Proceedings of the 2nd International Conference on Biological Engineering and Medical Science (ICBioMed 2022), Part I
DOI: 10.54254/2753-8818/3/20220152

AlphaFold2-based structure prediction and target study of PD-L1 protein

Yang Zhuoya[1,2]

[1]University of Chinese Academy of Sciences, 19, Yuquan Road, Shijingshan District, Beijing, China (A), 100043

[2]yangzy0221@gmail.com

Abstract.PD-L1 is an immune protein in human body that can play an important role in cancer immunotherapy. By binding to antibodies, the binding activity of PD-L1 and PD-1 is blocked, which in turn inhibits cancer cells. Thus the structure of PD-L1 is very important in studying the binding of antibodies to it. However, experimental methods to solve the structures of PD-L1 and numerous complexes are expensive and consuming. Thus, it is essential to exploit computational methods to help biologists figure out the structures and the underlying mechanisms. In this paper, we explore whether AlphaFold2 is able to accurately predict the structure of PD-L1 and whether we can use AlphaFold2 to capture the binding sites of PD-L1 when binding to different antibodies. Our results show that AlphaFold2 has high confident scores and accuracy in predicting the structure of PD-L1 and the binding sites with atezolizumab and durvalumab. For the interaction between PD-L1 and the antibodies, AlphaFold2 can capture most of the hydrogen bonds as well as the salt bridges. Our work suggests that AlphaFold2 can not only be used as a tool to predict the structure of proteins, but also serves as a useful tool for antibody discovery, e.g. providing high-quality predicted structures for downstreaming docking, which brings new hope for drug discovery.

Keywords: AlphaFold2, PD-L1, structure, antibody.

1. Introduction

Programmed cell death 1 ligand 1 (PD-L1) is one of the ligands of the programmed cell death protein-1 (PD-1) [1]. The Binding of PD-L1 and PD-1 induces expression of immune checkpoint proteins [2]. In many cases, PD-L1 expression enables tumor cells to evade immune surveillance [1,3]. PD-L1-blocking antibodies like tsr-042, atezolizumab, durvalumab and avelumab have been developed pharmacologically and clinically [4]. Recently, antibody-based PD-L1 blockade therapies have become the significant part of cancer immunotherapies with multiple clinical successes [1]. While a large number of clinical trials of PD-L1-blocking antibodies rapidly developing, the structural basis of their mechanisms has also continuously been studied. Understanding the structure of the binding of PD-L1 with PD-1 and its antibodies will help us understand the interactions between PD-L1 and PD-1/PD-L2, and therefore develop more effective PD-L1 targeting antibodies in the future.

The detailed complex structure information of several marketed PD-L1-blocking antibodies have already known. More specifically, PD-L1 contains two extracellular Ig domains: the N-terminal IgV domain and C-terminal immunoglobulin constant (IgC) domain [5]. Although PD-L1 blocking

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/3/20220152

antibodies bind to PD-L1 through different binding orientations and have different binding epitopes between them, they all generally interact with PD-L1 through five hotspot residues (Y56, E58, R113, M115, and Y123) on the central CC' FG β sheet of PD-L1, which also play pivotal roles in the combination of PD-1 and PD-L1[4]. Therefore, continuous researches and explorations on the structure of PD-L1 can help us find more effective blockers, and in addition to the existing biological structure analysis methods (X-ray, NMR, etc.), new computational methods are constantly proposed.

Among these computation methods, protein structure prediction has long been an important problem in structural biology. At the end of 2020, DeepMind unveiled AlphaFold2, a program that can predict three-dimentional (3D) structure of a protein from single sequence [6]. AlphaFold2 was credited with changing the field of structural biology, and in Critical Assessment of protein Structure Prediction (CASP), AlphaFold2 was significantly better than other protein prediction methods [7]. Furthermore, AlphaFold2 has be used to predict the structure of the UniProt [8] human reference proteome, with a maximum length of 2,700 residues [8,9]. The final data covered 98.5% of human proteins, of which 35.7% of residues could be predicted with high accuracy [10]. In collaboration with EMBL-EBI, DeepMind established AlphaFold2 Protein Structure Database [6], which contains about 350,000 protein structures, is freely available worldwide (https://alphafold.ebi.ac.uk/). This database can help to obtain proteins with unknown structures as well as to complement the correction of proteins, which has great implications for structural biology and drug discovery.

