Discovery of Potential PD-1 and PD-L1 Interaction Inhibitors Using Combined Molecular Modeling Approaches

Özlem ULUCAN[1,*]

[1]*İstanbul Bilgi University, Faculty of Engineering and Natural Sciences, Department of Genetics and Bioengineering, 34060, İstanbul, Türkiye*
ozlem.ulucan@bilgi.edu.tr, ORCID: 0000-0002-7442-5728

Abstract

Immune checkpoints are regulators of the immune system that maintain immune homeostasis and prevent autoimmunity. Cancer cells often manipulate immune checkpoint mechanisms to escape anti-tumor immune response by overexpressing the immune checkpoint ligands. Thus, the interactions between the immune checkpoint receptors and ligands attracted attention and were proven to be effective targets in treating cancer. In this study, combining several computational approaches, we discovered small molecules that effectively bind to the ligand Programmed Cell Death Ligand 1 (PD-L1) and have the potential to hamper its interaction with the negative immune checkpoint receptor Programmed Cell Death Protein-1 (PD-1). Different pharmacophore models were constructed using triple and quadruple combinations of the interface residues on PD-1, which were used later for scanning the ZINC15 database. 12714 small molecules were retrieved and virtually screened using molecular docking calculations. The complexes of promising small molecules with PD-L1 were further evaluated using energetic and structural analyses. Our results suggest that the three small molecules ZINC21075815, ZINC70692276, and ZINC64031730 retrieved from the ZINC15 database establish stable and energetically favorable interactions with PD-L1 at the hot region consisting of the residues Tyr

56, Glu 58, Arg 113, Met 115, and Tyr 123. These molecules can be used as a starting point to develop more effective and selective anti-PD-1/PD-L1 inhibitors.

Keywords: Inhibitory T cell receptors, immunomodulation; PD-1 receptor; computer-aided drug design; Molecular docking; Molecular dynamic simulations.

Moleküler Modelleme Yaklaşımları Kullanılarak Potansiyel PD-1 ve PD-L1 Etkileşim İnhibitörlerinin Keşfi

Öz

Bağışıklık kontrol noktaları, bağışıklık homeostazını sürdüren ve otoimmüniteyi önleyen bağışıklık sistemi düzenleyicileridir. Kanser hücreleri, anti-tümör bağışıklık yanıtından kaçmak için sıklıkla bağışıklık kontrol noktası ligandlarını aşırı ifade ederek bağışıklık kontrol noktası mekanizmalarını manipüle eder. Bu nedenle, bağışıklık kontrol noktası reseptörleri ve ligandları arasındaki etkileşimler araştırmacıların dikkatini çekmiş ve bu etkileşimin kanser tedavisinde etkili hedefler olduğu gösterilmiştir. Bu çalışmanın temel amacı, çeşitli hesaplamalı yaklaşımları birleştirilerek, ligand PD-L1'e etkili bir şekilde bağlanan ve negatif bağışıklık kontrol noktası reseptörü PD-1 ile etkileşimini engelleme potansiyeli taşıyan küçük moleküller keşfetmektir. Bu amaç doğrultusunda PD1/PD-L1 ara yüzünde PD-1 üzerindeki amino asitlerin üçlü ve dörtlü kombinasyonları kullanılarak farklı farmakofor modelleri oluşturuldu. Farmakofor modelleri kullanılarak ZINC15 veri tabanı tarandı. ZINC15 veri tabanından indirilen 12714 küçük molekül moleküller için moleküler kenetleme çalışmaları gerçekleştirildi. Moleküler kenetleme çalışmalarında öne çıkan küçük moleküllerin PD-L1 ile komplekslerinin moleküler dinamik simülasyonları gerçekleştirildi ve bu kompleksler enerjik ve yapısal analizler kullanılarak detaylıca incelendi. Bulgularımız, ZINC15 veri tabanından alınan ZINC21075815, ZINC70692276 ve ZINC64031730 adlı üç küçük molekülün, Tyr 56, Glu 58, Arg 113, Met 115 ve Tyr 123 amino asitlerinden oluşan sıcak bölgede PD-L1 ile kararlı ve enerjik açıdan uygun etkileşimler kurduğunu göstermektedir. Bu moleküller, daha etkili ve seçici anti-PD-1/PD-L1 inhibitörleri geliştirmek için bir başlangıç noktası olarak kullanılabilir.

Anahtar Kelimeler: İnhibitör T hücre reseptörleri; İmmünomodülasyon; PD-1 reseptörü; Bilgisayar destekli ilaç tasarımı; Moleküler kenetleme; Moleküler dinamik simülasyonları.

1. Introduction

The immune system's role in controlling tumor growth and spread is now well established. However, tumor cells may evolve mechanisms to escape immune surveillance and suppress immune response. One of the frequently used mechanisms by tumor cells is the exploitation of

co-inhibitory immune checkpoints via overexpression of immune checkpoint ligands such as PDL-1[1], CD155[2], and CD112[3]. Revealing how immune checkpoints work has paved the way for new strategies to fight against cancer. The primary focus in the rapidly emerging field of cancer immunotherapy is to control and redirect the host immune response to recognize and eliminate cancer cells. Developing immune checkpoint inhibitors, mainly antibodies, has provided a remarkable advancement in this field. Those inhibitors exhibit their effects by interrupting co-inhibitory signals and reactivating antitumor immune responses. PD-1/PD-L1 and CTLA-4 inhibitors have shown promising effects among the immune checkpoint inhibitors. Some have been approved for treating different types of cancers, while clinical trials for others are ongoing [4–6]. PD-1 (CD279), one of the well-studied immune checkpoint receptors, delivers a co-inhibitory signal when it binds to one of its ligands, PD-L1 and PD-L2. Several monoclonal antibodies targeting PD-1/PD-L1 interaction have been approved for monotherapy or combinatorial therapy with other therapeutic agents, such as other immune checkpoint inhibitors, radiation, and chemotherapy against different cancer types [7, 8].

Despite the success of monoclonal antibodies in immunotherapy against cancer, some of their disadvantages have led researchers to seek small-molecule or peptide alternatives [7, 9]. The production cost of monoclonal antibodies remains exceptionally high. While their high molecular weight prevents them from diffusing into the tumor, their high affinity for the target keeps them on the outer edge of the tumor. Additionally, their unintended interactions with various cell surface receptors prolong their retention in circulation and delay their arrival at the target site or excretion from the body [10, 11]. When administered as a monotherapy, monoclonal antibodies reported common side effects are fatigue, rash, diarrhea, colitis, endocrine and hepatic toxicities, pneumonitis, neurological syndromes, and ocular toxicities [12]. For the anti-PD-1 monoclonal antibodies, severe and sometimes fatal lung-related autoimmune adverse effects have been reported [13, 14]. This is partly due to PD-1 inhibition simultaneously disrupting the binding of both ligands, PD-L1 and PD-L2, which play an essential role in maintaining immune homeostasis in the lung. Targeting PD-L1 has some advantages over targeting PD-1[15]. When PD-L1 is targeted, PD-L2 remains active, reducing the risk of developing severe inflammatory toxicity in organs such as the lungs [16]. Moreover, targeting PD-L1 also disrupts the interaction of PD-L1 with B7-1, the function of which is to inhibit T-cell activation and cytokine production [17, 18].

