

Integrating Dynamic Network Analysis with AI for Enhanced Epitope Prediction in PD-L1:Affibody Interactions.

Diego E. B. Gomes,^{†,§} Byeongseon Yang,^{‡,¶,§} Rosario Vanella,^{‡,¶} Michael A.
Nash,^{*,‡,¶,||} and Rafael C. Bernardi^{*,†,||}

[†]*Department of Physics, Auburn University, Auburn, AL, 36849*

[‡]*Institute of Physical Chemistry, Department of Chemistry, University of Basel, Basel,
4058, Switzerland.*

[¶]*Department of Biosystems Science and Engineering, ETH Zurich, Basel, 4058,
Switzerland.*

[§]*These authors contributed equally to this work.*

^{||}*co-corresponding authors.*

E-mail: michael.nash@unibas.ch; rcbernardi@auburn.edu

Abstract

Understanding binding epitopes involved in protein-protein interactions and accurately determining their structure is a long standing goal with broad applicability in industry and biomedicine. Although various experimental methods for binding epitope determination exist, these approaches are typically low throughput and cost intensive. Computational methods have potential to accelerate epitope predictions, however, recently developed artificial intelligence (AI)-based methods frequently fail to predict epitopes of synthetic binding domains with few natural homologs. Here we have developed an integrated method employing generalized-correlation-based dynamic network

analysis on multiple molecular dynamics (MD) trajectories, initiated from AlphaFold2 Multimer structures, to unravel the structure and binding epitope of the therapeutic PD-L1:Affibody complex. Both AlphaFold2 and conventional molecular dynamics trajectory analysis were ineffective in distinguishing between two proposed binding models, parallel and perpendicular. However, our integrated approach, utilizing dynamic network analysis, demonstrated that the perpendicular mode was significantly more stable. These predictions were validated using a suite of experimental epitope mapping protocols, including cross-linking mass spectrometry and next-generation sequencing-based deep mutational scanning. Conversely, AlphaFold3 failed to predict a structure bound in the perpendicular pose, highlighting the necessity for exploratory research in the search for binding epitopes and challenging the notion that AI-generated protein structures can be accepted without scrutiny. Our research underscores the potential of employing dynamic network analysis to enhance AI-based structure predictions for more accurate identification of protein-protein interaction interfaces.

Introduction

Immunotherapies are a potentially transformative class of cancer treatment, with checkpoint inhibition emerging as a leading strategy.¹ The efficacy of these therapies hinges on achieving precise molecular interactions between a therapeutic biomolecule and the correct epitope on the checkpoint protein. For Programmed Death receptor Ligand 1 (PD-L1) targeted therapies (Fig. 1A) that have shown efficacy against a range of cancers,²⁻⁵ anti-PD-L1 antibodies must bind to the extracellular domain of PD-L1 (Fig. 1B, UniProt Q9NZQ7) in a manner that blocks the native PD-1 ligand binding, effectively disrupting cancer cells' evasion of the immune response.^{6,7} The development of novel therapeutics thus relies on achieving precise and specific binding to a particular surface epitope.⁸ This underscores the urgent need for efficient methods to rapidly and accurately identify binding epitopes of therapeutic biomolecules on the surfaces of immune checkpoint proteins.

Full-length IgG monoclonal antibodies are a mainstay of cancer immunotherapy, but many IgG formulations suffer from limitations such as limited tissue penetration, suboptimal pharmacokinetics, and complex post-translational modifications that must be optimized during production.⁹ These challenges have spurred research into alternative protein scaffolds for PD-L1 targeted therapies, including diabody, DARPIn, and affibody scaffolds.^{10–12} Such non-antibody scaffold proteins can be engineered using *in vitro* directed evolution to isolate high affinity binding domains against a range of molecular targets.^{13–15} In particular, the highly stable affibody triple α -helix bundle^{16,17} (Fig. 1C) with a molecular weight of ~ 6.5 kDa offers a promising alternative scaffold.

A variety of experimental methods have been employed for epitope mapping of antibodies and non-antibody scaffolds, including X-ray crystallography, NMR, mass spectrometry, peptide arrays, and deep mutational scanning-based approaches.^{18–26} While these methods are highly informative, they can be time-consuming and costly. Computational methods offer a faster and more cost-effective alternative. Tools like AIMS (Automated Immune Molecule Separator)²⁷ facilitate comprehensive characterization of TCR-MHC interactions and antibody polyreactivity, providing key biophysical insights into immunology. Studies on the biophysical compatibility in TCR-MHC interactions²⁸ and polyreactivity in antibodies²⁹ have demonstrated how AI-based methods can enhance our understanding of immune responses and protein interactions.

However, even advanced AI-based protein structure prediction tools such as AlphaFold and RoseTTAFold^{30–33} have faced challenges in accurately predicting protein-protein interactions.^{34–37} While there are successful cases, such as the model of a PD-L1/CD80 complex,³⁸ large screening studies have shown a low number of highly accurate antibody-antigen complex predictions.^{39,40} Similarly, traditional computational methods often fail to accurately identify binding epitopes, sometimes producing ambiguous results that suggest binding at multiple locations (Fig. 1D).

Here we present a comprehensive methodology for predicting and analyzing the structure and interactions of a PD-L1 complex,^{41,42} with a focus on epitope map-

ping. Our approach leverages recent advances in computational biophysics methodologies and the power of the latest generation of supercomputers.⁴³ Initially, we employ an integrated computational approach that utilizes AlphaFold,⁴⁴ and Zdock⁴⁵ for complex structure prediction. These predicted structures are then refined through molecular dynamics simulations performed with QwikMD⁴⁶ and GPU-accelerated NAMD.⁴⁷ Further insights are obtained through dynamic network analysis.⁴⁸ We prioritize computational methods because they are faster and more cost-effective compared to experimental approaches. While various experimental methods for binding epitope determination exist, they are typically low throughput and cost-intensive. By using advanced computational techniques first, we can generate and refine accurate models of the PD-L1 complex, providing a detailed understanding of the potential interaction interfaces efficiently and economically. These computational predictions then guide the subsequent experimental validation, making the process more targeted and effective. Here, experimental validation of our computational models is conducted using a suite of techniques including site-specific mutagenesis, biochemical assays, chemical cross-linking mass spectrometry, and next-generation sequencing (NGS)-based deep mutational scanning (DMS).⁴⁹ These experimental methods confirm the computationally predicted protein-protein interaction interface and identify specific residues involved in the interactions.

Results

We began by investigating the native PD-1:PD-L1 interaction interface, localized within the IgV-like domain of PD-L1.⁵⁰ An exhaustive analysis of bound PD-L1 structures in the Protein Data Bank (PDB) revealed insights into its interactions with other proteins and structural nuances of the complex assembly. By analyzing available PDB structures of PD-L1 bound to other proteins and focusing on amino acid residues within a 4Å radius of the binding interfaces, we found diverse sets of native contacts across PD-L1’s surface (Fig. 2A). This analysis highlights the complexity in defining the Af-

fibody binding epitope on PD-L1 (Fig. 1D), and shows PD-L1’s promiscuity and the challenges faced by traditional blind-docking methods.

A potential alternative to traditional blind-docking approaches is provided by AI-based modeling tools, such as AlphaFold2, specifically its Multimer version.^{30,31} In this study, we carried out a comparative analysis of PD-L1 binding orientations using three distinct prediction methods: AlphaFold2 Multimer, ClusPro, and ZDock, with ClusPro and ZDock utilizing existing structural data for binding space constraints and AlphaFold2 Multimer operating in a "blind-mode." Additionally, we used the recently launched AlphaFold3,⁵¹ also operating in this "blind-mode." With the exception of AlphaFold3, the top three predicted structures (see Fig.2B) from each method revealed two main binding orientations for the Affibody: a parallel and a perpendicular orientation of Affibody relative to PD-L1’s beta sheets (Fig.2C). AlphaFold3 only presented an ensemble of conformations around the parallel orientation. Although dual binding modes have been reported for certain receptor complexes,^{52,53} it is rather an exceptional case, and the parallel and perpendicular binding orientations for PD-L1 complexes were considered ostensible. It is important to note that what we describe here as a parallel or perpendicular mode is actually an ensemble of conformations that can be clustered into two main groups: one where the beta-sheets of PD-L1 are parallel to the Affibody alpha-helix, and another where the beta-sheets of PD-L1 are perpendicular to the Affibody alpha-helix, as illustrated in Fig.2C. Hence, our nomenclature in this article. The results presented in Fig.2B show that the four static computational methods could not confidently determine the binding epitope on PD-L1, prompting us to further analyze the structures to determine which orientation is more probable.

To investigate the stability of the two potential PD-L1:Affibody binding orientations, we next conducted molecular dynamics (MD) simulations using GPU-resident NAMD 3.0,⁴⁷ running 16 independent 100 *ns* replicas for each orientation. Initial visual assessments of the conformational clusters using VMD⁵⁴ indicated greater stability in the perpendicular orientation (Fig. S1). For a more rigorous evaluation, we analyzed the final 25 *ns* of each trajectory using three methods. Firstly, we measured

the root-mean-square deviation (RMSD) of the Affibody from its initial conformation, finding similar mean RMSD values for both orientations but with larger fluctuations in the parallel mode (Fig.2D). Secondly, the Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) method was used to calculate binding free energy, yielding comparable results but greater variability for the parallel orientation (Fig.2E). Lastly, leveraging a technique developed by our group we employed generalized-correlation-based Dynamic Network Analysis⁴⁸ to derive relative binding strengths.⁵⁵ The results, displayed in Fig.2F and Fig. S2, reveal the perpendicular mode as notably more stable. To assess the reliability of the results, a bootstrap analysis with a sample size of 16 was conducted, resampling the data 1,000 times with replacement. The BCa method was employed to construct 95% confidence intervals for the mean difference between the perpendicular and parallel configurations. The analysis revealed a significant difference in the sum of correlations (see Fig. S3). These combined findings suggest the perpendicular binding conformation as the most probable for the PD-L1:Affibody interaction.