Although experimental data on the structure of PD-L1 and its binding compounds are now available, the PD-L1 structure predicted by AlphaFold2 is still of high reference value considering the accuracy of structure prediction of AlphaFold2. For the design of small molecules and antibodies, AlphaFold2 can help understand the structure of ligand complexes, and comparative analysis of target proteins with the AlphaFold2 model of similar proteins can be used to generate more specific antibodies. By comparing the structure of PD-L1 predicted by AlphaFold2 with experimental ones, more information about the structure can be explored, and existing experimental structures can also be verified or modified.

In this paper, we first aligned the predicted structure by AlphaFold2 with the experimental structure of PD-L1, and the structures of PD-L1 in two complexes with atezolizumab and durvalumab were also aligned with the predicted structure, respectively. We also evaluated the ability of AlphaFold2 to capture the detailed interaction information between PD-L1 and the two antibodies. The results validate the predictive accuracy, and enhance the ability for AlphaFold2 to predict the binding sites of PD-L1 with specificity. Furthermore, AlphaFold2 can capture most of the non-covalent bonds between PD-L1 and the two antibodies. These results show that AlphaFold2 can not only be used as a tool to predict the structure of proteins, but can also be used as a useful tool to find antibodies, which facilitate the process of drug discovery.

2. Methods

2.1. Experimental structures of PD-L1

PD-L1 is a ligand of PD-1, and the interaction between PD-L1 and PD-1 can induce inhibitory signals and thus reduce the activity of T cells [1]. Experimental structures of PD-L1 were taken from RCSB PDB [11]. The accession code of human PD-L1 in UniProt [8] is Q9NZQ7, and there are 41 corresponding experimental structures in RCSB PDB database. We chose 3 of the 41 structures: 6NP9 [12], 5X8L [13] and 5X8M [13]. 6NP9 is a representative structure of PD-L1 which shows the PD-L1 IgV domain V76T with fragment in homo sapiens. IgV domain can interact with several reported antibodies like atezolizumab and durvalumab [4]. 6NP9 is obtained by X-ray diffraction and the resolution is down to 1.27 Å [12]. 5X8L and 5X8M are two crystal structure of the complex of PD-L1 interacted with atezolizumab and durvalumab, respectively. They are also obtained by X-ray diffraction with the resolution down to respectively 2.66 Å and 2.9 Å [13]. These two therapeutic antibodies can directly target PD-L1 and activate T cell immunity against tumor cells.

DOI: 10.54254/2753-8818/3/20220152

2.2. Structure of PD-L1 modeling with AlphaFold2

AlphaFold2 uses primary amino acid sequence and aligned sequences of homologues from UniProt as inputs. Through embedded multiple sequence alignments (MSAs) and pairwise features, AlphaFold2 learns from unlabelled protein sequences using self-distillation and self-estimates of accuracy and generates accurate end-to-end structure prediction[9]. The final dataset of AlphaFold2 covers 98.5% of human proteins with full chain predictions[9]. AlphaFold2 also produces a per-residue confidence metric call predicted local distance difference test (pLDDT) on a scale from 0 to 100 which indicates the degree of agreement between the predicted and the experimental structures based on the local distance difference test C α (IDDT-C α). According to this, pLDDT>90 is identified as a critical value for high accuracy with 80% correct, and pLDDT>70 is identified as generally correct backbone prediction. The 3D structure of PD-L1 predicted by AlphaFold2 was retrieved from the AlphaFold Protein Structure Database[9] by the UniProt accession code. The total length of the predicted sequence was 290 amino acids. The pLDDT of 210 amino acids is above 90 and of 245 amino acids is above 70.

2.3. Related analytically methods

We used alignment methods to analyze the identity and difference between the predicted structure and the experimental structure. The alignment of the predicted structure with experimental structure and the calculation of RMSD was accomplished by Pymol[14]. First, we used Pymol to label the predicted structure with different colors in terms of different pLDDT. Then the two structures are aligned by the alignment function of Pymol and labeled with different colors. Next, we showed the local align results of aa1-18 and aa231-190, and labeled the alignment structure with different colors based on pLDDT. Finally, we demonstrated the binding regions of atezolizumab and durvalumab with PD-L1 and showed the alignment results of the experiment and predicted structure.

