



OPEN

## Communication pathways bridge local and global conformations in an IgG4 antibody

Thomas Tarenzi<sup>1,2,3</sup>, Marta Rigoli<sup>1,2,3</sup> & Raffaello Potestio<sup>1,2</sup>✉

The affinity of an antibody for its antigen is primarily determined by the specific sequence and structural arrangement of the complementarity-determining regions (CDRs). Recent evidence, however, points toward a nontrivial relation between the CDR and distal sites: variations in the binding strengths have been observed upon mutating residues separated from the paratope by several nanometers, thus suggesting the existence of a communication network within antibodies, whose extension and relevance might be deeper than insofar expected. In this work, we test this hypothesis by means of molecular dynamics (MD) simulations of the IgG4 monoclonal antibody pembrolizumab, an approved drug that targets the programmed cell death protein 1 (PD-1). The molecule is simulated in both the apo and holo states, totalling 4  $\mu$ s of MD trajectory. The analysis of these simulations shows that the bound antibody explores a restricted range of conformations with respect to the apo one, and that the global conformation of the molecule correlates with that of the CDR. These results support the hypothesis that pembrolizumab features a multi-scale hierarchy of intertwined global and local conformational changes. The analysis pipeline developed in this work is general, and it can help shed further light on the mechanistic aspects of antibody function.

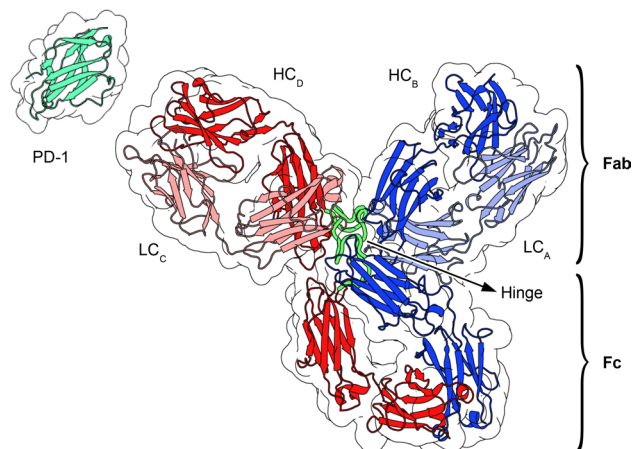
The number of monoclonal antibodies (mAbs) employed for therapeutic applications dramatically increased in the recent years: from 1997 to 2013, 34 mAbs-based pharmaceuticals were approved in US or EU, while from 2014 to 2020, in only 7 years, the number of approved mAbs was 61<sup>1</sup>. mAbs have been developed to treat a large variety of conditions, including cancer, autoimmune diseases and, very recently, COVID-19<sup>2,3</sup>.

Even in the case of naked immunoglobulin drugs, which do not involve conjugation with radionuclides or small molecules<sup>4,5</sup>, engineering of the antibody sequence is routinely performed to optimize its therapeutic efficacy for a given function, through successive steps of humanization, affinity maturation, and modifications aimed at overcoming challenges in stability and manufacturing<sup>6</sup>. On the one side, the selection of the isotype, and therefore those structural/dynamical features of the constant region that come with it, leads to different immune responses, and is thus performed on the basis of the planned application<sup>7,8</sup>; on the other side, modifications of single residues can determine a higher therapeutic efficacy, as in the case of those mutations introduced in the Fc domain to enhance effector function and recruitment of additional proteins<sup>9–11</sup>. A remarkable example is the single-residue mutation that, in the hinge of IgG4 antibodies, prevents Fab-arm exchange<sup>12–14</sup>.

Modifications of this type, which may be distributed throughout the whole antibody sequence, are usually introduced for reasons that are not directly linked to antigen affinity modulation. Engineering efforts intended to increase specificity and affinity are in fact mostly focused on the residues of the six loops comprising the complementarity-determining region (CDR), because of their preeminent role in antigen binding<sup>15</sup>. However, the CDR loops are not the only possible loci of intervention; it is experimentally shown that both mutations near and far from the antigen-binding site can affect affinity<sup>16</sup>. Such mutations act by modulating the interdomain conformational dynamics of the antigen-binding fragment, which eventually reflects on the paratope, namely the antigen binding site. On a similar note, NMR relaxation dispersion experiments allowed researchers to detect important fluctuating residues that are not located in the CDR, whose point mutation can nonetheless increase antigen-antibody affinity<sup>17</sup>.

Empirical experimental optimization of binding affinity can be laborious and costly, both in terms of time and resources<sup>18</sup>. Molecular dynamics (MD) simulations, on the other hand, offer a valuable tool for the investigation of the interplay between stabilizing interactions, fluctuation correlations, and conformational variability at different levels of resolution and experimental conditions<sup>19–22</sup>. In silico structural investigation of

<sup>1</sup>Department of Physics, University of Trento, Via Sommarive 14, 38123 Trento, Italy. <sup>2</sup>INFN-TIFPA, Trento Institute for Fundamental Physics and Applications, 38123 Trento, Italy. <sup>3</sup>These authors contributed equally: Thomas Tarenzi and Marta Rigoli. ✉email: raffaello.potestio@unitn.it



**Figure 1.** Graphical representation of the starting configuration of pembrolizumab employed in the MD simulations. LC stands for *light chain*, while HC stands for *heavy chain*. The antigen PD-1 is represented in the vicinity of the CDR.

immunoglobulins and antigen-antibody complexes has been successfully employed with different objectives, which include the detailed description of the dynamics of CDR loops and the transitions between their conformational states<sup>23–25</sup>, as well as the comprehension of structural rearrangements and allosteric modifications following antigen binding<sup>26–28</sup>.

Here, we employ atomistic MD simulations to investigate the internal dynamics of full-length pembrolizumab, a humanized IgG4 antibody used in immunotherapy, whose full structure has been experimentally solved<sup>29</sup> (Fig. 1). Pembrolizumab, whose commercial success rate is expected to profoundly impact pharmaceutical market in the next years<sup>30</sup>, is approved for the treatment of melanoma, lung cancer, head and neck cancer, Hodgkin's lymphoma and stomach cancer<sup>31–33</sup>. Its mechanism of action consists in binding to the programmed cell death protein 1 (PD-1), a 288-residues long receptor located on the membrane of T cells, B cells, and natural killer cells<sup>34</sup>. PD-1 promotes apoptosis of the lymphocyte when activated by the programmed cell death receptor ligands PD-L1 and PD-L2, whose expression is upregulated in malignant cells<sup>35</sup>. The large contact area between pembrolizumab and PD-1 hinders the binding of PD-L1 and PD-L2, thus preventing down-regulation of the anti-tumor activity of T cells<sup>36–38</sup>.

We performed a total of 4  $\mu$ s of dynamics of the deglycosylated antibody, both in presence and in absence of the antigen PD-1. The analysis pipeline, which combines structural analysis, investigation of chemical interactions, and information theory-based measures of correlation, allowed us to rationalize at a residue-level the observed conformational dynamics. Our simulations highlighted the particular role of key residues of the hinge, resulting in an asymmetric behavior of the two hinge segments in pembrolizumab; moreover, an interplay between large-scale conformation and binding state is observed, and the residues allowing such information routing from the paratope throughout the antibody structure are identified. We believe that the residue-level elucidation of the complex dynamics of potent and highly selective mAbs—of which pembrolizumab is a notable example—in relation to their binding state may assist in the process of antibody engineering, for the rational design of novel, optimized therapeutic agents.

## Results

Pembrolizumab is a 1324-residues long therapeutic antibody, belonging to the IgG4 class. On both hinge segments, it presents the typical S228P mutation, introduced to prevent fab-arm exchange in IgG4 antibodies<sup>12–14</sup>. A schematic representation of the protein is given in Fig. S3, where an index is associated to each structural domain. The four chains forming the antibody are labelled A, B, C and D, and correspond to chains A, B, F and G of the original PDB file, respectively<sup>29</sup>. We label Fab1 the Fab domain formed by chains A and B, and Fab2 the one including chains C and D; similarly, Fc1 includes domains CH2 and CH3 from chain B, while Fc2 includes those from chain D. Throughout the manuscript, all references to residue indices follow the EU numbering scheme<sup>39</sup>.