Small molecules have numerous key advantages, such as being cheap, being administered orally, having low immunogenic potential, and kinetic advantages over significant antibodies [19]. However, the discovery process for these molecules faces several difficulties, especially for protein-protein interaction inhibitors [20]. This is due to the lack of natural small-molecule

binders that can be used as starting points, the relatively large surface area buried by both protein partners, the flatness of the interfaces, and the absence of narrow, deep cavities at the interfaces. Fortunately, we also know that the small-molecule effectors do not need to cover the entire interface due to the small number of interface residues that contribute substantially to binding affinity [21]. Yet, the protein-protein interaction modulators have been reported to be larger on average than classical drug molecules and tend to be more hydrophobic and rigid. They have fewer hydrogen bonding groups than classical drug-like molecules [22]. PD-1/PD-L1 interaction forms a typical example of protein-protein interfaces. Zak et al. [23] reported that the relatively flat interface involves polar and nonpolar interactions and has a moderately large, buried surface area. Their structural analysis showed three major hotspots on PD-L1 that can be important for drug design. The first hotspot is a pocket of predominantly hydrophobic character that accommodates Ile 134 of PD-1. The second hotspot is a neighbor of the first one and accommodates Ile 126 of PD-1. Finally, the third hot spot is an extended groove where the PD-1 residues Tyr 68, Gln 75, and Thr 76 of PD-1 bind.

In this study, we combined several computational approaches to discover new small-molecule candidates that target the PD-1/PD-L1 interaction. We scanned the ZINC15 database and retrieved 12714 small molecules that were further evaluated using molecular docking and molecular dynamic (MD) simulations. Detailed energetic and structural analyses suggested three molecules that form stable and energetically favorable interactions with PD-L1 at the region consisting of the residues Tyr 56, Glu 58, Arg 113, Met 115, and Tyr 123.

2. Materials and Methods

2.1. Generation of Small Molecule Set

The atomic coordinates of the PD-1/PD-L1 complex were retrieved from the PDB databank with the PDB ID 4ZQK. PD-1 (chain B) was used to generate structure-based pharmacophore models, whereas PD-L1 (chain A) was used as the target protein in docking calculations.

All possible triple combinations of the interfacial residues were generated. Then, the triple combinations with a distance larger than 7 Å between any of their two residues were eliminated, resulting in 63 triple combinations in total. These triple residue combinations were used to construct pharmacophores, which were used later to screen the ZINC15 database [24]. 12714 small molecules were retrieved from the ZINC15 database using these pharmacophores. The pharmacophore models that resulted in a vast number of hits in the ZINC15 database were reconstructed by including another residue within 7 Å distance to the original residues. Thus, 8 more pharmacophore models were constructed using quadruple combinations of interface

residues. PocketQuery was combined with ZINCPharmer to construct pharmacophore models and search the ZINC15 database [25].

2.2. Molecular Docking Calculations

We carried out two sets of molecular docking simulations using the molecular docking software AutoDock 4.2 [26]. In the first set of calculations, we docked 12714 small molecules to the target protein PD-L1 using the standard Lamarckian genetic algorithm protocol, where the size of the initial population and the maximum number of energy evaluations were set to 300 and 5 million, respectively. Afterward, we ranked the small molecules based on their free binding energies to PD-L1. The top 600 molecules were selected and evaluated further in the second set of docking calculations. In the second set of calculations, the maximum number of energy evaluations and the number of independent runs were increased gradually. Here, we aimed to generate a single densely populated cluster with low energy to achieve convergence in the results. The results are based on 100 independent runs, with a maximum of 30 million energy evaluations set for each run.

The target protein PD-L1 and the small molecules were prepared for docking using the AutoDock Tools version 1.5.6 [26]. The torsions of the small molecules were determined by employing the AutoTors function of AutoDock Tools. Amide and ring torsions were kept rigid, while all other torsions were treated as flexible. Gasteiger atomic charges were assigned to both the protein and the small molecules. The non-polar hydrogen atoms were treated implicitly. Grid maps were generated with 0.375Å spacing by the AutoGrid program. The grid dimensions were chosen to ensure that all of the residues on PD-L1 at the PD-1/PD-L1 interface were included. Thus, the grid dimensions were 60Å x 80Å x 60Å. The initial position of the ligand was set randomly in both sets of docking calculations.

We assumed that a docking calculation was converged when 20% of the 100 independent runs resulted in the same binding conformation. Subsequently, we ranked the molecules that fulfill the convergence criterion according to their binding free energies and selected the top 20 molecules for further analysis.

2.3. Parameterization of Small Molecules

Topologies and parameters for the small molecules selected for molecular dynamic (MD) simulations were obtained using the Antechamber Python Parser Interface (ACPYPE) [27]. The spatial coordinates of each small molecule were obtained from the top-ranked conformation within the dominant cluster. The AM1-BCC charges were derived after the optimization of the

structure [28]. Force constants and equilibrium values for the bond lengths, angles, and dihedrals of the compounds were borrowed from the General Amber Force Field (GAFF) [29].

2.4. Molecular Dynamics (MD) Simulations

In this work, we performed 11 ns conventional MD simulations of 20 protein-ligand complexes solvated in explicit water, utilizing Gromacs version 5.1.2 [30]. For the MD simulation of each complex, we used the highest-ranking conformation within the dominant cluster derived from molecular docking runs as starting structures. The complex was inserted in a cubic box filled with TIP3P water molecules, and the whole system was neutralized by adding counterions. The Amber force field ff99sb-ildn [31] was utilized for modeling the interactions of protein and ions. The short-range non-bonded interactions were cut at 1.2 nm. The electrostatic interactions were treated using the particle mesh Ewald method. A dispersion correction was implemented for pressure and energy, while periodic boundary conditions were enforced in every direction.

Energy minimization was performed using a combination of algorithms, performing 1000 steps of the conjugate gradient method, with one steepest descent step inserted every 10 steps. After minimization, we equilibrated each system by running 100 ps of molecular dynamics (MD) in the NVT ensemble. During equilibration, the non-hydrogen atoms of the protein were harmonically restrained with a force constant of 1000 kJ mol^{-1} nm^{-2}. Solute and solvent atoms were independently coupled to temperature baths where the temperature was maintained at 300 K with a coupling time of 0.5 ps. A leap-frog algorithm was used to integrate Newton's motion equations, with a time step set to 2 fs. We used an NPT ensemble for the production phase without applying any restraints. The pressure was kept at 1 bar with the help of the Parrinello-Rahman barostat [32], setting the time constant to 0.5 ps. To regulate the temperature, we employed the Nose-Hoover temperature coupling method [33, 34], setting the reference temperature to 300 K and the time constant to 0.5 ps. The generated trajectories were used for free energy calculations.