We also used Protein-Ligand Interaction Profiler (PLIP), a novel web service for fully automated detection and visualization of relevant non-covalent protein-ligand contacts in 3D structures[15] to analyze the non-covalent bonding forces between the PD-L1 and the atezolizumab or durvalumab.

2.4. Structure identity analysis

The information about pLDDT of PD-L1 is acquired from the AlphaFold Protein Structure Database[9]. We first calculated the average pLDDT based on the information, and calculated the average RMSD, the RMSD of the fragment with pLDDT greater than 90, the fragment with pLDDT between 70 and 90, and the RMSD of the extracellular topological domain of PD-L1. Then we chose five hotspot residues (Y56, E58, R113, M115 and Y123) on the central CC’ FG β sheet within PD-L1 because they all involved in the interaction of PD-L1 with atezolizumab and durvalumab and calculated the RMSD. The residues for whose RMSD is below the specified threshold (0.5Å) are accepted as predicted accurately. Next, we got the information of binding forces between PD-L1 and atezolizumab/durvalumab from the literature[13].

2.5. Structure difference analysis

We first calculated the average RMSD of the fragment with pLDDT below 70, the RMSD of the transmembrane and cytoplasmic topological domain of PD-L1, and the average pLDDT of the fragment without experimental structures. Then we got protein bonding information of PD-L1 with atezolizumab and durvalumab[13].

3. Results

3.1. Structure identity analysis

The pLDDT of the predictive structure was shown in Figure 1A. The average pLDDT of the model confidence is 88.24, and the median pLDDT is 95.46. There are 209 residues with pLDDT above 90, accounting for 72.07% of the total residues. This indicates that a high confidence level is achieved for most of the structural predictions of PD-L1 by AlphaFold2. And the number of the residues with pLDDT

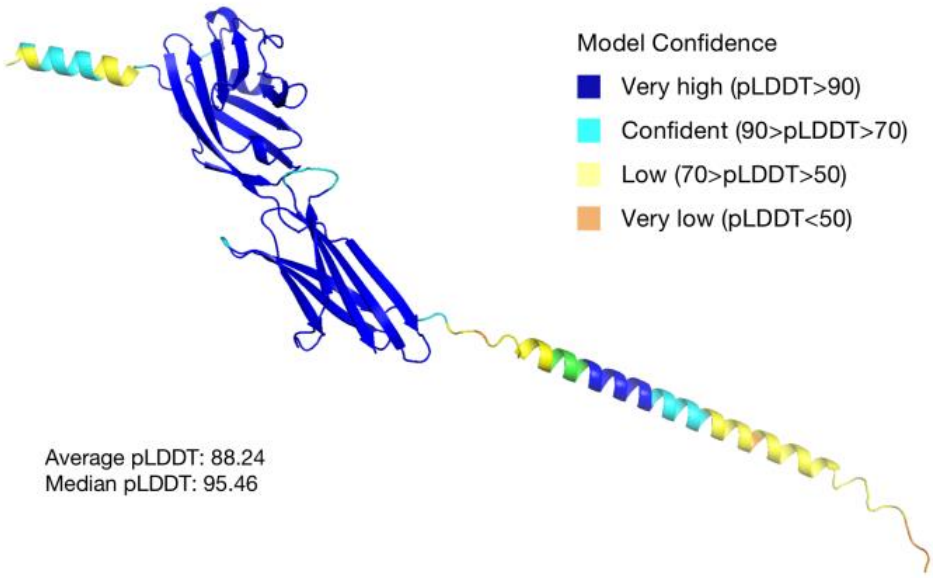
DOI: 10.54254/2753-8818/3/20220152

between 70 and 90, the pLDDT between 50 and 70, and the pLDDT below 50 is 36, 39 and 6 respectively, which account for 12.41%, 13.45%, 2.07% of the total residues. As for the extracellular topological domain (aa19-238), the transmembrane (aa239-259) and the cytoplasmic topological domain (aa260-290), the average pLDDT is 94.31, 86.77 and 56.13 respectively, which shows a comparative high confidence in predicting the extracellular topological domain and the transmembrane, and a low confidence in predicting the cytoplasmic topological domain of PD-L1 by AlphaFold2.