We focused on the analysis of conformations and the rationale behind domain motion in dependence of the binding state; to this aim, the results presented here are based on a comparison between the deglycosylated antibody alone (*apo* form) and bound to its antigen, the ectodomain of the protein PD-1 (*holo* form). The latter, comprising 114 residues, corresponds to the region of PD-1 whose structure has been experimentally solved in complex with pembrolizumab Fab<sup>40</sup>. To allow for an unbiased comparison, the starting structure of the holo system was built on the same initial antibody conformation of the *apo* case; specifically, Fab2 of pembrolizumab (formed by chains C and D) was replaced with the Fab-PD1 complex (PDB ID: 5GGS), after structural alignment on the antibody domain.

Both *apo* and *holo* systems were simulated in four independent replicas, with a duration of 500 ns each. Relaxation of the starting structure to an equilibrated conformation was assessed by computing the root-mean-square deviation (RMSD) of  $C_{\alpha}$  atoms of the full antibody (Figs. S4 and S5); in addition, the RMSD was computed

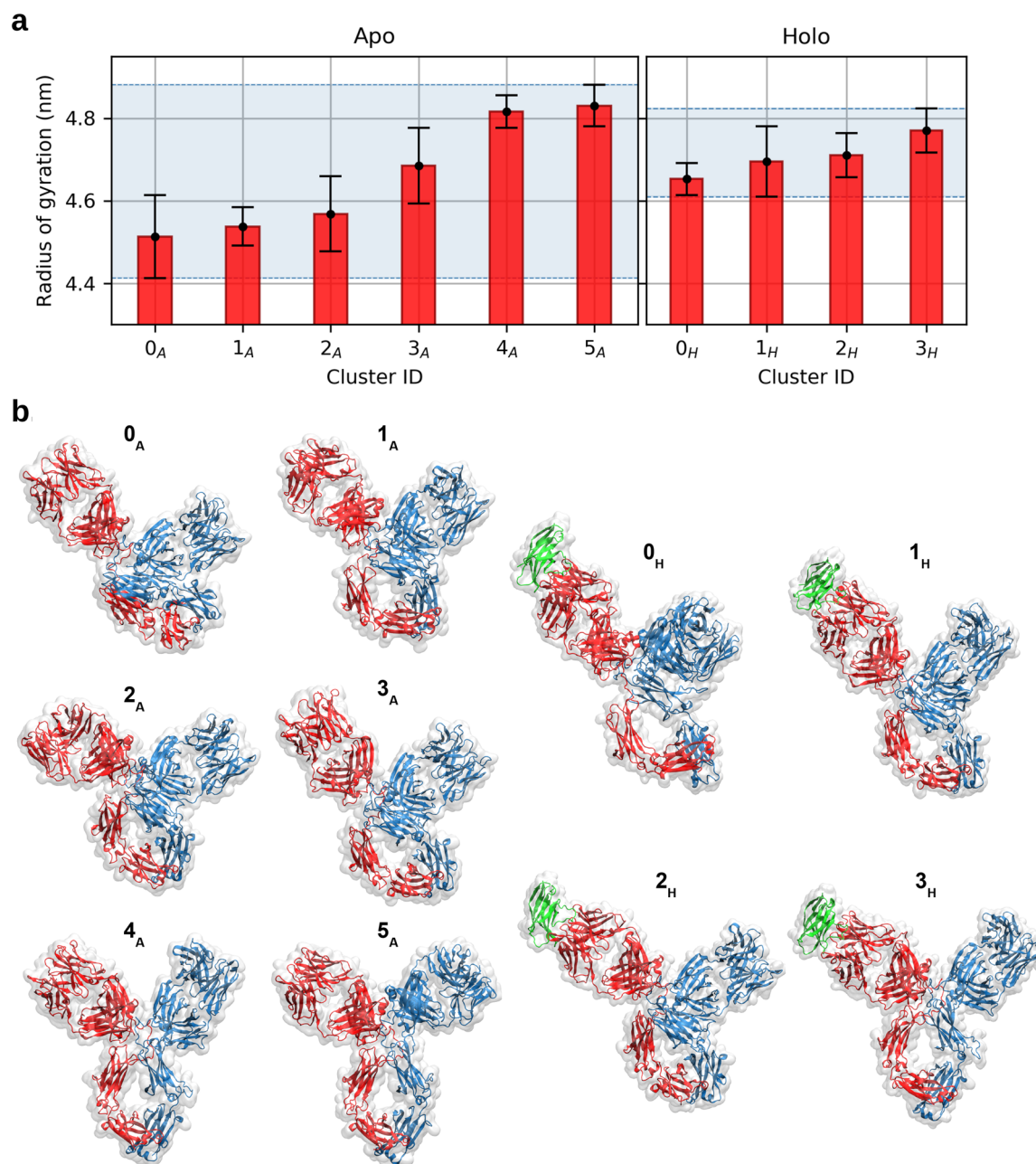
for the of  $C_\alpha$  atoms of Fab1, Fab2 and Fc independently, to verify the relaxation of each structural domain from the crystallographic structure (Figs. S6 and S7).

**The presence of the bound antigen restricts the range of conformations of the full-length pembrolizumab.** To facilitate the analysis of antibody flexibility, the conformations sampled from MD simulations are collected in clusters on the basis of their structural similarity (Fig. S8), as measured by the RMSD matrix of  $C_\alpha$  atoms. The cluster analysis is performed separately and independently on the two systems, in order to highlight possible differences; in both cases, an RMSD threshold of 1.2 nm is used. In the simulations of pembrolizumab alone, 6 conformational clusters are identified, while in the case of the holo system the same clustering protocol leads to the identification of only 4 clusters; in the presence of the antigen, a limited conformational variability is indeed observed, as apparent from the narrow range of radii of gyration spanned by these clusters when compared to the apo case (Fig. 2). This is consistent with the visual inspection of the sampled conformations, whose representative structures are reported in Fig. 2. A similar population shift towards a more uniform conformational distribution upon antigen binding was observed through MD simulations of an IgG1 antibody<sup>27</sup>.

For each system (apo or holo), clusters are indexed according to increasing radii of gyration; in this way, clusters  $0_A$  and  $0_H$  are the ones grouping the most compact conformations of the apo and holo systems, respectively. At a large scale, conformational clusters differ mainly in the relative arrangement of the Fab and Fc domains, reflecting the extent of packing. Clusters  $3_A$  and  $2_H$  include the equilibrated experimental structures of the apo and holo systems, respectively. These clusters are intermediate, in terms of radius of gyration, between clusters  $0_A/1_A/2_A$  and  $4_A/5_A$  for apo, and  $0_H/1_H$  and  $3_H$  for holo; this means that the simulations captured both tendencies of the antibody to shrink and to expand. Interestingly enough, the conformations that are not sampled by the holo system correspond to particularly low radii of gyration, namely those found in clusters  $0_A$ ,  $1_A$ , and  $2_A$ . Compact conformations, particularly clusters  $0_A$ ,  $2_A$  and  $0_H$ , present a higher structural stability, as measured in terms of RMSD with respect to the representative structure of the cluster (Fig. S9).

