SCIENTIFIC REPORTS

OPEN

Inhibition of the checkpoint protein PD-1 by the therapeutic antibody pembrolizumab outlined by quantum chemistry

Ana Beatriz M. L. A. Tavares, José X. Lima Neto (iD), Umberto L. Fulco & Eudenilson L. Albuquerque (iD)

Much of the recent excitement in the cancer immunotherapy approach has been generated by the recognition that immune checkpoint proteins, like the receptor PD-1, can be blocked by antibody-based drugs with profound effects. Promising clinical data have already been released pointing to the efficiency of the drug pembrolizumab to block the PD-1 pathway, triggering the T-lymphocytes to destroy the cancer cells. Thus, a deep understanding of this drug/receptor complex is essential for the improvement of new drugs targeting the protein PD-1. In this context, by employing quantum chemistry methods based on the Density Functional Theory (DFT), we investigate in silico the binding energy features of the receptor PD-1 in complex with its drug inhibitor. Our computational results give a better understanding of the binding mechanisms, being also an efficient alternative towards the development of antibody-based drugs, pointing to new treatments for cancer therapy.

The World Health Organization (WHO) labels cancer as one of the most pressing health challenges in the world. Statistical data from the National Cancer Institute (NCI) show that one in two (three) men (women) will develop this disease. Cancer is a leading cause of death, accounting for over 8.8 million casualties and 14.1 million cases diagnosed worldwide in 2015, numbers that are expected to increase in the coming years[1]. It is characterized by an uncontrolled growth of cells in the body, with the potential to invade or spread to its other parts with the formation of metastases. It is induced by mutations in the genome of a cell population, which changes the normal function of various classes of protein families, such as cytokines, cell surface receptors, signal transducers and transcription factor, making it one of the most difficult and complex disease to treat[2–4]. The more widely the cancer spreads, the harder it becomes to eradicate. In 2013, oncology was ranked on the top therapeutic class by worldwide sales, amounting to $73 billion[5]. However, though highly expensive, the treatment suffers several fails due to a defense mechanisms developed by some malignant cells, mainly those related to the activation of drug resistance processes[6].

Although these mutations provide a selective vantage to populations of cancer cells, they also increase their divergences from the normal one, which can allow the recognition by the immune system cells, such as the T lymphocytes (T cells) and B cells[7]. Nevertheless, tumors also evolved to deceive immune cells, including the ability to activate co-inhibitory signaling pathways on T cells by immune checkpoint proteins, such as the cytotoxic T-lymphocyte protein 4 (CTLA-4) and the programmed cell death protein 1 (PD-1), leading to a state of immune tolerance[8,9].

The molecular identification of cancer antigens helped the creation of new approaches for effective therapies, giving rise to a new era of treatment in which our own immune system evade the block created by malignant cells and fights against them. Among these new treatments, the immune checkpoint therapy has been clinically validated as an effective treatment for many cancer types[10]. The blockade of the ligand-receptor interaction of these immune checkpoint molecules can directly increase the function of the T-cells, whereas it has been shown as a way to release the immune system to unleash anti-tumor immune response[11].

The receptor PD-1 (*Pdcd1* gene on chromosome 2) is an immune cell-specific surface inhibitor, mainly expressed in the late effector phase on activated CD4$^+$/CD8$^+$ T cells, B cells, monocytes, natural killer T cells,

Departamento de Biofísica e Farmacologia, Universidade Federal do Rio Grande do Norte, 59072-970, Natal, RN, Brazil. Correspondence and requests for materials should be addressed to E.L.A. (email: eudenilson@gmail.com)

and antigen-presenting cells (APC), including dendritic cells[12–14]. It is a transmembrane glycoprotein of the Ig superfamily, accounting for 288 amino-acids, displaying four domains including a single V-set immunoglobulin superfamily (IgSF) domain, a stalk, a transmembrane domain and a cytoplasmic domain, which contains two tyrosine-based immunoreceptor signaling motifs: the inhibitory motif (ITIM) and the switch one (ITSM)[15,16]. Due to its relevance for immune system maintenance, a number of crystallographic structures have been published related to the human extracellular domain of PD-1, in apo form or bound to several ligands, making possible to recognize its overall structural architecture[17–19]. A comparison among them allow us to describe the PD-1 extracellular domain in a canonical IgV-type topology containing a β-sandwich arrangement formed by front and back β-strand faces of interacting domains, GFCC' and ABED β-sheets, respectively (for a review see ref.[20]).

Together with its ligands PD-L1/PD-L2, the receptor PD1 form a family of immune checkpoint proteins that act as co-inhibitory factors, which can halt or limit the development of the T cell response. Through interactions with them, the protein PD-1 affects negatively the function of the T and B cells by inhibiting their immune check, ensuring that the immune system is activated only at the appropriate time in order to minimize the possibility of chronic autoimmune inflammation, leading to a decrease of the cytokine production and antibody formation. On the other hand, tumor cells exploit this immune-checkpoint pathway as a mechanism to evade detection and inhibit the immune response, preventing the immune system from killing cancer cells. Many types of cells can express PD-L1, including tumor cells and immune cells after exposure to cytokines, while PD-L2 is expressed mainly on dendritic cells in normal tissues, making the PD-1/PD-L1 interaction more suitable to cancer therapies[21]. Thus, the search for antibodies that block the interaction between the receptor PD-1 and its ligands has been seen as a major therapeutic advancement in oncology, once it shows response often durable without causing serious toxicity effects in most people[22].

Currently, a number of cancer immunotherapy agents that target the PD-1 receptor have been developed. Among them, two antibody-based agents targeting PD-1 were approved by the US Food and Drug Administration (US-FDA) and other agencies around the world, namely the nivolumab and pembrolizumab drugs. Pembrolizumab is a humanized monoclonal antibody, belonging to the immunoglobulin IgG4 subclass with potential immune checkpoint inhibitory and antineoplastic activities. Upon administration, it blocks the interaction between the programmed cell death protein PD-1 on T cells with its ligands which results in the activation of T-cell-mediated immune responses against tumor cells[23]. It was the first anti-PD-1 agent to be approved by the FDA, being nowadays used to treat some cancer types where cells express PD-L1, including advanced melanoma, advanced non-small cell lung cancer (NSCLC), recurrent or metastatic head and neck squamous cell carcinoma, classical Hodgkin lymphoma (cHL), and counting, the list being increasing with the release of new clinical trials[24].

Although these antibodies are associated with substantial benefits, the immune checkpoint blockade can lead to inflammatory side-effects. Besides, there is a necessity for new drugs targeting PD-1[24,25]. Thus, obtain a deep understanding of the human PD-1/therapeutic antibody complex is essential for our knowledge about its inhibition mechanism and the design of improved anti-PD-1 therapeutics. In this sense, it has been shown that patients making use of pembrolizumab has presented better results and fewer side-effects than those using the nivolumab drug, turning the former a suitable candidate to provide new compounds based on its structure.

To push this field forward, we intend here to investigate the pembrolizumab recognition surface on PD-1 under a quantum mechanical-based point of view. Our main goal is the supply of enough information at the binding interface that substantially contribute to the affinity and specificity between the receptor (PD-1) and the therapeutic antibody (pembrolizumab). For this task, we take into account the crystallographic structure of the antigen-binding fragment (Fab) of pembrolizumab in complex with the extracellular domain of human PD-1 (PD-1$_{ECD}$) determined at a resolution of 2.00 Å[26], as depicted in Fig. 1, with the structural segments of the receptor PD-1 labeled with different letters. Afterwards, it is submitted to a fragmentation scheme based on the molecular fractionation with conjugate caps (MFCC) method[27,28]. Our results are the first attempt to outline the pembrolizumab/PD-1 interaction based in quantum mechanics, given an individual residue-residue interaction energy, which could be useful for the development of new pharmaceutical drugs targeting PD-1, and a key tool to avoid the onset of a number of cancerous tumors.

Results

The programmed cell death protein 1 (PD-1) is an important regulator for the immune tolerance and T cell exhaustion, being recently emerged as a key target in the treatment of several types of cancer. It is expressed after the T cell activation and binds to the ligands PD-L1 and PD-L2, suppressing immune response against autoantigens and playing an important role in the maintenance of peripheral immune tolerance[13]. However, the ligand PD-L1 is often overexpressed in different tumor including lymphoma, melanoma, non-small-cell lung cancer and other types, making the PD-1/PD-L1 signaling pathway crucial in dampening the immune surveillance of the tumor[29].