Experimental Validation of Predicted Epitope

To validate the predicted binding mode, we first used VMD⁵⁴ for visual inspection and AlphaFold2 Multimer scores to identify mutations to PD-L1 that would potentially disrupt Affibody binding. Ultimately, six point mutations were selected for both computational and experimental testing: I54A, Q66D, V68Y, M115A, Y123D, and R125E. These six amino acid residues were selected based on the following criteria: two residues where contacts were shared by both orientations (V68Y, M115A), two residues where contacts were predominantly found in the perpendicular orientation (Y123D and R125E), forming relevant hydrogen bonds along the trajectory, and two residues where contacts were predominantly found in the parallel orientation (I54A, Q66D). Except for the mutations V68Y and M115A, which showed a smaller reduction of score for both the parallel and perpendicular binding modes, the other four mutations caused AlphaFold2 to generate only one type of binding mode. Therefore,

from the computational results, I54A, Q66D, Y123D, and R125E were expected to be central to the definition of the binding epitope (Fig. 3A,B).

One mutation, in particular, caught our attention due to the position of the amino acid Q66 (see Fig.3D,E). The mutation Q66 lies on the C' sheet and has rare contacts with the Affibody in the perpendicular orientation during MD simulations (Fig.3D). However, it is close enough to eventually form hydrogen bonds with the Affibody. Meanwhile, in the parallel orientation, this residue is surrounded by many other amino acids, forming contacts with at least three of them, as shown in Fig. 3E). We then investigated the impact of Q66 mutations in more detail. A computational mutational scanning analysis at position 66 showed that the original amino acid Gln (Q) favored binding, while positively charged amino acids such as Arg (R) and His (H) were the most detrimental (Fig. 3C). Interestingly, Ile (I) and Val (V), which are hydrophobic amino acids, also impaired binding according to AlphaFold2.

These PD-L1 variants were recombinantly produced and their binding affinities were evaluated using native poly(acrylamide) gel electrophoresis and single bead-based cytometry (Fig.4A, S5, S6). The results indicated moderate binding loss for I54A, Q66D, and M115A, and minimal change for R125E, suggesting the binding epitope is located on the GFCC' face of the IgV-like domain (Fig.4D). Notably, Y123D's complete loss of binding suggested the perpendicular binding mode compared to V68Y with moderate affinity loss. Comparing the AlphaFold2 scores with the binding affinity data (Fig.3A,B vs. Fig.4A, Fig. S6) revealed that while AlphaFold2 effectively captured major affinity differences, it was less effective in detecting subtle variations. The observed reduction in binding affinity caused by the Q66D mutation also supported the perpendicular mode. As discussed previously, according to our MD simulations, mutations at position Q66 could impact the perpendicular mode but are expected to be more detrimental to the parallel mode.

Site-directed mutagenesis and affinity measurements indicated that the perpendicular binding mode was more likely to be the correct binding mode. However, the goal of this study is to confirm the computational results with maximum confidence. Since the

site-directed mutagenesis results did not fully align with the AlphaFold2 predictions, we considered them insufficient to definitively distinguish between the perpendicular and parallel binding modes. Therefore, we turned to crosslinking-coupled mass spectrometry (XL-MS), which identifies the epitope by determining the proximity of cross-linkable amino acid side chains within the complex. This technique involved exposing a sample of the bound PD-L1 complex to a mass-labeled crosslinker, followed by enzymatic fragmentation and MS/MS analysis to locate cross-linking sites. This approach revealed both the epitope and the binding orientation.

The XL-MS analysis revealed ten different proximity-dependent crosslinks between the Affibody and PD-L1, distributed across six PD-L1 amino acids: H69, Y81, K105, Y112, R113, and Y123. These residues form a triangular shape on the GFCC' face of the IgV-like domain (Fig.4E; Blue). The amino acids within this triangle, including I54, Y56, E58, D61, and others, were therefore implicated as being part of the binding epitope by XL-MS, however again XL-MS alone was not capable of completely excluding either parallel or perpendicular modes.

To provide even higher resolution for experimental epitope mapping, we next turned to deep mutational scanning (DMS) with next-generation sequencing (NGS) by yeast display. This offered a high-throughput approach to epitope mapping and involved displaying a mutated PD-L1 variant library on the surface of yeast cells, and quantifying both PD-L1 variant expression level and Affibody binding strength using fluorescent activated single cell sorting (FACS)/flow cytometry coupled with high-throughput DNA sequencing. The PD-L1 variant library contained all possible single amino acid substitutions within the PD-L1 sequence constructed as a scanning one codon-one amino acid library (Fig. S7, S8). The FACS gating strategy is shown in (Fig.4B). Sorting and deep-sequencing yeasts displaying PD-L1 variants that were highly expressed but which exhibited reduced or no Affibody binding activity (Fig.4B, regions P1 and P2) allowed us to identify the binding epitope with high resolution. The Illumina sequencing read counts of sorted PD-L1 variants were visualized as a heatmap (Fig.4C) and included several surface-exposed residues along with others within the β -sandwich of

the IgV-like domain, potentially destabilizing PD-L1’s structure (Fig. S9, S10). The binding epitope identified by DMS covering all possible single mutations, showed higher precision compared to the other methods (Fig.4F).

When we compared the computationally suggested perpendicular and parallel modes with the DMS data, both on the primary protein sequence of PD-L1 and its crystal structure with PD-1⁵⁶ (Fig.4G-H), the implicated residue positions validated the perpendicular binding mode, characterized by more significant involvement of the G β -strand and less of the BC loop, CC’ loop, C β -strand, and C’ β -strand. Our findings for the experimental validation of binding epitope and orientation are summarized in Figure 4I. Furthermore, in Fig. 4, we compared the binding epitopes determined by deep mutational scanning (DMS) and next-generation sequencing (NGS) to those suggested by AlphaFold2 and our correlation analysis. The results demonstrated a high degree of similarity between the experimental and computational methods.

Discussion

Alternative binding scaffolds such as Affibodies are being explored for anti-(PD-L1) therapies. In order to elicit an effective immune response, Affibody binding must target the correct epitope on PD-L1, therefore correctly identifying binding epitopes and providing high resolution methods to compare subtle differences in modes of binding for different binders is crucial. Computational methods offer significant potential in this regard, but frequently they provide ambiguous results.

Here we introduced a methodology combining computational techniques and experimental validation to predict and analyze the structure and interactions of a PD-L1:Affibody complex. Through an integrated approach utilizing AlphaFold2, network analysis of molecular dynamics simulations, and deep mutational scanning we were able to understand and validate the critical interaction regions of a high affinity PD-L1:Affibody complex.

Our findings indicated that AlphaFold2 Multimer, ClusPro, and ZDock yield anal-

ogous results, and suggested two potential binding modes referred to as parallel and perpendicular (Fig. 2B,C). We subjected the structures of both modes to MD simulations and used Dynamic Network Analysis of the trajectories to discover that the perpendicular mode exhibited higher stability (Fig. 2F, Fig. S2). Notably, traditional MD trajectory analyses such as RMSD and MM-PBSA struggled to distinguish the binding differences between the two modes (Fig. 2D,E). The advantage of the network analysis lies in its foundation on the generalized correlation of atomic motions, making it more adept than standard chemical descriptors in predicting interface interactions.⁵⁷

While conducting *in silico* targeted mutations for epitope mapping, we observed that AlphaFold2 adeptly captured experimentally measured trends in affinity of mutants versus the WT, but failed with small affinity variations (Fig. 3A,B, Fig.S6). These results highlight the important role MD simulations can play in successfully refining the AlphaFold2 predictions and elevating the perpendicular binding mode as the most probable predicted mode. Further experimental validation by high-resolution cross-linking mass spectrometry and deep mutational scanning confirmed the perpendicular mode as the accurate binding conformation (Fig. 4). Notably, using just AlphaFold2 scoring, less favorable structural conformations, such as the parallel binding mode, still garnered relatively high prediction scores.

During the review process of this work, AlphaFold3⁵¹ was released as a web server, promising significant improvements in prediction accuracy. As shown in Fig.2B, we explored its capabilities and found that AlphaFold3 predicted only an ensemble of conformations around the parallel orientation. This prediction conflicted with our molecular dynamics-based results and did not align with the experimental validation presented here. Given this discrepancy, we limited the use of AlphaFold3 to this initial search of conformations presented in Fig.2B and did not use these structures for further refinement using molecular dynamics simulations. We also did not employ AlphaFold3 for the mutational scanning presented in Fig. 3. We believe that AlphaFold3 predictions in our comparison would not contribute substantively to the findings presented in this study and could actually be misleading. Instead, we have focused on validating our

experimental results with orthogonal techniques and believe they provide a robust basis for the conclusions drawn.

As discussed in this work, various experimental methods for binding epitope determination exist, but they are typically low throughput and cost-intensive. Our approach demonstrates the efficacy of combining AlphaFold2 with MD simulations and Dynamic Network Analysis to characterize binding interfaces efficiently. The throughput of this computational methodology is generally higher than experimental approaches, though it varies based on system preparation, MD simulation time, and computational resources. For example, the PD-L1 system can be analyzed using AlphaFold Multimer 2.3 in under 3 hours per run on Nvidia DGX-A100 hardware. Concurrent execution of different mutations enhances throughput, allowing deep mutational scanning within the same 3-hour timeframe. MD simulations require careful setup and equilibration, but a 100-nanosecond simulation on modern graphical processing units (GPUs), such as the Nvidia H100, takes less than 5 hours, compared to 15 hours on older GPUs (Nvidia V100). Larger systems with more complex interactions may require more extensive computational resources and time. This highlights the impact of hardware advancements on improving throughput.

The low throughput in this work was primarily due to the experimental techniques used for validation. Our goal was to employ advanced experimental methods to validate our computational methodology, which involved complex and time-consuming techniques. However, our approach can be adapted to require less experimental validation — as discussed, the binding affinities analyzed by flow cytometry were mostly sufficient — or to use advanced experiments in combination with simulations for a feedback loop between computational and experimental work. For instance, the results presented in Fig. 4 could serve as constraints in the simulations, providing a more accurate depiction of the interface. This iterative approach could potentially aid in the development of better binders.

In summary, our approach demonstrates the efficacy of combining AlphaFold2 with MD simulations and Dynamic Network Analysis in characterizing binding interfaces.

The PD-L1:Affibody complex serves as an illustrative case study, especially considering the multitude of binding regions identified for PD-L1 binders (Fig. 2A). Our computational approach proved robust and was corroborated by an array of experimental techniques. The methodology holds promise for exploring other bimolecular interactions, as it can be easily implemented for other binding interfaces. Moreover, within cancer therapeutics, our technique marks a significant step towards correct epitope targeting for the design of enhanced cancer treatments.