The relative cluster population (quantified by the fraction of frames relative to the total number, Fig. S10), shows a large unbalance among conformational states. Despite the different population, each cluster nonetheless includes a sufficient sampling of the conformational basin, as proved by the convergence of the root-mean-square fluctuations (RMSF) of the antibody  $C_\alpha$  atoms (Figs. S11 and S12). Comparison of the cluster sizes reveals that the unbound antibody evolves towards a compact conformation for most of the simulation time, while the holo has a more symmetric distribution of configurations, associated to larger and smaller radii of gyration compared to the initial structure. Also in the holo case, however, the time spent in open conformations is a small fraction of the full simulation time. This result is in agreement with the experimental observation that the compact conformation is the most populated for IgG4 molecules in solution, as observed in <sup>15</sup>N TROSY NMR and SAXS experiments<sup>29,41</sup>. This conformational preference has been explained on the basis of the short hinge of IgG4 antibodies with respect to other IgG subclasses; IgG4 hinge has indeed a three amino acid deletion, compared to the one of IgG1<sup>42</sup>. Another factor that can favour the compact arrangement is the tilted conformation of the CH2 domain in the Fc<sup>29</sup>, which displays a rigid-body rotation of  $\approx 120^\circ$  relative to the position observed in the crystal structure of a truncated IgG4 Fc (PDB ID: 4C54). In the most compact conformations, the superposition of the pembrolizumab Fc and the one of 4C54 would lead to an overlap of CH2 and CL in Fab1 (Fig. S13). The peculiar CH2 conformation of pembrolizumab, which is present in the crystal structure, persists during the course of the simulations. The hypothesis that such CH2 domain rotation is found in solution was supported by a recent experimental observation of the reduced cleavage of pembrolizumab, mediated by the immunoglobulin-degrading enzyme from *Streptococcus pyogenes*, which binds the CH2–CH3 interface<sup>43</sup>. The asymmetric orientation of the two CH2 domains can also explain the contact area between Fabs and Fc, reported in Fig. S14. Here, the range of variability of the Fab1-Fc contact area greatly exceeds that of the Fab2-Fc pair, which remains low in the vast majority of the clusters; moreover, a significantly larger contact surface between Fab1 and Fc is observed in clusters  $0_A$  and  $0_H$ , with respect to the other cases. The rotated CH2 forms indeed a cavity that gives Fab1 a greater freedom to move, minimizing steric clashes with Fc. This is reflected also in clusters  $4_A$ ,  $5_A$  and  $3_H$ , where the symmetric arrangement of Fabs with respect to Fc leads to a slightly larger Fab2-Fc contact area than Fab1-Fc; and in cluster  $2_A$ ,  $3_A$  and  $1_H$ , where, although the Fc is clearly tilted toward the Fab1, the Fab1-Fc contact area is only slightly larger than the Fab2-Fc case. The large possibility of rearrangements of Fab1 with respect to the Fc is therefore one of the main determinants of the overall shape of pembrolizumab. We do not rule out the possibility of a conformational change of the CH2 domain in chain B, leading either to a symmetric FC conformation, or to a switching of the conformational properties of chains B and D; however, given the short hinge of pembrolizumab, we expect that prior breaking of the disulfide bonds in the hinge is necessary for the CH2 rotation to take place. Such an event, in turn, is unlikely to happen even in vitro or in vivo, given the stabilization of the disulfide bonds brought by the S228P mutation<sup>13</sup>.

A comparison between the conformational states in apo/holo clusters is facilitated by the RMSD matrix in Fig. S15. As expected, cluster  $3_H$  shares a low RMSD with the open conformations of clusters  $4_A$  and  $5_A$ , while clusters  $1_H$  and  $2_H$  are closer to the compact conformations in clusters  $1_A$ ,  $2_A$ ,  $3_A$ . The extremely compact conformation of cluster  $0_A$  and the peculiar cluster  $0_H$  are the ones sharing the lowest similarity with all the other clusters.  $0_H$  appears indeed as a singular conformation, where the major axis of Fab1 is perpendicular to the plane of Fab2 and Fc; once again, this movement is permitted by the rotated CH2 domain. Although such bent conformation of Fab1 is not directly observed in the apo case, visual inspection of the trajectory reveals that the variable region of Fab 1 in cluster  $4_A$  has the tendency to bend in a direction perpendicular to the plane of Fc-Fab2. This is confirmed by a principal components analysis (PCA) of the dynamics, performed on the  $C_\alpha$  atoms of the structures grouped in the conformational clusters (Section S2.1, Figs. S16–S35). The entity of the conformational change, however, is larger in  $0_H$  than in any apo cluster.



**Figure 2.** (a) Mean and standard deviation of the antibody radii of gyration in the apo forms (*left*) and holo forms (*right*), averaged within each conformational cluster. In the holo case, the radius is computed for the antibody alone. The error bars correspond to the standard deviation within each cluster. The shaded areas correspond to the range of variability, which is greatly reduced in the holo case. (b) Representative structures of pembrolizumab in the apo (A) and holo (H) forms, for each conformational cluster. Chains AB are in blue, chains CD in red, and the antigen PD-1 in green.

**Role of key residues in stabilizing the conformational states.** The determinants of the observed conformations are identified, at a finer level of detail, through the calculation of correlations between residues, as quantified by the mutual information (MI) of  $C_{\alpha}$  atom fluctuations. MI, which captures both linear and non-linear contributions to amino acid displacements from a reference position, is used to build inter-residue networks of information pathways, which can be interpreted in the light of the inter-domain non-bonded interactions established in the course of the simulations. Tables S3 and S4 list the residues involved in inter-domain contacts, including hydrogen bonds, salt bridges, and hydrophobic interactions. These contacts are highlighted as important channels for information transfer by the results of the network analysis, and in particular by the calculation of edge betweenness (Figs. S36 and S37), which measures the centrality of a graph edge as the number of shortest paths crossing it<sup>44</sup>.

The most open conformations (clusters  $4_A$  and  $5_A$ ) are stabilized by a persistent electrostatic interaction between side chains of ASP<sup>126</sup>, located in the CL domain of Fab2, and LYS<sup>221</sup>, located between CH1 of Fab1 and the hinge, which results in a highly central edge of the interaction network; we can expect that mutation of one of these residues would destabilize the open state. Despite this common feature, stable contacts between Fab1 and Fc are absent in cluster  $5_A$ , while in  $4_A$  the two domains interact through electrostatic interactions. Moreover, the comparison of the non-bonded interactions between Fab1/Fab2 and Fc confirms the asymmetry of the contact distributions in the extended conformations: even though the molecule adopts an overall symmetric Y-shape, the larger number of interactions between Fc and Fab2 (with respect to Fab1) in clusters  $4_A$ ,  $5_A$  and  $3_H$  supports the hypothesis that further approaching of the two domains would be impaired by steric clashes.

In the most compact conformations (clusters  $0_A$ ,  $1_A$  and  $0_H$ ), a key stabilizing role is played by ARG<sup>217</sup> in the CH1 domain of Fab1, which emerges as a central network node and is involved in electrostatic interactions with several Fab2 residues (ASP<sup>126</sup> and LYS<sup>130</sup> in particular, located in the CL domain). The network representations of closed conformations reveal also the large number of edges crossing the Fab1-Fc contact surface in the compact arrangements. While the residues involved in these key interactions belong mainly to the CH2 domains of chain B, in the case of cluster  $2_A$  the surface of interaction is extended to distant regions of the Fc domain, namely the CH3 domain of chain B (through the backbone of MET<sup>428</sup> and the side-chain of HIS<sup>429</sup>) and the CH2 domain of chain D (through the side-chains of THR<sup>335</sup> and LYS<sup>334</sup>). The peculiar conformations in cluster  $0_H$  are stabilized by a number of high betweenness contacts taking place outside of the hinge; fundamental for the stability of this compact and highly interconnected state are the electrostatic interactions between side chains of LYS<sup>194</sup>-GLU<sup>294</sup>, in the CL of Fab2 and in the CH2 of chain B, respectively, and between LYS<sup>246</sup>/ARG<sup>255</sup> in the AB loop of chain B CH2<sup>45</sup> and light-chain residues of Fab1 (specifically, hydrogen bonds LYS<sup>246</sup>-ASP<sup>174</sup>, ARG<sup>255</sup>-GLU<sup>17</sup>, and ARG<sup>255</sup>-PRO<sup>15</sup>).

