



**HAL**  
open science

# Mesoscale computational protocols for the design of highly cooperative bivalent macromolecules

Suman Saurabh, Francesco Piazza

► **To cite this version:**

Suman Saurabh, Francesco Piazza. Mesoscale computational protocols for the design of highly cooperative bivalent macromolecules. *Scientific Reports*, 2020, 10 (1), 10.1038/s41598-020-64646-5 . hal-02930878

**HAL Id: hal-02930878**

**<https://hal.science/hal-02930878>**

Submitted on 25 Nov 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



OPEN

# Mesoscale computational protocols for the design of highly cooperative bivalent macromolecules

Suman Saurabh<sup>1</sup>✉ & Francesco Piazza<sup>1,2</sup>

The last decade has witnessed a swiftly increasing interest in the design and production of novel multivalent molecules as powerful alternatives for conventional antibodies in the fight against cancer and infectious diseases. However, while it is widely accepted that large-scale flexibility (10–100 nm) and free/constrained dynamics (100 ns -  $\mu$ s) control the activity of such novel molecules, computational strategies at the mesoscale still lag behind experiments in optimizing the design of crucial features, such as the binding cooperativity (a.k.a. avidity). In this study, we introduced different coarse-grained models of a polymer-linked, two-nanobody composite molecule, with the aim of laying down the physical bases of a thorough computational drug design protocol at the mesoscale. We show that the calculation of suitable potentials of mean force allows one to apprehend the nature, range and strength of the thermodynamic forces that govern the motion of free and wall-tethered molecules. Furthermore, we develop a simple computational strategy to quantify the encounter/dissociation dynamics between the free end of a wall-tethered molecule and the surface, at the roots of binding cooperativity. This procedure allows one to pinpoint the role of internal flexibility and weak non-specific interactions on the kinetic constants of the nanobody-wall encounter and dissociation. Finally, we quantify the role and weight of rare events, which are expected to play a major role in real-life situations, such as in the immune synapse, where the binding kinetics is likely dominated by fluctuations.

Single-domain antibodies, also known as *nanobodies* (NB)<sup>1</sup>, are found naturally in camelids and represent an intriguing alternative to design and build novel multivalent and multi-specific immunotherapy agents for targeting tumors and viral infections<sup>2</sup>. In addition to leading to an increase in affinity, coupling two (or more) different *binders* helps in marginalising the effect of mutation and polymorphism of the target. Linked anti-CD16 nanobodies, C21 and C28<sup>3</sup>, for example, have been shown to be effective against treating breast cancer involving a low HER2 expression, which is resistant to the therapeutic antibody trastuzumab<sup>4</sup>.

As the conventional antibodies are large structures, they may not be effective when it comes to situations like a partially exposed tumor<sup>5,6</sup>. In addition, an uneven distribution of the anti-tumor antibody in the entire tumor region may lead to tumor regrowth<sup>7</sup>. In such situations, therapeutic agents with smaller structures are desirable<sup>8</sup>. A good solution is to use engineered structures composed of antigen-specific nanobodies linked with flexible linkers<sup>9</sup>, e.g. realized via a polymer such as [Gly<sub>4</sub> Ser]<sub>n</sub>. Such structures will be small and will thus have better ability to penetrate into the tumor microenvironment and reinforce the formation of the immune synapse. Such structures are easy to produce and are more efficient when compared to the bulky conventional antibodies, and can be designed to be multivalent and multispecific<sup>10,11</sup>.

An important property of an antibody is the strength of bivalent binding that it demonstrates, a property known in immunology as *avidity*. Although it has no unique quantitative definition (let alone whether it is more appropriate to regard it as an equilibrium or a kinetic parameter), avidity can be thought of as the cooperative gain in affinity afforded by double binding. Notably, such measure is intimately connected to the internal flexibility of the multi-domain molecules and to the geometric configuration of binding epitopes<sup>12</sup>. For example, virions employ various strategies to evade the action of antibodies of the immune system. Some virions, like HIV-1, have a very rapid rate of mutation. Experiments reveal that the enhanced antibody evasion capability of HIV-1 is based not just in its capability to mutate but is a combined effect of mutation along with the spike structure and low spike density<sup>13–15</sup>. While mutations reduce the affinity of natural antibodies towards the target spikes, the spike

<sup>1</sup>Centre de Biophysique Moléculaire, CNRS, Rue Charles Sadron, 45071, Orléans, France. <sup>2</sup>Université d'Orléans, Pôle de Physique, 45071, Orléans, France. ✉e-mail: [suman.saurabh@cnrs-orleans.fr](mailto:suman.saurabh@cnrs-orleans.fr)

structure and low density preclude intra- (i.e. multiple epitopes on the same target) and inter-spike cross-linking, thus preventing bivalent binding and avidity<sup>16,17</sup>.

Experiments have demonstrated that, linking two Fab domains of an antibody through an extended polymer like a DNA oligomer, leads to an increase in bivalent binding for an optimal linker length, and a resulting increase in the efficacy of the antibodies. Wu *et al.*<sup>18</sup> performed experiments on respiratory syncytial virus, which has a very high Env spike density, and showed that the affinity of low-affinity bivalent Fabs was 2–3 orders of magnitude higher as compared to their monovalent counterparts and the efficiency was not affected by mutations that increased the off-rates nearly 100-fold. The results show that multivalent structures made of polymer-linked nanobodies would lead to higher degree of avidity and thus higher efficiency. Galimidi *et al.*<sup>19</sup> performed experiments on linked Env (HIV-1 envelope glycoproteins) binders and showed that linking can lead to an increase in potency by 2–3 orders of magnitude. Jähnichen *et al.*<sup>20</sup> developed two different single domain nanobodies that could bind to different sites on the extracellular domain of the CXCR4 coreceptor. They found that joining the two nanobodies with protein linkers resulted in a 27-fold increase in CXCR4 affinity. With further analysis they concluded that the effect is pure avidity resulting from the heterobivalent linking of the two nanobodies. Zhang *et al.*<sup>21</sup> developed multimers of nanobodies leading to an increase in affinity and several orders of magnitude decrease in the rates of dissociation. Yang *et al.*<sup>22</sup> performed dissociation rate calculations for the binding between a bivalent antibody and hapten ligands as a function of the ligand density. They found cooperative binding as the hapten density increased and bivalent binding set in. They could determine two different dissociation constants for the double-step antibody-hapten binding process with one dissociation constant being 3-orders of magnitude larger than the other.

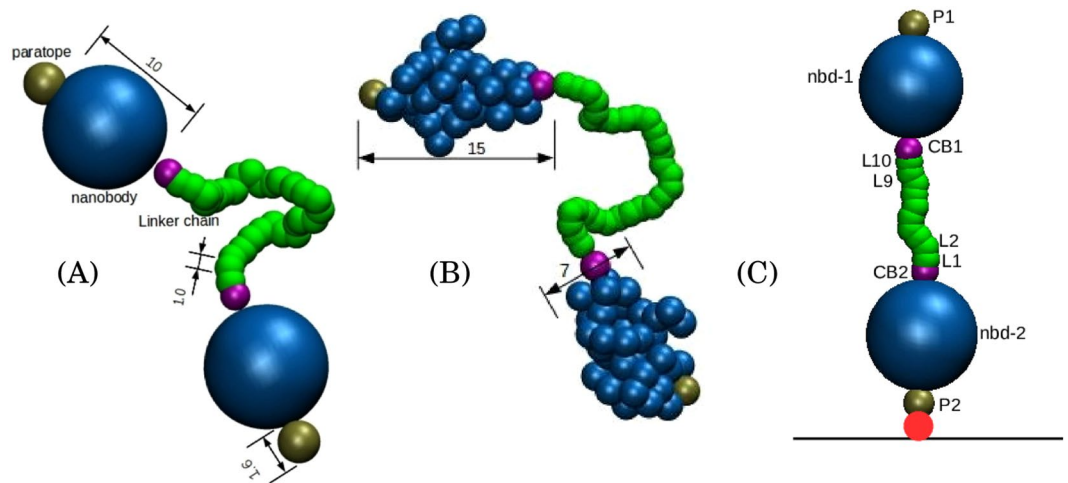
When one of the nanobodies in a two-NB construct binds to the receptor at its binding site (epitope), the other linked ligand spends more time in the vicinity, leading to a larger probability of the latter unit to bind to another similar or different epitope, on the same or on a facing surface, depending on whether the system is mono-specific or multi-specific. By hindering free diffusion of the ligands, linking can lead to an increase in rebinding events and strengthen the interaction between interfaces<sup>23</sup>. Further, Bongrand *et al.*<sup>24</sup> showed that the fraction of bivalent attachments between an antibody-coated microsphere and a mono- or bivalent ligand-coated surface, that resisted a force of 30 pN for a minimum of 5 seconds, was ~4 times higher than the number of monovalent attachments.

While the choice of the nanobody depends not only on its affinity towards the target epitope, but also on the nature of the bond it forms with the epitope<sup>25</sup>, the choice of linker would depend on the geometry of the epitope distribution/configurations. Experimental methods have been developed to create linkers of given stiffness and extension, out of a combination of proteins and peptides<sup>26–29</sup>. Among the most common are the (Gly<sub>4</sub> Ser)<sub>n</sub> linkers. Protein linkers have been accommodated into the hinge regions of natural antibodies, thus enabling intra-spike linking to viral receptors<sup>28</sup>. Other biocompatible polymers like PEG are also good candidates to be used to link the nanobodies. The properties of the linker are very important in determining the degree of avidity. Depending on the epitope density on a tumor cell or the distribution of the antibody binding sites on the viral envelope, a linker that is either too flexible or too stiff can lead to under-performance.

With the improvements in computational resources and speed, molecular dynamics (MD) simulations in recent days have been playing an important role in drug discovery<sup>30</sup> and also in unraveling fundamental mechanisms involving very large biological complexes, such as chromatin<sup>31</sup>. MD simulations can be important in determining the optimal properties of the linkers that would lead to an efficient multivalent binding for a particular target. In addition, simulations of a group of linked nanobodies can give important insights into their epitope-binding kinetics as a function of the linker structural properties and other important parameters, like the paratope-epitope binding energy. While the engineered nanobody-linker-nanobody systems are much smaller as compared to the conventional antibodies, simulating a significantly large group of them in atomistic detail would be computationally expensive and cannot be done routinely. Thus, some degree of coarse-graining is important to perform kinetics analysis using MD simulation as a tool.