The alignment results are shown in Fig.1. The average pLDDT of the alignment sequence (aa22-130) is 96.32, which indicates that for the fraction with experimental structure, the prediction by AlphaFold2 has a very high confidence level. The RMSD of the experimental structure 6NP9 and the predictive structure is 0.346Å. The predictive structure is very similar to the experimental structure.

3.2. Figures

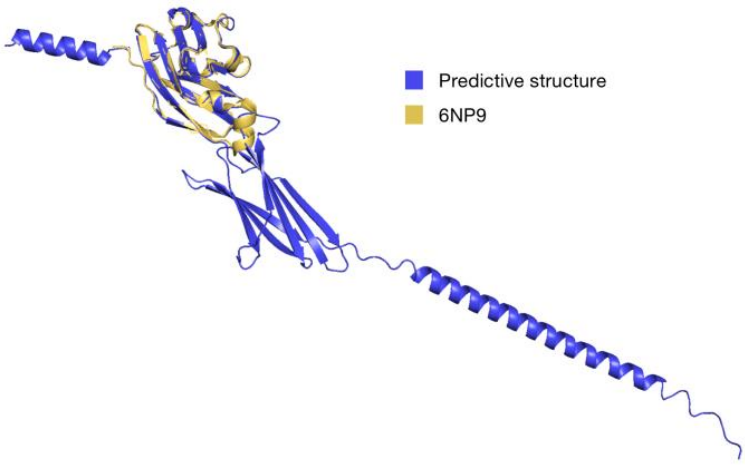
Then we chose five hotspot residues (Y56, E58, R113, M115 and Y123) on the central CC' FG β sheet within PD-L1 since they all involved in the interaction of PD-L1 with atezolizumab and durvalumab (Figure 2). The pLDDT of the five residues is 98, 97.46, 98.15, 98.12, 97.47, respectively. The average pLDDT is up to 97.84 and very close to 100. For these five important residues, the predictions of AlphaFold2 have a very high confidence level.



(a)

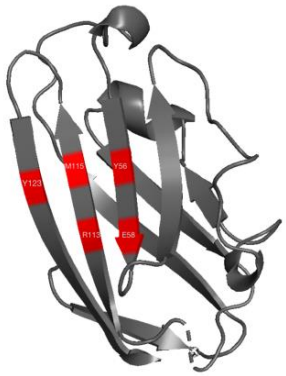
Figure 1. (a) Figure 1a shows the predictive structure and the pLDDT distribution in different colors. The average pLDDT is 88.24 and the median pLDDT is 95.46.

DOI: 10.54254/2753-8818/3/20220152



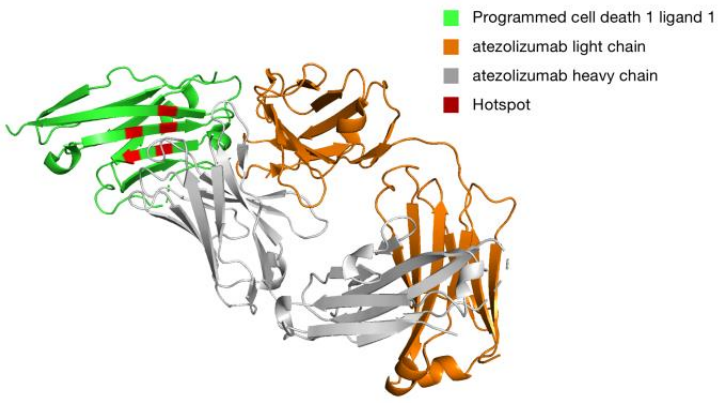
(b)

Figure 1. (b) Figure 1b shows the alignment of 6NP9 and the predictive structure. The predictive structure is shown in purple, and the experimental structure of 6NP9 is shown in yellow.