Methods

In summary, here we employed AlphaFold2 Multimer^{30,31} to predict the structure of the PD-L1:Affibody complex. To investigate the most probable conformation from the AlphaFold2 predictions, we prepared molecular dynamics (MD) simulations using QwikMD⁴⁶ and carried them out using the GPU-resident version of NAMD 3.0.⁴⁷ Metrics including RMSD, MM-PBSA, and Dynamic Network Analysis⁴⁸ were applied to evaluate the stability of the Affibody:PD-L1 complex structure.

To further confirm the correct conformation, we once more employed AlphaFold2 Multimer for a *in silico* mutational scanning. For native PAGE, bead-based cytometry and XL-MS analysis, we produced Affibody and PD-L1 variants in *E. coli*. For DMS, a scanning NNK codon library covering the entire length of PD-L1 was displayed in EBY100 yeast through an Aga2p anchor system. NGS was carried out using paired end 300 bp reads on an Illumina NextSeq2000.

A detailed exploration of the methodologies is presented below.

PD-L1:Affibody AlphaFold models

The aminoacid sequences for the mature domain of PD-L1 (UniProt Q9NZQ7, 18-234) and the Affibody⁵⁸ (1-60) were used as a single FASTA input. The models for the PD-L1:Affibody complex were predicted using AlphaFold-Multimer^{30,31} version 2.3.2, using all 5 available v3 multimer parameter sets, to generate 5 models each, resulting

in a total of 25 predictions per sequence pair. The QwikFold VMD's⁵⁴ plugin was used to set the experiments and post-process the results, and calculations were run using the Cybershuttle⁵⁹ Research Environment deployed at the SDSC Expanse⁶⁰ supercomputer. QwikFold was used to align the models for visual inspection, addressing per-residue confidence as measured by pLDDT.⁶¹ The predicted aligned error (PAE) matrices were additionally inspected to assess the confidence in the relative position and orientation of the two major binding conformations: perpendicular or parallel.

Using AlphaFold for epitope mapping

For *in silico* mutagenesis aimed at epitope mapping the PD-L1:Affibody complex by mutating the PD-L1 sequence, the selected mutations I54A, V68Y, M115A, Y123D, R125E were used following the same procedure as above. Position Q66 was further scanned for all the twenty common amino acids. The iptm+ptm score for the best ranking model for the perpendicular orientation was plotted as bar graph (Fig. 3A-B). The predicted template modelling (pTM)³⁰ score and the interface predicted template modelling (ipTM)³¹ score are extensions of the template modelling (TM)⁶² score, which evaluates the accuracy of a protein's global structure, minimizing the impact of localized inaccuracies.⁶³ Both provide valuable insights, the ipTM score is particularly useful for evaluating the relative positions of subunits within a complex. High ipTM scores generally correlate with accurate predictions of the entire complex. The pTM score estimates how accurately AlphaFold-Multimer predicts the overall structure of a protein complex. It is calculated by comparing the predicted structure with the hypothetical true structure. A pTM score above 0.5 indicates that the predicted fold might resemble the true structure, whereas a score below 0.5 suggests that the predicted structure is likely incorrect. On the other hand, the ipTM score focuses on the accuracy of the predicted relative positions of the subunits within the protein complex. This metric is crucial for assessing the interactions between proteins in a complex. An ipTM score above 0.8 signifies a confident, high-quality prediction, while a score below 0.6 indicates a likely failed prediction. Scores between 0.6 and 0.8 fall into a grey zone

where predictions might be accurate or erroneous.⁶⁴

Molecular Dynamics Simulations

The perpendicular and parallel AlphaFold2 predicted models for the PD-L1 complex, with PD-L1 trimmed to its IgV domain (residues 18-115), were subjected to refinement and conformational sampling by molecular dynamics (MD) simulations following standard protocols.⁴⁶ The CHARMM36 force field^{65,66} and TIP3P water model⁶⁷ were used for all systems, with sodium ions as counter-ions randomly arranged in the solvent. Each system comprised approximately 40,000 atoms. All MD simulations were performed using the GPU-accelerated NAMD 3 package,⁴⁷ assuming periodic boundary conditions in the NpT ensemble. Temperature was maintained at 300 K using Langevin dynamics, and pressure was kept at 1 bar using the Langevin piston method. A cut-off distance of 12.0 Å was applied to short-range, non-bonded interactions, and long-range electrostatic interactions were treated using the particle-mesh Ewald (PME) method.⁶⁸ The equations of motion were integrated using the r-RESPA multiple time step scheme,⁶⁹ updating Lennard-Jones interactions every step and electrostatic interactions every two steps, with a time step of 2 fs for all simulations. The first 2 ns of each simulation were used for system equilibration, followed by 100 ns production runs performed in 16 replicas.

Molecular Dynamics Simulations Analysis

All analyses of MD trajectories were carried out employing VMD,⁵⁴ its plugins and TCL scripts,⁷⁰ unless stated differently. Analysis outputs were post-processed to generate graphs using Python3⁷¹ libraries, including Matplotlib,⁷² Pandas,^{73,74} and Seaborn,⁷⁵ unless stated differently. Contacts were calculated using PyContact⁷⁶ using a protocol adapted from previous work.^{77,78}

Binding Affinity Prediction

MM-PBSA is a computationally efficient method for estimating the binding free energy (ΔG_{bind}) of protein-protein complexes.^{79–81} Here, to characterize PD-L1:Affibody coupling, we performed the effective free energy (ΔG_{eff} , neglecting the configurational entropic contribution) binding affinity prediction by MM-PBSA using the CaFE plugin.⁸² Molecular Mechanics (MM) was computed with NAMD3,⁴⁷ Solvent accessible area with VMD,⁵⁴ and Poisson-Boltzmann term was computed with APBS.⁸³

Dynamic Network Analysis

The dynamical network analysis python package⁴⁸ was used to extract correlations of motion from PD-L1:Affibody simulations. A network was defined as a set of nodes, all $C\alpha$, with connecting edges.⁸⁴ Edges connect pairs of nodes if corresponding monomers are in contact, and 2 nonconsecutive monomers are said to be in contact if they fulfill a proximity criterion, namely any heavy atoms (non-hydrogen) from the 2 monomers are within 4.5Å of each other for at least 75% of the frames analyzed.⁸⁵ Following previously defined protocols, the dynamical networks were constructed from 20 ns windows of the total trajectories sampled every 200 ps.^{86–88}

Cluster analysis

Clustering analysis was performed using the unsupervised k-means algorithm as implemented in CPPTRAJ,⁸⁹ with the parameters “randompoint maxit” set to 500 and “sieving” set to 10. The distance metric utilized for clustering was the RMSD of the complex C, N, O, $C\alpha$ and $C\beta$ atoms, and the clustering process concluded when the number of clusters reached 5. The analysis was applied to the last quarter of each of the 16 replica trajectories for each orientation (perpendicular and parallel), after fitting to the complex backbone, excluding the 5 residues from both Affibody’s termini.

Native contacts for PDB structures

To reveal insights into the PD-L1:Affibody complex assembly we performed an analysis of PD-L1 (UniProt Q9NZQ7) structures found in the Protein Data Bank (PDB), with interactions with other proteins.⁷⁶ The native contacts identified after superposition of available structures, with contacts defined by amino acid residues within a 4Å radius of the PD-L1. Residues with contacts are highlighted in Fig.2. The structures considered were: 3BIK, 3FN3, 3SBW, 4Z18, 4ZQK, 5GGT, 5GRJ, 5IUS, 5J89, 5J8O, 5JDR, 5JDS, 5N2D, 5N2F, 5NIU, 5O45, 5O4Y, 5X8L, 5X8M, 5XXY, 6NM7, 6NM8, 6NNV, 6NOJ, 6NOS, 6PV9, 6R3K, 6RPG, 6VQN, 6YCR, 7BEA, 7C88, 7CZD, 7DY7, 7NLD, 7OUN, 7SJQ, 7TPS.

Plasmids available on addgene:

Addgene plasmid #157674: pET28a-ybbR-His-ELP(MV7E2)3-FLN-SpyCatcher.

Cloning of PD-L1-HIS and PD-L1-HIS-SpyTag

The DNA sequence of the extracellular domain of human PD-L1 was chemically synthesized with optimal codons for production in *E. coli* (GeneArt, Thermo Fisher Scientific) and introduced into a pET28a vector via NdeI and XhoI restriction sites to generating a new vector pET28a-PD-L1-HIS. The plasmid contents were confirmed by further Sanger DNA sequencing analysis. A SpyTag was further introduced at the C-terminus of PD-L1 by PCR using primers #1 and #2 (Table S2) based on the plasmid pET28a-PD-L1-HIS and following Gibson assembly with master mix (NEB) generating a new vector pET28a-PD-L1-ECD-HIS-SpyTag, which was confirmed by further DNA sequencing analysis.

Cloning of PD-L1-HIS-SpyTag Mutants

For the mutational analysis, six point-mutations (I54A, Q66D, V68Y, M115A, Y123D and R125E) were designed and incorporated by site directed mutagenesis using the

Q5® Site-Directed Mutagenesis kit (NEB) with primers #3 and #4 for **I54A**, #5 and #6 for **Q66D**, #6 and #7 for **V68Y**, #8 and #9 for **M115A**, #10 and #11 for **Y123D**, and #11 and #12 for **R125E** (Table S2), generating new plasmids pET28-PD-L1-SpyTag-**I54A**, **-Q66D**, **-V68Y**, **-M115A**, **-Y123D** and **-R125E**, which was confirmed by further DNA sequencing analysis.

Expression, Refolding, and Purification of PD-L1 Variants

The plasmid with the sequence of the PD-L1 variant was introduced into competent *E. coli* BL21(DE3) strain. Recombinant cells were cultured in 5 ml of Luria-Bertani (LB) medium with 50 $\mu\text{g ml}^{-1}$ kanamycin at 37 °C overnight. The culture was transferred to 50 mL of Terrific broth (TB) medium with 50 $\mu\text{g ml}^{-1}$ kanamycin and cultivated at 37 °C and 200 rpm until an optical density at 600 nm (OD600) of 0.8-1.0 was reached. The expression of recombinant protein was induced by the addition of 1.0 mM isopropyl- β -D-thio-galactopyranoside (IPTG) and the culture was further incubated at 37 °C and 200 rpm for 9 hrs. The cells were harvested by centrifugation at 4,000 g for 20 min at 4 °C. The harvested cell pellet was resuspended in a denaturing lysis buffer (10 mM Tris-Cl, and 8M urea; pH 8). Resuspended cells were placed on ice and disrupted for 15 min using a sonic dismembrator using a 3 s on: 5 s off pattern to allow cooling between each pulse. The lysate was centrifuged at 14,000 g for 20 min at 4 °C. The supernatant was collected and incubated with Ni-NTA resin for 30 min at room temperature to allow the His6-tagged proteins to bind to the Ni-NTA resin. Then, the mixture was loaded onto a column. The resin was washed with 10–20 resin volumes of wash buffer (20 mM imidazole, 10 mM Tris-Cl, and 8M urea; pH 8). Recombinant proteins were eluted in the elution buffer (500 mM imidazole, 10 mM Tris-Cl, and 8M urea; pH 8). The eluted protein solution was serially dialyzed to 8 M, 4 M, 2 M, and 0 M Urea with 5% glycerol, 5% sucrose, 1% arginine, 0.5 mM NaCl in 20 mM Tris-Cl (pH 7.4), and finally to 1x PBS buffer. Precipitation during dialysis was removed by centrifugation at 14,000 g for 20 min at 4 °C and supernatant was further purified by SEC column.