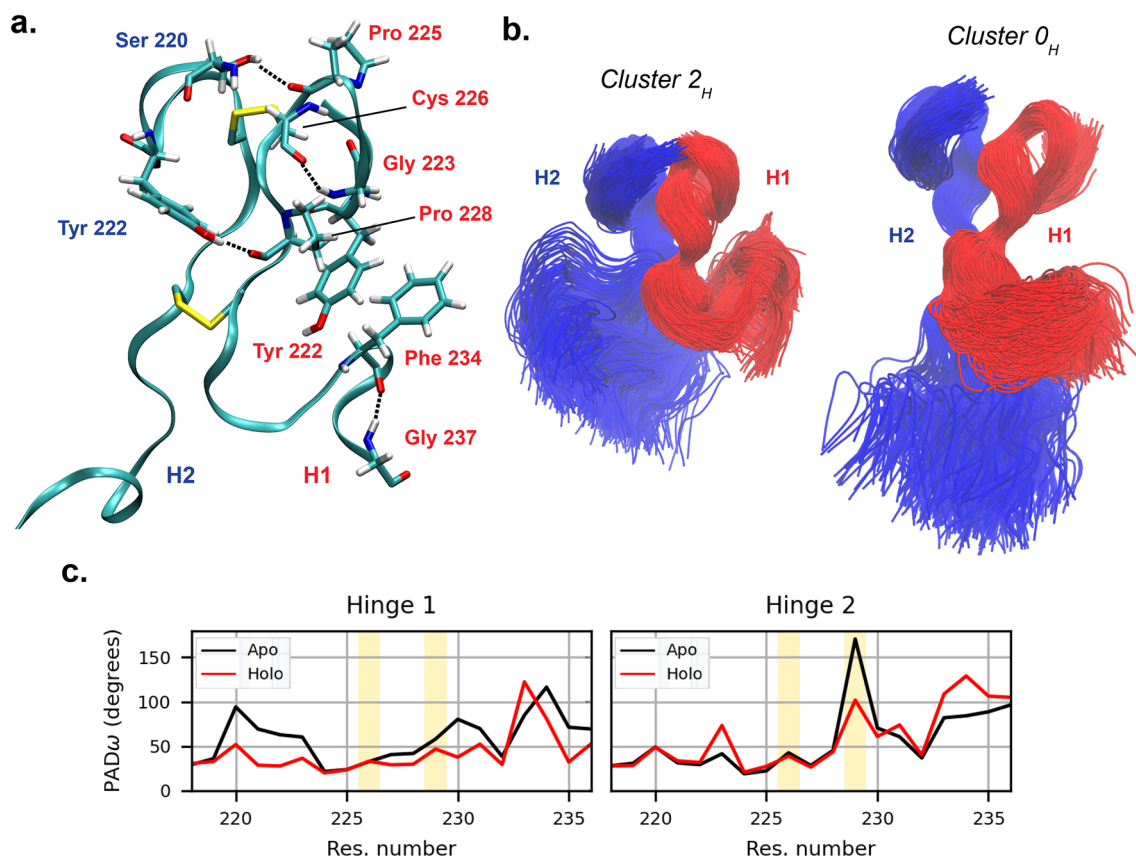
These observations are strengthened by the analysis of communities in the MI-based network. The latter is divided into substructures (communities) with dense internal correlations, but sparse inter-community connections (see Section S1). Although the optimal community distribution closely reflects the natural subdivision of the antibody in structural domains, residues previously identified as promoting inter-domain connections fall within the same community (Figs. S38 and S39), thus supporting results from the investigations of non-bonded interactions and edge centrality. In addition, community analysis serves to detect, at a first level, connections between the hinge and the neighbouring domains; in this regard, a significant consistency is observed in the intermediate and fully open conformations of both apo and holo cases, where the hinge belongs to the same community as the CH1 domain of Fab2. Building on the high level of connection between the hinge and the rest of the molecule, a more detailed investigation of the role of the hinge in the overall conformational variability of the antibody is explored in the next section.

**Role of the hinge in the observed conformational variability.** As noted above, the hinge displays significant correlations with other domains, both in the apo and holo systems, as apparent from the MI matrices (Figs. S40 and S41); such correlations appear particularly strong in pembrolizumab when compared to those reported in a previous study of an IgG1 antibody<sup>46</sup>. In pembrolizumab, in fact, the hinge retains a complex network of interactions with nearby domains, thus functioning as more than just a flexible linker; a detailed account on the interactions involving the hinge residues, as emerging from network analysis and inspection of non-bonded contacts, is reported in Section S2.2. Given its role as interaction hub, in addition to allowing large-scale rearrangements<sup>47,48</sup>, the hinge region deserves therefore particular attention.

During our simulations, the hinge shows relatively small variations in the radius of gyration (Fig. S42). Such result is expected on the basis of the presence of disulfide bonds between the two hinge segments, and of a number of transient intra- and inter-chain non-bonded interactions; among them, the most stable ones are the hydrogen bonds between the sidechain of SER<sup>220</sup> in hinge 2 and the backbone of PRO<sup>225</sup> in hinge 1, and between the sidechain of TYR<sup>222</sup> in hinge 2 and the backbone of PRO<sup>228</sup> in hinge 1 (Fig. 3).

The two hinge segments from chain B and chain D are highly asymmetric (Fig. S43), with hinge 1 assuming significantly bent conformations (Fig. 3). Although they are mostly attributable to the numerous non-bonded Fab1-Fc interactions, which prevent an extended conformation of chain B (see previous section), they are also stabilized by a number of non-bonded interactions, such as the stacking interaction between the side-chains of TYR<sup>222</sup> and PHE<sup>234</sup> and the hydrogen bonds between PHE<sup>234</sup>-GLY<sup>237</sup> and GLY<sup>223</sup>-CYS<sup>226</sup>. Hinge 2 shows instead the largest degree of variability, and it appears to be the main determinant of the overall hinge shape.

Results from the previous sections highlighted a smaller range of conformations of the pembrolizumab molecule in presence of the antigen. In order to have a closer look at the residues responsible for the increased rigidity in the holo structures, we computed the PAD<sub>ω</sub> (protein angular dispersion) parameter of the angle ω, the latter being the sum of the backbone torsional angles Φ and Ψ<sup>49</sup>. PAD<sub>ω</sub> measures the per-residue backbone plasticity through the variance of its torsional angles (Fig. 3). It ranges from 0 to 180°; a higher value corresponds to a higher backbone flexibility. PAD<sub>ω</sub> values of hinge 2 (connecting Fab2, where the antigen is bound, and Fc) are very similar in the apo and holo cases, and for a few residues are even slightly higher in the bound state (the average PAD<sub>ω</sub><sup>H2</sup> is 60° for apo and 62° for bound). However, hinge 1 (connecting Fab1 and Fc) is, on average, more rigid in the holo than in the apo form (the average PAD<sub>ω</sub><sup>H1</sup> is 61° for apo and 44° for holo). Since the flexibility of hinge 1 is responsible for the relative movements of Fab1 and Fc, the calculated PAD<sub>ω</sub> values are in line with the restricted conformational variability observed in the bound conformations. RMSF computed on the hinge C<sup>α</sup> atoms confirms the reduction of flexibility in hinge 1 with respect to hinge 2 when going from the apo to the bound state (Fig. S44). Therefore, we suggest that the limited conformational variability of pembrolizumab in presence of the bound antigen is linked to a higher rigidity in the hinge region with respect to the apo case. Moreover, hinge 1 shares a higher MI with domains CH1 and CL of Fab2 than hinge 2 (Figs. S40 and S41), despite