(a)

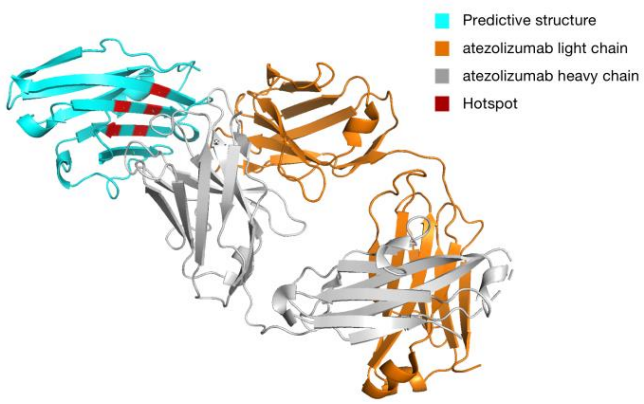
Figure 2. (a) Figure 2a shows the predictive structure of the aa22-130. And the five hotspots are shown in red color and labeled.



(b)

Figure 2. (b) Figure 2b shows the experimental structure of 5X8L. The five hotspots are shown in red color. The chain of PD-L1 is shown in green, the light chain of atezolizumab is shown in orange and the heavy chain of atezolizumab is shown in gray.

DOI: 10.54254/2753-8818/3/20220152



(c)

Figure 2. (c) Figure 2c shows the complex of predictive structure and the atezolizumab. The five hotspot is shown in red color. The chain of predictive structure is colored in blue, the atezolizumab light chain is colored in orange and the atezolizumab heavy chain is in colored gray.

The average RMSD of the predictive structure and 5X8L is 0.481Å. This is much higher than the RMSD generated by the alignment of PD-L1 monomer and the predicted structure. The result that RMSD becomes larger is reasonable considering that PD-L1 may change its structure when binding to atezolizumab, and the predicted structure is about PD-L1 monomer.

We used PLIP [15] to analyze the non-covalent bond forced in binding residues between the atezolizumab and durvalumab complex. However, the results are negative and no non-covalent bonds can be detected by PLIP in the two complexes. This may because PLIP is more suitable for detecting non-covalent bonds in small molecule complexes. For larger antibodies such as atezolizumab and durvalumab, PLIP may not be able to predict the bonding information very well [15].

So we analyzed the non-covalent bond based on the literature "Molecular mechanism of pd-1/pd-l1 blockade via anti-pd-l1 antibodies atezolizumab and durvalumab" [13]. In total, 23 residues of PD-L1 participate in the interaction with atezolizumab. The interaction of hydrogen bonds and salt-bridge interactions are shown in the table1. And the real distance in the 5X8L complex and the distance between predictive structure and atezolizumab are also shown in Table 1. The distance means the distance of alpha carbon between two amino acid residues. The average difference between the distance of the predicted complex and the distance of the experiment complex is 0.85Å. For most interaction between PD-L1 and atezolizumab, AlphaFold2 is well predicted, suggesting that the predictive structure by AlphaFold2 could helps us to better understand the binding of PD-L1 and other antibodies. For a few predicted sites, such as the hydrogen bonds between PD-L1 E45 and lightS30, the difference between the predicted distance and the actual distance is comparative large. It may because that the distance we measured is between the two alpha carbons of the amino acid residues, but hydrogen bonds do not form between the two carbons, which may cause errors to the distance measurement.

DOI: 10.54254/2753-8818/3/20220152

3.3. Tables

Table 1. The predicted distance and experimental distance between PD-L1 and atezolizumab.

PD-L1	atezolizumab light chain	atezolizumab heavy chain	interaction	complex distance (Å)	predicted distance (Å)
E58		S52	hydrogen bonds	5.2	5.5
E58		S57	hydrogen bonds	6.0	7.2
Q66		T58	hydrogen bonds	4.9	5.8
V111		Y54	hydrogen bonds	9.4	9.4
E45	S30		hydrogen bonds	6.2	9.7
D49	Y93		hydrogen bonds	9.5	9.2
R113		D31	hydrogen bonds and salt-bridge interactions	9.1	9.3
R125		D31	hydrogen bonds and salt-bridge interactions	9.0	9.4

In total, 16 residues of PD-L1 participate in the interaction with durvalumab through hydrogen bonds, salt bridges, and hydrophobic interactions. And we analyzed some of the interactions, the details and the two distances are shown in the Table 2.