2.5. Binding Free Energy Calculations

The end-point MM(PB/GB)SA method integrates gas-phase and solvation energy contributions derived using an implicit solvent model for the ligand, receptor, and complex. Additionally, solute entropy contributions are included to refine the total free energy.

In this work, we used the single trajectory approach to derive the binding free energies for the selected compounds. We treated the first 1 ns of each trajectory as the equilibrium phase. Before the free energy calculations, water molecules and ions were removed from the trajectory files. The computations were performed with the MMPBSA.py script of AmberTools18 [35]. For

the calculation of the gas phase free energy contributions, the Sander program in AmberTools18 was utilized.

The polar contribution of solvation-free energy was derived using the continuum solvent model GB-OBC introduced by Onufriev et al. [36]. The values 1.0, 0.8, and 4.85 were used for the parameters α, β, and γ. As Onufriev et al. [37] suggested, we utilized mbondi2 radii for effective Born radii. The LCPO method [38] was used to approximate the nonpolar components of solvation-free energy. Translational, rotational, and vibrational entropies were calculated individually and then summed to estimate the solute configurational entropy.

We computed the translational and rotational entropies from their gas phase partition functions. Frequencies for the vibrational modes were derived by normal-mode analysis following energy minimization. The entropy was averaged using 200 equally spaced snapshots taken from the trajectory.

To calculate the standard error of the binding free energy, we combined the standard errors for the effective energy from the MMGBSA method and the solute entropy using quadrature. The resulting standard error was then used to create a 95% Welch-Satterthwaite confidence interval for the binding free energy.

As positive controls, two well-characterized PD-L1 dimerizers (BMS-202 [39] and A1D9R [40]) and two compounds that bind PD-L1 in its monomeric form (CA-170 [41] and the cyclic peptide, Peptide 104, [42]) were used. For BMS-202 (5J89, chain C) and A1D9R (9INU, chain A), docked models and available crystal structures were used as starting structures for the simulations. For Peptide 104, only the crystal structure (PDB ID: 7OUN) was used, as the number of torsions exceeded the maximum allowed for docking. In the case of CA-170, a representative docked conformation was obtained as explained in the Materials and Methods, as no crystal structure was available.

3. Results and Discussion

3.1. Generation of Small-molecule Dataset

The interface residues of the PD-1/PD-L1 complex were retrieved from the PDBsum database. PD-1 and PD-L1 contribute to the interface with 17 residues, corresponding to an interface area of 1557 Å^2 that forms 2 salt bridges, 11 hydrogen bonds, and many other non-bonded close contacts. Using the interface residues on PD-1, 71 different pharmacophore models were constructed to be used later for scanning the ZINC15 database. Some of the pharmacophores did not match with any small molecules, while others resulted in millions of molecules. We

modified the pharmacophore models, which resulted in millions of molecules, including another nearby residue that fulfilled our distance criterion. More than 12000 small molecules were retrieved from the ZINC15 database using the pharmacophore models constructed from the PD-1 interfacial residues. Based on the residue combinations used for pharmacophore modeling, the number of retrieved molecules ranges from 1 to 4367.

3.2. Virtual Screening by Molecular Docking

We selected the molecules based on their binding free energies in the first set of docking calculations. This calculation set resulted in 600 promising small molecules further evaluated in the second set of calculations, where a more extensive search was performed. The molecules were filtered based on the convergence criterion (see the Methods Section) and ranked according to their binding free energies. Table 1 lists the top 20 molecules together with some of their molecular properties and their binding free energies obtained from molecular docking. As seen in Table 1, the most promising 20 molecules have quite different molecular properties. Among those molecules, there exists one tripeptide (ZINC83308150) and 2 macrocycles (ZINC94303267 and ZINC94303139). The molecular weight, logP values, and the charges for those molecules range from 350.5 to 765, -6.9 to 7.1, and 0 to 5, respectively (see Table 1).

We examined the docked conformations of the top 20 molecules and found that all of them, although they are of different molecular characters, were bound to the same site on PD-L1 (see Fig. 1b). This site covers two of the three hotspots reported before by Zak et al. (see Fig. 1a). Even though the grid box used in docking calculations spanned the whole interface on PD-L1, the site preference of the top 20 molecules suggests that the two hotspots (HS-1 and HS-2) may, in fact, be druggable.

Table 1: The top 20 promising molecules came out of molecular docking calculations. The reported binding free energies were obtained from the second set of molecular docking calculations. The molecular weight and logP values were taken from the ZINC15 database.

Molecule ID	Molecular Weight	LogP	Charge	Binding free energy (kcal/mol)
ZINC77271764	539.6	-6.9	5	-17.3
ZINC77271775	539.6	-6.9	5	-15.8
ZINC67910521	666.8	1.8	0	-14.2
ZINC83308150	426.6	-1.7	0	-12.1
ZINC94303267	720.9	1.3	2	-11.9
ZINC67903231	682.8	0.4	0	-11.7
ZINC77257248	678.9	3.3	0	-11.4
ZINC22048461	449.7	2.2	2	-11.2
ZINC36047071	433.4	-1.8	0	-11.2
ZINC39500620	412.5	1.2	1	-11.0