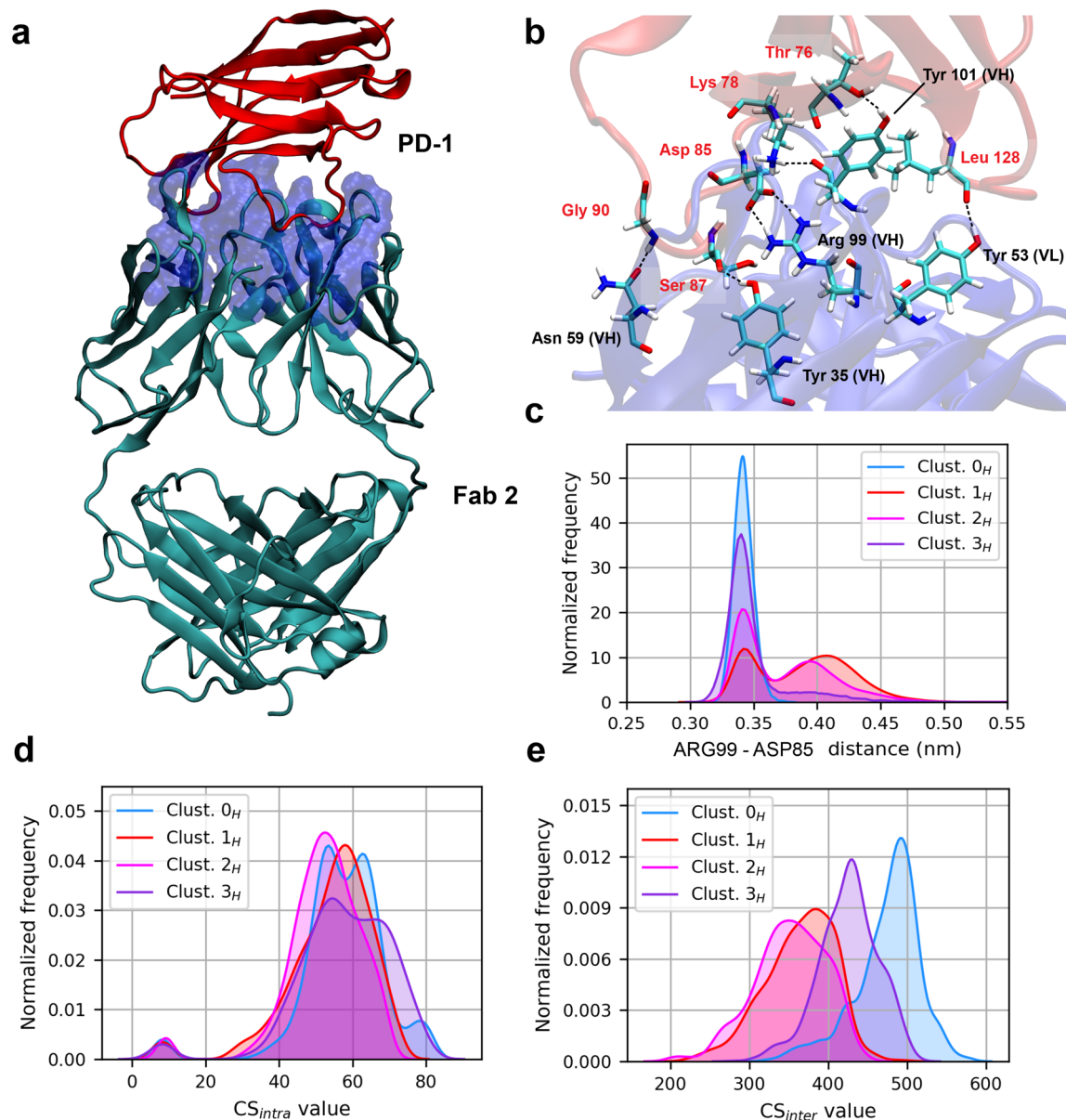


**Figure 3.** (a) Interaction stabilizing the hinge conformation; residues belonging to hinge 1 are labelled in red, while those belonging to hinge 2 are in blue. The disulfide bonds between the two CYS<sup>226</sup> of chains B and D, and the CYS<sup>229</sup> of chains B and D, are also shown. (b) Conformational ensembles of the hinge in its most compact and most extended conformations (clusters 2<sub>H</sub> and 0<sub>H</sub> respectively), after structural alignment on the cysteine residues. Chain B is represented in red, chain D in blue. The overall hinge shape is determined by the conformation of H2. (c) Per-residue flexibility of hinge backbone, as quantified by PAD<sub>ω</sub> parameter. The yellow-shaded areas correspond to the cysteine residues forming inter-chain disulfide bonds. Particularly high is the backbone plasticity of the second cysteine in the hinge 2 of the apo antibody, which might result from the torsional stress imposed by the high conformational variability of the Fc relative to the Fab in the apo simulations.

the latter being the direct extension of Fab2. Results from network analysis corroborate this observation: hinge 2 in the holo system does not include residues with high centrality, as opposed to the apo case (Tables S5 and S6).

The hinge is not the only region of the holo antibody where a decrease in backbone plasticity is observed, with respect to the apo case; from the PAD<sub>ω</sub> analysis, this is true also for the binding site of Fab2 (as expected for the presence of the antigen) and residues 242–252 and 338 in the CH2 domain of chain B (Fig. S45). Interestingly enough, this domain is indeed involved in high-betweenness contacts, as evidenced by the network analysis (Figs. S36 and S37), through highly central paths that put it in communication with Fab2. We notice here that a concurrent reduction in flexibility in the binding site, in hinge 1, and in the CH2 domain is not in contradiction with the observation that changes of conformational flexibility in antibodies follow Le Chatelier's principle<sup>50,51</sup>: upon binding, counteracting changes in rigidity and flexibility occur at distant sites. Figure S45 reveals indeed that a reduction in the value of PAD<sub>ω</sub> in the paratope of Fab2 is associated to an increased backbone plasticity in distant regions of the holo system, especially in the chain D of the Fc domain (residues 252–255, 271, and 323–330).

**Binding modes of PD-1 elicit different degrees of correlation within the antibody.** The results from the previous section suggest the presence of a communication channel between the antigen-bound Fab2 and hinge 1, leading to a transfer of information that might modulate an interplay between binding and conformational state. In this regard, different conformational clusters correspond to different binding site conformations, as quantified by the distributions of the RMSD between C<sub>α</sub> atoms in the paratope of the simulated system and the bound crystal structure<sup>40</sup> (Fig. S46). At a first level of distinction, the presence of the antigen largely reduces the displacements in the binding site, as expected; moreover, in the apo case, a correspondence is observed between the entity of the displacements and the compactness of the conformation, with more compact clusters shifted toward larger values of RMSD. On the other hand, the displacements of the binding site in Fab1



**Figure 4.** (a) Pembrolizumab Fab2 (cyan) in complex with PD-1 protein (red). The paratope is represented as a blue surface. (b) Residues involved in electrostatic interactions at the antigen/antibody interface; antigen residues are labelled in red. (c) Distance distributions between ARG99 in VH and ASP85 of PD-1. (d) Distributions of the intra- and inter- (e) domain correlation coefficients of antibody residues, for each conformational cluster of the holo state.

are not affected by the presence of PD-1 on Fab2; their RMSD distributions are similar in the apo and holo case (Fig. S47).

In order to better understand the interplay between the binding site on Fab2 and the rest of the pembrolizumab structure, we performed an investigation of the stabilizing interactions between the antibody and the antigen in each holo cluster, followed by a study of the correlations elicited by the ligand within the antibody.

The complex formed by Fab2 and the protein PD-1 is shown in Fig. 4. The number of hydrogen bonds between the two molecules varies slightly among the distinct conformational clusters (Fig. S48); large deviations are instead observed in the persistence of the salt bridge formed between side-chains of ARG<sup>99</sup> in the VH domain of chain D and ASP<sup>85</sup> of PD-1 (Fig. 4), which has been identified as a key contact for the stabilization of the complex<sup>40</sup>. For all electrostatic interactions, the binding is strengthened when the antibody is in cluster 0<sub>H</sub>, resulting in large and stable values of the contact surface area between pembrolizumab and PD-1 (Fig. S48). The RMSF of C<sub>α</sub> atoms in the antigen molecule, after structural alignment of the antibody variable region, reflects the strength of the interactions, with fluctuations in 0<sub>H</sub> that are approximately the half of those in the other clusters (Fig. S49). Furthermore, the distribution of the antigen RMSD in each cluster, calculated with respect to the experimental structure, reveals a lower dispersion in cluster 0<sub>H</sub> (Fig. S49). This particular stability might be correlated to the rigidity of the antibody structure in this conformation (Fig. S9), and the corresponding

absence of large-scale conformational changes perturbing the Fab/antigen complex. These observations are in agreement with the calculated values of binding enthalpies, obtained with the MM/PBSA approach<sup>52</sup> and reported in Table S7. The values shown are relative to the simulation with the sole Fab in complex with the antigen, taken as the reference. In addition, for all of the conformational clusters, the simulation of the full-structure antibody results in a more stabilizing energy than in the simulation of the Fab alone, thus suggesting the importance of maintaining the full antibody structure in the studies of antibody/antigen binding free energies through MD simulation; a similar stabilizing effect was already observed upon inclusion in the simulation set-up of the constant region, with respect to the sole variable region<sup>53</sup>.

An ordering of conformational clusters similar to the one emerging from the strength of the binding interaction is reflected in the intensity of the residue-residue correlations within the antibody, as quantified by the generalized correlation coefficient (GCC)<sup>54</sup>; the latter corresponds to a normalized form of MI (Section S1), ranging from 0 (no correlation) to 1 (perfect correlation). Distributions of GCC values are shifted toward higher values in those clusters with the tightest binding (Fig. S50), particularly in cluster  $0_H$ . For a more detailed inspection, we followed the example set in Palermo et al.<sup>55</sup>, and we computed the correlation score (CS) for each residue  $i$  as the sum of GCC values with all the other antibody residues. This allowed us to distinguish *intra*-domain scores (Fig. S51), where the summation extends to the residues belonging to the same structural domain, and *inter*-domain scores (Fig. S52), which take into account only residues belonging to the other structural domains, excluding the one that includes residue  $i$ . From the distributions of intra- and inter-CS values shown in Fig. 4, it is apparent that the change in correlations among the different clusters stems from an increased inter-correlation between the antibody domains in clusters  $0_H$  and, to a smaller extent, in cluster  $3_H$ . On the opposite, the intra-correlations do not show significant variations. In cluster  $0_H$ , in particular, the inter-correlations are significantly higher in Fab2 and in the chain B of Fc, with respect to the other clusters (Fig. S52). This result can be explained on the basis of the tight binding between antigen and antibody in cluster  $0_H$ , and of its largely distributed communication network, as highlighted above.