Table 2. The predicted distance and experimental distance between PD-L1 and durvalumab.

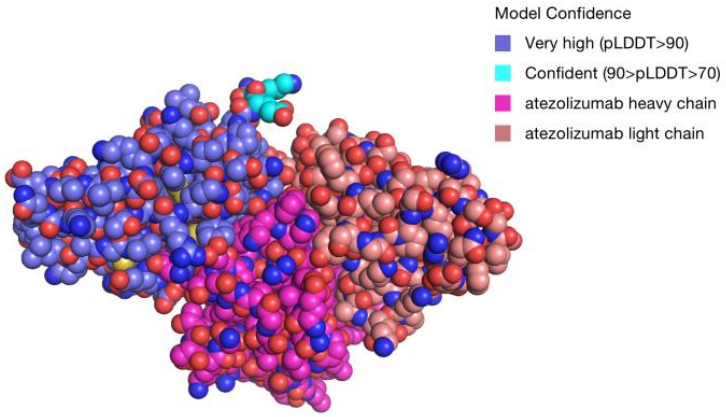
PD-L1	durvalumab light chain	durvalumab heavy chain	interaction	complex distance (Å)	predicted distance (Å)
R113		E57	salt-bridge interactions	9.0	9.1
R125		F103	hydrogen bonds	9.3	9.6
R125	Y92		hydrogen bonds	9.6	9.6

The average difference between the distance of the predicted complex and the distance of the experiment complex is 0.13Å, which indicates that for most non-covalent bonds between PD-L1 and durvalumab, AlphaFold2 can predict the interactions with a relative accuracy.

In the analysis of the two complex interactions, for the salt-bridge interactions, the average of the distance differences is 0.23Å. For the hydrogen bonds, the average of the distance differences is 0.82Å. Such results indicate that AlphaFold2 is more accurate for the prediction of salt bridges than hydrogen bonds. We also assume that another possible reason may be that the hydrogen bonds are formed between atoms that are highly electronegative and hydrogen atoms. However, the distance we measured is between the two alpha carbons of the amino acid residues so there may be more errors in the analysis of the hydrogen bonding distance, which does not account well for the poor predictive performance of AlphaFold2.

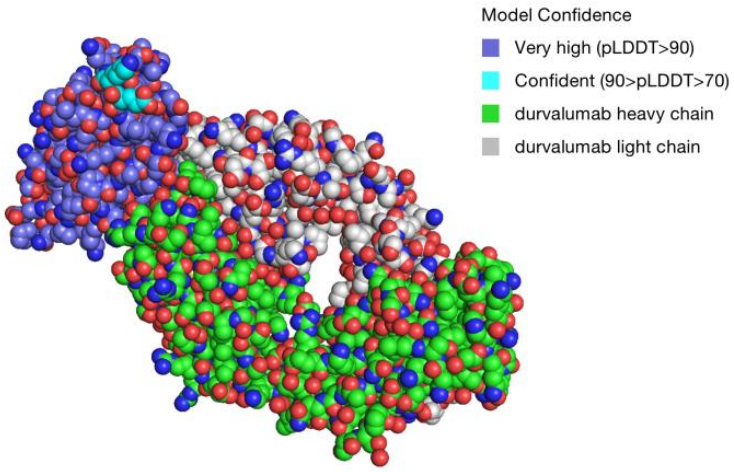
DOI: 10.54254/2753-8818/3/20220152

Figure 3 shows the binding areas of predictive structure with atezolizumab and durvalumab. The binding area of PD-L1 with atezolizumab and durvalumab both have a very high pLDDT in the AlphaFold2 prediction. This results highly verify that AlphaFold2 has a high confidence in predicting binding sites in PD-L1. For the estimation of binding sites of the protein similar to PD-L1, we can also optimistically speculate that AlphaFold2 will have a high predictive accuracy. Plus, although the binding sites of atezolizumab and durvalumab with PD-L1 are different, AlphaFold2 has high confidence in predicting the binding regions for both of them. This may indicate that AlphaFold2 has the ability to specifically predict the binding region.