Keeping the above discussion in mind, here we perform MD simulation of a coarse-grained system consisting of two nanobodies connected by a linker (referred to as a diabody) and study its structural and dynamical properties. We use different levels of coarse-graining schemes to represent the diabody. In the simplest representation, we perform simulations of two extended (rigid) spheres connected by a flexible bead-string linker (see Fig. 1(A)). A similar model has been used previously to study the dynamics of a polymer tethered to a surface with a big bead at the free end<sup>32</sup>. In addition to this, with a future aim to study the dynamics of diabodies in the presence of their target receptors (such as HER2), where a nanobody represented by a hard sphere will be incapable of representing the important features of the diabody-target interaction properly, we perform a finer coarse-graining of the nanobody (see Fig. 1(B)). In this scheme, we represent the nanobody using the shape-based coarse-graining (SBCG) scheme developed by Schulten *et al.*<sup>33</sup> (see methods). Using Umbrella Sampling (US) simulations we calculate the free energy profiles on which various conformations of diabodies tethered to a surface lie. From the MD trajectories we calculate the flight and residence times of the free end of the tethered diabody in a region close to the tethering wall. Comparing the results from the two different models we demonstrate how the coarse-graining scheme could affect the results and, notably, we highlight the role of the “shape” in the wall-domain dynamics.

The paper is organized as follows. First, we describe and comment our results, mainly concerning the different potentials of mean force (PMF) and the detailed analyses of the encounter and dissociation kinetics of the free NB of a wall-tethered molecule and discuss them in relation to their relevance to therapeutic applications. In the conclusion, we wrap up our discussion and provide some final comments on this work and its perspectives. In the methods section we provide the details of our computational methods, models and simulations.



**Figure 1.** The two coarse-grained models of diabody studied in this work. Two nanobodies are linked through a polymer. The nanobody beads are represented in blue and the polymer in green. The maroon beads are the connector beads, which connect the nanobody to the polymer, while the tan beads are the paratopes. **(A)** The SPH model uses single spherical beads to represent the nanobodies. In this model, the nanobody, along with the corresponding paratope and the connector bead are treated as a single rigid body. **(B)** The SBCG model uses a shape-based coarse-graining algorithm for the nanobodies (see methods). **(C)** A 10-mer SPH diabody with the labels of different beads. The same naming scheme is used for all linker lengths and models. The tethering point is represented by the red bead. The bead sizes are in the reduced units (see methods).

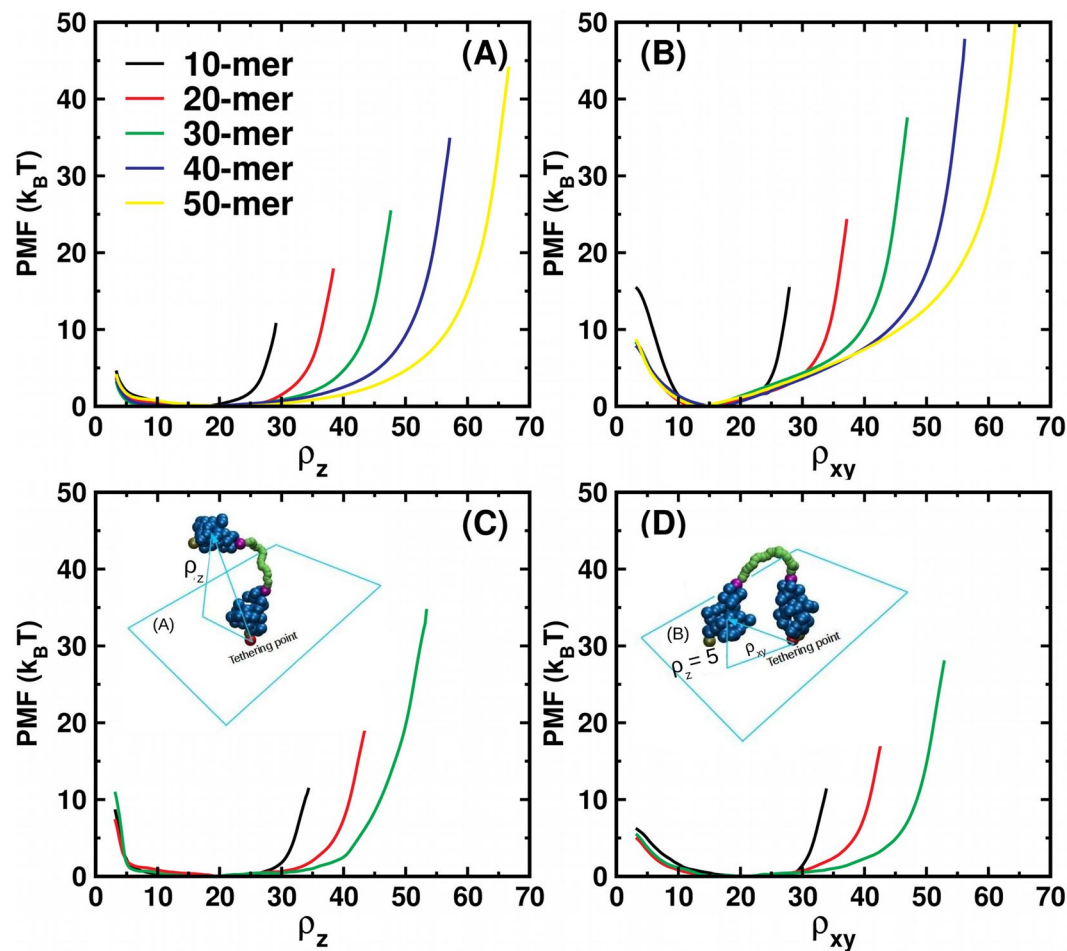
## Results and discussion

All the simulations in this work have been performed on diabodies tethered to the fixed lower  $z$ -wall of a cuboidal simulation box periodic in  $x$  and  $y$ . The symbols used to refer to different beads in the tethered diabody are shown in Fig. 1(C). In the rest of the article, the free nanobody is referred to as nbd-1 while the nanobody tethered to the wall is referred to as nbd-2. The connector beads (beads linking the linker polymer to the nanobodies) corresponding to nbd-1 and nbd-2 are named CB1 and CB2 respectively, while the paratopes are referred to as P1 and P2 respectively.

**Free energy profiles.** When one of the nanobodies of the diabody (nbd-2) binds to an epitope on a surface, the attachment of the other nanobody (nbd-1) to another epitope while nbd-2 is still attached depends on many factors, like the nature of the linker and the interaction of the nanobodies with the wall. The effect of these properties on the binding kinetics can be captured well by the free energy profile as a function of the height of the free nanobody above the wall. Another factor that influences the ability of bivalent attachment is the density of the epitopes on the target wall. In particular, it is desirable that the most probable radial distance between the paratopes of the two nanobodies roughly matches the average inter-epitope distance. For a particular linker one can estimate the most probable value of radial distance from the free energy profile as a function of the radial distance. Keeping this in mind, we calculate the PMF profiles for the spherical (SPH) and SBCG diabodies, with one of the nanobodies tethered on the  $x - y$  plane to a repulsive wall as a function of two collective variables (reaction coordinates (RC)), named  $\rho_z$  and  $\rho_{xy}$ . The former is the height of the center of mass of nbd-1 above the wall. The latter is the projection of the vector joining the tethering point to the center of mass of nbd-1 on the  $x - y$  plane, constrained to  $\rho_z = 5 \sigma$  (see Fig. 2). All distances are reported in reduced units, as a multiple of monomer size of the linker polymer,  $\sigma$  (see methods).

The PMF profiles as a function of  $\rho_z$  are shown in Fig. 2(A,C) for the SPH and SBCG diabodies, respectively. In spite of the PMF profiles having similar shape, there are a few notable differences. For the 10-mer system, for example, the PMF profile is flat in  $15 \sigma \leq \rho_z \leq 25 \sigma$  for the SBCG system, while the same region for the SPH system occurs at shorter distances, extending between  $8 \sigma$  and  $20 \sigma$ . In addition, for small values of RC, the profiles for the SBCG system have a larger slope as compared to the SPH system, which means that the SBCG nanobody faces a larger (entropic) repulsive force from the fixed wall than the SPH nanobody. While the repulsive force will also have an enthalpic component from the deformation in the shape of the SBCG nanobodies, the enthalpic contribution is expected to be weaker given their compact conformation. The derivative of the PMF profiles is shown in Fig. S2 of the supplementary information (SI). These differences suggest that the different coarse-graining schemes highlight differences in the dynamics near the tethering wall. In particular, an excessively simple spherical model seems to fall short of capturing important features of the interactions with a boundary.

The PMFs are portrayed as a function of  $\rho_{xy}$  in Fig. 2(B,D). The profiles for the two models appear very different. While the PMFs for the SPH systems increase steadily in a linear fashion after  $\rho_{xy} \sim 15 \sigma$ , the profiles for the SBCG systems are flatter and start rising at a much later stage. Figure 3(A) shows the average of all values of (normalized)  $\rho_{xy}$  for which the  $\text{PMF} \leq k_B T$ , plotted as a function of the linker length. The error bars gauge the flatness of the PMF profiles, which is seen to vary markedly in the two models. More precisely, the relative flatness of the

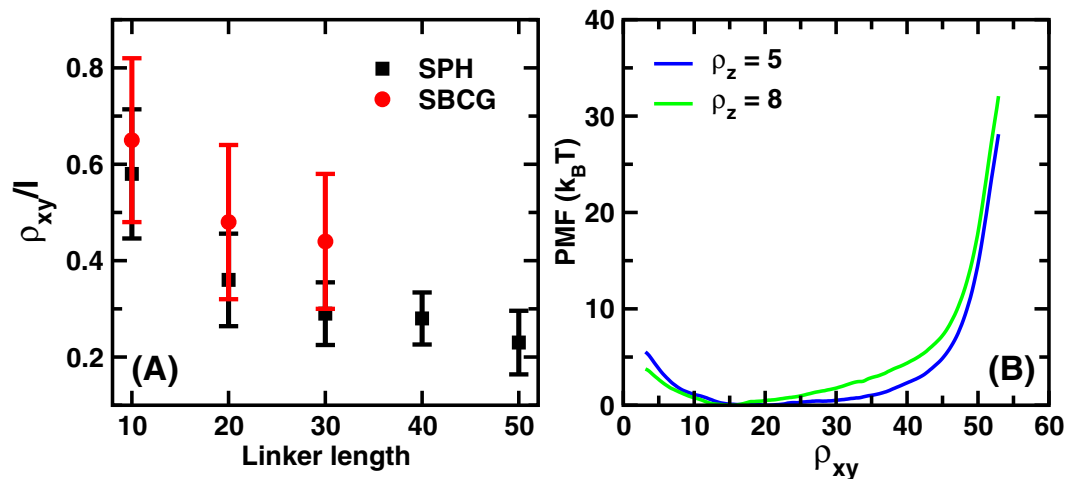


**Figure 2.** PMF profiles for tethered diabodies. PMF profiles as a function of the height above the wall of the center of mass of the free nanobody ( $\rho_z$ ) for the (A) SPH and (C) SBCG systems. PMF profiles as a function of the projection on the  $x - y$  plane of the vector joining the tethering point to the center of mass of the free nanobody ( $\rho_{xy}$ ), with the free NB restrained to stay at a height  $z = 5 \sigma$  for the (B) SPH and (D) SBCG systems along with a schematic representation of (Inset A)  $\rho_z$  and (Inset B)  $\rho_{xy}$ . The red bead is the tethering point. The  $x$ -axis is represented in units of  $\sigma$ .