In this context, the target of the PD-1/PD-L1 interaction with monoclonal antibodies has demonstrated to be an important strategy for the control and eradication of several types of cancers. The treatment with pembrolizumab was approved by US-FDA in 2014 for advanced melanoma, and recently it has also been approved to an increasing number of cancer types, such as Hodgkin lymphoma and non-small-cell lung cancer. It is believed that the weak frequency in humans and the induction of cell activation, characteristics of the members of IgG4 subclass, are some of the main basis for the success of this therapeutic compound[23]. As we can see from Fig. 2, the pembrolizumab recognition surface on PD-1 is filled by intermolecular direct and water-mediated hydrogen bonds (HBs and wHBs, respectively), non-conventional hydrogen bonds (nHBs), salt bridges (SB) and hydrophobic contacts. Therefore, mapping the relevant interactions among the pembrolizumab/PD-1 surface complex is highly important to the rationale drug development of new compounds.
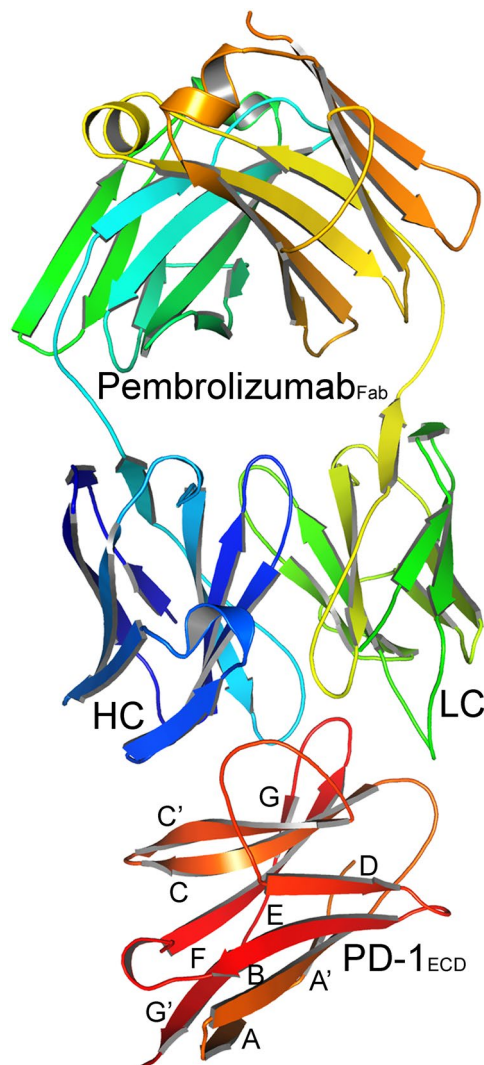
Figure 1. Structural representation of the antigen-binding fragment (Fab) of pembrolizumab (PDB ID:5GGS) in complex with the extracellular domain of human PD-1 receptor. Here, HC and LC means heavy-chain and light-chain, respectively.

In this work, we employed the MFCC/DFT scheme to assess the relative energetic contribution of the individual as well as pair-interaction contribution of each residue from the pembrolizumab/PD-1 complex. From now on we will describe all the steps in this recognition process.

Pembrolizumab/PD-1 recognition surface. Through a structural analysis of the binding site, 408 pair of interactions were detected within a range of 8.0 Å. Figure 3 shows the sum of all interaction energies from each individual residue of the pembrolizumab drug (R_i), taking into account a dielectric constant $\varepsilon = 20$ (ε_{20}) and 40 (ε_{40}) depicted in Fig. 3a and b, respectively. One can note that only two residues (one residue) in the pembrolizumab heavy-chain (HC) energy spectrum are (is) repelled in $\varepsilon = 20$ ($\varepsilon = 40$), while four (three) residues are repelled in its light-chain (LC) one, namely: $Y32_{HC}$ (ε_{20}: 0.04 kcal mol^{-1}; ε_{40}: -0.10 kcal mol^{-1}), $R98_{HC}$ (ε_{20}: 1.80 kcal mol^{-1}; ε_{40}: 0.85 kcal mol^{-1}), $G33_{LC}$ (ε_{20}: 0.008 kcal mol^{-1}; ε_{40}: -0.04 kcal mol^{-1}), $Y57_{LC}$ (ε_{20}: 6.06 kcal mol^{-1}; ε_{40}: 6.17 kcal mol^{-1}), $G61_{LC}$ (ε_{20}: 0.01 kcal mol^{-1}; ε_{40}: -0.002 kcal mol^{-1}) and $H94_{LC}$ (ε_{20}: 0.82 kcal mol^{-1}; ε_{40}: 0.64 kcal mol^{-1}). For instance, the energy value 0.04 kcal mol^{-1} calculated for Y32 (HC and LC designated residues, respectively) represents the sum of the binding energies of the four amino-acids that compose the receptor PD-1 and are within a radius distance of 8.0 Å from it, namely: $K78$, $P89$, $K131$ and $A132$. Thus, one can see that the repulsive energies are almost insignificant, excluding the residue $Y57$ that shows the highest one for both dielectric constant values ε_{20} and ε_{40}.

All other residues shown attractive interactions (negative energies), with the most intense one being observed for the heavy-chain amino-acids $Y33_{HC}$ (ε_{20}: -23.32 kcal mol^{-1}; ε_{40}: -22.84 kcal mol^{-1}), $Y101_{HC}$ (ε_{20}: -18.49 kcal mol^{-1}; ε_{40}: -18.07 kcal mol^{-1}), $R102_{HC}$ (ε_{20}: -21.58 kcal mol^{-1}; ε_{40}: -20.03 kcal mol^{-1}) and $F103_{HC}$ (ε_{20}: -13.91 kcal mol^{-1}; ε_{40}: -13.62 kcal mol^{-1}), and light-chain residues $Y36_{LC}$ (ε_{20}: -7.41 mol^{-1}; ε_{40}: -7.16 kcal mol^{-1}), $E59_{LC}$ (ε_{20}: -11.69 kcal mol^{-1}; ε_{40}: -9.69 kcal mol^{-1}) and $S95_{LC}$ (ε_{20}: -6.99 kcal mol^{-1}; ε_{40}: -6.85 kcal mol^{-1}).
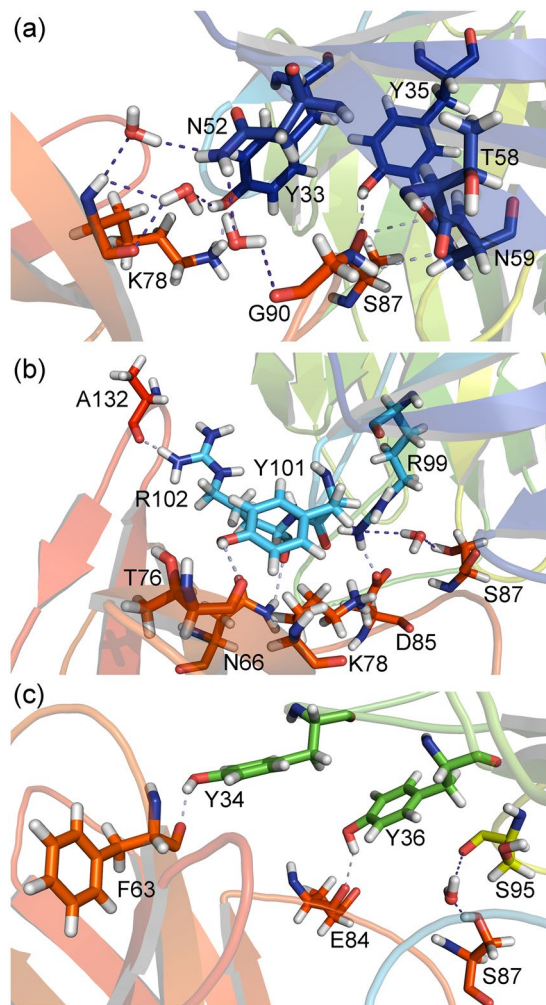
Figure 2. Interaction patterns of pembrolizumab/PD-1 recognition surface involving the amino-acid residues and the water molecules. (**a**) and (**b**) heavy-chain (HC) residues; (**c**) light-chain (LC) residues.

Pembrolizumab heavy-chain/PD-1 receptor interaction energy. Among the 408 residue-residue pairs analyzed here, 260 were pembrolizumab heavy-chain/PD-1 interactions. It is a reflection of the proximity between the HC/LC pembrolizumab with the receptor PD-1 represented in crystallographic structures, where pembrolizumab $_{Fab}$ heavy-chain fragment is closer to the PD-1 receptor than the light-chain one[26,30]. Besides a higher number of pairs, the sum of the energetic interaction between the pembrolizumab heavy-chain and the PD-1 receptor amino-acids shows also the higher value, accounting for -142.50 kcal mol^{-1} (-138.33 kcal mol^{-1}) for the dielectric constant ε_{20} (ε_{40}).