(a)

Figure 3. (a) Figure 3a shows the binding of predictive structure and atezolizumab. The model confidence is shown is the figure labeled in corresponding color. The heavy chain of atezolizumab is in pink and the light chain of atezolizumab is in orange.



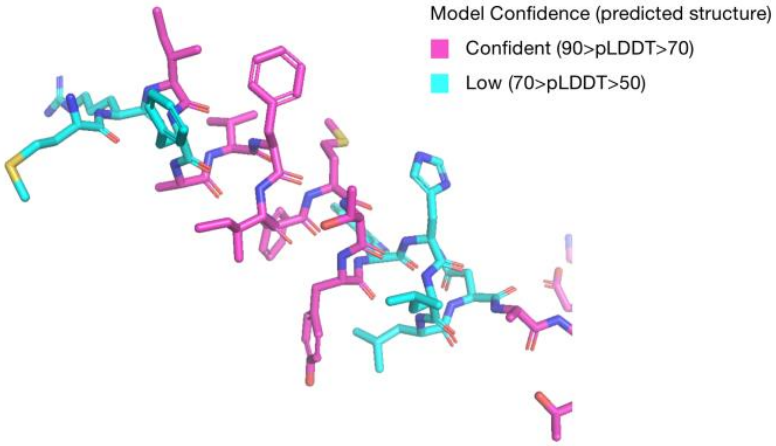
(b)

Figure 3. (b) Figure 3b shows the binding of predictive structure and durvalumab. The model confidence is shown is the figure labeled in corresponding color. The heavy chain of durvalumab is in green and the light chain of durvalumab is in gray.

DOI: 10.54254/2753-8818/3/20220152

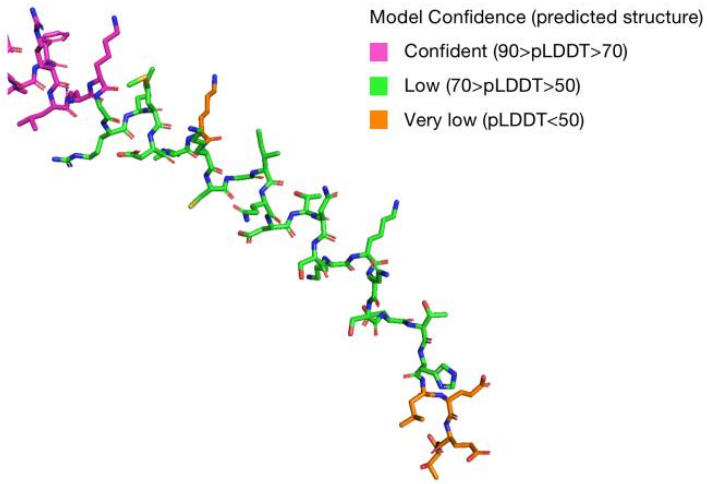
3.4. Structure difference analysis

The fragments of aa1-18 and aa260-290 have no experimental structures. The pLDDT of the two fragments are shown in Fig. 4 The average pLDDT of the fragment aa1-18 is 71.05, which shows that the predicting performance of N terminal of PD-L1 is comparative bad. It may because this fragment is relatively disorder so the AlphaFold2 can not give a high-confident prediction(9). Considering that these residues do not form any functional domain, the accuracy of its prediction has little impact on the analysis of the whole protein function. The average pLDDT of the fragment aa260-290 is 56.13 at a very low level. And this fragment forms the functional domain of cytoplasmic topological domain. The predicted structure may not be very helpful for probing the intracellular function of PD-L1. However, the cytoplasmic topological domain of PD-L1 is not its main functional domain compared to extracellular topological domain. So the low predictive confidence have limited impacts.



(a)

Figure 4. (a) Figure 4a shows the predictive structure and the corresponding pLDDT of the fragment of aa1-18 of PD-L1.



(b)

Figure 4. (a) Figure 4b shows the predictive structure and the corresponding pLDDT of the fragment of aa260-290 of PD-L1.