Cloning of AFF-HIS and ybbR-HIS-ELP-FLN-Anti-PD-L1-AFF (L-AFF)

DNA sequence of Anti-PD-L1 Affibody (AFF) was chemically synthesized based on the codon usage of *E. coli* (GeneArt, Thermo Fisher Scientific) and introduced into pET28a vector via NdeI and XhoI restriction sites generating a new vector pET28a-Anti-PD-L1-AFF-HIS (for preparation of AFF-HIS), which was confirmed by further DNA sequencing analysis. To immobilize AFF on the PS beads for flow cytometry analysis, ybbR tag and linker was introduced at the N-terminus of AFF (for preparation of L-AFF) by PCR using primers #13 and #14 based on the plasmid pET28a-Anti-PD-L1-AFF-HIS and using primers #15 and #16 based on the plasmid #157674 (Addgene) (Table S2). Two PCR products were assembled into a new vector pET28a-ybbR-HIS-ELP-FLN-Anti-PD-L1-AFF by Gibson assembly with master mix (NEB), which was confirmed by further DNA sequencing analysis.

Expression and Purification of AFF-His and L-AFF

The plasmid with the sequence of AFF-His or L-AFF was introduced into competent *E. coli* BL21(DE3) strain. Recombinant cells were cultured in 5 ml of Luria-Bertani (LB) medium with 50 $\mu\text{g ml}^{-1}$ kanamycin at 37 °C overnight. The culture was transferred to 50 mL of Terrific broth (TB) medium with 50 $\mu\text{g ml}^{-1}$ kanamycin and cultivated at 37 °C and 200 rpm until an optical density at 600 nm (OD₆₀₀) of 0.8-1.0 was reached. The expression of recombinant protein was induced by the addition of 0.5 mM IPTG and the culture was further incubated at 20 °C and 200 rpm for 9 hrs. The cells were harvested by centrifugation at 4,000 g for 20 min at 4 °C. The cells were harvested by centrifugation at 4,000 g for 20 min at 4 °C. The harvested cell pellet was resuspended in lysis buffer (50 mM Tris, 50 mM NaCl, 0.1% Triton X-100, 5 mM MgCl₂; pH 8.0), and disrupted with a sonic dismembrator. The lysate was centrifuged at 14,000 g for 20 min at 4 °C. The supernatant was collected and incubated with Ni-NTA resin, loaded onto a column, washed with wash buffer (1x PBS with 20 mM imidazole; pH 7.4), and

eluted in elution buffer (1x PBS with 500 mM imidazole; pH 7.4). The eluted protein solution was further purified by SEC column.

Native-PAGE Analysis

Binding behavior between AFF and PD-L1 mutants (WT, I54A, Q66D, V68Y, M115A, Y123D and R125E) were screened by Native-PAGE. 5 μ L of 10 μ M L-AFF was mixed with 5 μ L of each 10 μ M PD-L1 mutants, incubated several hours at RT, and then total solution was run in Native-PAGE. Protein bands were visualized by Coomassie staining. Bound and unbound fraction of PD-L1 were calculated based on the intensity of stained protein bands.

Flow Cytometry Analysis

The binding affinity between Anti-(PD-L1)-AFF and PD-L1 mutants was analyzed using the Attune NxT (Thermo Fisher Scientific) flow cytometer equipped with a 488 nm and a 561 nm laser. L-AFF was immobilized onto the surface of amine-functionalized PS beads via ybbR Tag. The amine groups reacted to a NHS group from sulfo-succinimidyl 4-(N-maleimidomethyl)cyclohexane-1-carboxylate (sulfo-SMCC; Thermo Fischer Scientific) in 50 mM HEPES buffer pH 7.5 for 30 min. The thiol group from Coenzyme A (CoA, 200 μ M) reacted to a maleimide group from sulfo-SMCC in coupling buffer (50 mM sodium phosphate, 50 mM NaCl, 10 mM EDTA, pH 7.2) for 2 hrs. Finally, the ybbR-tagged protein L-AFF was immobilized onto the surface using SFP-mediated ligation to CoA in Mg²⁺ supplemented 1x PBS buffer. This resulted in covalent immobilization of AFF to PS beads. Protein-immobilized beads were extensively washed and kept in 1x PBS buffer prior to immediate use. GFP-labeled PD-L1 mutants were prepared by conjugating SpyTag of PD-L1 mutants to GFP-SpyCatcher. L-AFF immobilized beads were incubated in GFP-labeled PD-L1 mutants' solution with different concentrations ranging from 0.008 nM to 625 nM for 1-2 hrs at RT. After washing, shift of fluorescence from GFP was recorded and plotted against the

concentration of PD-L1 mutants to derive the dissociation constant between PD-L1 mutants and L-AFF.

XL-MS Epitope Mapping

The conformational epitope mapping for PD-L1 and Anti-PD-L1-AFF was performed by chemical cross-linking and high-resolution mass spectrometry (XL-MS; CovalX, Switzerland). 10 μ L of PD-L1 (6 μ M) was mixed with 10 μ L of AFF (6 μ M) to obtain PD-L1/AFF mix with a final concentration of 3 μ M. Then, 2 μ L of deuterated cross-linker disuccinimidyl suberate (DSS d0/d12, 2 mg/mL in DMF) was added to the protein mixture and the solution was incubated for 180 min at RT to complete the cross-linking reaction. After cross-linking the reaction was stopped with 20 mM ammonium bicarbonate, the samples were submitted to reduction, alkylation, and proteolysis with five different enzymes (trypsin, chymotrypsin, ASP-N, elastase, and thermolysin). After enrichment of the cross-linked peptides, the samples were analyzed by high-resolution mass spectrometry (nLC-Q-Exactive Orbitrap MS). The NHS groups of DSS reacts only with positively charged amino groups or hydroxyl groups including Arg, His, Lys, Ser, Thr, and Tyr. Specific amino acid residues that were cross-linked were identified by tandem MS/MS analysis. Based on these cross-linked amino acids and the crystal structure of PD-L1, possible binding epitope was proposed.

Cloning for Yeast Surface Display of PD-L1

For the mutational analysis, PD-L1 sequence was amplified by PCR using primers #17 and #18 (Table S2) based on the plasmid pET28a-PD-L1-ECD-HIS and introduced via NheI and BamHI restriction site into the yeast plasmid pYD1 for surface display (Addgene plasmid #73447) generating a new vector pCHA-HA-PD-L1-ECD-HIS-Xpress-Aga2p, which was confirmed by further DNA sequencing analysis.

Site-Saturation PD-L1 Libraries

Sequence of PD-L1-ECD was divided into two regions; IgV-like domain (F19 to A132; 114 aa) and IgC-like domain (P133 to R238; 106 aa). Site-saturation mutagenesis libraries spanning all 114 positions for IgV-like domain (Library A) and for all 106 positions for IgC-like domain (Library B) encoding all the 19 possible amino acid mutations were produced by Twist Bioscience. The backbone template was prepared by digestion of plasmid pCHA-HA-PD-L1-ECD-HIS-Xpress-Aga2p with NheI and BamHI. After agarose gel electrophoresis, purified backbone template was mixed with each of synthesized PD-L1 Library A and B with a ratio of 1:10. Then, the mixture was directly transformed into *Saccharomyces cerevisiae* EBY100 following a typical lithium acetate transformation procedure for introducing site=saturation mutagenesis library into backbone template via endogenous homologous recombination. Right after transformation, serial dilutions were plated on synthetic defined (SD) agar 2% (w/v) glucose plates lacking tryptophan (-Trp) to count the number of transformants. Remaining transformation reaction solution was further grown in liquid SD glucose -Trp medium for 48 h at 30 °C with continuous shaking at 200 rpm. Then, cultured yeast library was harvested, prepared as 50% glycerol stock and stored at -80 °C for further analysis.

Yeast Surface Display

Saccharomyces cerevisiae EBY100 transformants harboring the plasmid pCHA-HA-PD-L1-ECD-HIS-Xpress-Aga2p or PD-L1 yeast library A and B were cultivated in SD-TRP liquid medium with 2% glucose for 24 h at 30 °C with continuous shaking at 200 rpm. Protein expression and protein display were then induced by transferring the culture to a fresh pH-buffered liquid medium (0.1 M Potassium phosphate pH 7.0) lacking tryptophan containing 0.2% (w/v) glucose and 1.8% (w/v) galactose and by shaking for 24 h at 30 °C.

Cell Sorting with Dual Labelling

PD-L1 libraries are displayed with N-terminal HA tag. Therefore, the expression of PD-L1 was labelled via HA-tag and the binding of AFF-HIS was labelled via HIS-tag. Yeast cells displaying the PD-L1 libraries were incubated with 4 nM of AFF-HIS in 1x PBS containing 0.1% BSA for > 1 h at RT with shaking. After washing with 1x PBS containing 0.1% BSA, cells were incubated with a mixture of primary antibodies, HA tag recombinant rabbit monoclonal antibody (RM305; Thermo) and 6x-His tag monoclonal antibody (HIS.H8; Sigma) in 1x PBS containing 0.1% BSA for 1 h at RT. After washing, cells were further incubated with a mixture of secondary antibodies, Goat anti-rabbit IgG (H+L) cross-adsorbed secondary antibody, alexa Fluor™ 488 (Thermo) and Goat anti-Mouse IgG (H+L) highly cross-adsorbed secondary antibody, Alexa Fluor™ 594 (Thermo) in 1x PBS containing 0.1% BSA for 1 h on ice. Finally, dually labelled cells were washed with ice-cold 1x PBS containing 0.1% BSA for further analysis. Cells were sorted by FACSMelody™ Cell Sorter (BD Bioscience). Cells with high expression level with decreased or no binding were sorted, transferred to SD -TRP liquid medium with 2% glucose and 50 µg ml⁻¹ ampicillin, cultivated for 48 h at 30 °C, harvested, and prepared as 15% glycerol stock and stored at -80 °C for further analysis.