ZINC39362507	378.5	0.9	1	-10.8
ZINC21075815	464.6	4.7	0	-10.6
ZINC70692276	536.6	0.4	0	-10.5
ZINC20832674	487.6	2.1	0	-10.5
ZINC59676745	632.8	7.1	0	-10.4
ZINC40173059	505.5	5.8	0	-10.3
ZINC64031730	350.5	0.4	1	-10.3
ZINC77257358	620.8	3.1	0	-10.2
ZINC35456718	520.5	0.1	0	-10.1
ZINC94303139	765	1.9	1	-10.1
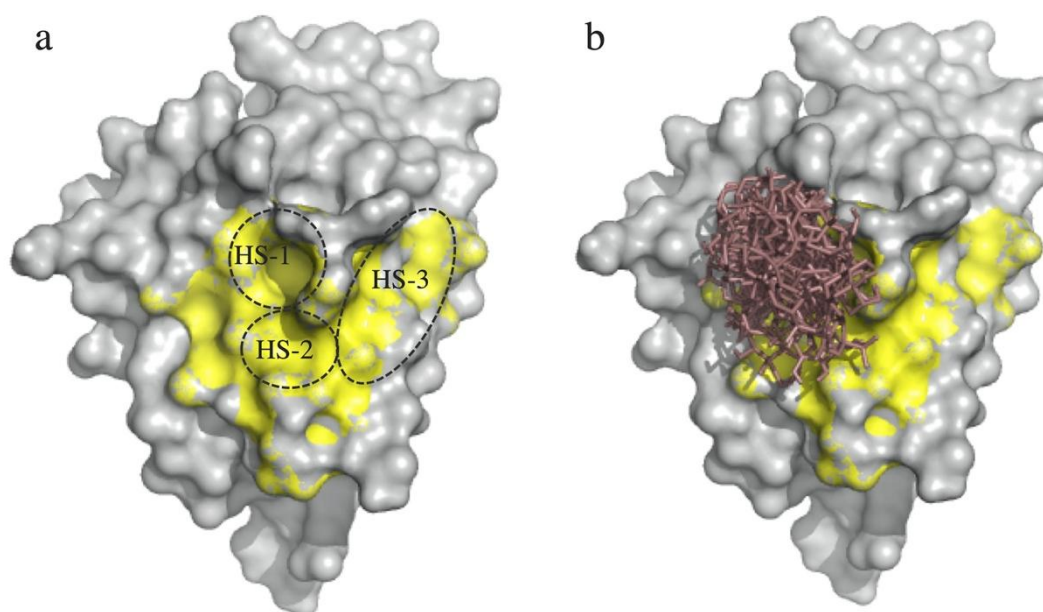
Figure 1: Site preference of the top 20 molecules on PD-L1. a) The 3 hotspots (HS-1, HS-2, and HS-3) on PD-L1 reported by Zak et al. b) Docked top 20 molecules shown with stick representation. The PD-L1 residues that contribute to the PD-1/PD-L1 interface are depicted in the yellow surface.

3.3. Energetic and Structural Evaluation of PD-L1/Small-molecule Complexes

We performed 11 ns MD simulations of 20 protein-ligand complexes derived from molecular docking calculations. For all analyses performed using MD simulation trajectories, the first 1 ns was treated as the equilibrium phase and discarded.

Table 2 summarizes the energy contributions of each term obtained from MMGBSA and entropy calculations. Free energy calculations put forward 6 molecules (ZINC77271775, ZINC67910521, ZINC67903231, ZINC21075815, ZINC70692276, and ZINC64031730) whose free energies of binding vary from -7.1 to -18.1 kcal/mol (see Table 2). Binding free energies for the other 14 molecules are higher than -5 kcal/mol, which indicates nonspecific binding or no binding. Therefore, we decided to further assess the complexes for the 6 promising small molecules using detailed structural analysis.

Table 2: MMGBSA and entropy contributions to the binding free energies for the top 20 molecules. VdW and EEL refer to the average Van der Waals and electrostatic interaction energies, respectively, while EGB and ESURF represent the average polar and nonpolar solvation-free energies. TOTAL indicates the MMGBSA effective energy, which includes the contribution of solute entropy. The values in parentheses represent the standard deviations, and the brackets indicate the 95% Welch-Satterthwaite confidence interval for the binding free energy.

Molecule ID	MMGBSA					Entropy	Binding free energy (kcal/mol)
	VdW	EEL	EGB	ESURF	TOTAL		
ZINC77271764	-16.1	-283.2	279.3	-3.1	-23.1 (3.9)	-19.8 (3.6)	-3.3 [-3.8, -2.8]
ZINC77271775	-2.5	-392.0	360.8	-2.7	-36.4 (6.5)	-18.3 (3.4)	-18.1 [-18.7, -17.5]
ZINC67910521	-32.9	-15.5	24.0	-4.1	-28.5 (3.9)	-19.3 (3.0)	-9.2 [-9.7, -8.7]
ZINC83308150	-22.2	-10.2	21.9	-3.2	-13.7(4.1)	-18.6 (3.9)	4.9 [4.3, 5.5]
ZINC94303267	-12.6	-80.0	81.6	-2.2	-13.2 (5.4)	-18.2 (3.1)	5.0 [4.5, 5.5]
ZINC67903231	-34.5	-12.1	24.9	-4.5	-26.2 (3.8)	-18.0 (3.4)	-8.5 [-8.7, -7.7]
ZINC77257248	-26.4	-16.6	22.1	-3.7	-24.6 (6.2)	-20.1 (3.3)	-4.5 [-5.0, -4.0]
ZINC22048461	-12.1	-162.3	153.3	-2.5	-23.6 (6.5)	-19.6 (3.7)	-4.0 [-4.6, -3.4]
ZINC36047071	-25.0	-36.7	45.1	-3.4	-20.0 (3.9)	-19.5 (2.5)	-0.5 [-0.9, -0.1]
ZINC39500620	-25.5	-32.4	43.2	-3.1	-17.8 (3.8)	-19.7 (2.9)	1.9 [1.5, 2.3]
ZINC39362507	-16.4	-76.7	78.3	-2.2	-17.0 (4.3)	-19.5 (3.4)	2.5 [2.0, 3.0]
ZINC21075815	-25.1	-37.9	39.3	-3.8	-27.5 (4.0)	-18.9 (3.0)	-8.6 [-9.1, -8.1]
ZINC70692276	-25.0	-40.4	43.9	-3.4	-24.9 (3.8)	-17.8 (2.1)	-7.1 [-7.4, -6.7]
ZINC20832674	-30.8	-13.2	25.5	-4.4	-22.9 (4.9)	-21.9 (4.2)	-1.0 [-1.6, -0.4]
ZINC59676745	-33.3	-8.9	21.4	-4.1	-24.9 (4.5)	-22.6 (3.0)	-2.3 [-2.8, -1.8]
ZINC40173059	-26.7	-9.5	18.3	-3.1	-21.0 (3.4)	-18.3 (2.2)	-2.7 [-3.0, -2.4]
ZINC64031730	-21.9	-75.0	72.6	-2.4	-26.7 (3.0)	-18.8 (2.8)	-7.9 [-8.3, -7.5]
ZINC77257358	-34.2	-10.2	22.7	-4.5	-26.2 (4.2)	-21.9 (3.2)	-4.3 [-4.8, -3.8]
ZINC35456718	-27.3	-20.3	29.3	-3.7	-22.0 (8.3)	-21.9 (3.5)	-0.1 [-0.7, 0.5]
ZINC94303139	-22.0	-26.8	39.0	-3.1	-12.9 (2.9)	-21.3 (3.5)	8.4 [7.9, 8.9]
Peptide 104	-49.9	-144.0	158.2	-7.9	-43.6 (10.5)	-31.0 (3.9)	-12.6 [-13.4, -11.8]