Although $Y33_{HC}$ has been shown to be the most energetic pembrolizumab amino-acid residue, it only interacts with 15 residues from the PD-1 receptor. Figure 4a depicts the highest interaction energies calculated to $Y33_{HC}$/PD-1 residues. As one can see, the strongest energy of $Y33_{HC}$ is mainly related to its binding with three residues: $K78$ (ε_{20}: -12.42 kcal mol^{-1}; ε_{40}: -11.99 kcal mol^{-1}), $Q88$ (ε_{20}: -4.45 kcal mol^{-1}; ε_{40}: -4.51 kcal mol^{-1}) and $P89$ (ε_{20}: -3.17 kcal mol^{-1}; ε_{40}: -3.15 kcal mol^{-1}). Meanwhile, $Y101_{HC}$, $R102_{HC}$ and $Y103_{HC}$ are involved with 23, 30 and 21 pairs, respectively. As shown in Fig. 4b, $Y101_{HC}$ has its most intense interaction with the residue $K78$ (ε_{20}: -9.43 kcal mol^{-1}; ε_{40}: -8.92 kcal mol^{-1}), being followed by the residue $T76$ (ε_{20}: -4.82 kcal mol^{-1}; ε_{40}: -4.71 kcal mol^{-1}). The residue $R102_{HC}$ does not show a strong individual interaction with none of the PD-1 residues (Fig. 4c), its high binding energy being due to a number of minor energies contributions, including those from the residues $A132$ (ε_{20}: -3.84 kcal mol^{-1}; ε_{40}: -3.44 kcal mol^{-1}), $I126$ (ε_{20}: -2.71 kcal mol^{-1}; ε_{40}: -2.87 kcal mol^{-1}), $K78$ (ε_{20}: -2.43 kcal mol^{-1}; ε_{40}: -3.12 kcal mol^{-1}) and $N66$ (ε_{20}: -2.53 kcal mol^{-1}; ε_{40}: -2.17 kcal mol^{-1}). Similar to $R102$, $Y103_{HC}$ binding energy is composed by the sum of a number of small interactions (Fig. 4d), with the strongest being associated with the residue $V64$ (ε_{20}: -3.49 kcal mol^{-1}; ε_{40}: -3.49 kcal mol^{-1}).

Figure 5a displays how close the residues $Y33_{HC}$, $Y101_{HC}$, $R102_{HC}$ and $Y103_{HC}$ are to the PD-1 receptor. Analyzing it one can understand the reason why these amino-acids present some of the largest number of pair-interaction and binding energies. All of them are involved in a network of hydrogen bonds with the PD-1 residues from CC'FG β-strands and some of its loops, mainly those belonging to the C'D loop, which was
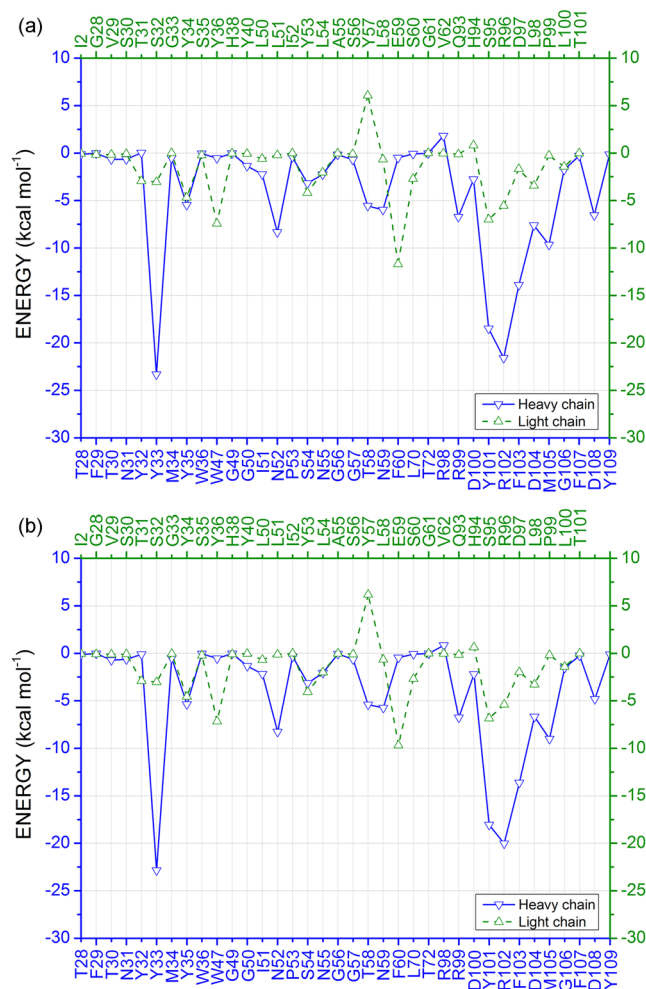
Figure 3. Energy profile for each pembrolizumab amino-acid residues in recognition surface. (**a**) This figure represents the sum of interaction energies of the pembrolizumab residues with each amino-acid from PD-1 within a radius of 8.0 Å, using the dielectric constant ε_{20}. (**b**) The same for ε_{40}. Blue solid line represents the heavy-chain energy spectrum, while the green dashed line is used to depict the light-chain one.

described to intrude into the complementary determining region (CDR) of the pembrolizumab drug, a variable portion present in some igG molecules responsible for the recognition of specific antigens[30]. Figure 5b depicts some of these interactions, starting with the residue $Y33_{HC}$. Here one can see that it forms two direct hydrogen bonds with the residues $K78$ charged amine group and $Q88$ side-chain oxygen. Besides, the residue $Y33_{HC}$ also makes a water-mediated hydrogen bond with the $K78$ carbonyl group of the main-chain, and is involved in non-polar contacts with $P89$. The two hydrogen bonds between the residues $Y33_{HC}$-$K78$ give the major contribution to their high energy.

The residues $Y101_{HC}$, $R102_{HC}$ and $Y103_{HC}$ interaction network is shown in Fig. 5c–e. Figure 5c depicts 3 directs, 2 water-mediated and 1 non-conventional hydrogen bonds. The strongest interaction of this residue is related to the pair $Y101_{HC}$-$K78$, where it creates a direct hydrogen bond with charged side-chain amine and a non-conventional hydrogen bond with amine of the main-chain. Three hydrogen bonds are formed with the residue $T76$, two with its oxygen atoms (carbonyl and hydroxyl) and a water-mediated with the hydroxyl group. It is also involved in one wHB with the residue $Y68$ side-chain hydroxyl and in some non-polar contacts with the residues $D77$ and $N66$. The residues $R102_{HC}$ and $Y103_{HC}$ are mainly involved in weak interactions. The latter form a σ-π interaction with the residue $V64$, a non-conventional hydrogen bond with the residue $P83$ and some non-polar contacts with the residues $L128$ and $D85$, while the former makes hydrogen bonds with $A132$ oxygen atom and $N66$ side-chain nitrogen, besides non-conventional hydrogen bonds with the side-chain of the residues $I126$ and $K78$.