Illumina Sequencing and Data Analysis

Plasmids were extracted from the sorted cells using zymolyase (Zymo Research, Irvine, USA) and GeneJET Plasmid Miniprep Kit (Thermo Scientific). Regions of the PD-L1 IgV- and IgC-like domain were amplified by the first PCR step and then indexes and adapters were added for Illumina sequencing by the second PCR step with the primer sets from IDT for Illumina–DNA/RNA UD Indexes Plate A and NEBNext® Ultra™ II Q5® Master Mix (NEB). Final products were purified using AMPure XP beads (Beckman Coulter) and Illumina sequencing was performed with the Illumina NextSeq2000 for paired end 300 bp (PE 2x 300 bp, 600 cycles) (Functional Genomics Center Zurich). The following data analyses were performed at sciCORE (<http://scicore.unibas.ch/>)

scientific computing center at University of Basel. Briefly, sequencing results from the paired end were combined after trimming reads presenting a quality under 20 and translated into amino acid sequences. Each of identical amino acid sequences only with one mutation (or wild type) were grouped and counted to calculate the ratio in the sorted population (high expression with low/no binding) compared to the initial population.

Acknowledgement

RCB and DEBG are supported by the National Science Foundation under Grant MCB-2143787, and the National Institute of General Medical Sciences (NIGMS) of NIH through the grant R24-GM145965. This work was supported by a grant from the Swiss State Secretariat for Education, Research and Innovation (SERI) to MN.

Supporting Information Available

Additional information is available in the Supporting Information section, including Supporting Tables (Cluster Analysis, Protein/Peptide Sequences) and Supporting Figures (Cluster Analysis, Correlation Analysis, Bootstrap Analysis, Binding Epitope Simulations/Experiments).

References

- (1) Topalian, S. L.; Drake, C. G.; Pardoll, D. M. Immune Checkpoint Blockade: A Common Denominator Approach to Cancer Therapy. *Cancer Cell* **2015**, *27*, 450–461.
- (2) Chen, L.; Han, X. Anti-PD-1/PD-L1 therapy of human cancer: past, present, and future. *Journal of Clinical Investigation* **2015**, *125*, 3384–3391.

- (3) Borghaei, H. et al. Nivolumab versus Docetaxel in Advanced Nonsquamous Non-Small-Cell Lung Cancer. *New England Journal of Medicine* **2015**, *373*, 1627–1639.
- (4) Robert, C. et al. Nivolumab in Previously Untreated Melanoma without BRAF Mutation. *New England Journal of Medicine* **2015**, *372*, 320–330.
- (5) Powles, T.; Eder, J. P.; Fine, G. D.; Braiteh, F. S.; Loriot, Y.; Cruz, C.; Bellmunt, J.; Burris, H. A.; Petrylak, D. P.; Leng Teng, S.; Shen, X.; Boyd, Z.; Hegde, P. S.; Chen, D. S.; Vogelzang, N. J. MPDL3280A (anti-PD-L1) treatment leads to clinical activity in metastatic bladder cancer. *Nature* **2014**, *515*, 558–562.
- (6) Patel, S. P.; Kurzrock, R. PD-L1 expression as a predictive biomarker in cancer immunotherapy. *Molecular cancer therapeutics* **2015**, *14*, 847–856.
- (7) Alsaab, H. O.; Sau, S.; Alzhrani, R.; Tatiparti, K.; Bhise, K.; Kashaw, S. K.; Iyer, A. K. PD-1 and PD-L1 checkpoint signaling inhibition for cancer immunotherapy: mechanism, combinations, and clinical outcome. *Frontiers in pharmacology* **2017**, *8*, 561.
- (8) Bryan, C. M.; Rocklin, G. J.; Bick, M. J.; Ford, A.; Majri-Morrison, S.; Kroll, A. V.; Miller, C. J.; Carter, L.; Goreshnik, I.; Kang, A.; others Computational design of a synthetic PD-1 agonist. *Proceedings of the National Academy of Sciences* **2021**, *118*, e2102164118.
- (9) Kholodenko, R. V.; Kalinovsky, D. V.; Doronin, I. I.; Ponomarev, E. D.; Kholodenko, I. V. Antibody fragments as potential biopharmaceuticals for cancer therapy: success and limitations. *Current Medicinal Chemistry* **2019**, *26*, 396–426.
- (10) Xiong, C.; Mao, Y.; Wu, T.; Kang, N.; Zhao, M.; Di, R.; Li, X.; Ji, X.; Liu, Y. Optimized expression and characterization of a novel fully human bispecific single-chain diabody targeting vascular endothelial growth factor165 and programmed death-1 in *Pichia pastoris* and evaluation of antitumor activity in vivo. *International Journal of Molecular Sciences* **2018**, *19*, 2900.

- (11) Jing, L.; Liu, J.; Cui, D.; Li, Y.; Liu, Z.; Tao, L.; Zhao, Q.; Diao, A. Screening and production of an affibody inhibiting the interaction of the PD-1/PD-L1 immune checkpoint. *Protein Expression and Purification* **2020**, *166*, 105520.
- (12) Foord, E.; Klynning, C.; Schoutrop, E.; Förster, J. M.; Krieg, J.; Mörtberg, A.; Müller, M. R.; Herzog, C.; Schiegg, D.; Villemagne, D.; others Profound functional suppression of tumor-infiltrating T-cells in ovarian cancer patients can be reversed using PD-1-blocking antibodies or DARPin® proteins. *Journal of Immunology Research* **2020**, 2020.
- (13) Grindel, B. J.; Engel, B. J.; Ong, J. N.; Srinivasamani, A.; Liang, X.; Zacharias, N. M.; Bast Jr, R. C.; Curran, M. A.; Takahashi, T. T.; Roberts, R. W.; others Directed Evolution of PD-L1-Targeted Affibodies by mRNA Display. *ACS Chemical Biology* **2022**, *17*, 1543–1555.
- (14) Wang, X.; Li, F.; Qiu, W.; Xu, B.; Li, Y.; Lian, X.; Yu, H.; Zhang, Z.; Wang, J.; Li, Z.; others SYNBP: synthetic binding proteins for research, diagnosis and therapy. *Nucleic Acids Research* **2022**, *50*, D560–D570.
- (15) Simeon, R.; Chen, Z. In vitro-engineered non-antibody protein therapeutics. *Protein Cell* **9**: 3–14. 2018.
- (16) Frejd, F. Y.; Kim, K.-T. Affibody molecules as engineered protein drugs. *Experimental Molecular Medicine* **2017**, *49*, e306–e306.
- (17) Ståhl, S.; Gräslund, T.; Karlström, A. E.; Frejd, F. Y.; Nygren, P.-Å.; Löfblom, J. Affibody molecules in biotechnological and medical applications. *Trends in biotechnology* **2017**, *35*, 691–712.
- (18) Schutkowski, M., Reineke, U., Eds. *Epitope Mapping Protocols*; Humana Press, 2009; Vol. 524.
- (19) Nilvebrant, J.; Rockberg, J. In *Epitope Mapping Protocols*; Rockberg, J., Nilvebrant, J., Eds.; Springer New York: New York, NY, 2018; pp 1–10.

- (20) King, M. T.; Brooks, C. L. *Epitope Mapping of Antibody-Antigen Interactions with X-Ray Crystallography*; 2018; pp 13–27.
- (21) Bardelli, M.; Livoti, E.; Simonelli, L.; Pedotti, M.; Moraes, A.; Valente, A. P.; Varani, L. Epitope mapping by solution NMR spectroscopy. *Journal of Molecular Recognition* **2015**, *28*, 393–400.
- (22) Renaud, J.-P.; Chari, A.; Ciferri, C.; ti Liu, W.; Rémy, H.-W.; Stark, H.; Wiesmann, C. Cryo-EM in drug discovery: achievements, limitations and prospects. *Nature Reviews Drug Discovery* **2018**, *17*, 471–492.
- (23) Qi, H. et al. Antibody Binding Epitope Mapping (AbMap) of Hundred Antibodies in a Single Run. *Molecular Cellular Proteomics* **2021**, *20*, 100059.
- (24) Infante, Y. C.; Pupo, A.; Rojas, G. A combinatorial mutagenesis approach for functional epitope mapping on phage-displayed target antigen. *mAbs* **2014**, *6*, 637–648.
- (25) Jethva, P. N.; Gross, M. L. Hydrogen deuterium exchange and other mass spectrometry- based approaches for epitope mapping. *Frontiers in Analytical Science* **2023**, *3*.
- (26) Van Blarcom, T.; Rossi, A.; Foletti, D.; Sundar, P.; Pitts, S.; Melton, Z.; Telman, D.; Zhao, L.; Cheung, W. L.; Berka, J.; Zhai, W.; Strop, P.; Pons, J.; Rajpal, A.; Chaparro-Riggers, J. Epitope Mapping Using Yeast Display and Next Generation Sequencing. *Methods Mol. Biol.* **2018**, *1785*, 89–118.
- (27) Boughter, C. T.; Meier-Schellersheim, M. An integrated approach to the characterization of immune repertoires using AIMS: An Automated Immune Molecule Separator. *PLOS Computational Biology* **2023**, *19*, e1011577.
- (28) Boughter, C. T.; Meier-Schellersheim, M. Conserved biophysical compatibility among the highly variable germline-encoded regions shapes TCR-MHC interactions. *Elife* **2023**, *12*, e90681.