Table 3 depicts some structural properties of the complexes formed by PD-L1 and 6 small molecules. The root mean square deviation (RMSD) values were computed using the docked

conformations of the molecules as reference structures after aligning the protein conformations. While RMSD values show how the molecular conformations during the simulations differed from the docked conformations, the standard deviations show how stable the binding conformations were throughout the simulations (See Table 3). Fig. 2 shows how the RMSD values for the promising 6 small molecules change over time. The compounds ZINC77271775 and ZINC67910521 adopted conformations that were different from their docked conformations during the simulations. The compound ZINC67903231 did not adopt a stable conformation over time (see Fig. 2). However, ZINC21075815, ZINC70692276, and ZINC64031730 preserved conformations close to their docked conformations throughout the simulations (see Fig. 2). The average RMSD values were 2.4 Å, 2.5 Å, and 3.1 Å for ZINC21075815, ZINC70692276, and ZINC64031730, respectively (see Table 3).
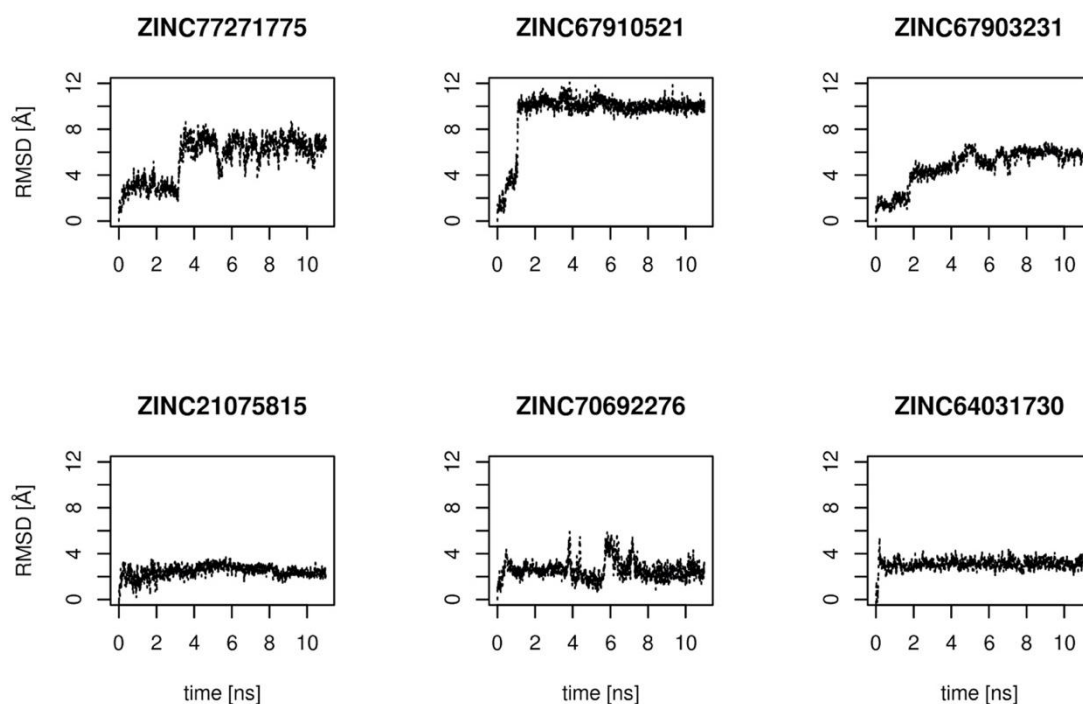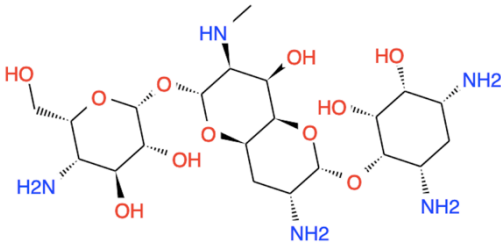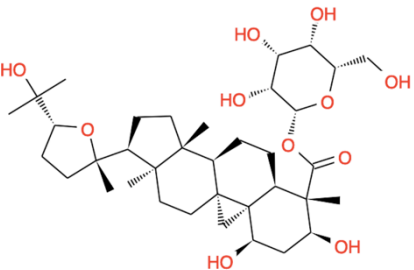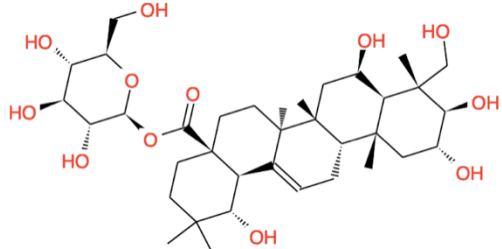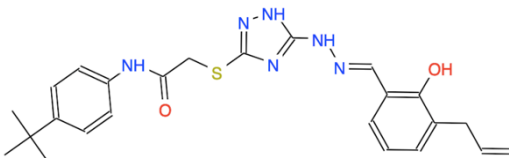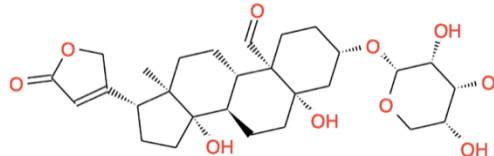

Figure 2: RMSD values for 6 promising molecules over time. The RMSD values were computed using the docked conformations of the molecules as reference structures after aligning the protein conformations.

The buried solvent-accessible surface area (BSASA) was calculated as the sum of the solvent-accessible surface area of PD-L1 and the small molecule minus the solvent-accessible surface area of the complex. Note that we did not assume interface symmetry and did not divide the total buried solvent accessible surface area by two. Hence, the values in Table 3 involve the surface contributions of both binding partners. The smallest (ZINC64031730) and the largest (ZINC77271775) BSASA values given in Table 3 account for 43% and 63% of the solvent-accessible surface area buried upon complexation of PD-1 and PD-L2 (1557 Å^2). Except for

ZINC77271775, our small molecules tend to form fewer hydrogen-bonding contacts than average drug-like molecules. The average BSASA values and the number of hydrogen-bonding contacts the molecules formed are comparable to those previously reported for TIMBAL molecules [22].

Table 3: Structural properties of 6 promising molecules and their complexes. The docked conformations of the molecules served as the reference points for RMSD calculations following the alignment of the protein conformations. Values given in parentheses stand for standard deviations. BSASA values and several hydrogen bonds were computed using the Gromacs routines gmx-sasa and gmx-hbond, respectively, with default settings.