PMF profiles (in the  $x - y$  plane) for short linkers is more pronounced, while it decreases with increasing linker length. This indicates that a network of nanobodies linked with short linkers may be more effective at inter-epitope binding even for configurations with a large standard deviation in the inter-epitope distances. The distribution of the position of P1 in the  $x - y$  plane for the SPH system, for different linker lengths is shown in Fig. S3 of SI, which reaffirms the results shown in Fig. 3(A).

To better understand the differences displayed by the two models for what concerns the thermodynamic forces exerted by the wall on the free nanobody, we computed the PMF along  $\rho_{xy}$  for different values of  $\rho_z$  for the SBCG diabodies with 30-mer linkers. Figure 3(B) shows the profiles obtained for  $\rho_z = 5 \sigma$  (nbd-1 in contact with the fixed tethering wall) and  $\rho_z = 8 \sigma$  (nbd-1 separated from the wall). For the larger value of  $\rho_z$ , the flatness of the PMF profile appears to be reduced, which implies a larger energy cost associated with extended conformations in the  $x - y$  plane. While the steepness of the PMF profile is still lower than that for the SPH system, we observe a reduction in the flatness of the PMF profile for the SBCG systems when the free nanobody moves away from the wall. This indicates that for epitope systems with targets present very close to the cell-membrane, which may act as a soft wall, one might have a larger chance of inter-epitope binding as the wall clearly has a considerable flattening effect on the SBCG PMF profiles. Overall, the differences between the PMF profiles of the two systems demonstrates that molecular shape may play a big role in controlling the interaction with the tethering wall.

One notices a steeper increase of the PMF as a function of  $\rho_{xy}$  at low distances for the 10-mer as compared to other linker lengths. This effect is much more pronounced for the spherical molecules (Fig. 2(B,D)). This indicates that when the linker length is comparable to the dimensions of the nanobodies, it is difficult to approach the epitopes close-by to the one to which the diabody is tethered. It is to be stressed that the slope is larger in the case of the SPH system as compared to the SBCG system, which indicates that the shape of the nanobody is expected to play a role when it comes to inter-epitope binding for a high target density, especially for low linker lengths.



**Figure 3.** Flatness of PMF profiles and the effect of the wall. **(A)** Average of  $\rho_{xy}/l$  over the flat region of the PMF profiles as a function of linker length. Here,  $L$  is the distance between the tethering point and P1. The error bars quantify the extent of the flat region. **(B)** PMF profiles for the SBCG diabodies with the 30-mer linker as a function of  $\rho_{xy}$  for two different values of  $\rho_z$ . The PMF profile for a wall-touching situation is different from the one where nbd-1 is farther from the fixed wall. In that case, the PMF profile for the SBCG diabody resembles that of the SPH diabody.

In the spirit of computationally aided molecular design, the PMF profiles as a function of  $\rho_{xy}$  can be used to estimate the length of the linker required to efficiently result in multi-epitope binding for a given target geometry. With a knowledge of the epitope size and average inter-epitope distances, one can estimate, with a knowledge of the position of minima of the PMF profiles and also their degree of flatness, what length (or range of lengths) of the linker polymer would result in higher avidity. Simulations with bending modulus of the linker matching different polymers used as linkers in practical situations, like various peptides or nucleic acids, can help predict the appropriate stiffness of the linker for a given geometrical arrangement of epitopes. Tumor receptors like HER2<sup>34</sup> have one epitope for natural antibodies while some engineered triple helix proteins called affibodies<sup>35,36</sup> are known to engage a different region on the opposite side<sup>37</sup>. With PMF calculations and knowledge of the position of minima as a function of suitable reaction coordinates, one can predict, using our MD scheme, the linker length that would maximize bispecific binding. At this point we would like to emphasize that, while the methodology would still be applicable, improvements on the current model at the level of the interaction between the nanobody and the plasma membrane would yield more realistic results when the deformation of the latter is also taken into account. To this regard, it should be observed that the simple model of the membrane as a rigid wall could introduce some unrealistic features biasing the comparison between the SPH and SBCG models. Despite these limitations, other factors like nanobody-nanobody interaction and linker properties have importance at small and intermediate values of  $\rho_{xy}$  respectively. Hence we expect that the present setup is good enough to provide a reliable zero-order set of answers regarding the effect of nanobody shape and the determination of appropriate linker lengths for therapeutic fusion proteins.

**Flight/residence times: quantifying the kinetics of the second binding.** As a general rule, the trajectory of the free nanobody can be decomposed into a series of consecutive stretches where it lies close to the wall (residence stretch) or away from it (flight stretch), as specified by some typical interaction distance. The statistics of the flight/residence times (F/RT) of the free NB are crucial observables, as they embody the kinetic determinants of the second binding, and hence can help quantify the avidity. More precisely, the F/RTs are defined as stretches of consecutive frames that the paratope of the free NB (bead P1) spends above/below, respectively, a fixed threshold height  $z_0$  from the tethering wall. Let us denote with  $P_f(t)$  and  $P_r(t)$  the equilibrium distributions of flight and residence times. The corresponding complementary cumulative distributions,  $S_f(t)$  and  $S_r(t)$ , defined as

$$S_f(t) = \int_t^\infty P_f(t') dt'$$

$$S_r(t) = \int_t^\infty P_r(t') dt' \quad (1)$$

coincide with the survival probabilities relating to the corresponding domains,  $\mathcal{D}_f = \{z \in [0, L] | z \geq z_0\}$  and  $\mathcal{D}_r = \{z \in [0, L] | z < z_0\}$ ,  $L$  being the side of the simulation box along the  $z$  direction. We note that the domain  $\mathcal{D}_f$  can be defined as the paratope-wall *interaction* domain. More precisely,  $S_f(t)$  represents the fraction of flight events whose duration exceeds  $t$ , hence this is nothing but the probability that the paratope be still in region  $\mathcal{D}_f$  after a time  $t$ , i.e. indeed the survival probability for domain  $\mathcal{D}_f$ . Analogously,  $S_r(t)$  represents the survival proba-

bility for domain  $\mathcal{D}_r$ . The corresponding distribution of exit times (in the sense of first passage times),  $\mathcal{P}_f(t)$  and  $\mathcal{P}_r(t)$ , can then be computed straightforwardly as<sup>38</sup>

$$\mathcal{P}_f(t) = -\frac{dS_f(t)}{dt}$$

$$\mathcal{P}_r(t) = -\frac{dS_r(t)}{dt} \quad (2)$$

The inverse mean-exit times can be considered as measures of the escape rate from the corresponding domains. Therefore, combining eqs. (1) and (2), it is possible to estimate the on- and off-rates directly from the series of flight and residence times, as

$$k_{on} = \left[ \int_0^\infty t \mathcal{P}_f(t) dt \right]^{-1} = \left[ \int_0^\infty S_f(t) dt \right]^{-1} \quad (3)$$

$$k_{off} = \left[ \int_0^\infty t \mathcal{P}_r(t) dt \right]^{-1} = \left[ \int_0^\infty S_r(t) dt \right]^{-1} \quad (4)$$

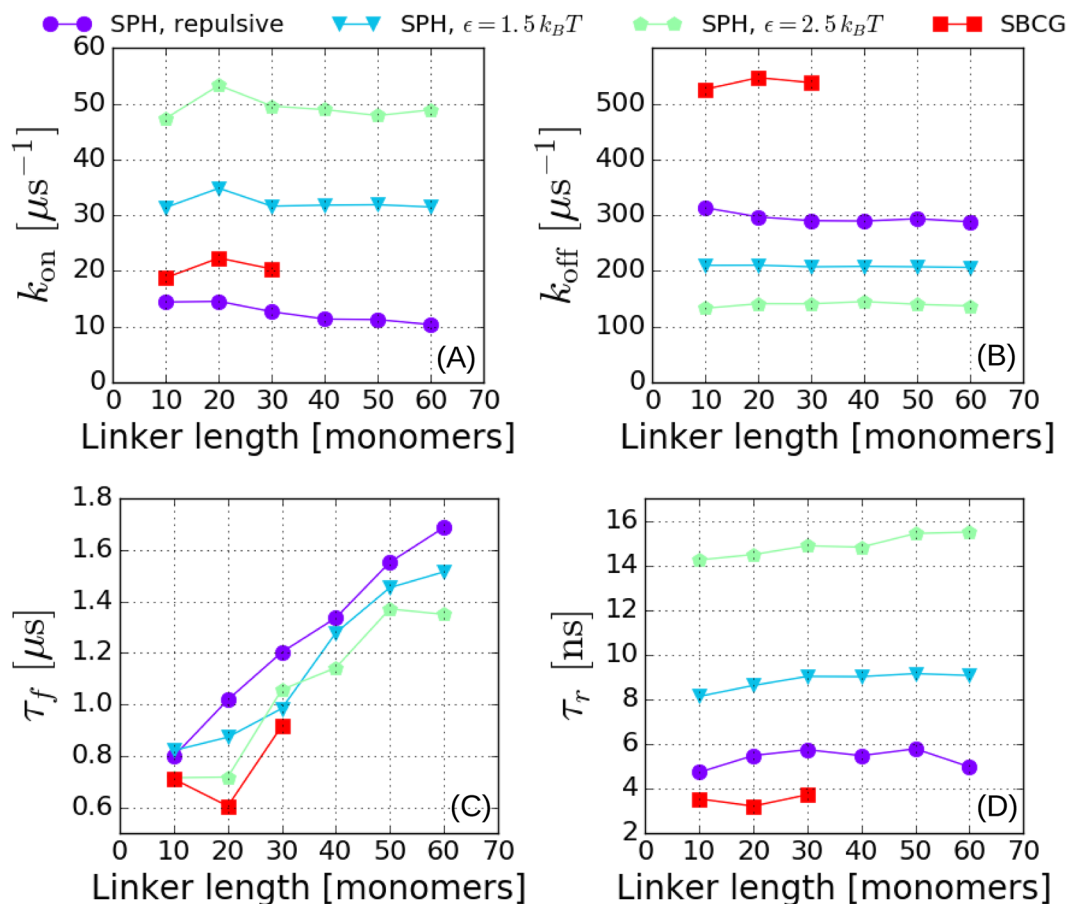
To calculate the flight ( $t_f$ ) and residence times ( $t_r$ ) numerically, the  $z$ -coordinate of the paratope was monitored through the simulation. A threshold  $z$ -height of  $z_0 = 3.5\sigma$  was set to define whether the paratope is in the flight or residence regions. If the paratope stayed above or below the threshold for at least 0.5 ns, the corresponding trajectory stretch was designated to correspond to a flight or residence event, respectively. This additional constraint was tailored specifically to avoid short recrossing events that could bias the statistics unphysically at short times. Data for such events were accumulated over 30  $\mu$ s-long trajectories of 25 non-interacting tethered diabodies. A set of 3 such simulations were performed for each linker length. In addition, the simulations were performed for both repulsive and attractive tethering walls (with attractive energies equal to  $1.5 k_B T$  and  $2.5 k_B T$ ), with the aim of assessing the effect on the paratope-wall kinetics of some weak non-specific attraction between the protein and the wall/membrane.