The MFCC scheme yields not only the individual pembrolizumab/PD-1 interactions but also important information from the pair-interactions that could be otherwise missing, namely: $R99_{HC}$-$D85$ (ε_{20}: $-9.75\,\mathrm{kcal\,mol^{-1}}$; ε_{40}: $-7.43\,\mathrm{kcal\,mol^{-1}}$), $Y35_{HC}$-$S87$ (ε_{20}: $-2.13\,\mathrm{kcal\,mol^{-1}}$; ε_{40}: $-2.08\,\mathrm{kcal\,mol^{-1}}$), $M105_{HC}$-$K131$ (ε_{20}: $-6.61\,\mathrm{kcal\,mol^{-1}}$; ε_{40}: $-6.06\,\mathrm{kcal\,mol^{-1}}$), and $D108_{HC}$-$K131$ (ε_{20}: $-6.56\,\mathrm{kcal\,mol^{-1}}$; ε_{40}: $-4.82\,\mathrm{kcal\,mol^{-1}}$).
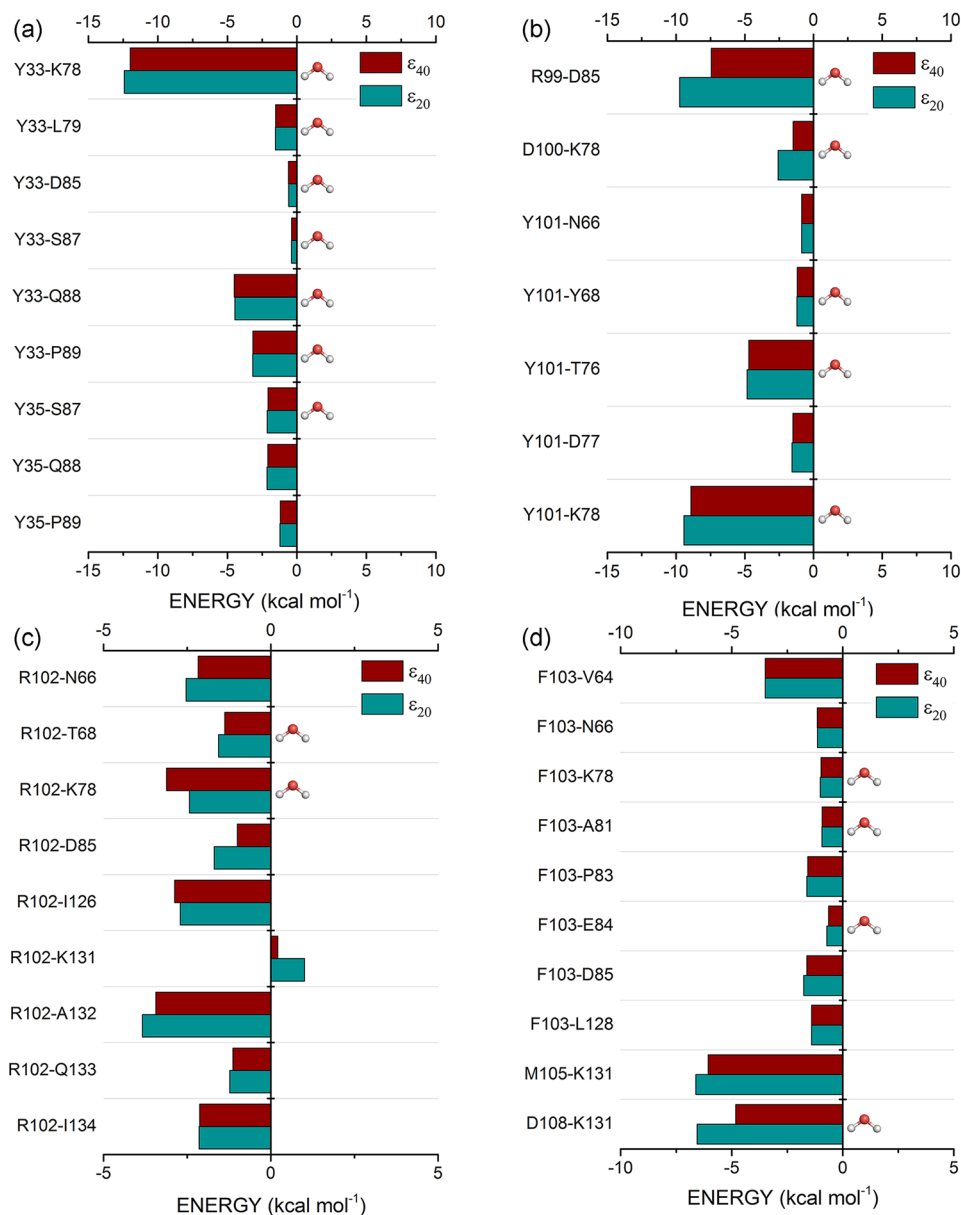
Figure 4. Graphical panel presenting the most relevant interactions involving the pembrolizumab heavy-chain residues. (**a**) $Y33$ and $Y35$; (**b**) $R99$, $D100$ and $Y101$; (**c**) $R102$ and (**d**) $F103$, $M105$ and $D108$, respectively.

The residues $M105_{HC}$ and $D108_{HC}$ make hydrogen bonds with the residue $K131$. The main-chain oxygen carbonyl of $M105_{HC}$ makes a direct hydrogen bonds with the charged amine of this lysine residue, while the residue $D108$ is engaged in a water-mediated hydrogen bond with the same molecular group, yet presenting its negatively charged carboxyl group ($D108$) in a closer distance to positively charged amine group (see Fig. 5f). The pair $Y35_{HC}$-$S87$ binding energy is governed by a hydrogen bond formed between the hydroxyl group of this tyrosine residue ($Y35_{HC}$) and the main-chain carbonyl from $S87$. The pair $R99_{HC}$-$D85$ is described for some authors as the only salt bridge formed between the pembrolizumab drug and the PD-1 receptor[26,30].

Pembrolizumab light-chain /PD-1 receptor interaction energy. The pembrolizumab light-chain is a little more distant of the PD-1 receptor than the heavy-chain one, as it can be inferred from its contribution to the total binding energy (ε_{20}: -44.67 kcal mol^{-1}; ε_{40}: -47.75 kcal mol^{-1}), as well as from the number of pairs formed with the receptor (148). This is consistent with previous crystallographic results[25,26]. According to Fig. 6, the most intense binding energies for the LC amino-acids is associated to the pair $E59_{LC}$-$K131$ (ε_{20}: -11.90 kcal mol^{-1}; ε_{40}: -9.70 kcal mol^{-1}), being followed by the pairs $S95_{LC}$-$S87$ (ε_{20}: -6.00 kcal mol^{-1}; ε_{40}: -5.83 kcal mol^{-1}), $Y53_{LC}$-$K131$ (ε_{20}: -4.15 kcal mol^{-1}; ε_{40}: -4.05 kcal mol^{-1}), $R96_{LC}$-$R86$ (ε_{20}: -1.47 kcal mol^{-1}; ε_{40}: -2.62 kcal mol^{-1}), and $Y36_{LC}$-$E84$ (ε_{20}: -2.62 kcal mol^{-1}; ε_{40}: -2.48 kcal mol^{-1}).
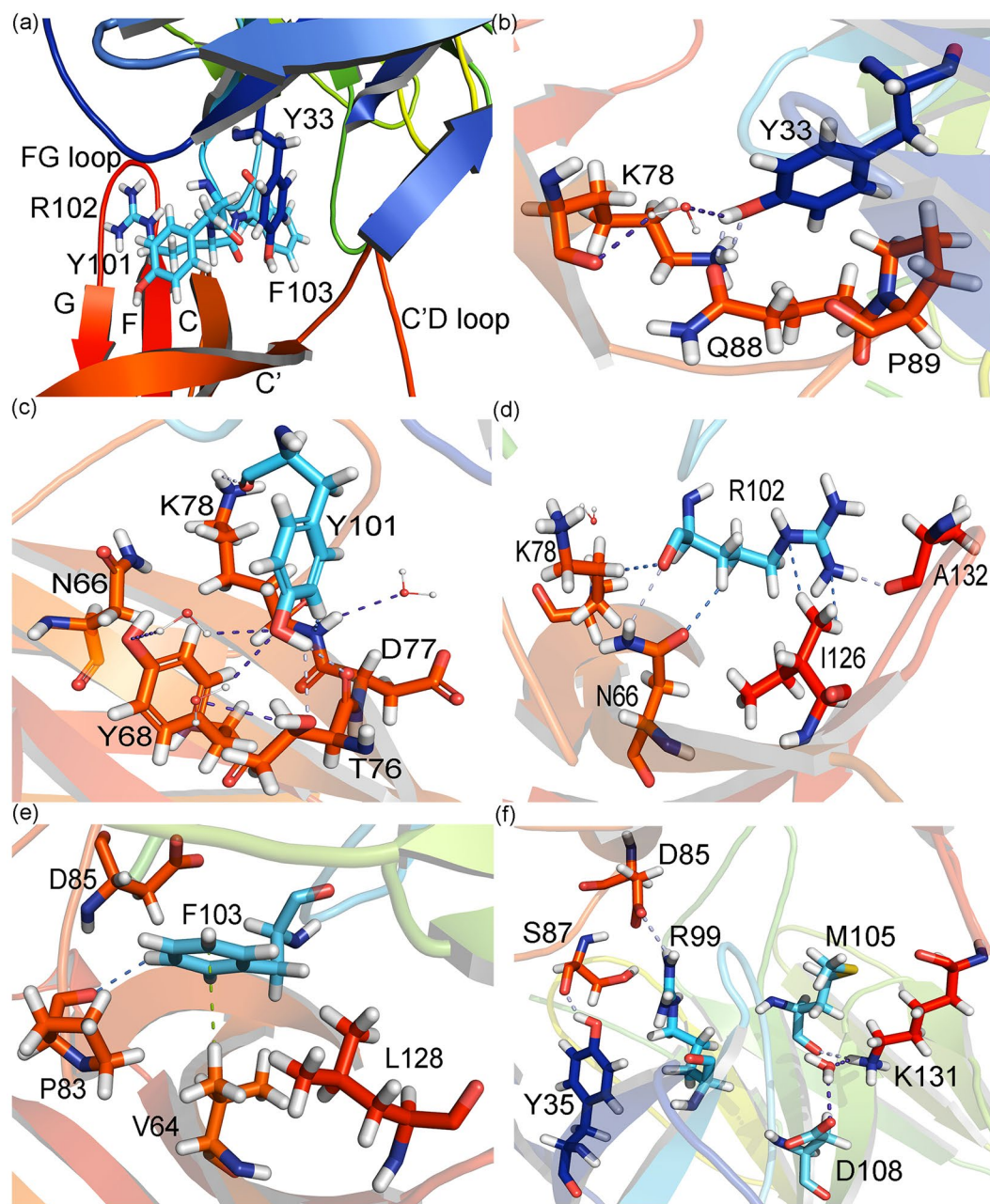
Figure 5. Intermolecular interactions of the most energetic pembrolizumab heavy-chain residues. (**a**) Structural representation of the residues $Y33_{HC}$, $Y101_{HC}$, $R102_{HC}$ and $F103_{HC}$ within the PD-1 binding site. (**b**)–(**f**) Interaction of these amino-acids with the most relevant residues of the receptor PD-1, respectively. Dashed lines in light blue (purple) (marine) represent direct (water-mediated) (non-conventional) hydrogen bonds, and the green lines represent σ-π interactions.