- (29) Borowska, M. T.; Boughter, C. T.; Bunker, J. J.; Guthmiller, J. J.; Wilson, P. C.; Roux, B.; Bendelac, A.; Adams, E. J. Biochemical and biophysical characterization of natural polyreactivity in antibodies. *Cell reports* **2023**, *42*.
- (30) Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596*, 583–589.
- (31) Evans, R. et al. Protein complex prediction with AlphaFold-Multimer. *bioRxiv* **2022**,
- (32) Baek, M. et al. Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **2021**, *373*, 871–876.
- (33) Yang, Z.; Zeng, X.; Zhao, Y.; Chen, R. AlphaFold2 and its applications in the fields of biology and medicine. *Signal Transduction and Targeted Therapy* **2023**, *8*, 115.
- (34) Marchand, A.; Hall-Beauvais, A. K. V.; Correia, B. E. Computational design of novel protein–protein interactions – An overview on methodological approaches and applications. *Current Opinion in Structural Biology* **2022**, *74*, 102370.
- (35) Chang, L.; Perez, A. Ranking Peptide Binders by Affinity with AlphaFold**. *Angewandte Chemie* **2023**, *135*.
- (36) Gomes, P. S. F. C.; Gomes, D. E. B.; Bernardi, R. C. Protein structure prediction in the era of AI: Challenges and limitations when applying to in silico force spectroscopy. *Frontiers in Bioinformatics* **2022**, *2*.
- (37) Gomes, P. S. F. C.; Forrester, M.; Pace, M.; Gomes, D. E. B.; Bernardi, R. C. May the force be with you: The role of hyper-mechanostability of the bone sialoprotein binding protein during early stages of Staphylococci infections. *Frontiers in Chemistry* **2023**, *11*.
- (38) Kang-Pettinger, T.; Walker, K.; Brown, R.; Cowan, R.; Wright, H.; Baravalle, R.; Waters, L. C.; Muskett, F. W.; Bowler, M. W.; Sawmynaden, K.; Coombs, P. J.;

- Carr, M. D.; Hall, G. Identification, binding, and structural characterization of single domain anti-PD-L1 antibodies inhibitory of immune regulatory proteins PD-1 and CD80. *Journal of Biological Chemistry* **2023**, *299*, 102769.
- (39) Yin, R.; Pierce, B. G. Evaluation of AlphaFold Antibody-Antigen Modeling with Implications for Improving Predictive Accuracy. *Protein Science* **2023**, e4865.
- (40) Hummer, A. M.; Abanades, B.; Deane, C. M. Advances in computational structure-based antibody design. *Current Opinion in Structural Biology* **2022**, *74*, 102379.
- (41) Trotter, D. E. G. et al. In vivo imaging of the programmed death ligand 1 by18F PET. *Journal of Nuclear Medicine* **2017**, *58*, 1852–1857.
- (42) Rubins, D. J.; Meng, X.; McQuade, P.; Michael Klimas, S.-A. L. B. M. C. S. S. O. H. H. M. P. L. G. M. H. J. L., Krista Getty; Eklund, P.; Ekblad, C.; Frejd, F. Y.; Hostetler, E. D.; Trotter, D. E. G.; Evelhoch, J. L. In Vivo Evaluation and Dosimetry Estimate for a High Affinity Affibody PET Tracer Targeting PD-L1. *Molecular Imaging and Biology* **2021**, *23*, 214–249.
- (43) Melo, M. C.; Bernardi, R. C. Fostering discoveries in the era of exascale computing: How the next generation of supercomputers empowers computational and experimental biophysics alike. *Biophysical Journal* **2023**,
- (44) Desta, I. T.; Porter, K. A.; Xia, B.; Kozakov, D.; Vajda, S. Performance and Its Limits in Rigid Body Protein-Protein Docking. *Structure* **2020**, *28*, 1071–1081.e3.
- (45) Pierce, B. G.; Hourai, Y.; Weng, Z. Accelerating Protein Docking in ZDOCK Using an Advanced 3D Convolution Library. *PLoS ONE* **2011**, *6*, e24657.
- (46) Ribeiro, J. V.; Bernardi, R. C.; Rudack, T.; Stone, J. E.; Phillips, J. C.; Fredolino, P. L.; Schulten, K. QwikMD—integrative molecular dynamics toolkit for novices and experts. *Scientific reports* **2016**, *6*, 26536.

- (47) Phillips, J. C. et al. Scalable molecular dynamics on CPU and GPU architectures with NAMD. *The Journal of Chemical Physics* **2020**, *153*.
- (48) Melo, M. C. R.; Bernardi, R. C.; de la Fuente-Nunez, C.; Luthey-Schulten, Z. Generalized correlation-based dynamical network analysis: a new high-performance approach for identifying allosteric communications in molecular dynamics trajectories. *The Journal of Chemical Physics* **2020**, *153*.
- (49) Hanning, K. R.; Minot, M.; Warrender, A. K.; Kelton, W.; Reddy, S. T. Deep mutational scanning for therapeutic antibody engineering. *Trends in Pharmacological Sciences* **2022**, *43*, 123–135.
- (50) wei Lin, D. Y.; Tanaka, Y.; Iwasaki, M.; Gittis, A. G.; Su, H.-P.; Mikami, B.; Okazaki, T.; Honjo, T.; Minato, N.; Garboczi, D. N. The PD-1/PD-L1 complex resembles the antigen-binding Fv domains of antibodies and T cell receptors. *Proceedings of the National Academy of Sciences* **2008**, *105*, 3011–3016.
- (51) Abramson, J. et al. Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **2024**, *630*, 493–500.
- (52) Liu, Z.; Liu, H.; Vera, A. M.; Bernardi, R. C.; Tinnefeld, P.; Nash, M. A. High force catch bond mechanism of bacterial adhesion in the human gut. *Nat. Commun.* **2020**, *11*, 1–12.
- (53) Jobst, M. A.; Milles, L. F.; Schoeler, C.; Ott, W.; Fried, D. B.; Bayer, E. A.; Gaub, H. E.; Nash, M. A. Resolving dual binding conformations of cellulosome cohesin-dockerin complexes using single-molecule force spectroscopy. *Elife* **2015**, *4*.
- (54) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *Journal of Molecular Graphics* **1996**, *14*, 33–38.
- (55) Bauer, M. S.; Gruber, S.; Hausch, A.; Melo, M. C.; Gomes, P. S.; Nicolaus, T.; Milles, L. F.; Gaub, H. E.; Bernardi, R. C.; Lipfert, J. Single-molecule force

- stability of the SARS-CoV-2–ACE2 interface in variants-of-concern. *Nature Nanotechnology* **2024**, *19*, 399–405.
- (56) wei Lin, D. Y.; Tanaka, Y.; Iwasaki, M.; Gittis, A. G.; Su, H.-P.; Mikami, B.; Okazaki, T.; Honjo, T.; Minato, N.; Garboczi, D. N. The PD-1/PD-L1 complex resembles the antigen-binding Fv domains of antibodies and T cell receptors. *Proceedings of the National Academy of Sciences* **2008**, *105*, 3011–3016.
- (57) Melo, M. C.; Gomes, D. E.; Bernardi, R. C. Molecular origins of force-dependent protein complex stabilization during bacterial infections. *Journal of the American Chemical Society* **2023**, *145*, 70–77.
- (58) Wahlberg, E.; Gunneriusson, E. New Polypeptide Having Affinity to PD-L1. WIPO Patent WO2017072280A1, May 4. 2017.
- (59) Marru, S. et al. Cybershuttle: An End-to-End Cyberinfrastructure Continuum to Accelerate Discovery in Science and Engineering. 2023; pp 26–34.
- (60) Strande, S. et al. Expanse: Computing without Boundaries. 2021; pp 1–4.
- (61) Mariani, V.; Biasini, M.; Barbato, A.; Schwede, T. IDDT: a local superposition-free score for comparing protein structures and models using distance difference tests. *Bioinformatics* **2013**, *29*, 2722–2728.
- (62) Zhang, Y.; Skolnick, J. Scoring function for automated assessment of protein structure template quality. *Proteins: Structure, Function, and Bioinformatics* **2004**, *57*, 702–710.
- (63) Xu, J.; Zhang, Y. How significant is a protein structure similarity with TM-score = 0.5? *Bioinformatics* **2010**, *26*, 889–895.
- (64) Jeppesen, M.; André, I. Accurate prediction of protein assembly structure by combining AlphaFold and symmetrical docking. *Nature Communications* **2023**, *14*, 8283.

- (65) MacKerell, A. D. et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *The Journal of Physical Chemistry B* **1998**, *102*, 3586–3616.
- (66) Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D. Optimization of the Additive CHARMM All-Atom Protein Force Field Targeting Improved Sampling of the Backbone ϕ and Side-Chain χ_1 and χ_2 Dihedral Angles. *Journal of Chemical Theory and Computation* **2012**, *8*, 3257–3273.
- (67) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics* **1983**, *79*, 926–935.
- (68) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An $N \log(N)$ method for Ewald sums in large systems. *The Journal of Chemical Physics* **1993**, *98*, 10089–10092.
- (69) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kalé, L.; Schulten, K. Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry* **2005**, *26*, 1781–1802.
- (70) Spivak, M.; Stone, J. E.; Ribeiro, J.; Saam, J.; Freddolino, P. L.; Bernardi, R. C.; Tajkhorshid, E. VMD as a platform for interactive small molecule preparation and visualization in quantum and classical simulations. *Journal of chemical information and modeling* **2023**, *63*, 4664–4678.
- (71) Van Rossum, G.; Drake, F. L. *Python 3 Reference Manual*; CreateSpace: Scotts Valley, CA, 2009.
- (72) Hunter, J. D. Matplotlib: A 2D graphics environment. *Computing in Science & Engineering* **2007**, *9*, 90–95.

- (73) development team, T. P. Pandas-dev/pandas: Pandas. 2020; <https://doi.org/10.5281/zenodo.3509134>.
- (74) Wes McKinney Data Structures for Statistical Computing in Python. Proceedings of the 9th Python in Science Conference. 2010; pp 56 – 61.
- (75) Waskom, M. L. seaborn: statistical data visualization. *Journal of Open Source Software* **2021**, *6*, 3021.
- (76) Scheurer, M.; Rodenkirch, P.; Siggel, M.; Bernardi, R. C.; Schulten, K.; Tajkhorshid, E.; Rudack, T. PyContact: rapid, customizable, and visual analysis of non-covalent interactions in MD simulations. *Biophysical journal* **2018**, *114*, 577–583.
- (77) dos Santos Natividade, R.; Koehler, M.; Gomes, P. S.; Simpson, J. D.; Smith, S. C.; Gomes, D. E.; de Lhoneux, J.; Yang, J.; Ray, A.; Dermody, T. S.; others Deciphering molecular mechanisms stabilizing the reovirus-binding complex. *Proceedings of the National Academy of Sciences* **2023**, *120*, e2220741120.
- (78) Seppälä, J.; Bernardi, R. C.; Haataja, T. J.; Hellman, M.; Pentikäinen, O. T.; Schulten, K.; Permi, P.; Yläanne, J.; Pentikäinen, U. Skeletal dysplasia mutations effect on human filamins’ structure and mechanosensing. *Scientific reports* **2017**, *7*, 4218.
- (79) Srinivasan, J.; Miller, J.; Kollman, P. A.; Case, D. A. Continuum Solvent Studies of the Stability of RNA Hairpin Loops and Helices. *Journal of Biomolecular Structure and Dynamics* **1998**, *16*, 671–682.
- (80) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; Donini, O.; Cieplak, P.; Srinivasan, J.; Case, D. A.; Cheatham, T. E. Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Accounts of Chemical Research* **2000**, *33*, 889–897.