Figure 4(A,B) illustrate the calculation of the on- and off-rates as described by formulas (3) and (4). The probability per unit time that the paratope enter the interaction domain appears of the order of tens of  $\mu s^{-1}$ , while the probability per unit time that it exit the same domain turns out to be about ten times higher. Interestingly, the SBCG model shows a higher on-rate than the spherical model (with a pure repulsive wall). At the same time, the exit probability is higher for the SBCG model.

Figure 4 also demonstrates that a weak non-specific attraction between the NBs and the wall of the order  $\varepsilon \simeq 1 - 2k_B T$  increases the on-rate (i.e. decreases the overall survival probability in the flight domain) and decreases the off-rate (i.e. makes journeys of the paratope in the interaction domains longer). More specifically, these data refer to a modified SPH-wall system with non-specific isotropic attraction (LJ) between the wall and the diabody elements nbd-1/2 and P1/2. It is interesting to observe that the gain afforded by a weak attractive wall in terms of on-rate, as gauged by  $k_{on}(\varepsilon, N)/k_{on}(0, N)$ , is found to increase with the linker length  $N$ . In fact, while,  $k_{on}(0, N)$  decreases with  $N$ , a weak attraction makes  $k_{on}(\varepsilon, N)$  nearly insensitive to variations in the linker length. This possibly reflects the fact that the entropic cost associated with entering the interaction domain decreases with increasing linker length in the presence of attraction between the paratope/NB system and the wall. Conversely, the ratio  $k_{off}(\varepsilon, N)/k_{off}(0, N)$  appears to remain constant as  $N$  increases. According to our definition,  $k_{off}$  measures the inverse average time required for a NB touching the wall to leave the surface. This is associated with large fluctuations in the  $z$ -direction (orthogonal to the wall). In such configurations, it is likely that a restoring entropic force exists acting parallel to the surface (in the  $x - y$  plane). It could be surmised that, to first order in the typical amplitude of such fluctuations, the vertical component of the entropic force along  $z$  is negligible, rendering the off-rates independent of the linker length.

It is instructive to inspect in more detail the survival probabilities. Fig. S4 of SI depicts the survival curves for both the flight and paratope-wall interaction domain. The short time behavior ( $t \lesssim 0.5 \mu s$ ) appears to follow an inverse power law with exponent 1/2 (see dashed lines in Fig. S4), irrespective of the linker length. This is the expected behaviour for the survival probability of a free random walk in three dimensions<sup>38</sup>. This means that short survival times in either domain are dominated by the unconstrained diffusion of the paratope. By contrast, longer survival events depend markedly on the length of the linker and are distributed exponentially. The inset in Fig. S4 makes this point very clear in the case of the function  $S_f(t)$  for the SPH model. We find that this is a general feature of the tails of the paratope survival probabilities in either region. As it shows from the figures, the tails of the flight times depend on the linker length, while those of the residence times much less so. In order to have some insight into the tail, rare-event region, we might reason as follows:

We can safely assume that rare events make uncorrelated time series. In this case, the statistics of return events will be specified by the configurational probability (independent of time) that the paratope be in the relevant regions, either  $z \geq z_0$  or  $z < z_0$ , for a given threshold height  $z_0$  above the wall. In turn, this will depend on the conformational statistics of the NB-polymer systems and, in the absence of an appropriate analytical model, can be easily determined from our PMF calculations (see Fig. 2). Let us denote with  $P_+$ ,  $P_-$  the equilibrium probability that the free NB be above or below the threshold  $z_0$ , respectively. In this case, the probability  $P_a(k)$  and  $P_b(k)$  of observing  $k$  consecutive sampled frames above (a) or below (b)  $z_0$ , respectively, can be computed as



**Figure 4.** Kinetic parameters of wall-paratope interaction. (A) On and (B) off-rates describing the kinetics of paratope entering and exiting the interaction domain  $\mathcal{D}_r$  (i.e.  $z < z_0$ ), computed according to the prescriptions (3) and (4). Time constants of the exponential tails of the survival probabilities in the (C) flight and (D) residence regimes. These can be thought of as the average time of the rare, very long events. The data for the SPH model refer to the simulation set #2 (see Table 1). The data for the SBCG model refer to a purely repulsive wall.

Simulation type	System	wall parameter (SPH nbd-1/2)	wall parameter (other beads)	$\epsilon$
Repulsive wall #1	SPH (10-, 20-, 30-, 40-, 50-, 60-mer)	$s_{wall} = 4.5/3.0$ $r_{cutwall} = 5.05/3.36$	$s_{wall} = 0.8$ $r_{cutwall} = 0.89$	0
	SBCG (10-, 20-, 30-mer)	n. a.	$s_{wall} = 0.8$ $r_{cutwall} = 0.89$	0
Repulsive wall #2	SPH (10-, 20-, 30-, 40-, 50-, 60-mer)	$s_{wall} = 4.5/4.5$ $r_{cutwall} = 5.05/5.05$	$s_{wall} = 0.8$ $r_{cutwall} = 0.89$	0
Attractive wall #1	SPH (10-, 20-, 30-, 40-, 50-, 60-mer)	$s_{wall} = 4.5/4.5$ $r_{cutwall} = 11.25/11.25$	$s_{wall} = 0.8$ $r_{cutwall} = 2.0$	$5k_B T$
Attractive wall #2	SPH (10-, 20-, 30-, 40-, 50-, 60-mer)	$s_{wall} = 4.5/4.5$ $r_{cutwall} = 11.25/11.25$	$s_{wall} = 0.8$ $r_{cutwall} = 2.0$	$5k_B T$

**Table 1.** Simulation details. Four sets of simulations were performed for calculating the flight/residence times, corresponding to different choices of the parameters describing the interaction between CG beads forming the NBs and the wall (see eqn. 16).

$$P_a(k) = P_{>}(1 - P_{<})^{k-1} P_{<} \simeq P_{<} e^{-P_{<} k} \tag{5}$$

$$P_b(k) = P_{<}(1 - P_{>})^{k-1} P_{>} \simeq P_{>} e^{-P_{>} k} \tag{6}$$



If for the sake of the argument we take the Gaussian tethered model for the  $z$ -distribution of the free end of a tethered polymer, given by:

$$P(z) = \frac{3z}{\langle d^2 \rangle} e^{-3z^2/2\langle d^2 \rangle} \quad (7)$$

where  $\langle d^2 \rangle$  is the average square end-to-end distance of the free polymer, it is readily seen that

$$P_{>} = \int_{z_0}^{+\infty} P(z) dz = e^{-3z_0^2/2\langle d^2 \rangle} \quad (8)$$

$$P_{<} = \int_0^{z_0} P(z) dz = 1 - e^{-3z_0^2/2\langle d^2 \rangle} \quad (9)$$

If we introduce the time decay constants  $\tau_f, \tau_r$  of the exponential tails of the survival probabilities in the flight and interaction domains, respectively, Eqs. (8) and (9) entail

$$\tau_f = \frac{\Delta t_c}{P_{<}} = \frac{\Delta t_c}{1 - e^{-3z_0^2/2\langle d^2 \rangle}} \simeq \Delta t_c \frac{2\langle d^2 \rangle}{3z_0^2} \quad (10)$$

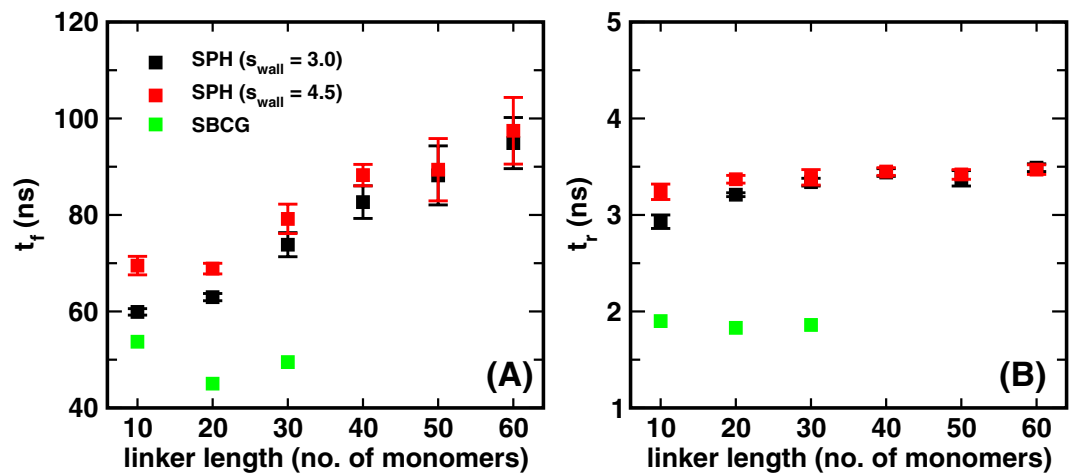
$$\tau_r = \frac{\Delta t_c}{P_{>}} = \Delta t_c e^{3z_0^2/2\langle d^2 \rangle} \simeq \Delta t_c \left( 1 + \frac{3z_0^2}{2\langle d^2 \rangle} \right) \quad (11)$$

where  $\Delta t_c$  is a time of the order of the typical correlation time of consecutive frames and in the last passages we have made use of the fact that  $z_0^2/\langle d^2 \rangle \ll 1$ .