Figure 7 depicts some interactions made by the pembrolizumab light-chain residues linked to the PD-1 receptor. As one can see, the residue $Y36_{LC}$ is surrounded by some charged residues ($E84$, $D85$ and $R86$) from the PD-1 receptor (Fig. 7a). However, it only makes a single direct hydrogen bond with the main-chain carbonyl of the residue $E84$, all the other interactions being non-conventional hydrogen bonds ($P83$ and $R86$), and water-mediated hydrogen bonds. It would be inconsistent with the position occupied by this residue among the most energetic one of the pembrolizumab light-chain, if the number of pairs formed with $Y36_{LC}$ (10) would not overcome it. The residue making the biggest number of pair-interaction is $Y34_{LC}$, namely 19. Notwithstanding, the residue $Y34_{LC}$ shows a binding energy less than the $Y36_{LC}$ one, due to an unfavorable interaction with the residue $F63$ (ε_{20}: 5.27 kcal mol^{-1}; ε_{40}: 5.30 kcal mol^{-1}).

Although only five pair-interactions were calculated with the residue $S95_{LC}$ within a radius of 8.0 Å, it shows one of the highest interaction energies of the drug's light-chain. Similar to the residue $Y36_{LC}$, the majority of the surrounding residues make weak interactions, excluding $S87$ which forms a water-mediated hydrogen bond
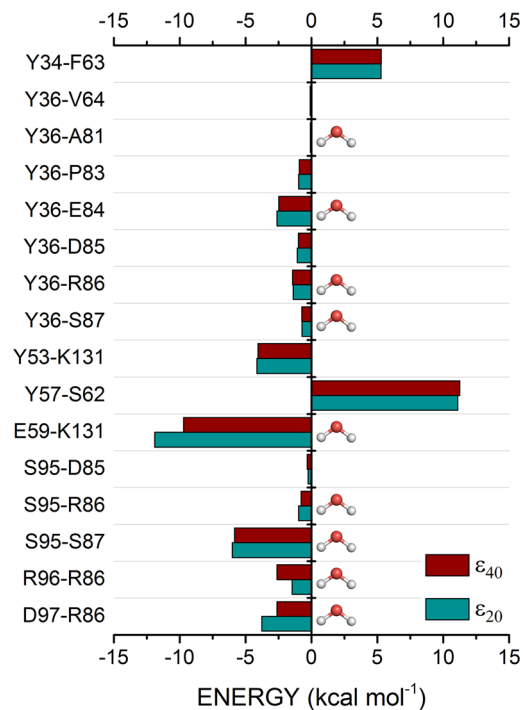
Figure 6. Graphical panel depicting the most relevant interactions involving the pembrolizumab light-chain residues Y34, Y36, Y53, Y57, E59, E95, R96 and D97.
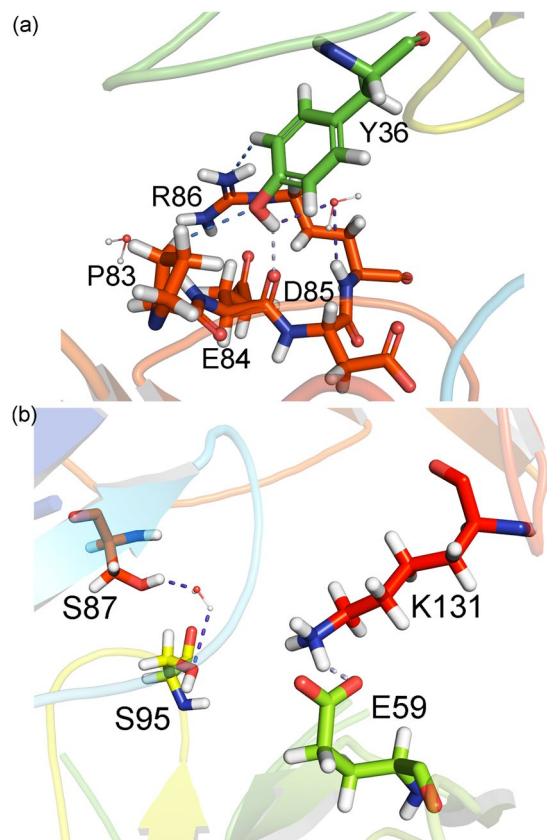

Figure 7. Intermolecular interactions of the most energetic pembrolizumab light-chain residues. (**a**) Interaction of the residue Y36 with the most relevant residues of the receptor PD-1. (**b**) The same for the residues E59 and S95. Dashed lines in light blue (purple) (marine) represent direct (water-mediated) (non-conventional) hydrogen bonds.
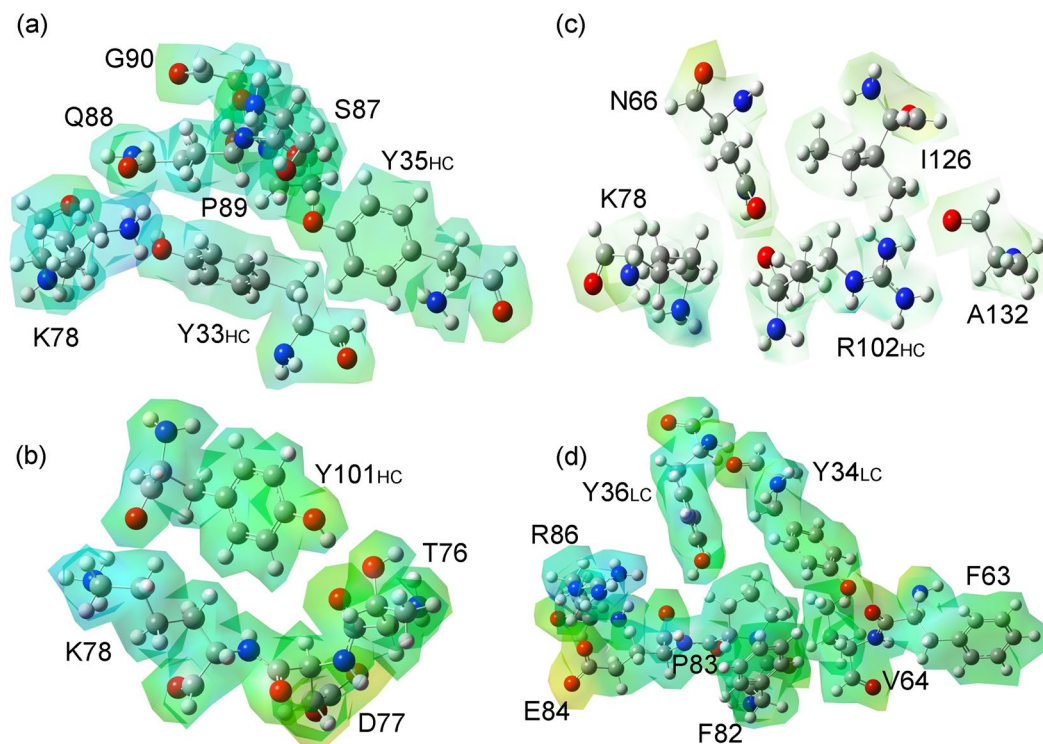
Figure 8. Electrostatic potential isosurfaces with the projected electron densities for some pembrolizumab residues interacting with the most attractive residues of the receptor PD-1.

through the hydroxyl group from both side-chains (see Fig. 7b). It is also important to notice the attractive binding energy found to the $R96_{LC}$-$R86$ pair, even with the positively charged guanidine group from both amino-acids, assuming a conformation where they are very close. These two arginine residues are in a T-shaped stacking interaction, which favors an attractive bind[30]. Finally, the $E59_{LC}$-$K131$ pair forms the second salt bridge of the complex drug/receptor, depicting the highest individual interaction energy in the drug's light-chain. This interaction is due to the negatively charged carboxyl group of the residue $E59_{LC}$, and the positively charged amine of the residue $K131$ side-chain. It gives us the idea of the dynamic process which govern the interaction between the pembrolizumab Fab fragment and the extracellular region of the PD-1 receptor, such as the formation of a new salt bridge between the residues $D108_{HC}$-$K131$ whose opposite charges from the side-chain are very close (3.6 Å).