Mutual information shows the highest correlation between the antigen molecule and all the antibody domains in cluster  $0_H$  (Fig. S53), especially in Fab1. To further investigate the interplay between paratope and such peculiar antibody conformation, simulations of pembrolizumab in the apo form were started from the representative structure of cluster  $0_H$ , after removal of PD-1. Three 100ns-long replicas were performed to allow residues in the paratope to relax to new equilibrium conformations, while the antibody retains an overall conformation closely similar to the starting one. As shown in Fig. S54, the RMSD distribution of the residues in the binding site overlaps with the one obtained from the apo cluster with the highest structural similarities, namely cluster  $1_A$  (Fig. S15); in addition to a large network of interactions, as observed in the case of the other compact apo clusters, the latter is characterized by particularly strong contacts between the Fab domains (Table S3). A similar transition of the binding site conformation is not obtained in the case of the other holo clusters, thus suggesting once again a complex interplay between the large-scale conformation of the antibody and the antigen binding site, as elicited by the highly extended correlation network of pembrolizumab in conformational states compatible with cluster  $0_H$ .

## Discussion and conclusions

While impressive progress has been made in the experimental characterisation and manipulation of antibodies, a detailed, atomistic investigation of their properties is still incomplete. This is particularly true for the interplay between the molecule structure and its dynamics, which is extremely rich and varied, as several studies have recently shown<sup>27,56–58</sup>.

If, on the one hand, the specific sequence of the CDR plays the most prominent role in the selectivity and binding affinity of the antibody, on the other hand the observation that a modulation of the binding strength can be effected though mutations in distal sites shows that the internal mechanics of these molecules can be extremely complex<sup>16,17</sup>. This is especially the case for IgG4 antibodies such as pembrolizumab, whose short hinge reduces the degree of flexibility and tightens the interactions among the various domains.

In this work, we have made use of atomistic MD simulations and information-theoretical analysis methods to elucidate the relation between the large-scale arrangement of pembrolizumab and the stability of the antigen binding. The analysis pipeline employed has allowed us to highlight a substantial conformational variability, a quality rather different from that one might observe in the case of loosely-connected rigid bodies. On the contrary, a complex pattern of structural arrangements, intramolecular communication pathways, and binding strength has emerged.

In accordance with experimental results, we observed that the antibody is prevalently found in a compact, asymmetric shape. This particular arrangement can be explained on the basis of the short hinge, which limits the conformational space accessible to Fab arms, as well as the rotated CH2 domain of chain B. Furthermore, the spectrum of structures sampled by the molecule is modulated by the binding state: the large conformational variability of the apo case is substantially restricted when the antibody is complexed with the antigen. Similarly, the flexibility of the hinge (hinge 1 in particular) is reduced in this latter case, with a high degree of correlation emerging between the binding site on Fab2 and hinge 1. The binding site was shown to bear strict ties with the rest of the molecule in general. The analysis of the distributions of the apo binding site RMSD, computed taking the experimentally resolved holo structure as a reference, highlighted a correlation between the binding site and the large-scale arrangement of the molecule as a whole.

The binding of the PD1 antigen to Fab2 is, in general, rather stable in all holo state simulations; however, the strength of the binding is different across the various conformational clusters. The trend follows the intensity of the inter-domain correlations not only within the antibody, but also between the antigen and the antibody itself; a particularly strong affinity is observed in a specific configuration in which Fab1 is bent toward Fab2 and Fc.



The tight binding and high correlations in this cluster suggest an interplay between the binding mode and the intensity of correlations, which manifest themselves in an extended network of interactions.

The results reported in this work return a picture of antibodies as extremely complex molecules, with a rich pattern of structural and dynamical features. The analysis protocol here applied to pembrolizumab is completely general, which enables its widespread application to other antibodies, with the objective of acquiring an ever deeper understanding of their inner life and, in turn, providing sharper tools for the manipulation of these molecules for medical and technological purposes.

## Methods

**System setup.** The crystallographic structure of pembrolizumab in the apo form (Ab1, PDB ID: 5DK3)<sup>29</sup> was used as a starting structure for the MD simulations. Missing residues (positions 230–232 of chain B and 230–235 of chain D, located in the hinge region) were modelled with Chimera<sup>59</sup>, using an IgG2a antibody (PDB ID: 1IGT<sup>60</sup>) as a template. Glycans were removed from the crystal structure, so as to isolate the role of structural features of the antibody itself in the emerging dynamics. The system in the holo form (Ab2) was obtained by replacing one Fab in Ab1 with the crystallized structure of Fab-PD1 complex (PDB ID: 5GGS)<sup>40</sup>, after structural alignment. The Fab-fragment alone bound to the PD-1 (Ab3) was also simulated, in order to assess possible differences in the interaction with the antigen with respect to the full-length antibody.

**Simulation protocol.** MD simulations were performed with the Gromacs 2018 software<sup>61,62</sup>. The amber99SB-ILDN force field<sup>63</sup> was used to define the topology of the systems. The protein was solvated in a box of TIP3P water molecules<sup>64</sup> and the charge was neutralized with Cl<sup>−</sup> and Na<sup>+</sup> ions at physiological concentration (150 mM). The final number of atoms was 571932 in Ab1, 618265 in Ab2, and 193009 in Ab3; of these, the number of protein atoms was 20218, 21976 and 8317, respectively. Energy minimization was performed until the maximum force reached the value of 500 kJ mol<sup>−1</sup> nm<sup>−1</sup>. NVT and NPT equilibrations were performed for 100 ps each using the Velocity-rescale thermostat<sup>65</sup> and the Parrinello-Rahman barostat<sup>66</sup>. Temperature was set at 300 K and pressure at 1 bar, and the time constants were set to  $\tau_t = 0.1$  ps and  $\tau_p = 2$  ps. Constraints were applied to hydrogen-containing bonds through the LINCS algorithm<sup>67</sup>. A cutoff of 10 Å was used for Van der Waals and short-range Coulomb interactions; long range electrostatics was treated with Particle Mesh Ewald. Ab1 and Ab2 were simulated for a total of 2  $\mu$ s each, in 4 independent replicas of 500 ns. Ab3 was simulated for 500 ns. Additional simulations of the apo state were started from the representative holo conformations, after removal of the antigen; the structures were solvated and equilibrated using the above-mentioned procedure, and three 100ns-long replicas were simulated for each system.

**Analyses.** Clustering of the simulation frames according to structural similarity was performed with an in-house script using hierarchical clustering (UPGMA)<sup>68</sup>. The distance used in the cluster definition is based on the root-mean-square deviation between C $\alpha$  of all the structure pairs sampled in the simulations, following superimposition with the mdtraj algorithm that translates and rotates the structures for optimal alignment<sup>69</sup>. Calculation of atomic displacement, residue fluctuations, interatomic distances and contact surface areas, were performed using Gromacs 2018 utilities, while calculation of MM/PBSA energy was performed with the software *g\_mmpbsa*<sup>52</sup>. Additional details on the MM/PBSA method and on the PAD $\omega$  calculation are reported in the SI. The number of hydrogen bonds was computed with VMD<sup>70</sup>, setting an angle cutoff of 25° and a maximum donor-acceptor distance of 3 Å.

**Principal component analysis.** Principal component analysis (PCA) was performed for each cluster using C $\alpha$  atoms positions. Calculation, diagonalization and analysis of the covariance matrices were performed using Gromacs tools *gmx covar* and *gmx ana eig*<sup>62</sup>. In order to visualize the direction of movements captured by the eigenvectors, porcupine plots were generated using the extreme projections on the first three principal component, and visualized with VMD; here, the direction of the arrow on each C $\alpha$  atom represents the direction of motion, while the length of the arrow characterizes the movement strength.

**Correlations.** Mutual information (MI) was calculated as a measure of inter-residue communications. MI is defined as follows:

$$MI_{ij} = \iint dx_i dx_j p(x_i, x_j) \log \frac{p(x_i, x_j)}{p(x_i)p(x_j)} \quad (1)$$

where  $x_i$  and  $x_j$  are the displacements of the C $\alpha$  atoms with respect to their average positions,  $p(x_i)$  and  $p(x_j)$  are the probability functions of finding the  $i$ -th or  $j$ -th atoms with a displacement equal to  $x_i$  or  $x_j$ ,  $p(x_i, x_j)$  is the joint probability function. The calculation was performed with in-house scripts, and the displacement of each atom was divided in 100 discrete bins covering the range between 0 and the largest distance of that atom from its equilibrium position as sampled during the simulations. From the mutual information, the generalized correlation coefficient (GCC) is computed as:

$$GCC_{ij} = \sqrt{1 - e^{-2MI_{ij}/3}} \quad (2)$$

GCC represents a measure of normalized MI, ranging from 0 (no correlation) to 1 (perfect correlation)<sup>54</sup>. Calculations were performed through in-house scripts.