Figure 4(C) indeed shows that  $\tau_f$  increases linearly with  $N$  ( $\langle d^2 \rangle \propto N$ ) for the spherical model, according to the prediction (10). It is interesting to observe that the slope does not seem to depend on the value of the weak attractive energy  $\varepsilon$ . This is expected, as rare, long flight events are dominated by the statistical weight of the configurations of the combined NB/linker molecule away from the wall. The simple calculation leading to Eq. (11) also correctly explains the observed reduced variability of residence times as the linker length is increased (see again Fig. 4(D)). However, a closer inspection reveals that the average duration of rare, long residence events increases with the linker length  $N$  in the presence of an attractive interaction between the paratope/NB system and the wall, even though with a much smaller slope than for the increase of  $\tau_f$ . Overall, we conclude that the effect of a weak non-specific interaction with the wall decreases the average duration of rare, long flights and residence events. It is interesting to observe that the SBCG model (with a repulsive wall) displays the shortest duration of rare long flights (Fig. 4(C)), even shorter than for the spherical model in the presence of the most attractive wall ( $\varepsilon = 2.5 k_B T$ ). Moreover, there seems to be an optimum (a pronounced dip) at a linker length of  $N = 20$ . The presence of this optimum could originate from the steric repulsion that nbd-1 experiences because of the presence of nbd-2 (see Fig. S6 of SI) in the case of smaller linkers (10-mer), resulting in an increase in the average length of flight stretches and an increase in the occurrence of very long flight events. Owing to its shape, the effect is more pronounced in the case of the SBCG model. While rare, long flight and residence times are on the  $\mu s$  and tens of ns scales (exponential tails), respectively, the average values  $\langle t_f \rangle = \frac{1}{k_{on}}$ ,  $\langle t_r \rangle = \frac{1}{k_{off}}$  are dominated by the short-time power-law behaviour. Figure 5 reveals that average flight times turn out to be of the order of about  $10^2$  ns, while residence times are about 50 times shorter, of the order of 2-3 ns.

In order to investigate further the details of the NB-wall interactions, we performed the simulations with repulsive tethering walls for two different values of the LJ repulsive length  $s_{wall}$  ( $4.5\sigma$  and  $3.0\sigma$ ) for nbd-2 for the spherical model (see methods). We observe that for smaller linker lengths (10–30-mers), flight times are shorter for a smaller value of  $s_{wall}$ . The effect is negligible for the longer linkers, and seems, in fact, inversely proportional to the linker length (see Fig. 5). The effect arises because the value of  $s_{wall}$  determines how nbd-2 (the tethered NB) would interact with the tethering wall and how much it is able to bend. This is expected to affect the value of  $\langle t_f \rangle$ , more so in presence of the external harmonic potential that restrains the value of the angle between P2-CB2-L1 (the axis of the tethered sphere) to  $180^\circ$  (see methods). One would then expect that how much nbd-2 can bend would depend on the shape of the nanobody, as the shape would then determine the dynamics and effective interaction of the nanobody with the tethering wall.

The bending propensity of the tethered unit can be gauged by calculating the height distribution of CB2 (see Fig. 1(C)) measured from the tethering wall surface for the SPH and the SBCG systems (see Fig. S7 of SI). From the plot we see that the very low heights have a significant probability in the case of the SBCG diabodies, which may contribute to the lower values of  $\langle t_f \rangle$ . By contrast, for the SPH diabodies there is an abrupt lower cutoff depending on the value of  $s_{wall}$ . From Fig. 5 we notice that the values of  $\langle t_f \rangle$  for the SBCG diabodies are substantially shorter as compared to the SPH diabodies. For the 30-mer linker, for example, the SPH system has  $\langle t_f \rangle \sim 80$  ns while for the SBCG system  $\langle t_f \rangle \sim 50$  ns. It thus seems that it is not an obvious task to reproduce the interaction with a wall within a model that preserves the shape of the atomistic structure of the NB via a simple spherical representation. Thus the SBCG model seems more appropriate to calculate relevant dynamical parameters, such as the on-rate for second binding, that are expected to rely substantially on the details of the interactions with a wall/membrane.



**Figure 5.** Dependence of flight and residence times on CG model and wall softness. Average flight times (A) and average residence times (B), computed over the trajectories of 25 non-interacting, tethered diabodies. A distance cutoff  $z_0 = 3.5\sigma$  between the tethering wall and the free paratope (P1) was used to calculate the values. A jump above (a plunge below) the cutoff was considered to be a flight (residence) event only if it lasted for more than 0.5 ns. The systems compared comprise two different variants of the SPH model, as described in Table 1, and the SBCG system.

A closer look at the average residence times reveals that, while  $\langle t_r \rangle$  shows negligible dependence on the linker length, the dependence on the model is noticeable. The residence time for the SPH diabodies is  $\sim 3.5$  ns, while the values for the SBCG diabodies is close to 2 ns. The difference can be attributed to the fact that the SBCG nanobody likely generates larger reaction forces from the tethering walls on behalf of its fluctuating structure (see Fig. S2 of SI), thus reducing the time it stays near the wall. By contrast, the SPH nanobody would face a smaller reaction from the wall, given its rigid and fixed surface. Here again one can appreciate the importance of the model being used. All the data for  $t_f$  and  $t_r$  for this set of simulations are reported in tabulated form in tables S1 to S5 in the SI.

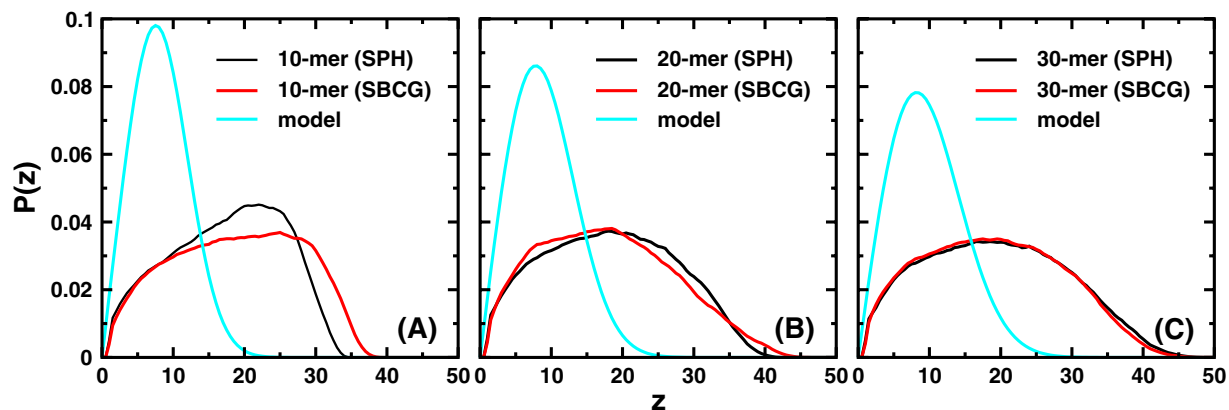
**z-distribution of the free paratope: comparison with models.** An important observable that is tightly related to the statistics of flight and residence times is the distribution of the  $z$ -coordinate of the paratope of the free nanobody (bead P1). We computed these distributions for all the systems from the simulations performed to calculate the flight and residence times (see methods). It is interesting to compare the data for the SPH and SBCG systems with a simple model. The expression for the normalized equilibrium  $z$ -distribution of the free end of a Gaussian polymer, tethered at a height  $z_t$  from a reflecting wall, reads<sup>39</sup>

$$P(z) = \frac{3}{\sqrt{2\pi N_k b^2}} \frac{1}{\operatorname{erf}(z_t/\sqrt{3/2 N_k b^2})} [e^{-3(z-z_t)^2/2N_k b^2} - e^{-3(z+z_t)^2/2N_k b^2}] \quad (12)$$

where  $N_k$  is the number of Kuhn monomers and  $b$  is the Kuhn length and  $z$  is the  $z$ -coordinate of the free end measured from the wall. In the limit  $z_t \rightarrow 0$ , the expression reduces to Eq. (7)<sup>40</sup>.

It is interesting to inquire whether the data for the composite two-NB models can be described by an *effective* polymer model. The most obvious choice would be a Gaussian polymer with the same flexibility as the diabody linker, contour length equal to the contour length of the diabody measured from CB2 to the free paratope (P1) and tethered at the average height of CB2 above the lower  $z$  wall. A reasonable guess for the length of the effective model is thus  $N_k = (N + 2r)/b$ , where  $N$  is the number of monomers in the linker,  $r$  is half the distance between P1 and CB1 and  $b$  is the Kuhn length of the linker polymer (see again Fig. 1). We set  $z_t = 7.5\sigma$ , which is the average height for CB2 measured from the wall for the SPH system (see Fig. S7 of SI). This would represent a polymer tethered at  $z_t = 7.5\sigma$  distance units above the tethering wall and having a contour length equal to the linker and nbd-1 combined. Our linker has a persistence length of  $\ell_p \simeq 1.5\sigma$  (see page S1 of SI), which leads to a Kuhn length  $b = 2\ell_p = 3\sigma$ .