For completeness, we display in Fig. 8 the electrostatic potential isosurface with projected electron densities for the pembrolizumab amino-acids bound to some of the most important residues at the binding pocket site.

Total binding energy. To evaluate the binding interaction energies through fragment-based quantum mechanics method, it is important to take into account every significant attractive and repulsive amino-acid residue which can influence this mechanism. Therefore, instead of taking an arbitrary region of the binding site, we performed a search for an optimal binding pocket radius (r) in which a variation less than 10% of the sequential pocket radius could be observed after a radius increase. For this task, the binding pocket radius r is varied from 2.0 Å (-32.04 kcal mol^{-1} for $\varepsilon = 20$ and -27.29 kcal mol^{-1} for $\varepsilon = 40$) to 8.0 Å (-187.17 kcal mol^{-1} for $\varepsilon = 20$ and -186.08 kcal mol^{-1} for $\varepsilon = 40$) in order to determine the best value of r, found to be 6.0 Å corresponding to an energy of -179.94 (-179.60) kcal mol^{-1} for $\varepsilon = 20$ (40), from which the convergence was achieved (see Fig. 9). This result is in agreement with previous works, stating that after 6.0-7.0 Å the molecular interactions are weak[31,32].

Discussion

Cancer immunotherapy research is trying to overcome the cancer's ability to resist the immune responses by stimulating the body's own mechanisms to remain effective against the disease. Early clinical results using blocking agents against the human cell surface receptor PD-1 and its ligants PDL-1 and PDL-2, point to unprecedented rates of long lasting anti-tumor activity in patients with metastatic cancers of different histologies[33]. From the pharmaceutical point of view, however, the clinical development of the PD-1 pathway blockers requires an understanding of the signals that induce expression of its ligands within the tumor.

However, there are some patients showing resistance to the blockage of the receptor PD-1[34]. Over stimulation of immune responses, on the other hand, may lead to damage normal, healthy tissue. Therefore, to be more efficient, the design and production of such a pharmaceutical drug require a precise knowledge not only of its biochemical structure, but also its binding mechanism with the receptor PD-1, leading to a better understanding of how this complex influences the cancer tumor and its anti-angiogenesis therapies as well as its microenvironment.
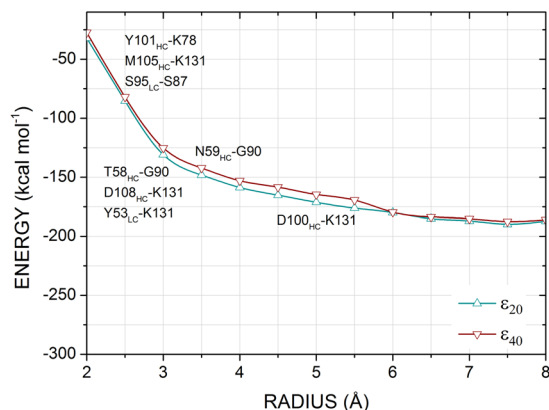
Figure 9. The total interaction energy as a function of the binding pocket radius r calculated using the GGA functional B97D. Pairs of amino-acids responsible for the regions of steepest negative and positive variation are highlighted.

Computational tools based on quantum chemistry may be a promising step in this route, further increased by combination with other anti-cancer therapies.

Within this context, we presented here an *in silico* quantum biochemistry calculation of the electronic structure of the complex pembrolizumab Fab drug/PD-1 receptor in order to map its recognition surface, searching for the binding interactions that stabilize it. The quantum MFCC approach used here is a route to investigate accurately large biological systems with low computational cost, and has been applied previously to describe molecular interactions at the quantum level related to the collagen stability[35], central nervous system disorders[36], and breast cancer[37], to cite just a few. There is no other quantum mechanical study relating the binding of the pembrolizumab drug at the PD-1 receptor, albeit its high pharmaceutical relevance.

After a convergence study on the size of the binding pocket sphere, we have considered all ligand-residue interactions within a pocket radius of 8.0 Å from the ligand. In addition, we have established the residues that play the most important roles on the binding affinity, namely: $Y33_{HC} > R102_{HC} > F103_{HC} > Y101_{HC} > E59_{LC} > M105_{HC} > N52_{HC} > Y36_{LC} > S95_{LC} > D104_{HC}$. We also highlighted the energetic and structural differences between the pembrolizumab's heavy-chain and light-chain besides unveiling the relevance to take into account the water molecules and weak interactions in the design of this therapeutic compounds. The strong attractive character observed in the amino-acid residues $Y101_{HC}$, $R102_{HC}$, $F103_{HC}$, $D104_{HC}$, $M105_{HC}$ and $D108_{HC}$ could give the base for the development of an improved drug targeting the PD-1/PD-L1 pathway by activating an immune resistance mechanism in response to endogenous anti-tumor activity, avoiding an immunosuppressive tumor microenvironment. It is also remarkable the relevance of the residues $K78$, $D85$, $P89$, $I126$ and $L128$ of the receptor PD-1. These amino-acids have been shown making pairs with almost all pembrolizumab residues, and mutational studies have displayed their relevance to the PD-1/ligand interaction[38].

Some of these residues binding energies spectra are consistent with previous experimental studies[26,39]. Our results show also that in the Pembrolizumab/PD-1 complex, the residue D85 is associated with the strongest interaction energy regime, forming pairs with the residues $R99_{HC}$ and $Y36_{LC}$ through a salt bridge and hydrophobic contacts, respectively, while the residue R86 has no interaction whatsoever with any of the residues from the ligand, in agreement with the experimental results[30]. As shown by Fessas *et al.*[40], one of the most relevant structural aspect of the Pembrolizumab/PD-1 complex is the interaction of the former with residues from the C'D loop of PD-1 (amino-acids P83-S93), emphasized in our previous section (Results), excluding the amino-acids $D100_{HC}$, $Y101_{HC}$, $Y53_{LC}$, $Y57_{LC}$ and $E59_{LC}$, commonly showing one energetically relevant pair. These achievements confirm the relevance of this loop for anchoring Pembrolizumab and, likely, its ligand-based compounds.

Despite the absence of mutation experimental results, outcomes based on the crystallographic structures of the PD-1 protein mutants bound to the receptors PD-L1 and PD-L2 could be an useful background to understand the energetic behavior depicted in this work. Recently, Lázár-Molnár performed a mutagens study taking into account the wild-type human PD-1 receptor and a mutant form of the protein A132L[41]. Mutation of the amino-acids K78A in both proteins (wild-type and mutant) was responsible for the complete loss of the PD-L1 and PD-L2 interaction bind energy. Notwithstanding, K78, a residue from the C' strand, is seen forming strong interaction pairs with a number of amino-acids from the Pembrolizumab, including the most energetic heavy-chain residues $Y33_{HC}$, $D100_{HC}$, $Y101_{HC}$, $R102_{HC}$ and $F103_{HC}$. Also, although the amino-acid I126 makes non-conventional hydrogen bonds only with the heavy-chain residue $R102_{HC}$, it is responsible for one of its strongest interaction energies. Meanwhile, the residues I134, L128 and E136, although not presenting a large number of energetically strong pairs, are related to $R102_{HC}$ and $Y103_{HC}$, which are the second and third most energetic Pembrolizumab heavy-chain residues, thus defining, along with the residue I126, a region of stabilization of the ligand. Taken together, it becomes clear the relevance of these residues for the inhibition mechanism of the Pembrolizumab.