DOI: 10.54254/2753-8818/3/20220152

4. Conclusion

In this paper, we have compared the predicted structure of PD-L1 by AlphaFold2 with the experimental structure 6NP9, 5X8L and 5X8M. Both for monolithic structure and binding sites of PD-L1, AlphaFold2 achieves a high prediction accuracy and confidence level. This result validates the predictive power of AlphaFold2 and provides new insights for future use of AlphaFold2 as a target for protein-antibody binding sites discovery. Although the accuracy of AlphaFold2 prediction may not be very high for certain hydrogen bonds in some of the complexes, this may be influenced by errors in the measurement method. For the search of protein ligands, future work could focus on predicting and evaluating the binding sites by AlphaFold2 for more proteins, and docking the predicted structures with other possible ligands aiming to discover potential drugs. All these results indicate the AlphaFold2 can be used as a powerful and useful tool to accurately predict structures of proteins, and it also has its potential advantages for future drug discovery.

References

[1] Doroshow, D. B., Bhalla, S., Beasley, M. B., Sholl, L. M., Kerr, K. M., Gnjatic, S., ... & Hirsch, F. R. (2021). PD-L1 as a biomarker of response to immune-checkpoint inhibitors. Nature Reviews Clinical Oncology, 18(6), 345-362.

[2] Alsaab, H. O. , †, Sau, S. , Alzhrani, R. , & Tatiparti, K. . PD-1 and PD-L1 Checkpoint Signaling Inhibition for Cancer Immunotherapy: Mechanism, Combinations, and Clinical Outcome.

[3] Munn, D. H. , & Bronte, V. . (2016). Immune suppressive mechanisms in the tumor microenvironment. Current Opinion in Immunology, 39, 1-6.

[4] Lee, H. T., Lee, S. H., & Heo, Y. S. (2019). Molecular interactions of antibody drugs targeting PD-1, PD-L1, and CTLA-4 in immuno-oncology. Molecules, 24(6), 1190.R. T.

[5] Tan, S., Chen, D., Liu, K., He, M., Song, H., Shi, Y., ... & Gao, G. F. (2016). Crystal clear: visualizing the intervention mechanism of the PD-1/PD-L1 interaction by two cancer therapeutic monoclonal antibodies. Protein & cell, 7(12), 866-877.

[6] Jumper,J., Evans, R., Pritzel, A. et al. Highly accurate protein structure prediction with AlphaFold. Nature 596, 583-589(2021).

[7] Thornton, J. M., Laskowski, R. A., & Borkakoti, N. (2021). AlphaFold heralds a data-driven revolution in biology and medicine. Nature Medicine, 27(10), 1666-1669.

[8] Consortium, U. P. . (2020). Uniprot: the universal protein knowledgebase in 2021. Nucleic Acids Research(D1), D1.

[9] Tunyasuvunakool, K., Adler, J., Wu, Z., Green, T., Zielinski, M., Žídek, A., ... & Hassabis, D. (2021). Highly accurate protein structure prediction for the human proteome. Nature, 596(7873), 590-596.10 Mullard, A. (2021).

[10] What does AlphaFold mean for drug discovery?. Nature reviews. Drug discovery.

[11] Ormo, M., Cubitt, A.B., Kallio, K., Gross, L.A., Tsien, R.Y., Remington, S.J. (1996) Crystal structure of the Aequorea victoria green fluorescent protein Science 273: 1392-1395 doi: 10.1126/science.273.5280.1392

[12] Perry, E. , Mills, J. J. , Zhao, B. , Wang, F. , Sun, Q. , & Christov, P. P. , et al. (2019). Fragment-based screening of programmed death ligand 1 (pd-l1). Bioorganic & Medicinal Chemistry Letters.

[13] Lee, H. T. , Lee, J. Y. , Lim, H. , Lee, S. H. , Moon, Y. J. , & Pyo, H. J. , et al. (2017). Molecular mechanism of pd-1/pd-l1 blockade via anti-pd-l1 antibodies atezolizumab and durvalumab. Rep, 7(1), 5532.

[14] Delano, W. L. . (2002). PyMOL: An Open-Source Molecular Graphics Tool.

[15] Sebastian, S. , Sven, S. , Joachim, H. V. , Adasme, M. F. , & Michael, S. . (2015). Plip: fully automated protein–ligand interaction profiler. Nucleic Acids Research(W1), W443.