Figure 6 shows the comparison of the effective model with the data for the SPH and the SBCG systems. It is apparent that the distribution for the diabodies describes a higher representation of large  $z$  values and a reduced representation of small  $z$  values when compared to the tethered polymer. This is an expected consequence of the higher entropic repulsive force exerted by the wall on the free bead, a mechanism akin to the entropic pulling force demonstrated in the disassembling and translocating action of Hsp70s chaperones<sup>41</sup>. More precisely, the conformational space near the wall is restricted more severely for the diabodies, due to presence of the nanobodies at the two ends of the polymer, than for a *bare* polymer with the same contour length and flexibility. The extent of the difference between the simulation and the model suggests extended, composite molecules such as our diabodies, belong to a different universality class altogether. Note that the results would be different if we added convexity to the wall. With an increase in the convexity of the wall (as will be the case when mimicking cells of



**Figure 6.**  $z$ -distribution of P1.  $z$ -distribution of the free paratope (bead P1) for (A) 10- (B) 20- and (C) 30-mer SPH and SBCG systems plotted along with the equivalent Gaussian polymer approximation (12) (see text).

increasingly smaller sizes) the distributions are expected to approach those corresponding to the end-to-end distance of the Gaussian chain. One can, however, use Eqs. (12) or (7) to determine the Kuhn length  $N_k^{eff}$  of an effective polymer with the same persistence length  $\ell_p$  as the linker in the diabodies. For this, we fit the distributions of the 40-, 50- and 60-mer SPH systems with Eq. (7) (see Fig. S8 of SI) with  $b = 3\sigma$ , which leads to  $N_k^{eff} \sim 108, 115$  and  $119$  for 40-, 50- and 60-mer diabodies respectively.

It is interesting to note that the distributions for the SBCG and the SPH systems differ to a highest extent when the linker length matches the dimensions of the linked NBs, i.e. for the shortest linker (10-mer), while they approach each other rapidly as the linker length increases and almost overlap for the 30-mer linker. This means that, as the statistics of the vertical coordinate above the wall is concerned, for linkers longer than approximately the size of the NBs, an equivalent SPH model can be used, entailing considerable simplification of the simulations. Having said that, we would expect that the SBCG model would still be a better model to represent the NBs. The therapeutic molecules that we are trying to model may, in certain situations, interact with the surface of trans-membrane protein complexes away from the plasma membrane (the molecules modeled for intra-spike binding to the gp120 trimer on the HIV surface for example). In such a situation, the shape-based model would better represent the subtle details of the nanobody-target interaction. In addition, as we have discussed earlier, after one of the nanobodies attaches to the target, the (in-plane) reach of the free nanobody depends on how much the bound nanobody can bend. The magnitude of this bending will strongly depend on the shape of the nanobody and hence the shape-based model would be more appropriate. These effects will be observed even in presence of a more realistic plasma membrane with undulations on its surface, a detail that we have not considered in the present study.

## Conclusions and perspectives

In this work we have introduced two different coarse-grained models of polymer-linked, two-nanobody molecules, as a simple but paradigmatic example of novel immunotherapy agents that are increasingly being developed in a variety of contexts. More precisely, while the linker has been modeled invariably as a bead-and-spring system (stretching and bending), the nanobodies have been represented either as single large rigid spheres, or as collections of small spheres suitably connected by springs. Such representation was parameterized in a bottom-up philosophy directly from atomistic simulations in explicit solvent. The latter scheme led to binding units displaying the same *shape* and large-scale flexibility as the atomistic systems.

The aim of this work was to lay the bases of coarse-grained, computationally aided drug design in this area from the firm standpoint of statistical and computational physics. In the spirit of the accepted model of sequential binding of bivalent molecules<sup>42</sup>, whereby bivalent agents first bind to a target-covered surface from the bulk, and subsequently dynamically explore the surroundings for a second target within reach (either on the same or on a facing surface), our main focus was to elucidate the physics of the latter kinetic step. For this purpose, we have mainly focussed on the kinetic and equilibrium properties of a molecule tethered to a wall through one of its binding units (NB), in order to investigate the main dynamical and structural determinants of the second binding. In particular, we aimed at investigating (i) to what extent the degree of coarse-graining may impact the dynamics of the combined linker/free NB system and (ii) the interaction dynamics of the free paratope (the active binding site carried by the NB) with the surface.

In the first part of this work, we have shown that the calculation of potentials of mean force (PMF) along specific, one-dimensional collective coordinates can provide considerable insight into the nature, strength and range of the thermodynamic forces that govern the motion of the free paratope. Furthermore, such calculations may constitute a precious tool to investigate the role of such forces in the dynamics of the paratope-wall encounter, which, in turn, governs the kinetics of the second binding. We aim at illustrating this aspect in a forthcoming publication. For example, the PMF can be fruitfully used as an effective potential in approaches based on the Smoluchowski equation or on first-passage processes<sup>38</sup>.

In the second part of the paper, we have delved into the kinetics of the paratope-wall encounters. To this end, we have developed a general strategy based on dissecting an equilibrium trajectory of the free paratope in flight and residence stretches, depending on whether the active site on the free NB was found above or below, respectively, a critical interaction threshold close to the wall in the vertical direction. We have shown that the encounter and escape kinetics with respect to the wall are simply related to the survival probability of the paratope in the flight and interaction domains, respectively, which can be simply computed from the series of flight and residence times observed over a long MD trajectory. We have illustrated how this method allows one to estimate the kinetic constants of the second binding,  $k_{on}$  and  $k_{off}$ , in the presence of a purely repulsive wall and with weak, non-specific interactions between the free NB and the wall. Our simple method not only allows one to quantify the role of factors such as the linker length and flexibility (not considered in this study) and non-specific interactions on the average flight/residence times, but also makes clear and quantifies the role and weight of rare, long-duration events that show up in the exponential tails of the survival probabilities.

It is worth stressing that the statistics of rare events is by no means a secondary issue in this context, as in many real-life situations the binding kinetics of such molecules is expected to be dominated by fluctuations, e.g. due to low copy numbers or tiny reaction volumes. For example, this is the case of novel bivalent and bispecific diabodies engineered to bind within the immune synapse, i.e. at the interface of two facing membranes, on the effector cell (NK or B-cell) on one side and on epitopes on a tumor cell on the other side. The synapse covering an area of the order of  $100\ \mu\text{m}^2$  for a cell-cell separation of about  $15\ \text{nm}$ <sup>24,25</sup>, the role of fluctuations in the number of bridging molecules is expected to be important, which likely makes the statistics of rare events a major determinant of the binding kinetics.

As mentioned before, while the methodology described in this work can be used to derive inputs for the design of bivalent/bispecific therapeutic molecules, there are some critical details that one can incorporate to obtain results that could be quantitatively more accurate. One of the improvements could be in the description of the plasma membrane. While we have approximated the membrane as a repulsive/attractive wall, it would be interesting to understand how a more detailed description in terms of its surface charge, shape and structural anisotropy (spatial and temporal) could affect the results, and would be a useful extension to the present work.

## Methods

**Coarse-graining schemes and simulations.** All the data reported in this work were generated from Langevin dynamics simulations performed with LAMMPS<sup>43</sup> in the Lennard-Jones (LJ) unit system, with the unit of length being  $\sigma = 3.5\ \text{\AA}$  (the size of one monomer of the linker) and the unit of energy being  $\varepsilon = 100\ \text{K}$ . All the simulations were conducted via Langevin dynamics in the overdamped regime.

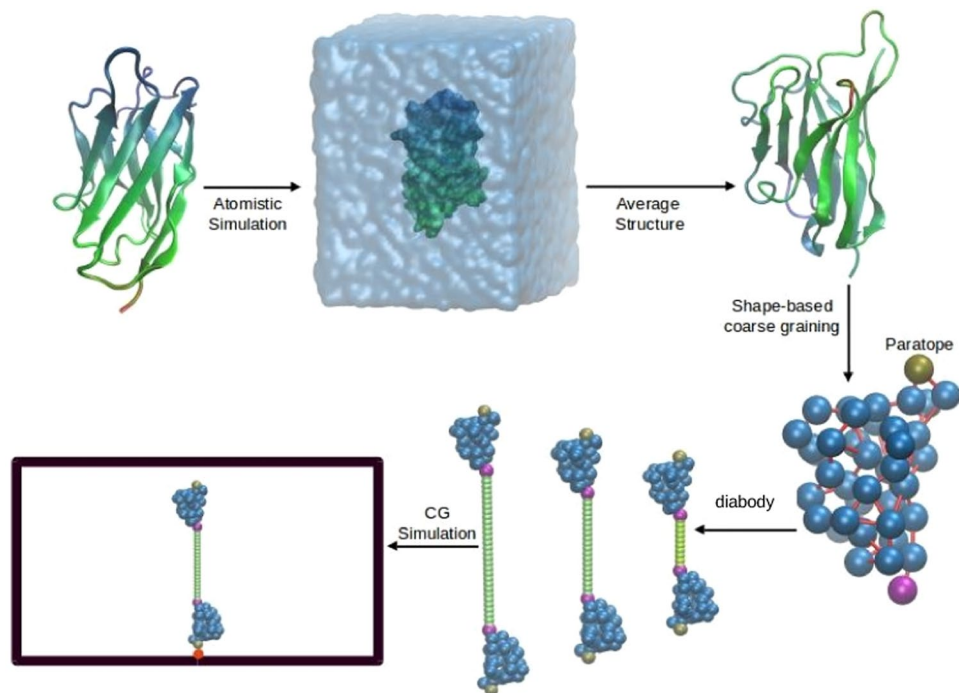
In this work, we considered two different coarse-grained representations of a two-NB molecule joined by a flexible linker. In the first approach (SPH model), the NBs were modeled as rigid spheres of diameter 10 (in units of linker monomers), decorated with two smaller fixed spheres at two diametrically opposite ends, representing the NB-linker connecting unit and the paratope, respectively (see Fig. 1). The paratope bead has diameter 1.6 and the connector bead has diameter 1. A bead-spring polymer bridges the two connector beads. The beads constituting the polymer linker have unit diameter. This model has been referred to in the text as the SPH model.

The second model was meant to reproduce the *shape* and large-scale flexibility of the NBs. For that we used the shape-based coarse-graining scheme developed in Schulten's group<sup>33</sup>. This procedure requires a trajectory from an atomistic equilibrium MD simulation to be sampled and fed as an input. The crystal structure with pdb id: 1qd0<sup>44</sup> was used as the starting structure for the MD simulation to generate the input structure. This is a camelid heavy chain variable (VHH) domain, in complex with a RR6 dye dimer.