To close this section, we would like to propose two pharmacophores that may be useful for the rational design of ligands targeting the receptor PD-1, looking for the development of antibody-based drugs. As one can note,

the Pembrolizumab/PD-1 binding surface is mainly governed by a set of hydrophobic contacts strengthened by strong hydrogen bonds. Five residues from the PD-1 receptor can be observed to form two hot-spots required for the inhibition of the PD-1/PD-ligands. The first one is in the C'D loop and C' sheet, while the second one is in the FG loop and F sheet, respectively. The former is composed by the residues K78, D85 and P89, with their most distinctive feature being the steric complementarity between the residue K78 and the hydrophobic/hydrophilic heavy-chain amino-acids $Y33_{HC}$, $Y101_{HC}$, $R102_{HC}$ and $F103_{HC}$. They are coupled to flexibilize the C'D loop, permitting the residue D85 to interact with the light-chain amino-acids $S36_{LC}$ and $S95_{LC}$. This pocket is ruled by hydrogen bonds and is, likely, one of the first docking points on the PD-1 surface due to its more external position, when compared to other motifs. The latter is located just nearby and has a most hydrophobic character, being composed by the residues I126 and L128. These residues are seen to accommodate the heavy-chain amino-acids $R102_{HC}$ and $F103_{HC}$ by anchoring them through hydrophobic interactions and non-conventional h-bonds. These proposed pharmacophores should provide a solid point for the design of new antibodies or small molecules targeting the receptor PD-1.

In summary, we have provided a quantum mechanics description of the Pembrolizumab/PD-1 complex at a molecular and energetic level by decomposing both proteins into amino-acid fragments and calculating the residue-residue interaction pair through an MFCC-based method within the density functional theory framework. The use of quantum mechanics in our calculations provided a good characterization of important molecular features by including an accurate electronic description and enhancing the differences of interaction intensity among the amino-acid pairs. Since Pembrolizumab has shown great improvement in the treatment of several cancer types, our results could favor the development of new compounds targeting the PD-1/PD-ligands, in special those based on the hot-spots, by avoiding structural features which are in minor relationship with this binding scenario. Notwithstanding, the quantum chemistry computational methods used in this work emerged as a simple and efficient alternative to unveil the drug's amino-acids residues that play the most important role on the binding affinity of the receptor-ligand complex. No doubt, taking into account the small cost/benefit of the operation, *in silico* approach, not only in clinical oncology but also in pharmaceutical medicine in general, is an important initial step toward the development of efficient cancer pharmaceutical drugs.

Methods

The calculations performed in this study have taken full advantage of the first X-ray crystal structure of human programed cell death protein 1 (PD-1) solved in complex with a Fab fragment of pembrolizumab (PDB ID: 5GGS) at 2.00 Å of resolution. The protonation state of the receptor PD-1 and its ligand was adjusted according to the results obtained from the PROPKA 3.1 package[42]. Afterwards, atoms not resolved by X-ray diffraction and therefore absent in the crystallographic files, as hydrogens and some amino-acid side-chains, were added to the structures and submitted to a classical geometry optimization fixing the other atoms. This optimization was performed using the classical force field CHARMm22 (Chemistry at Harvard Molecular Mechanics 22 forcefield), which is especially parametrized for organic molecules[43]. The calculations were carried out with convergence tolerances set to 10^{-5} kcal mol^{-1} (total energy variation), 0.001 kcal mol^{-1}Å^{-1} (RMS gradient) and 10^{-5} Å (maximum atomic displacement).

The interaction energy between the amino-acids from the pembrolizumab (at position R$_i$) and PD-1 (at position R$_j$) was calculated by using the formalism of MFCC fragmentation scheme proposed by Zhang and Zhang[44,45], and modified by Rodrigues and co-workers to calculate protein-protein interactions[28], within a DFT framework. For each amino acid residue of interest at position R$_i$, we draw an imaginary sphere with radius equal to 8.0 Å, and evaluate the interaction energy EI(R$_i$-R$_j$) with each residue at position R$_j$, considering at least one atom inside the sphere. Hence, the scheme to compute the residue-residue interaction energies follows the MFCC approach, in which the interaction energy between the residue R$_i$ and the other residues R$_j$ was calculated according to:

$$
\begin{aligned}
EI(R_i - R_j) = \; & E(C_i R_i C_{i_*} - C_j R_j C_{j_*}) - E(C_i R_i C_{i_*} - C_j C_{j_*}) \\
& - E(C_i C_{i_*} - C_j R_j C_{j_*}) + E(C_i C_{i_*} - C_j C_{j_*})
\end{aligned}
\tag{1}
$$

where the C$_i$ and C$_{i*}$ (C$_j$ and C$_{j*}$) caps are respectively the neighbour residues R$_{i-1}$ and R$_{i+1}$ (R$_{j-1}$ and R$_{j+1}$) attached to the reference residue R$_i$ (R$_j$). The term E(C$_i$ R$_i$ C$_{i*}$ − C$_j$ R$_j$ C$_{j*}$) corresponds to the total energy of the fragment comprises both residues as well as their capped residues. The second (third) term, E(C$_i$ R$_i$ C$_{i*}$ − C$_j$ C$_{j*}$) [E(C$_i$ C$_{i*}$ − C$_j$ R$_j$ C$_{j*}$)], gives the total energy of the system formed by the capped residue R$_i$ (R$_j$) and the hydrogenated caps of R$_j$ (R$_i$). E(C$_i$ C$_{i*}$ − C$_j$ C$_{j*}$) is the total energy of the system formed by the caps only.

In order to achieve the structural stability of the complex promoted by interactions with extended hydration network, all water molecules forming hydrogen bonds with a particular residue were included for completeness in the fragments. The analysis of the binding scenario were based in the criteria adopted by Bezerra *et al.*[35].

Energetic calculations for each residue were performed using the Gaussian G09 code[46], within the density functional theory (DFT) formalism, with the generalized gradient approximation (GGA) functional B97D (for a review see ref.[47]). It was proposed by Grimme[48] and include dispersion terms in order to improve the description of the non-covalent interactions, being recently used together with the D3 dispersion correction under vacuum and COSMO continuum solvation model to calculate the binding energy of some protein ligand complexes within a MFCC scheme[49]. We selected the 6–311 + G(d,p) basis set for the calculations of the expanded Kohn-Sham orbitals for all electrons.

The representation of molecular environment is a necessary step for the theoretical study of molecular properties. Recently, models of implicit solvation have been used in the MFCC scheme by Ourique *et al.*[50] and Dantas *et al.*[51] to calculate binding affinity in protein-ligand complexes, by assuming different values of the dielectric constant (ε) to take into account the effect of the protein and solvent environment in the evaluation of the

electrostatic energies. Here, we have employed the MFCC approach together with the CPCM (Conductor-like Polarizable Continuum Model) continuum solvation model[52] with dielectric constants ε considered to be 20 (ε_{20} and 40 (ε_{40}), respectively) to increase the similarity with the protein environment and estimate energy effects, such as the electrostatic polarization promoted by the solvent.

Data availability.
All other relevant data are available from the corresponding authors upon reasonable request.