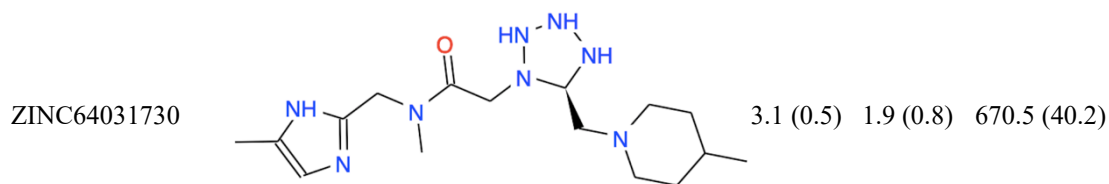| Molecule ID	Chemical Structure	RMSD	H-bonds	BSASA (Å^2)
ZINC77271775		5.5 (1.8)	6.2 (1.0)	979.0 (63.0)
ZINC67910521		9.4 (2.3)	1.6 (0.6)	922.3 (52.2)
ZINC67903231		4.7 (1.6)	0.7 (0.8)	950.9 (56.4)
ZINC21075815		2.4 (0.5)	1.6 (1.2)	772.0 (50.4)
ZINC70692276		2.5 (0.8)	2.3 (0.8)	942.4 (48.3)

ZINC64031730   3.1 (0.5)   1.9 (0.8)   670.5 (40.2)

Figure 3 depicts the binding conformations of the three molecules (ZINC21075815, ZINC70692276, and ZINC64031730) that preserved conformations close to their docked conformations. As evident from Fig. 3, the PD-L1 surface exhibits high structural plasticity; therefore, the three hotspots shown in Fig.1a are no longer readily distinguishable. The rearrangement of the side chains of the contact residues (the RMSD value for the backbone of PD-L1 was smaller than 2 Å throughout the simulations) enabled stable interactions with the small molecules. We utilized the program LigPlot+ [43] to identify the PD-L1 residues forming contact with the small molecules. Even though the binding region on PD-L1 adopts a different conformation for each molecule, the contact residues are common (Tyr 56, Glu 58, Arg 113, Met 115, and Tyr 123) for all molecules. The common contact residues Tyr 56, Glu 58, Arg 113, Met 115, and Tyr 123 constitute the first hotspot, a pocket of predominantly hydrophobic character, and accommodates the PD-1 residue Ile 134 [23]. It is well-established that the free energy of binding is not evenly distributed across the interface; instead, a small subset of interface residues (enriched in tryptophan, tyrosine, and arginine) contributes substantially to the free energy of binding [21].
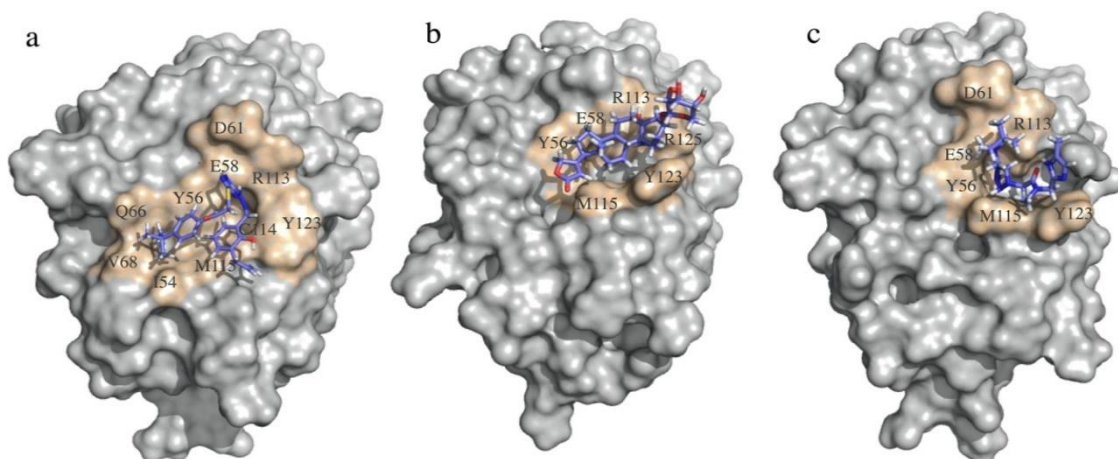

Figure 3: Conformations adopted by the small-molecules ZINC21075815 (a), ZINC70692276 (b), and ZINC64031730 (c) on PD-L1. The small molecules are shown in stick representation. The PD-L1 residues that contribute to binding are depicted in the orange surface. The conformations represent the last snapshots of the corresponding MD trajectories.

Moreover, it was shown that those energetically critical residues are not randomly distributed but rather form clusters termed hot regions [44]. A recent study that addresses hot regions at the PD-1/PD-L1 interfaces defined two hot regions, which comprise the residues Tyr

56, Glu 58, and Asn 66 and the residue Met 115, respectively. The first region (Tyr 56, Glu 58, and Asn 66) was important in antibody and small-molecule binding. In contrast, the second region (Met 115) was found to be necessary only in small-molecule binding [45]. Considering the site preference of the top 20 molecules in molecular docking calculations and the structural and energetic analysis of the MD trajectories of the PD-L1/Small-molecule complexes together with previously reported findings suggests that the PD-L1 residues Tyr 56, Glu 58, Arg 113, Met 115 and Tyr 123 comprise the hot region on PD-L1 at the PD-1/PD-L1 interface and this hot region is the most likely region to accommodate a small-molecule inhibitor of protein-protein interactions.

3.4. Comparisons with Known PD-L1 Inhibitors

Most well-known small-molecule PD-L1 inhibitors like BMS-202, A1D9R, and their analogs work by inducing PD-L1 dimerization, which sterically blocks PD-1 binding [46]. However, our aim in this study was to discover non-dimerizing small molecules that inhibit PD-1 binding. There are limited non-dimerizing PD-L1 inhibitors with well-defined structural and mechanistic evidence, and all are either antibodies or larger peptide/macrocyclic compounds that fall outside the typical small-molecule criteria. In fact, an analysis of all 67 human PD-L1 structures deposited in the Protein Data Bank further supported this, revealing the lack of well-characterized non-dimerizing small molecules.

Since no non-dimerizing small molecule with strong structural or mechanistic evidence for PD-L1 binding was available, we performed simulations of several reported compounds to evaluate their stability and binding behavior. Simulations of BMS-202 and A1D9R with monomeric PD-L1 showed unstable binding. None of the ligands kept their initial conformations, whether starting from docking poses or crystal structures. The average RMSD values exceeded 12 Å and 8 Å for BMS-202 and A1D9R, respectively, indicating substantial deviation from crystal structures. This is not surprising as these compounds bind to the dimeric form of PD-L1 in their crystal structures and interact with residues from both monomers.

CA-170 is an orally available small-molecule antagonist targeting PD-L1/PD-L2 and VISTA, proposed to disrupt PD-1/PD-L1 interaction without inducing PD-L1 dimerization. Although it has advanced to Phase I clinical trials, its direct binding to PD-L1 remains debated [41, 47]. CA-170 detached from PD-L1 within the first 3 ns of simulation, suggesting a lack of stable interaction. This observation aligns with the study reporting no direct interaction with PD-L1, supporting the view that CA-170 may act through an indirect or alternative mechanism [46].