The dye was removed from the complex and the remaining protein was solvated in TIP3P water with a 20  $\text{\AA}$  buffer, leading to a system size of 40298 atoms, with 13284 water molecules. The solvated system was then neutralized by adding 5  $\text{Cl}^-$  ions to generate the starting configuration for the MD simulation (see Fig. 7). The system was minimized for 10000 steps using the conjugate gradient method. During minimization, all the atoms in the protein were constrained to their starting positions. This allowed water molecules to re-organize and eliminate unfavorable contacts with the protein. After minimization, the atoms were assigned velocities generated from a Maxwell distribution at 300 K. The particle mesh Ewald (PME) method with a real space cut-off of  $12\ \text{\AA}$  was used to estimate the energy component from the long-range electrostatic interaction. The system was simulated for 100 ns in the NPT ensemble. A Langevin thermostat was used<sup>45</sup> with a temperature coupling constant of  $5\ \text{ps}^{-1}$ , while the pressure was regulated using the Langevin barostat with a pressure coupling constant of  $50\ \text{ps}^{-1}$ . The simulation was performed using NAMD<sup>46</sup> and the CHARMM 27<sup>47</sup> force field was used to describe the protein. The MD trajectories were visualized using VMD<sup>48</sup>.

An average structure was generated from the snapshots from the last 10 ns of the MD trajectory (see Fig. 7). The average structure of the nanobody was used to generate a shape-based coarse grained (SBCG) model using the procedure formulated by Schulten *et al.*<sup>33</sup>. This scheme uses a topology-conserving algorithm developed for neural networks to generate a coarse-grained representation that reproduces the shape of the protein. In a trade-off between the system size and a good representation of the protein shape, we used 40 beads to represent the 126 residue protein. To generate the system to be simulated, two of the CG proteins were connected by a polymer linker with monomer diameter 1, similar to the one used in the SPH model (see Fig. 1). It is to be noted that the diameter of the spherical bead that represents the nanobody in the SPH model is nearly equal to the geometric average of the major and minor axes of the roughly spheroidal SBCG nanobody. In this sense the two models are equivalent and comparable.

**Interaction parameters.** In general, the total interaction potential of our coarse-grained models had bond, angle and van der Waals (vdW) terms, given by



**Figure 7.** The coarse-graining procedure. The crystal structure (PDB id. 1qd0) is used as the starting configuration for the explicit solvent atomistic simulation. Snapshots from the last 10 ns of the trajectory are used to generate an average structure. The average structure is then coarse-grained using the shape-based coarse-graining method. A pair of the coarse-grained structures are connected by a linker (10-, 20- and 30-mer), enclosed in a box with  $x$ - $y$  periodicity, aligned along the  $z$ -direction and tethered to the box base, thus generating the starting configuration for CG simulations.

$$V_{ij}^{bond} = \frac{1}{2}k_b(r_{ij} - r_0)^2 \quad (13)$$

$$V_{ijk}^{angle} = \frac{1}{2}k_\theta(\theta_{ijk} - \theta_0)^2 \quad (14)$$

$$V_{ij}^{vdW} = 4\varepsilon_{ij} \left[ \left( \frac{s_{ij}}{r_{ij}} \right)^{12} - \left( \frac{s_{ij}}{r_{ij}} \right)^6 \right] \quad (15)$$

where  $k_b$  and  $k_\theta$  are the bond force constant and the angle bending energy, respectively,  $r_0$  and  $\theta_0$  are the equilibrium bond length and angle, respectively,  $\varepsilon_{ij}$  is the LJ interaction energy and  $s$  is the inter-bead distance at which the LJ potential becomes repulsive (referred to as *repulsive* length in the following), which depends on the combined radii of the interacting beads. In Eqs. (13), (14) and (15),  $i$ ,  $j$  and  $k$  may be beads belonging to a coarse-grained nanobody (including the connector and paratope beads) or to the linker.

**Interaction parameters for the SBCG nanobodies and the linker.** All the CG beads were kept neutral. The connectivity and spring constants for the bonds between the beads were used as generated by the SBCG scheme. It is to be noted that the bonds are not set by a distance-based cut-off scheme, but are in accordance with the bonds present in the atomistic system. This helps in maintaining the flexibility of the nanobody and providing the required flexibility to the loop regions, which may play a defining role in the kinetics of the diabody. Repulsive vdW interactions among the beads were also introduced to ensure that the shape is maintained during the course of the simulation. The vdW radii of the protein beads were generated by comparing the masses of the protein beads to that of the PEG monomer. The repulsion between the constituent beads of the nanobody was represented by a Weeks-Chandler-Andersen potential (WCA)<sup>49</sup>, i.e. a shifted LJ potential cut off at the minimum, i.e.  $r_{cut} = 2^{1/6}\sigma$ . The angle parameters for the nanobody beads were used as generated by the SBCG scheme.

The linker is represented as a freely-jointed chain with two-body harmonic bond-stretching and three-body harmonic bond-bending potentials. The masses and diameters of the polymer beads were set to 1. The vdW interaction between the bead pairs was set to purely repulsive as represented by a WCA potential. The force constant of the (stiff) bond-stretching potential was set to 54 N/m, which corresponds to an average fluctuation of the bonds at room temperature of about 2% of the equilibrium length. In order to estimate the appropriate value of

the bending rigidity  $k_\theta$ , it is expedient to refer to the calculation of the persistence length  $\ell_p$  for the freely jointed chain with angle-bending interactions in the hypothesis of zero correlation between bending and torsion degrees of freedom. Referring to published data for PEG<sup>50</sup>, we fixed  $k_\theta = 1.8k_B T$  as the bending coefficient for the linker (see page S1 of SI for a detailed discussion).

**Interaction of the beads with the box walls.** The walls of the simulation box in the  $x$  and  $y$  directions have periodic boundary condition, while the  $z$ -walls are fixed. The beads interact with the  $z$ -walls via a LJ (12–6) potential given by:

$$V_{wall}^{vdW} = 4\varepsilon \left[ \left( \frac{s_{wall}}{r_{wall}} \right)^{12} - \left( \frac{s_{wall}}{r_{wall}} \right)^6 \right] \quad (16)$$

Here  $r_{wall}$  is the distance of the center of any bead from the wall. The potential energy (Eq. 16) is cut off at  $r = r_{cutwall}$ , which defines the nature of the wall ( $r_{cutwall} = 2^{1/6} s_{wall}$  for a purely repulsive wall,  $r_{cutwall} = 2.5 s_{wall}$  for an attractive wall). The wall interaction parameters,  $s_{wall}$ ,  $r_{cutwall}$  and  $\varepsilon$  define the nature of interaction between different beads and the wall. While the interaction of the nanobodies and paratopes with the  $z$ -walls was set to be either repulsive or attractive in different simulations, the linker and connector beads always had repulsive interaction with the walls.

**Preparing the systems for simulation.** The first set of simulations reported involve the calculation of potentials of mean force (PMF) for the diabodies as a function of various reaction coordinates. To perform the PMF calculations, the diabodies were enclosed in cuboidal boxes which were periodic in the  $x$  and  $y$  directions, while the  $z$ -walls were fixed and repulsive. The diabodies were tethered to the lower  $z$ -wall by imposing an attractive LJ interaction between the paratope of one of the nanobodies and a fixed epitope bead attached to the lower  $z$ -wall. The simulations were performed for linker lengths of 10, 20, 30, 40 and 50 monomers for the SPH system and 10, 20 and 30 monomers for the SBCG system.

The second kind of simulations were performed on a collection of diabodies to calculate dynamical parameters.  $N = 25$  diabodies were tethered to the lower  $z$ -wall of a cuboidal box. The tethering points were arranged in a  $5 \times 5$  lattice (see Fig. S9 of SI). Again, the  $x$  and  $y$  directions had periodic boundary conditions, while the lower  $z$ -wall was fixed and either perfectly reflecting or attractive. The linker lengths for different systems were similar to those considered for the PMF calculation, with an additional linker length of 60-mer for the SPH system.

In addition to the various bonded and non-bonded interactions described in the previous section, various extra angle restraints were introduced. For example, for the 10-mer diabody, the angles L10-CB1-P1, P2-CB2-L1, L2-L1-CB2 and L9-L10-CB1 (see Fig. 1(C)) were restrained to  $180^\circ$  using a harmonic angle bending potential of the form of Eq. (14) with  $\theta_0 = 180^\circ$  employing stiffer bending coefficients as compared to the angles corresponding to the linker. Similar restraints were employed for diabodies with all other linker lengths using the equivalent beads. On top of this, for the SBCG systems, the nanobodies were restrained from rotating about their respective long axes by restraining a dihedral formed by L10-CB1 (L1-CB2) and two beads of nbd-1 (nbd-2) to their starting values throughout the simulation. The extra restraints were introduced to mimic the fact that, in real-life systems, the bonds between the linker and the nanobody restrict the angular motion of the nanobody about the CB1/2-P1/2 axis.

**Calculation of the potentials of mean force.** The PMF calculations have been performed using the umbrella sampling (US) technique. The values of the reaction coordinate (RC) varied from 3 to 30, 40, 50, 60 and 70  $\sigma$  for the 10-, 20- and 30-, 40- and 50-mer linker systems, respectively, with windows at gaps of  $0.5\sigma$ , leading to 55, 75, 95, 115 and 135 simulation windows for the SPH systems. For the SBCG systems, the RC values varied from 5 to 35, 45 and 55  $\sigma$  for the 10-, 20- and 30-mer linker systems, leading to 61, 81 and 101 windows. The US simulations in different windows were performed in parallel. To generate the starting structure for each window, a short simulation was performed with the free nanobody being dragged from the initial to the final value of the reaction coordinate with an equilibration time of  $5 \times 10^5$  steps at each value, and the final snapshots at each value were used as the starting structures for different windows. Starting from the structures thus generated, simulations lasting for  $5 \times 10^7$  time steps were performed in each US window. The weighed histogram analysis method (WHAM)<sup>51</sup> was used to generate the PMF profiles from the biased distributions of RCs obtained from the US windows.

**Calculation of flight/residence times.** The different kinds of simulations performed to calculate the flight and residence times are listed in Table 1. We simulated a system consisting of a group of 25 diabodies arranged in a  $5 \times 5$  square lattice, tethered to the lower  $z$ -wall (see Fig. S9 of SI). The distance between two neighboring diabodies was set such that no interaction would be possible between them at any time. An initial simulation of 15  $\mu s$  was performed. The final structure was used to perform three independent simulations lasting for 30  $\mu s$  for the SPH system (linker lengths: 10–60). For the SBCG system (linker lengths: 10–30) one single simulation lasting 15  $\mu s$  was performed for each linker length for comparison. In these simulations, the  $z$ -walls were repulsive. Additional simulations were performed for the SPH system for all different linker lengths, with slightly attractive  $z$ -walls. The value of  $\varepsilon$  for these simulations was set to  $1.5 k_B T$  and  $2.5 k_B T$  for two different sets of simulations.