References

1. Torre, L. A. *et al*. Global cancer statistics, 2012. *CA Cancer J. Clin.* **65**, 87–108 (2015).
2. Begley, C. G. & Ellis, L. M. Drug development: Raise standards for preclinical cancer research. *Nature* **483**, 531–533 (2012).
3. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011).
4. Giancotti, F. G. & Ruoslahti, E. Integrin signaling. *Science* **285**, 1028–1032 (1999).
5. Bach, P. B. Indication-specific pricing for cancer drugs. *J. Am. Med. Assoc.* **312**, 1629–1630 (2014).
6. Chen, D. S. & Mellman, I. Elements of cancer immunity and the cancer-immune set point. *Nature* **541**, 321–330 (2017).
7. Sharma, P. & Allison, J. P. The future of immune checkpoint therapy. *Science* **348**, 56–61 (2015).
8. Byun, D. J., Wolchok, J. D., Rosemberg, L. M. & Girotra, M. Cancer immunotherapy - immune checkpoint blockade and associated endocrinopathies. *Nat. Rev. Endocrinol.* **13**, 195–207 (2017).
9. Viricel, C., Ahmed, M. & Barakat, K. Human pd-1 binds differently to its human ligands: A comprehensive modeling study. *J. Mol. Graph. Model.* **57**, 131–142 (2015).
10. Melero, I. *et al*. Evolving synergistic combinations of targeted immunotherapies to combat cancer. *Nat. Rev. Cancer* **15**, 457–472 (2015).
11. Zou, W., Wolchok, J. D. & Chen, L. PD-L1 (B7-H1) and PD-1 pathway blockade for cancer therapy. *Transl. Med.* **8**, 328rv4 (2016).
12. Ishida, Y., Agata, Y., Shibahara, K. & Honjo, T. Induced expression of PD-1, a novel member of the immunoglobulin gene superfamily, upon programmed cell death. *EMBO J.* **11**, 3887–3895 (1992).
13. Nishimura, H. *et al*. Development of lupus-like autoimmune diseases by disruption of the PD-1 gene encoding an ITIM motif-carrying immunoreceptor. *Immunity* **11**, 141–151 (1999).
14. Nishimura, H. *et al*. Autoimmune dilated cardiomyopathy in PD-1 receptor-deficient mice. *Science* **291**, 319–322 (2001).
15. Iwai, Y., Hamanishi, J., Chamoto, K. & Honjo, T. Cancer immunotherapies targeting the PD-1 signaling pathway. *J. Biomed. Sci.* **24**, 26 (2017).
16. Yamane, H. *et al*. Programmed cell death protein 1 and programmed death-ligand 1 are expressed on the surface of some small-cell lung cancer lines. *Am. J. Cancer Res.* **5**, 1553–1557 (2015).
17. Zhang, X. *et al*. Structural and functional analysis of the costimulatory receptor programmed death-1. *Immunity* **20**, 337–347 (2004).
18. Lázár-Molnár, E. *et al*. Crystal structure of the complex between programmed death-1 (PD-1) and its ligand PD-L2. *Proc. Natl. Acad. Sci. USA* **105**, 10483–10488 (2008).
19. Cheng, X. *et al*. Structure and interactions of the human programmed cell death 1 receptor. *J. Biol. Chem.* **288**, 11771–11785 (2013).
20. Zak, K. M. *et al*. Structural biology of the immune checkpoint receptor PD-1 and its ligands PD-L1/PD-L2. *Structure* **25**, 1163-11-747 (2017).
21. Keir, M. E., Butte, M. J., Freeman, G. J. & Sharpe, A. H. PD-1 and its ligands in tolerance and immunity. *Annu. Rev. Immunol.* **26**, 677–704 (2008).
22. Khoja, L. *et al*. Pembrolizumab. *J. Immunother. Cancer* **3**, 36 (2015).
23. Scapin, G. *et al*. Structure of full-length human anti-PD1 therapeutic IgG4 antibody pembrolizumab. *Nat. Struct. Mol. Biol.* **22**, 953–958 (2015).
24. Ivashko, I. N. & Kolesar, J. M. Pembrolizumab and nivolumab: PD-1 inhibitors for advanced melanoma. *Am. J. Health-Syst. Pharm.* **73**, 193–201 (2016).
25. Horita, S. *et al*. High-resolution crystal structure of the therapeutic antibody pembrolizumab bound to the human PD-1. *Sci. Rep.* **6**, 35297 (2016).
26. Lee, J. Y. *et al*. Structural basis of checkpoint blockade by monoclonal antibodies in cancer immunotherapy. *Nat. Commun.* **7**, 13354 (2016).
27. da Costa, R. F. *et al*. Explaining statin inhibition effectiveness of HMG-CoA reductase by quantum biochemistry computations. *Phys. Chem. Chem. Phys.* **14**, 1389–1398 (2012).
28. Rodrigues, C. R. F. *et al*. Quantum biochemistry study of the T3-785 tropocollagen triple-helical structure. *Chem. Phys. Letters* **559**, 88–93 (2013).
29. Okazaki, T., Chikuma, S., Iwai, Y., Fagarasan, S. & Honjo, T. A rheostat for immune responses: the unique properties of PD-1 and their advantages for clinical application. *Nat. Immunol.* **14**, 1212–1218 (2013).
30. Na, Z. *et al*. Structural basis for blocking PD-1-mediated immune suppression by therapeutic antibody pembrolizumab. *Cell Res.* **27**, 147–150 (2017).
31. Ryde, U. & Söderhjelm, P. Ligand-binding affinity estimates supported by quantum-mechanical methods. *Chem. Rev.* **116**, 5520–5566 (2016).
32. Stierand, K. & Rarey, M. Drawing the PDB: protein-ligand complexes in two dimensions. *ACS Med. Chem. Lett.* **1**, 540–545 (2010).
33. Shin, D. S. *et al*. Primary resistance to PD-1 blockade mediated by JAK1/2 mutations. *Cancer Discov.* **7**, 188–201 (2017).
34. Zaretsky, J. M. *et al*. Mutations associated with acquired resistance to PD-1 blockade in melanoma. *N. Engl. J. Med.* **375**, 819–829 (2016).
35. Bezerra, K. S. *et al*. Quantum binding energy features of the T3-785 collagen-like triple-helical peptide. *RSC Adv.* **7**, 2817–2828 (2017).
36. Lima Neto, J. X. *et al*. A quantum biochemistry investigation of willardiine partial agonism in AMPA receptors. *Phys. Chem. Chem. Phys.* **17**, 13092–13103 (2015).
37. Mota, K. B. *et al*. A quantum biochemistry model of the interaction between the estrogen receptor and the two antagonists used in breast cancer treatment. *Comp. Theor. Chem.* **1089**, 21–27 (2016).
38. Pabon, N. A. & Camacho, C. J. Probing protein flexibility reveals a mechanism for selective promiscuity. *Elife* **6**, e22889 (2017).
39. Pascolutti, R. *et al*. Structure and dynamics of PD-L1 and an ultra-high-affinity PD-1 receptor mutant. *Structure* **24**, 1719–1728 (2016).
40. Fessas, P. *et al*. A molecular and preclinical comparison of the PD-1-targeted T-cell checkpoint inhibitors nivolumab and pembrolizumab. *Semin. Oncol.* **44**, 136–140 (2017).
41. Lázár-Molnár, E. *et al*. Structure-guided development of a high-affinity human programmed cell death-1: Implications for tumor immunotherapy. *EBioMedicine* **17**, 30–44 (2017).
42. Sondergaard, C. R., Olsson, M. H. M., Rostkowski, M. & Jensen, J. H. Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pK$_a$ values. *J. Chem. Theory Comput.* **7**, 2284–2295 (2011).
43. MacKerell, A. D. *et al*. All-atom empirical potential for molecular modeling and ynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586–3616 (1998).

44. Zhang, D. W. & Zhang, J. Z. H. Molecular fractionation with conjugate caps for full quantum mechanical calculation of protein-molecule interaction energy. *J. Chem. Phys.* **119**, 3599–3605 (2003).
45. Zhang, D. Quantum mechanical calculation of nanomaterial-ligand interaction energies by molecular fractionation with conjugated caps method. *Sci. Rep.* **7**, 44645 (2017).
46. Frisch, M. J. *et al. Gaussian 09*. (Gaussian Inc., Wallingford CT, 2009).
47. Albuquerque, E. L. *et al.* DNA-based nanobiostructured devices: The role of quasiperiodicity and correlation effects. *Phys. Rep.* **535**, 139–209 (2014).
48. Grimme, S. Semiempirical GGA-type density functional constructed with a long-range dispersion correction. *J. Comput. Chem.* **27**, 1787–1799 (2006).
49. Antony, J. & Grimme, S. Fully ab initio protein-ligand interaction energies with dispersion corrected density functional theory. *J. Comput. Chem.* **33**, 1730–1739 (2012).
50. Ourique, G. S. *et al.* A quantum chemistry investigation of a potential inhibitory drug against the dengue virus. *RSC Adv.* **6**, 56562–56570 (2016).
51. Dantas, D. S. *et al.* Quantum molecular modelling of ibuprofen bound to human serum albumin. *RSC Adv.* **5**, 49439–49450 (2015).
52. Cossi, M., Rega, N., Scalmani, G. & Barone, V. Energies, structures, and electronic properties of molecules in solution with the C-PCM solvation model. *J. Comput. Chem.* **24**, 669–681 (2003).

Acknowledgements

Author Contributions

A.B.M.L.A.T. and E.L.A. conceived the idea of the paper. U.L.F., J.X.L.N. and A.B.M.L.A.T. contributed to the acquisition and analysis of the available data. J.X.L.N. and U.L.F. carried out the computer simulations. E.L.A. and J.X.L.N. discussed the numerical results. E.L.A., A.B.M.L.A.T. and J.X.L.N. prepared the draft. E.L.A. and U.L.F. coordinated the project, and handled the financial support. All authors critically revised the manuscript and gave final approval of the version to be published.

Additional Information

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.