GCC was in turn employed to calculate the correlation score (CS). For each residue  $i$ ,  $CS_i$  is computed as the sum of the generalized correlation coefficient values between residue  $i$  and the other protein residues:

$$CS_i = \sum_{j \neq i} GCC_{ij} \quad (3)$$

In the case of the *intra*-domain CS, the sum extends to the residues belonging to the same structural domain as residue  $i$ ; in the case of the *inter*-domain CS, the summation takes into account only residues belonging to all the other structural domains, excluding that of residue  $i$ .

**Network analysis.** Networks were defined as sets of interconnected nodes centered on the C $^\alpha$  atoms. The total number of nodes of each system corresponds therefore to the number of residues. A couple of nodes is considered connected by an edge if any heavy atoms of the two residues is within a distance of 4.5 Å for at least 75% of the simulation time. These cutoffs were selected after a convergence study based on the Community Repartition Difference (see SI).

Each edge is weighted according to the generalized correlation coefficient measure; specifically, the weight of the edge between nodes  $i$  and  $j$  is defined as:

$$w_{ij} = -\log [GCC_{ij}] \quad (4)$$

where  $GCC_{ij}$  is the generalized correlation coefficient (equation 2).

The network analysis was performed with the Python implementation found in Melo et al.<sup>71</sup>, which makes use of the NetworkX package<sup>72</sup> and is optimized using Cython<sup>73</sup> and Numba<sup>74</sup>. In the weighted networks, the sets of communities were identified using the Girvan-Newman algorithm<sup>75,76</sup>. The importance of the edge for communication within the network was measured in terms of edge betweenness, which is defined as the number of shortest pathways that cross the edge; this quantity was calculated with the commonly used Floyd–Warshall algorithm<sup>77,78</sup>. For a detailed description of the methods used for network analysis, the reader is referred to the Supplementary Information.

## Data availability

The raw data produced and analysed in this work are freely available on the Zenodo repository <https://doi.org/10.5281/zenodo.5018432>.

Received: 12 July 2021; Accepted: 15 November 2021

Published online: 01 December 2021

## References

- Kaplon, H. & Reichert, J. M. Antibodies to watch in 2021. *MAbs* **13**, 1860476 (2021) (PMID: 33459118).
- Lu, R.-M. *et al.* Development of therapeutic antibodies for the treatment of diseases. *J. Biomed. Sci.* **27**, 1 (2020).
- Scourfield, D. O. *et al.* The role and uses of antibodies in covid-19 infections: A living review. *Oxford Open Immunol.* **2**, iqab003 (2021).
- Beck, A., Goetsch, L., Dumontet, C. & Corvaia, N. Strategies and challenges for the next generation of antibody-drug conjugates. *Nat. Rev. Drug Discov.* **16**, 315–337 (2017).
- Larson, S. M., Carrasquillo, J. A., Cheung, N.-K.V. & Press, O. W. Radioimmunotherapy of human tumours. *Nat. Rev. Cancer* **15**, 347–360 (2015).
- Dumet, C., Pottier, J., Gouilleux-Gruart, V. & Watier, H. Insights into the igg heavy chain engineering patent landscape as applied to igg4 antibody development. *MAbs* **11**, 1341–1350 (2019).
- Beers, S. A., Glennie, M. J. & White, A. L. Influence of immunoglobulin isotype on therapeutic antibody function. *Blood* **127**, 1097–1101 (2016).
- Samsudin, F., Yeo, J. Y., Gan, S.K.-E. & Bond, P. J. Not all therapeutic antibody isotypes are equal: The case of igm versus igg in pertuzumab and trastuzumab. *Chem. Sci.* **11**, 2843–2854 (2020).
- Wang, X., Mathieu, M. & Brezski, R. J. Igg fc engineering to modulate antibody effector functions. *Protein Cell* **9**, 63–73 (2018).
- Saunders, K. O. Conceptual approaches to modulating antibody effector functions and circulation half-life. *Front. Immunol.* **10**, 1296 (2019).
- Liu, Z. *et al.* Asymmetrical fc engineering greatly enhances antibody-dependent cellular cytotoxicity (adcc) effector function and stability of the modified antibodies. *J. Biol. Chem.* **289**, 3571–3590 (2014).
- Handlogten, M. W. *et al.* Prevention of fab-arm exchange and antibody reduction via stabilization of the igg4 hinge region. *MAbs* **12**, 1779974 (2020).
- Silva, J.-P., Vetterlein, O., Jose, J., Peters, S. & Kirby, H. The s228p mutation prevents in vivo and in vitro igg4 fab-arm exchange as demonstrated using a combination of novel quantitative immunoassays and physiological matrix preparation. *J. Biol. Chem.* **290**, 5462–5469 (2015).
- Herbener, P. *et al.* Functional relevance of in vivo half antibody exchange of an igg4 therapeutic antibody-drug conjugate. *PLoS ONE* **13**, e0195823 (2018).
- Chiu, M. L. & Gilliland, G. L. Engineering antibody therapeutics. *Curr. Opin. Struct. Biol.* **38**, 163–173 (2016).
- Koenig, P. *et al.* Mutational landscape of antibody variable domains reveals a switch modulating the interdomain conformational dynamics and antigen binding. *Proc. Natl. Acad. Sci.* **114**, E486–E495 (2017).
- Yanaka, S., Moriwaki, Y., Tsumoto, K. & Sugase, K. Elucidation of potential sites for antibody engineering by fluctuation editing. *Sci. Rep.* **7**, 1–10 (2017).
- Cannon, D. A. *et al.* Experimentally guided computational antibody affinity maturation with de novo docking, modelling and rational design. *PLoS Comput. Biol.* **15**, e1006980 (2019).
- Patel, J. S. & Ytreberg, F. M. Fast calculation of protein–protein binding free energies using umbrella sampling with a coarse-grained model. *J. Chem. Theory Comput.* **14**, 991–997 (2018).
- Lees, W. D., Stejskal, L., Moss, D. S. & Shepherd, A. J. Investigating substitutions in antibody-antigen complexes using molecular dynamics: A case study with broad-spectrum, influenza a antibodies. *Front. Immunol.* **8**, 143 (2017).