Since well-known small molecules do not form stable interactions with monomeric PD-L1, we used the cyclic Peptide 104 as a positive control. This peptide maintained stable binding to monomeric PD-L1 throughout the simulation, with an average RMSD below 2.0 Å. The binding free energy computed using our protocol was −12.6 kcal/mol, consistent with the reported moderate binding affinity (120 nM) of the peptide [42], considering that MM/GBSA calculations often overestimate binding affinity [48]. Peptide 104 keeps its key interactions with residues Ile 54, Tyr 56, Met 115, and Tyr 123 in the simulation, which aligns with the binding regions observed for our most promising compounds, emphasizing the importance of this surface region for non-dimerizing PD-L1 binding.

4. Conclusion

In this study, using triple and quadruple combinations of the interface residues on PD-1, we constructed several pharmacophore models to scan the ZINC15 database. More than 12000 small molecules were retrieved from the ZINC15 database. The retrieved small molecules were virtually screened using molecular docking calculations. We further evaluated the complexes of promising small molecules energetically and structurally. Overall, our results suggest that the three small molecules ZINC21075815, ZINC70692276, and ZINC64031730 establish stable and energetically favorable interactions with PD-L1 at the hot region consisting of the residues Tyr 56, Glu 58, Arg 113, Met 115, and Tyr 123. This surface region undergoes local structural rearrangements and provides sufficient space and interaction surface to accommodate relatively large ligands such as ours. The positive control, Peptide 104, forms also key interactions with this surface region, underscoring the relevance of this region for non-dimerizing binding.

Our study highlights the relevance of incorporating molecular dynamics in drug discovery, as it provides a deeper understanding of the interactions and behavior of molecules in a dynamic environment. The molecules ZINC21075815, ZINC70692276, and ZINC64031730 have moderate estimated binding affinities (-8.6, -7.1, and -7.9 kcal/mol, respectively) for PD-L1 and can be used as starting points to develop more effective and selective non-dimerizing anti-PD-1/PD-L1 inhibitors.

References

[1] Wang, X., Teng, F.F., Kong, L., Yu, J.M., *PD-L1 expression in human cancers and its association with clinical outcomes*, Oncotargets and Therapy, 9, 5023–5039, 2016.

[2] Li, Y.C., Zhou, Q., Song, Q.K., Wang, R.B., Lyu, S.Z., Guan, X.D., et al., *Overexpression of an Immune Checkpoint (CD155) in Breast Cancer Associated with Prognostic Significance and Exhausted Tumor-Infiltrating Lymphocytes: A Cohort Study,* Journal of Immunology Research, 3948928, 2020.

[3] Pardoll, D.M., *The blockade of immune checkpoints in cancer immunotherapy*, Nature Reviews Cancer, 12, 252–264, 2012.

[4] Gong, J., Chehrazi-Raffle, A., Reddi, S., Salgia, R., *Development of PD-1 and PD-L1 inhibitors as a form of cancer immunotherapy: a comprehensive review of registration trials and future considerations*, Journal for Immunotherapy of Cancer, 6, 8, 2018.

[5] Darvin, P., Toor, S.M., Nair, V.S., Elkord, E., *Immune checkpoint inhibitors: recent progress and potential biomarkers*, Experimental and Molecular Medicine, 50, 1–11, 2018.

[6] Ai, L., Chen, J., Yan, H., He, Q., Luo, P., Xu, Z., et al., *Research Status and Outlook of PD-1/PD-L1 Inhibitors for Cancer Therapy*, Drug Design, Development and Therapy, 14, 3625–3649, 2020.

[7] Guzik, K., Tomala, M., Muszak, D., Konieczny, M., Hec, A., Blaszkiewicz, U., et al., *Development of the Inhibitors That Target the PD-1/PD-L1 Interaction: A Brief Look at Progress on Small Molecules, Peptides and Macrocycles*, Molecules, 24, 2071, 2019.

[8] Guo, L.B., Wei, R., Lin, Y., Kwok, H.F., *Clinical and Recent Patents Applications of PD-1/PD-L1 Targeting Immunotherapy in Cancer Treatment-Current Progress, Strategy, and Future Perspective*, Frontiers in Immunology, 11, 1508, 2020.

[9] Konstantinidou, M., Zarganes-Tzitzikas, T., Magiera-Mularz, K., Holak, T.A., Dömling, A., *Immune Checkpoint PD-1/PD-L1: Is There Life Beyond Antibodies ?,* Angewandte Chemie-International Edition, 57, 4840–4848, 2018.

[10] Chames, P., Van Regenmortel, M., Weiss, E., Baty, D., *Therapeutic antibodies: successes, limitations and hopes for the future,* British Journal of Pharmacology, 157, 220–233, 2009.

[11] Ryman, J.T., Meibohm, B., *Pharmacokinetics of Monoclonal Antibodies,* CPT Pharmacometrics and Systems Pharmacology, 6, 576–588, 2017.

[12] Hansel, T.T., Kropshofer, H., Singer, T., Mitchell, J.A., George, A.J.T., *The safety and side effects of monoclonal antibodies,* Nature Reviews Drug Discovery, 9, 325–338, 2010.

[13] Topalian, S.L., Hodi, F.S., Brahmer, J.R., Gettinger, S.N., Smith, D.C., McDermott, D.F., et al., *Safety, Activity, and Immune Correlates of Anti-PD-1 Antibody in Cancer,* New England Journal of Medicine, 366, 2443–2454, 2012.

[14] McDermott, J., Jimeno, A., *Pembrolizumab: PD-1 Inhibition as a Therapeutic Strategy in Cancer*, Drugs of Today, 51, 7–20, 2015.

[15] Chen, D.S., Irving, B.A., Hodi, F.S., *Molecular Pathways: Next-Generation Immunotherapy-Inhibiting Programmed Death-Ligand 1 and Programmed Death-1*, Clinical Cancer Research, 18, 6580–6587, 2012.

[16] Brahmer, J.R., Tykodi, S.S., Chow, L.Q.M., Hwu, W.J., Topalian, S.L., Hwu, P., et al., *Safety and Activity of Anti-PD-L1 Antibody in Patients with Advanced Cancer*, New England Journal of Medicine, 366, 2455–2465, 2012.

[17] Yang, J., Riella, L.V., Chock, S., Liu, T., Zhao, X., Yuan, X., et al., *The Novel Costimulatory Programmed Death Ligand 1/B7.1 Pathway Is Functional in Inhibiting Alloimmune Responses In Vivo,* Journal of Immunology, 187, 1113–1119, 2011.

