A (anti-PD-L1) treatment leads to clinical activity in metastatic bladder cancer, *Nature*, 2014, **515**(7528), 558–562.

51 S. L. Topalian, F. S. Hodi, J. R. Brahmer, S. N. Gettinger, D. C. Smith, D. F. McDermott, J. D. Powderly, R. D. Carvajal, J. A. Sosman, M. B. Atkins, P. D. Leming, D. R. Spigel, S. J. Antonia, L. Horn, C. G. Drake, D. M. Pardoll, L. Chen, W. H. Sharfman, R. A. Anders, J. M. Taube, T. L. McMiller, H. Xu, A. J. Korman, M. Jure-Kunkel, S. Agrawal, D. McDonald, G. D. Kollia, A. Gupta, J. M. Wigginton and M. Sznol, Safety, activity, and immune correlates of anti-PD-1 antibody in cancer, *N. Engl. J. Med.*, 2012, **366**(26), 2443–2454.

52 R. S. Herbst, J. C. Soria, M. Kowanetz, G. D. Fine, O. Hamid, M. S. Gordon, J. A. Sosman, D. F. McDermott, J. D. Powderly, S. N. Gettinger, H. E. K. Kohrt, L. Horn, D. P. Lawrence, S. Rost, M. Leabman, Y. Y. Xiao, A. Mokatrin, H. Koeppen, P. S. Hegde, I. Mellman, D. S. Chen and F. S. Hodi, Predictive correlates of response to the anti-PD-L1 antibody MPDL3280A in cancer patients, *Nature*, 2014, **515**(7528), 563–567.

53 F. Zhang, H. Wei, X. Wang, Y. Bai, P. Wang, J. Wu, X. Jiang, Y. Wang, H. Cai, T. Xu and A. Zhou, Structural basis of a novel PD-L1 nanobody for immune checkpoint blockade, *Cell Discov.*, 2017, **3**, 17004.

54 T. G. Phan, G. V. Long and R. A. Scolyer, Checkpoint inhibitors for cancer immunotherapy. Multiple checkpoints on the long road towards cancer immunotherapy, *Immunol. Cell Biol.*, 2015, **93**(4), 323–325.

55 S. L. Topalian, C. G. Drake and D. M. Pardoll, Targeting the PD-1/B7-H1(PD-L1) pathway to activate anti-tumor immunity, *Curr. Opin. Immunol.*, 2012, **24**(2), 207–212.

56 T. Zarganes-Tzitzikas, M. Konstantinidou, Y. Gao, D. Krzemien, K. Zak, G. Dubin, T. A. Holak and A. Domling, Inhibitors of programmed cell death 1 (PD-1): a patent review (2010-2015), *Expert Opin. Ther. Pat.*, 2016, **26**(9), 973–977.

57 B. R. Huck, L. Kotzner and K. Urbahns, Small Molecules Drive Big Improvements in Immuno-Oncology Therapies, *Angew. Chem., Int. Ed. Engl.*, 2018, **57**(16), 4412–4428.
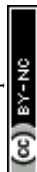
58 Y. Han, Y. Gao, T. He, D. Wang, N. Guo, X. Zhang, S. Chen and H. Wang, PD-1/PD-L1 inhibitor screening of caffeoylquinic acid compounds using surface plasmon resonance spectroscopy, *Anal. Biochem.*, 2018, **547**, 52–56.

59 K. Guzik, K. M. Zak, P. Grudnik, K. Magiera, B. Musielak, R. Torner, L. Skalniak, A. Domling, G. Dubin and T. A. Holak, Small-Molecule Inhibitors of the Programmed Cell Death-1/Programmed Death-Ligand 1 (PD-1/PD-L1) Interaction via Transiently Induced Protein States and Dimerization of PD-L1, *J. Med. Chem.*, 2017, **60**(13), 5857–5867.

60 K. M. Zak, R. Kitel, S. Przetocka, P. Golik, K. Guzik, B. Musielak, A. Domling, G. Dubin and T. A. Holak, Structure of the Complex of Human Programmed Death 1, PD-1, and Its Ligand PD-L1, *Structure*, 2015, **23**(12), 2341–2348.

61 Y. Chen, P. Liu, F. Gao, H. Cheng, J. Qi and G. F. Gao, A dimeric structure of PD-L1: functional units or evolutionary relics?, *Protein Cell*, 2010, **1**(2), 153–160.