- (81) Wang, C.; Greene, D.; Xiao, L.; Qi, R.; Luo, R. Recent Developments and Applications of the MMPBSA Method. *Frontiers in Molecular Biosciences* **2018**, *4*.
- (82) Liu, H.; Hou, T. CaFE: a tool for binding affinity prediction using end-point free energy methods. *Bioinformatics* **2016**, *32*, 2216–2218.
- (83) Jurrus, E. et al. Improvements to the APBS biomolecular solvation software suite. *Protein Science* **2018**, *27*, 112–128.
- (84) Schoeler, C.; Bernardi, R. C.; Malinowska, K. H.; Durner, E.; Ott, W.; Bayer, E. A.; Schulten, K.; Nash, M. A.; Gaub, H. E. Mapping mechanical force propagation through biomolecular complexes. *Nano letters* **2015**, *15*, 7370–7376.
- (85) Bernardi, R. C.; Durner, E.; Schoeler, C.; Malinowska, K. H.; Carvalho, B. G.; Bayer, E. A.; Luthey-Schulten, Z.; Gaub, H. E.; Nash, M. A. Mechanisms of nanonewton mechanostability in a protein complex revealed by molecular dynamics simulations and single-molecule force spectroscopy. *Journal of the American Chemical Society* **2019**, *141*, 14752–14763.
- (86) Bauer, M. S.; Gruber, S.; Hausch, A.; Gomes, P. S.; Milles, L. F.; Nicolaus, T.; Schendel, L. C.; Navajas, P. L.; Procko, E.; Lietha, D.; others A tethered ligand assay to probe SARS-CoV-2: ACE2 interactions. *Proceedings of the National Academy of Sciences* **2022**, *119*, e2114397119.
- (87) Sedlak, S. M.; Schendel, L. C.; Gaub, H. E.; Bernardi, R. C. Streptavidin/biotin: Tethering geometry defines unbinding mechanics. *Science advances* **2020**, *6*, eaay5999.
- (88) Sedlak, S. M.; Schendel, L. C.; Melo, M. C.; Pippig, D. A.; Luthey-Schulten, Z.; Gaub, H. E.; Bernardi, R. C. Direction matters: Monovalent streptavidin/biotin complex under load. *Nano letters* **2018**, *19*, 3415–3421.

- (89) Roe, D. R.; Cheatham, T. E. PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *Journal of Chemical Theory and Computation* **2013**, *9*, 3084–3095.
- (90) Xia, B.; Vajda, S.; Kozakov, D. Accounting for pairwise distance restraints in FFT-based protein–protein docking. *Bioinformatics* **2016**, *32*, 3342–3344.
- (91) Kozakov, D.; Beglov, D.; Bohnuud, T.; Mottarella, S. E.; Xia, B.; Hall, D. R.; Vajda, S. How good is automated protein docking? *Proteins: Structure, Function, and Bioinformatics* **2013**, *81*, 2159–2166.
- (92) Kozakov, D.; Hall, D. R.; Xia, B.; Porter, K. A.; Padhorny, D.; Yueh, C.; Beglov, D.; Vajda, S. The ClusPro web server for protein–protein docking. *Nature Protocols* **2017**, *12*, 255–278.
- (93) Vajda, S.; Yueh, C.; Beglov, D.; Bohnuud, T.; Mottarella, S. E.; Xia, B.; Hall, D. R.; Kozakov, D. New additions to the ClusPro server motivated by CAPRI. *Proteins: Structure, Function, and Bioinformatics* **2017**, *85*, 435–444.

Figures

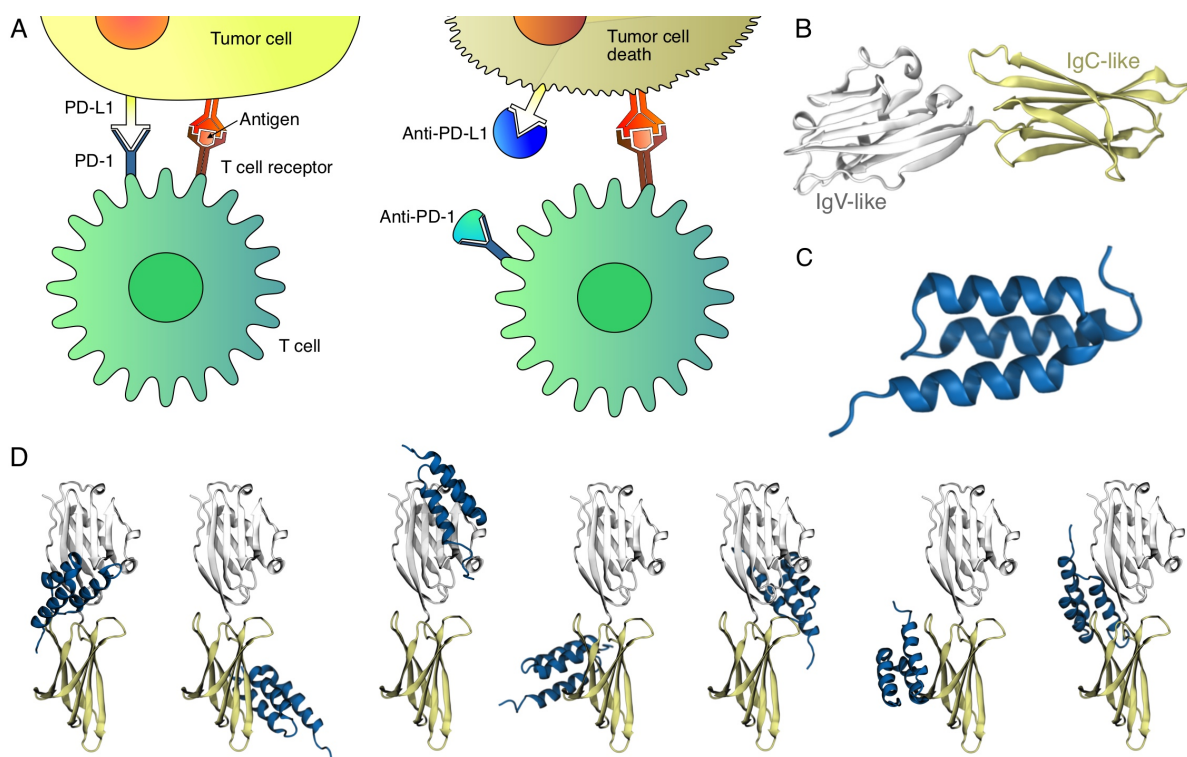


Figure 1: Overview of PD-L1 cancer targeting and lack of reliability of blind docking. A) (Left) By binding to PD-1, tumor cells overexpressing PD-L1 are able to inhibit T-cell activation, thereby avoiding clearance. (Right) Interfering with PD-L1 interactions using an immune checkpoint inhibitor (e.g., anti-(PD-L1) or anti-(PD-1) antibody) empowers T-cells to effectively eliminate tumor cells. B) Structure of PD-L1 (UniProt: Q9NZQ7) with its two domains: IgV-like (white) and IgC-like (yellow). C) Cartoon representation of the 3-helix bundle Affibody. D) Blind docking predictions for the complex using ClusPro 2.0 server^{44,90–93} show a diverse set well ranked solutions.