21. Codina, N. *et al.* An expanded conformation of an antibody fab region by X-ray scattering, molecular dynamics, and smfret identifies an aggregation mechanism. *J. Mol. Biol.* **431**, 1409–1425 (2019).
22. Yoshida, K. *et al.* Exploring designability of electrostatic complementarity at an antigen-antibody interface directed by mutagenesis, biophysical analysis, and molecular dynamics simulations. *Sci. Rep.* **9**, 4482 (2019).
23. Fernández-Quintero, M. L., Heiss, M. C., Pomarici, N. D., Math, B. A. & Liedl, K. R. Antibody cdr loops as ensembles in solution vs. canonical clusters from X-ray structures. *MAbs* **12**, 1744328 (2020).
24. Fernández-Quintero, M. L., Kraml, J., Georges, G. & Liedl, K. R. Cdr-h3 loop ensemble in solution-conformational selection upon antibody binding. *MAbs* **11**, 1077–1088 (2019).
25. Fernández-Quintero, M. L., Kroell, K. B., Hofer, F., Riccabona, J. R. & Liedl, K. R. Mutation of residue 71 h results in different paratope states in solution. *Front. Immunol.* **12**, 243 (2021).
26. Al Qaraghuli, M. M., Kubiak-Ossowska, K., Ferro, V. A. & Mulheran, P. A. Antibody-protein binding and conformational changes: Identifying allosteric signalling pathways to engineer a better effector response. *Sci. Rep.* **10**, 1–10 (2020).
27. Zhao, J., Nussinov, R. & Ma, B. Antigen binding allosterically promotes fc receptor recognition. *MAbs* **11**, 58–74 (2019).
28. Zhao, J., Nussinov, R. & Ma, B. Mechanisms of recognition of amyloid- $\beta$  monomer, oligomer, and fibril by homologous antibodies. *J. Biol. Chem.* **292**, 18325–18343 (2017).
29. Scapin, G. *et al.* Structure of full-length human anti-pd1 therapeutic igg4 antibody pembrolizumab. *Nat. Struct. Mol. Biol.* **22**, 953–958. <https://doi.org/10.1038/nsmb.3129> (2015).
30. Urquhart, L. Top companies and drugs by sales in 2019. *Nat. Rev. Drug Discov.* **19**, 228–229 (2020).
31. Patnaik, A. *et al.* Phase i study of pembrolizumab (mk-3475; anti-pd-1 monoclonal antibody) in patients with advanced solid tumors. *Clin. Cancer Res.* **21**, 4286–4293 (2015).
32. Nghiem, P. T. *et al.* Pd-1 blockade with pembrolizumab in advanced merkel-cell carcinoma. *N. Engl. J. Med.* **374**, 2542–2552 (2016).
33. Goldberg, S. B. *et al.* Pembrolizumab for patients with melanoma or non-small-cell lung cancer and untreated brain metastases: Early analysis of a non-randomised, open-label, phase 2 trial. *Lancet Oncol.* **17**, 976–983 (2016).
34. Sharpe, A. H. & Pauken, K. E. The diverse functions of the pd1 inhibitory pathway. *Nat. Rev. Immunol.* **18**, 153 (2018).
35. Ahmadzadeh, M. *et al.* Tumor antigen-specific cd8 t cells infiltrating the tumor express high levels of pd-1 and are functionally impaired. *Blood* **114**, 1537–1544 (2009).
36. Pardoll, D. M. The blockade of immune checkpoints in cancer immunotherapy. *Nat. Rev. Cancer* **12**, 252–264 (2012).
37. Ivashko, I. N. & Kolesar, J. M. Pembrolizumab and nivolumab: PD-1 inhibitors for advanced melanoma. *Am. J. Health Syst. Pharm.* **73**, 193–201 (2016).
38. Chen, Y. *et al.* Looking for the optimal pd-1/pd-l1 inhibitor in cancer treatment: A comparison in basic structure, function, and clinical practice. *Front. Immunol.* **11**, 1–10 (2020).
39. Edelman, G. M. *et al.* The covalent structure of an entire  $\gamma$ g immunoglobulin molecule. *Proc. Natl. Acad. Sci.* **63**, 78–85 (1969).
40. Lee, J. Y. *et al.* Structural basis of checkpoint blockade by monoclonal antibodies in cancer immunotherapy. *Nat. Commun.* **7**, 13354. <https://doi.org/10.1038/ncomms13354> (2016).
41. Tian, X. *et al.* In-depth analysis of subclass-specific conformational preferences of igg antibodies. *IUCr* **2**, 9–18 (2015).
42. Aalberse, R. C. & Schuurman, J. Igg4 breaking the rules. *Immunology* **105**, 9–19 (2002).
43. Deveuve, Q., Lajoie, L., Barrault, B. & Thibault, G. The proteolytic cleavage of therapeutic monoclonal antibody hinge region: More than a matter of subclass. *Front. Immunol.* **11**, 168 (2020).
44. Freeman, L. C. A set of measures of centrality based on betweenness. *Sociometry* **1**, 35–41 (1977).
45. Lefranc, M.-P. & Lefranc, G. Imgt<sup>®</sup> and 30 years of immunoinformatics insight in antibody v and c domain structure and function. *Antibodies* **8**, 29 (2019).
46. Kortkhonjia, E. *et al.* Probing antibody internal dynamics with fluorescence anisotropy and molecular dynamics simulations. *MAbs* **5**, 306–322 (2013).
47. Roux, K. H., Strelets, L. & Michaelsen, T. E. Flexibility of human igg subclasses. *J. Immunol.* **159**, 3372–3382 (1997).
48. Dangi, J. L. *et al.* Segmental flexibility and complement fixation of genetically engineered chimeric human, rabbit and mouse antibodies. *EMBO J.* **7**, 1989–1994 (1988).
49. Caliandro, R., Rossetti, G. & Carloni, P. Local fluctuations and conformational transitions in proteins. *J. Chem. Theory Comput.* **8**, 4775–4785 (2012).
50. Li, T. *et al.* Redistribution of flexibility in stabilizing antibody fragment mutants follows Le Chatelier's principle. *PLoS ONE* **9**, e92870 (2014).
51. Li, T. *et al.* Rigidity emerges during antibody evolution in three distinct antibody systems: Evidence from qsfr analysis of fab fragments. *PLoS Comput. Biol.* **11**, e1004327 (2015).
52. Kumari, R., Kumar, R., Consortium, O. S. D. D. & Lynn, A. G<sub>1</sub>\_mmpbsa: A gromacs tool for high-throughput mm-pbsa calculations. *J. Chem. Inf. Model.* **54**, 1951–1962 (2014).
53. Knapp, B., Dunbar, J., Alcalá, M. & Deane, C. M. Variable regions of antibodies and t-cell receptors may not be sufficient in molecular simulations investigating binding. *J. Chem. Theory Comput.* **13**, 3097–3105 (2017).
54. Lange, O. F. & Grubmüller, H. Generalized correlation for biomolecular dynamics. *Proteins Struct. Funct. Bioinform.* **62**, 1053–1061 (2006).
55. Palermo, G. *et al.* Protospacer adjacent motif-induced allostery activates crispr-cas9. *J. Am. Chem. Soc.* **139**, 16028–16031 (2017).
56. Adhikary, R. *et al.* Adaptive mutations alter antibody structure and dynamics during affinity maturation. *Biochemistry* **54**, 2085–2093 (2015).
57. Janda, A., Bowen, A., Greenspan, N. S. & Casadevall, A. Ig constant region effects on variable region structure and function. *Front. Microbiol.* **7**, 22 (2016).
58. Yanaka, S. *et al.* Dynamic views of the fc region of immunoglobulin g provided by experimental and computational observations. *Antibodies* **8**, 1–10 (2019).
59. Pettersen, E. F. *et al.* Ucsf chimera: A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
60. Harris, L. J., Larson, S. B., Hasel, K. W. & McPherson, A. Refined structure of an intact igg2a monoclonal antibody. *Biochemistry* **36**, 1581–1597 (1997) (PMID: 9048542).
61. Bekker, H. *et al.* Gromacs: A parallel computer for molecular dynamics simulations. *Phys. Comput.* **92**, 252–256 (1993).
62. Abraham, M. J. *et al.* Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1**, 19–25 (2015).
63. Lindorff-Larsen, K. *et al.* Improved side-chain torsion potentials for the amber ff99sb protein force field. *Proteins Struct. Funct. Bioinform.* **78**, 1950–1958 (2010).
64. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).
65. Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**(1), 1 (2007).
66. Parrinello, M. & Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).
67. Hess, B., Bekker, H., Berendsen, H. J. C. & Fraaije, J. G. E. M. Lincs: A linear constraint solver for molecular simulations. *J. Comput. Chem.* **18**, 1463–1472 (1997).

68. Sokal, R. R. & Michener, C. D. A statistical method for evaluating systematic relationships. *Univ. Kans. Sci. Bull.* **28**, 1409–1438 (1958).
69. McGibbon, R. T. *et al.* Mdrtraj: A modern open library for the analysis of molecular dynamics trajectories. *Biophys. J.* **109**, 1528–1532 (2015).
70. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
71. Melo, M. C., 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. *J. Chem. Phys.* **153**, 134104 (2020).
72. Hagberg, A., Swart, P. & Chult, D. Exploring network structure, dynamics, and function using networkx. Tech. Rep. (Los Alamos National Lab. (LANL), 2008).
73. Behnel, S. *et al.* Cython: The best of both worlds. *Comput. Sci. Eng.* **13**, 31–39 (2011).
74. Lam, S. K., Pitrou, A. & Seibert, S. Numba: A llvm-based python jit compiler. In *Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC*, 1–6 (2015).
75. Girvan, M. & Newman, M. E. Community structure in social and biological networks. *Proc. Natl. Acad. Sci.* **99**, 7821–7826 (2002).
76. Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008**, P10008 (2008).
77. Floyd, R. W. Algorithm 97: Shortest path. *Commun. ACM* **5**, 345. <https://doi.org/10.1145/367766.368168> (1962).
78. Warshall, S. A theorem on Boolean matrices. *J. ACM* **9**, 11–12. <https://doi.org/10.1145/321105.321107> (1962).

## Acknowledgements

The authors thank Emiliano Biasini and Attilio Vittorio Vargiu for an insightful reading of the manuscript. This project received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (Grant 758588).

## Author contributions

R.P. designed the project; T.T. and M.R. performed the simulations and the analyses; all authors contributed to the interpretation of the results and to the writing of the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-02323-x>.

**Correspondence** and requests for materials should be addressed to R.P.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021