62 D. Y. W. Lin, Y. Tanaka, M. Iwasaki, A. G. Gittis, H. P. Su, B. Mikami, T. Okazaki, T. Honjo, N. Minato and D. N. Garboczi, The PD-1/PD-L1 complex resembles the antigen-binding Fv domains of antibodies and T cell receptors, *Proc. Natl. Acad. Sci. U.S.A.*, 2008, **105**(8), 3011–3016.

63 X. W. Zhang, J. C. D. Schwartz, X. L. Guo, S. Bhatia, E. H. Cao, L. P. Chen, Z. Y. Zhang, M. A. Edidin, S. G. Nathenson and S. C. Almo, Structural and functional analysis of the costimulatory receptor programmed death-1, *Immunity*, 2004, **20**(3), 337–347.

64 Y. Yan, M. Yang, C. G. Ji and J. Z. H. Zhang, Interaction Entropy for Computational Alanine Scanning, *J. Chem. Inf. Model.*, 2017, **57**(5), 1112–1122.

65 L. Q. Qiu, Y. N. Yan, Z. X. Sun, J. N. Song and J. Z. H. Zhang, Interaction entropy for computational alanine scanning in protein-protein binding, *Wiley Interdiscip. Rev.: Comput. Mol. Sci.*, 2018, **8**(2), e1342.

66 L. Duan, X. Liu and J. Z. Zhang, Interaction Entropy: A New Paradigm for Highly Efficient and Reliable Computation of Protein-Ligand Binding Free Energy, *J. Am. Chem. Soc.*, 2016, **138**(17), 5722–5728.

67 D. Huang, Y. Qi, J. Song and J. Z. H. Zhang, Calculation of hot spots for protein-protein interaction in p53/PMI-MDM2/MDMX complexes, *J. Comput. Chem.*, 2019, **40**(9), 1045–1056.

68 A. Waterhouse, M. Bertoni, S. Bienert, G. Studer, G. Tauriello, R. Gumienny, F. T. Heer, T. A. P. de Beer, C. Rempfer, L. Bordoli, R. Lepore and T. Schwede, SWISS-MODEL: homology modelling of protein structures and complexes, *Nucleic Acids Res.*, 2018, **46**(W1), W296–W303.

69 D. A. Case, V. Babin, J. T. Berryman, R. M. Betz, Q. Cai, D. S. Cerutti, T. E. Cheatham, T. A. Darden, R. E. Duke, H. Gohlke, A. W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T. S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K. M. Merz, F. Paesani, D. R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C. L. Simmerling, W. Smith, J. Swails, Walker, J. Wang, R. M. Wolf, X. Wu and P. A. Kollman, *Amber 14*, University of California, San Francisco, 2014.

70 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, Comparison of Simple Potential Functions for Simulating Liquid Water, *J. Chem. Phys.*, 1983, **79**(2), 926–935.

71 J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB, *J. Chem. Theory Comput.*, 2015, **11**(8), 3696–3713.

72 T. Darden, D. York and L. Pedersen, Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems, *J. Chem. Phys.*, 1993, **98**(12), 10089–10092.

73 J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, *J. Comput. Phys.*, 1977, **23**(3), 327–341.

74 R. Zwanzig, Nonlinear generalized Langevin equations, *J. Stat. Phys.*, 1973, **9**(3), 215–220.

75 H. J. C. Berendsen, J. P. M. Postma, W. F. Vangunsteren, A. Dinola and J. R. Haak, Molecular-Dynamics with Coupling to an External Bath, *J. Chem. Phys.*, 1984, **81**(8), 3684–3690.

76 P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. H. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case and T. E. Cheatham, Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models, *Acc. Chem. Res.*, 2000, **33**(12), 889–897.

77 J. Srinivasan, T. E. Cheatham, P. Cieplak, P. A. Kollman and D. A. Case, Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate-DNA helices, *J. Am. Chem. Soc.*, 1998, **120**, 9401–9409.

78 D. Shi, S. Zhou, X. Liu, C. Zhao, H. Liu and X. Yao, Understanding the structural and energetic basis of PD-1 and monoclonal antibodies bound to PD-L1: A molecular modeling perspective, *Biochim. Biophys. Acta*, 2018, **1862**(3), 576–588.

79 M. Ahmed and K. Barakat, The Too Many Faces of PD-L1: A Comprehensive Conformational Analysis Study, *Biochemistry*, 2017, **56**(40), 5428–5439.

80 W. Liu, B. Huang, Y. Kuang and G. Liu, Molecular dynamics simulations elucidate conformational selection and induced fit mechanisms in the binding of PD-1 and PD-L1, *Mol. Biosyst.*, 2017, **13**(5), 892–900.