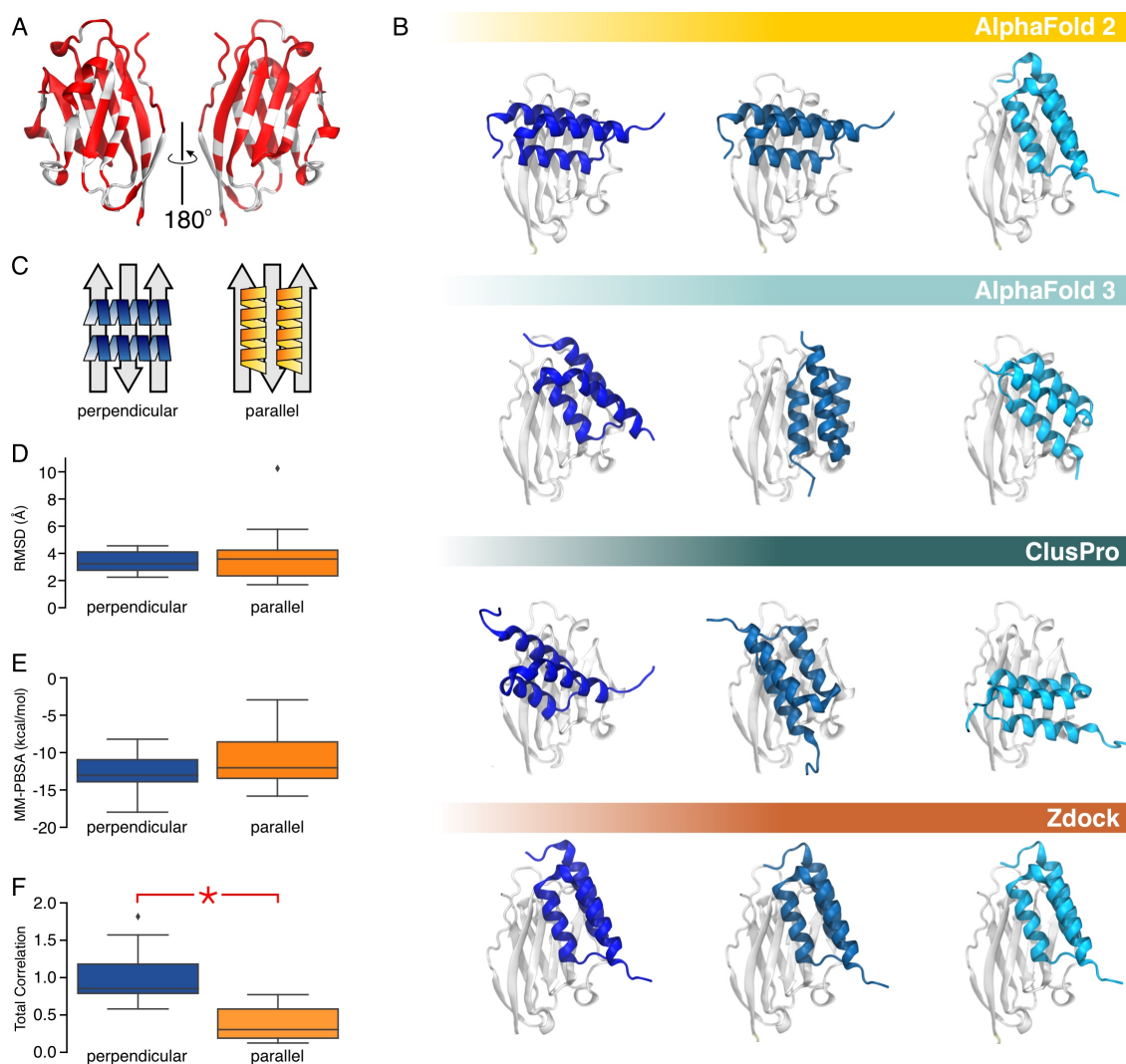


Figure 2: Results of structural predictions of the PD-L1:Affibody complex. A) Native contacts of PD-L1 complexes found in the PDB highlighted in red B) Comparative Analysis of PD-L1 and Affibody Binding Orientations: AlphaFold2 Multimer, AlphaFold3 ClusPro, and ZDock Predictions. The top 3 ranked structures from each prediction method depict the binding orientations of PD-L1 and Affibody. Two dominant orientations were consistent across all predictions: one parallel to the beta sheet and another perpendicular to it. This highlights the convergence of computational approaches in capturing putative binding modes of this complex. C) Illustration of the two proposed orientations. D-F) Boxplots for independent properties obtained from sixteen 100-nanosecond all atom MD simulation replicates of the parallel and perpendicular orientations. D) Affibody RMSD after fitting to PD-L1. E) Estimation of binding free energy using the Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) method. F) Sum of the correlation of interfacing residues normalized by the perpendicular average. The combined analysis of these MD-based metrics strongly supports the perpendicular binding mode between PD-L1 and Affibody as the most probable configuration, underscoring the robustness and consensus of computational assessment methods. The * indicates a significant difference between the perpendicular and parallel configurations in the total correlation, as determined by bootstrap analysis.

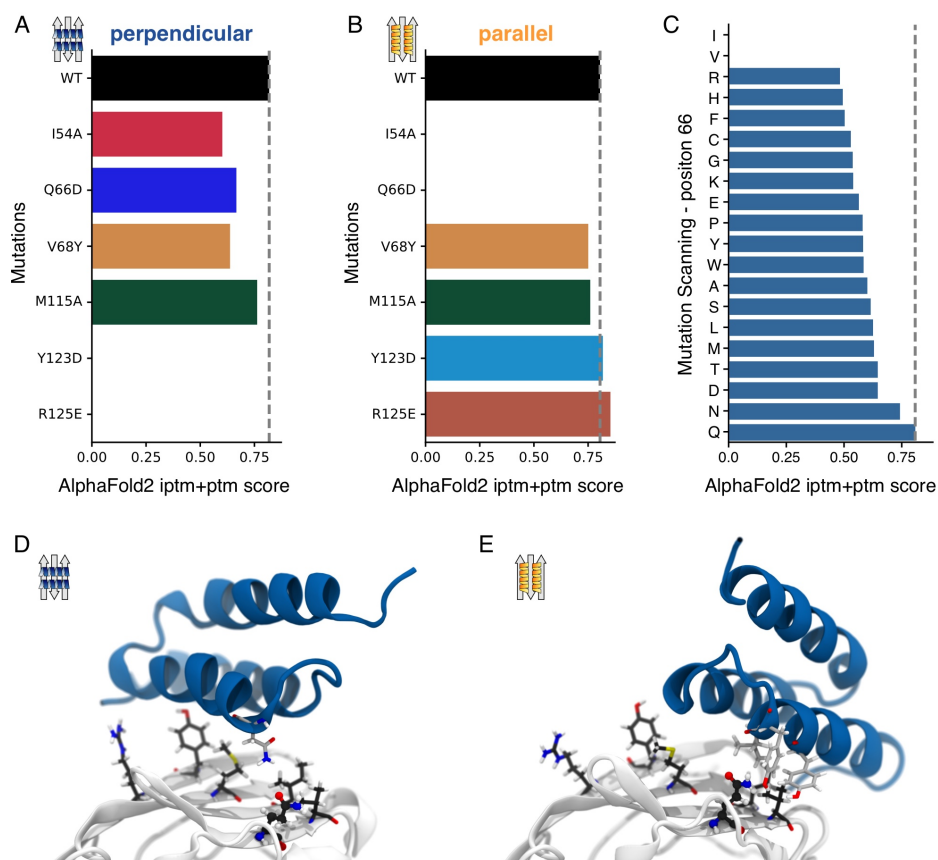


Figure 3: *In silico* site-specific mutagenesis to find PD-L1 residues involved in each of the binding modes. A-C) AlphaFold2's iptm+ptm score for the best ranking model of the complex. A-B) Screening of selected PD-L1 mutations within the predicted binding interface: I54A, Q66D, V68Y, M115A, Y123D, and R125E. C) Computational mutational scanning for PD-L1 at position 66 for the perpendicular mode. D-E) Wild type model for the two major orientations displaying position of residues selected for mutations. PD-L1 residues are showed as dark gray sticks while Affibody as silver sticks. Q66 is represented as ball-and-stick. For the perpendicular binding mode the nearest affibody residue to Q66 is N25, while for the parallel form Q66 is surrounded Y15, L18, and Y19.

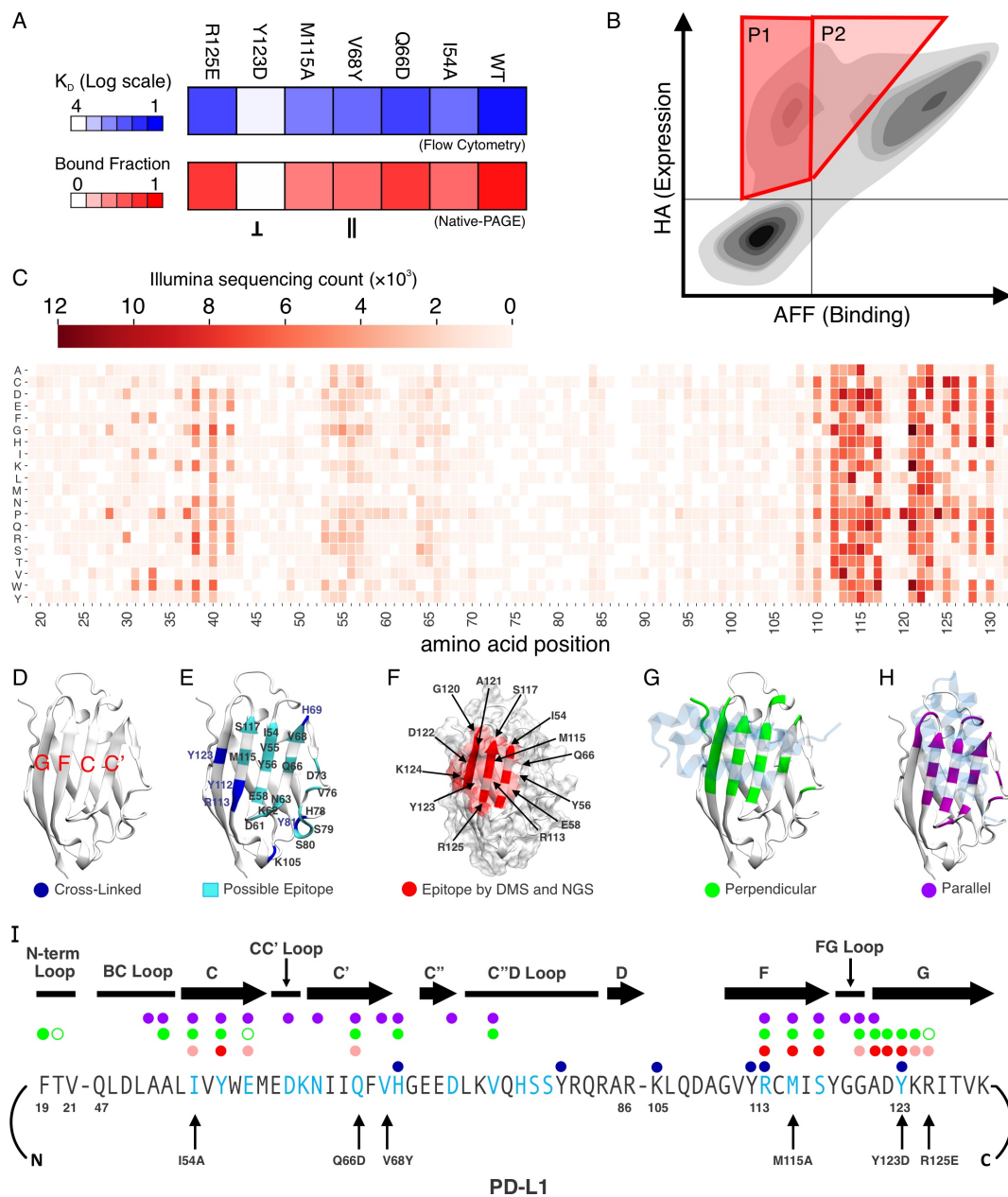


Figure 4: Experimental validation of binding epitope and perpendicular orientation. A) Site-specific PD-L1 mutagenesis and quantification of relative binding activity by flow cytometry and Native-PAGE. B) Cell labelling and sorting for strategy for epitope mapping by deep mutational scanning using a yeast display PD-L1 variant library. C) Heatmap of positions found among PD-L1 variants with reduced/no binding activity. D-G) Epitope mapping analysis results mapped onto the crystal structure of PD-L1. Labels are color coded to match the property analyzed, and indicated along the sequence in I. D) β -sheet labels. E) Binding epitope determined by XL-MS analysis (possible epitope in cyan, crosslinked residues in blue). F) Binding epitope determined by DMS and NGS (red). G) Analysis of binding orientation for the perpendicular orientation (green) H) Analysis of binding for the parallel orientation (violet). I) PD-L1 sequence with markings in corresponding colors to figures D-H.

TOC Graphic