[18] Paterson, A.M., Brown, K.E., Keir, M.E., Vanguri, V.K., Riella, L.V., Chandraker, A., et al., *The Programmed Death-1 Ligand 1: B7-1 Pathway Restrains Diabetogenic Effector T Cells In Vivo,* Journal of Immunology, 187, 1097–1105, 2011.

[19] Dranitsaris, G., Amir, E., Dorward, K., *Biosimilars of Biological Drug Therapies Regulatory, Clinical and Commercial Considerations,* Drugs, 71, 1527–1536, 2011.

[20] Wells, J.A., McClendon, C.L., *Reaching for high-hanging fruit in drug discovery at protein-protein interfaces,* Nature, 450, 1001–1009, 2007.

[21] Bogan, A.A., Thorn, K.S., *Anatomy of hot spots in protein interfaces,* Journal of Molecular Biology, 280, 1–9, 1998.

[22] Higueruelo, A.P., Schreyer, A., Bickerton, G.R.J., Pitt, W.R., Groom, C.R., Blundell, T.L., *Atomic Interactions and Profile of Small Molecules Disrupting Protein-Protein Interfaces: the TIMBAL Database,* Chemical Biology and Drug Design, 74, 457–467, 2009.

[23] Zak, K.M., Kitel, R., Przetocka, S., Golik, P., Guzik, K., Musielak, B., et al., *Structure of the Complex of Human Programmed Death 1, PD-1, and Its Ligand PD-L1,* Structure, 23, 2341–2348, 2015.

[24] Sterling, T., Irwin, J.J., *ZINC 15-Ligand Discovery for Everyone,* Journal of Chemical Information and Modeling, 55, 2324–2337, 2015.

[25] Koes, D.R., Camacho, C.J., *PocketQuery: protein-protein interaction inhibitor starting points from protein-protein interaction structure,* Nucleic Acids Research, 40, W387–W392, 2012.

[26] Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S., et al., *AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility,* Journal of Computational Chemistry, 30, 2785–2791, 2009.

[27] Sousa da Silva, A.W., Vranken, W.F., *ACPYPE - AnteChamber PYthon Parser interfacE*, BMC Res Notes, 5, 367, 2012.

[28] Jakalian, A., Bush, B.L., Jack, D.B., Bayly, C.I., *Fast, efficient generation of high-quality atomic Charges. AM1-BCC model: I. Method,* Journal of Computational Chemistry, 21, 132–146, 2000.

[29] Wang, J., Wolf, R., Caldwell, J., Kollman, P., Case, D., *Development and testing of a general amber force field,* Journal of Computational Chemistry, 25, 1157–1174, 2004.

[30] Pronk, S., Pall, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., et al., *GROMACS 4.5: a high-throughput and highly parallel open-source molecular simulation toolkit,* Bioinformatics, 29, 845–854, 2013.

[31] Lindorff-Larsen, K., Piana, S., Palmo, K., Maragakis, P., Klepeis, J.L., Dror, R.O., et al., *Improved side-chain torsion potentials for the Amber ff99SB protein force field,* Proteins-Structure Function and Bioinformatics, 78, 1950–1958, 2010.

[32] Parrinello, M., Rahman, A., *Polymorphic Transitions in Single-Crystals - A New Molecular-Dynamics Method,* Journal of Applied Physics, 52, 7182–7190, 1981.

[33] Nose, S., *A Unified Formulation of the Constant Temperature Molecular-Dynamics Methods.* Journal of Chemical Physics, 81, 511–519, 1984.

[34] Hoover, W.G., *Canonical Dynamics - Equilibrium Phase-Space Distributions.* Physical Review A, 31, 1695–1697, 1985.

[35] Case, D.A., Betz, R.M., Cerutti, D.S., Cheatham, T.E. III, Darden, T.A., Duke, R.E., et al., *Amber 18,* University of California, San Francisco, 2018.

[36] Onufriev, A., Bashford, D., Case, D.A., *Modification of the generalized Born model suitable for macromolecules,* Journal of Physical Chemistry B, 104, 3712–3720, 2000.

[37] Onufriev, A., Bashford, D., Case, D.A., *Exploring protein native states and large-scale conformational changes with a modified generalized born model,* Proteins-Structure Function and Bioinformatics, 55, 383–394, 2004.

[38] Weiser, J., Shenkin, P.S., Still, W.C., *Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO),* Journal of Computational Chemistry, 20, 217–230, 1999.

[39] Zak, K.M., Grudnik, P., Guzik, K., Zieba, B.J., Musielak, B., Dömling, A., et al., *Structural basis for small molecule targeting of the programmed death ligand 1 (PD-L1),* Oncotarget, 7, 30323–30335, 2016.

[40] Sun, C., Cheng, Y., Dong, J., Hu, L., Zhang, Y., Shen, H., et al., *Novel PD-L1/VISTA Dual Inhibitor as Potential Immunotherapy Agents,* Journal of Medicinal Chemistry, 68, 156–173, 2025.

[41] Sasikumar, P.G., Sudarshan, N.S., Adurthi, S., Ramachandra, R.K., Samiulla, D.S., Lakshminarasimhan, A., et al., *PD-1 derived CA-170 is an oral immune checkpoint inhibitor that exhibits preclinical anti-tumor efficacy,* Communications Biology, 4, 699, 2021.

[42] Zyla, E., Musielak, B., Holak, T.A., Dubin, G., *Structural Characterization of a Macrocyclic Peptide Modulator of the PD-1/PD-L1 Immune Checkpoint Axis*, Molecules, 26, 4848, 2021.

[43] Laskowski, R.A., Swindells, M.B., *LigPlot+: multiple ligand-protein interaction diagrams for drug discovery,* Journal of Chemical Information and Modeling, 51, 2778-86, 2011.

[44] Cukuroglu, E., Gursoy, A., Keskin, O., *HotRegion: a database of predicted hot spot clusters,* Nucleic Acids Research, 40, D829–D833, 2012.

[45] Lim, H., Chun, J., Jin, X., Kim, J., Yoon, J., No, K.T., *Investigation of protein-protein interactions and hot spot region between PD-1 and PD-L1 by fragment molecular orbital method*, Scientific Reports, 9, 16727, 2019.

[46] Wu, Q., Jiang, L., Li, S.C., He, Q.J., Yang, B., Cao, J., *Small molecule inhibitors targeting the PD-1/PD-L1 signaling pathway,* Acta Pharmacologica Sinica, 42, 1–9, 2021.

[47] Musielak, B., Kocik, J., Skalniak, L., Magiera-Mularz, K., Sala, D., Czub, M., et al., *CA-170 - A Potent Small-Molecule PD-L1 Inhibitor or Not ?,* Molecules, 24, 2804, 2019.

[48] Alibay, I., Magarkar, A., Seeliger, D., Biggin, P.C., *Evaluating the use of absolute binding free energy in the fragment optimisation process,* Communications Chemistry, 5, 105, 2022.