The  $z$ -coordinate of P1 was recorded and a threshold  $z_{th} = 3.5\sigma$  between P1 and the tethering wall was set to distinguish between flight and residence. More precisely, a series of consecutive simulation frames during which P1 stayed below the threshold was considered to be a residence event, while flight events (and corresponding times) were associated with consecutive frames where the paratope remained above the threshold. Taking an average over the 25 nanobodies (nbd-1) and three independent simulations, we computed the average flight and residence times and also calculated the corresponding distributions.

To ascertain that the length of the simulations was sufficient to arrive at converged values of  $\langle t_f \rangle$  and  $\langle t_r \rangle$ , we performed the calculation for different durations of the simulations, and checked how the calculated values changed as a function of the simulation length. For the 30-mer linker, for example,  $\langle t_f \rangle$  started from a value of  $85 \pm 22$  ns for a simulation length of  $3.75 \mu\text{s}$  and slowly converged to the reported value of  $\sim 77$  ns for a simulation length of  $15 \mu\text{s}$  and stayed close to that for longer simulation lengths, suggesting that our simulation length of  $30 \mu\text{s}$  is appropriate for producing converged results. The values of  $\langle t_f \rangle$  as a function of simulation length for the SPH systems with  $N = 30$  and  $N = 60$  linkers are shown in Fig. S10 of SI.

Received: 27 October 2019; Accepted: 7 April 2020;

Published online: 14 May 2020

## References

- Desmyter, A., Spinelli, S., Roussel, A. & Cambillau, C. Camelid nanobodies: killing two birds with one stone. *Curr. Opin. Struct. Biol.* **32**, 1–8 (2015).
- Chames, P., Regenmortel, M. V., Weiss, E. & Baty, D. Therapeutic antibodies: successes, limitations and hopes for the future. *Br. J. Pharmacol.* **157**, 220–233 (2009).
- Behar, G. *et al.* Isolation and characterization of anti-Fc $\gamma$ RIII (CD16) llama single-domain antibodies that activate natural killer cells. *Protein Eng. Des. and Sel.* **21**, 1–10 (2008).
- Turini, M., Chames, P., Bruhns, P., Baty, D. & Kerfele, B. *Oncotarget* **5**(14), 5304–5319 (2014).
- Bannas, P. *et al.* Validation of nanobody and antibody based *in vivo* tumor xenograft NIRF-imaging experiments in mice using *ex vivo* flow cytometry and microscopy. *J. Vis. Exp.* **98**, e52462 (2015).
- Bannas, P. *et al.* Molecular imaging of tumors with nanobodies and antibodies: Timing and dosage are crucial factors for improved *in vivo* detection. *Contrast Media Mol. Imaging* **10**(5), 367–78 (2015).
- Kijanka, M., Dorresteijn, B. & Oliveira, S. & van Bergen en Henegouwen, P. M. Nanobody-based cancer therapy of solid tumors. *Nanomedicine (Lond)* **10**, 161–74 (2015).
- Bannas, P. *et al.* Molecular imaging of tumors with nanobodies and antibodies: Timing and dosage are crucial factors for improved *in vivo* detection. *Contrast Media Mol. Imaging* **10**, 367–78 (2015).
- Kostelny, S. A., Cole, M. S. & Tso, J. Y. Formation of a bispecific antibody by the use of leucine zippers. *J. Immunol.* **148**, 1547–53 (1992).
- Kimchi-Sarfaty, C. *et al.* Building better drugs: Developing and regulating engineered therapeutic proteins. *Trends Pharmacol. Sci.* **34**(10), 534–548 (2013).
- Lagassé, H. A. *et al.* Recent advances in (therapeutic protein) drug development. *F1000Res.* **6**, 113 (2017).
- Galanti, M., Fanelli, D. & Piazza, F. Conformation-controlled binding kinetics of antibodies. *Sci. Rep.* **6**, 18976 (2016).
- Chertova, E. *et al.* Envelope glycoprotein incorporation, not shedding of surface envelope glycoprotein (gp120/SU), is the primary determinant of SU content of purified human immunodeficiency virus type 1 and simian immunodeficiency virus. *J. Virol.* **76**, 5315–5325 (2002).
- Liu, J., Bartesaghi, A., Borgnia, M. J., Sapiro, G. & Subramaniam, S. Molecular architecture of native HIV-1 gp120 trimers. *Nature* **455**, 109–113 (2008).
- Zhu, P. *et al.* Distribution and three-dimensional structure of AIDS virus envelope spikes. *Nature* **441**, 847–852 (2006).
- Klein, J. S. & Bjorkman, P. J. Few and far between: how HIV may be evading antibody avidity. *PLoS Pathog.* **6**, e1000908 (2010).
- Mouquet, H. *et al.* Polyreactivity increases the apparent affinity of anti-HIV antibodies by heterooligation. *Nature* **467**, 591–595 (2010).
- Wu, H. *et al.* Ultra-potent antibodies against respiratory syncytial virus: effects of binding kinetics and binding valence on viral neutralization. *J. Mol. Biol.* **350**, 126–144 (2005).
- Galimidi, R. P. *et al.* Intra-spike crosslinking overcomes antibody evasion by HIV-1. *Cell* **160**, 433–446 (2015).
- Jähnichen, S. *et al.* CXCR4 nanobodies (VHH-based single variable domains) potently inhibit chemotaxis and HIV-1 replication and mobilize stem cells. *Proc. Natl. Acad. Sci. USA* **107**, 20565–20570 (2010).
- Zhang, J. *et al.* Pentamerization of single-domain antibodies from phage libraries: a novel strategy for the rapid generation of high-avidity antibody reagents. *J. Mol. Biol.* **335**, 49–56 (2004).
- Yang, T., Baryshnikova, T., Mao, O. K., Holden, H. & Cremer, M. A. P. S. Investigations of bivalent antibody binding on fluid-supported phospholipid membranes: the effect of hapten density. *J. Am. Chem. Soc.* **125**, 4779–4784 (2003).
- Fasting, C. *et al.* Multivalency as a chemical organization and action principle. *Angewandte Chemie* **51**(42), 10472–10498 (2012).
- Schiavom, V. L., Robert, P., Limozin, L. & Bongrand, P. Quantitative Modeling Assesses the Contribution of Bond Strengthening, Rebinding and Force Sharing to the Avidity of Biomolecule Interactions. *PLoS One* **7**, e44070 (2012).
- González, C. *et al.* Nanobody-CD16 Catch Bond Reveals NK Cell Mechanosensitivity. *Biophys. J.* **116**, 1516–1526 (2019).
- van Rosmalen, M. & Maarten, M. K. Tuning the Flexibility of Glycine-Serine Linkers To Allow Rational Design of Multidomain Proteins. *Biochemistry* **56**, 6565–6574 (2017).
- Xiaoying, C., Jennica, Z. & Wei-Chiang, S. Fusion Protein Linkers: Property, Design and Functionality. *Adv. Drug Deliv. Rev.* **65**, 1357–1369 (2013).
- Klein, J. S., Jiang, S., Galimidi, R. P., Keeffe, J. R. & Bjorkman, P. J. Design and characterization of structured protein linkers with differing flexibilities. *Protein Eng. Des. Sel.* **27**, 325–330 (2014).
- Arai, R., Ueda, H., Kitayama, A., Kamiya, N. & Nagamune, T. Design of the linkers which effectively separate domains of a bifunctional fusion protein. *Protein Eng. Des. Sel.* **14**, 529–532 (2001).
- Zhao, H. & Caflisch, A. Molecular dynamics in drug design. *Eur. J. Med. Chem.* **91**, 4–14 (2015).
- Jung, J. *et al.* Scaling molecular dynamics beyond 100,000 processor cores for large-scale biophysical simulations. *J. Comput. Chem.* **40**, 1919–1930 (2019).
- Windisch, B., Bray, D. & Duke, T. Balls and Chains – A Mesoscopic Approach to Tethered Protein Domains. *Biophys. J.* **91**, 2383–2392 (2006).
- Arkhipov, A., Roos, W. H., Wuite, G. J. L. & Schulten, K. Elucidating the Mechanism behind Irreversible Deformation of Viral Capsids. *Biophys. J.* **97**, 2061–2069 (2009).
- Gutierrez, C. & Schiff, R. HER2: biology, detection, and clinical implications. *Arch. Pathol. Lab. Med.* **135**, 55–62 (2011).

35. Frejd, F. Y. & Kim, K. T. Affibody molecules as engineered protein drugs. *Exp. Mol. Med.* **49**, e306 (2017).
36. Löfblom, J. *et al.* Affibody molecules: engineered proteins for therapeutic, diagnostic and biotechnological applications. *FEBS Lett.* **584**, 2670–2680 (2010).
37. Eigenbrot, C., Ultsch, M., Dubnovitsky, A., Abrahmsén, L. & Härd, T. Structural basis for high-affinity HER2 receptor binding by an engineered protein. *Proc. Natl. Acad. Sci. USA* **107**, 15039–15044 (2010).
38. Redner, S. *A guide to First-Passage processes*, (Cambridge University Press, Cambridge 2001).
39. Dill, K. A. & Bromberg, S. *Molecular Driving Forces: Statistical Thermodynamics in Biology, Chemistry, Physics, and Nanoscience*, 2nd ed., 2010 (New York: Garland Science, 2010).
40. DiMarzio, E. A. Proper Accounting of Conformations of a Polymer Near a Surface. *J. Chem. Phys.* **42**, 2101 (1965).
41. De Los Rios, P., Ben-Zvi, A., Slutsky, O., Azem, A. & Goloubinoff, P. Hsp70 chaperones accelerate protein translocation and the unfolding of stable protein aggregates by entropic pulling. *Proc. Natl. Acad. Sci. USA* **103**(16), 6166–6171 (2006).
42. De Michele, C., Rios, D. L., Foffi, P. & Piazza, G. F. Simulation and Theory of Antibody Binding to Crowded Antigen-Covered Surfaces. *PLoS Comput. Biol.* **12**, e1004752 (2016).
43. Plimpton, S. Fast Parallel Algorithms for Short-Range Molecular. *Dynamics. J. Comp. Phys.* **117**, 1–19 (1995).
44. Spinelli, S. *et al.* Camelid heavy-chain variable domains provide efficient combining sites to haptens. *Biochemistry* **39**, 1217–1222 (2000).
45. Allen, M. P. & Tildesley, D. J. *Computer simulation of liquids* (Oxford university press: New York, 1991).
46. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26**, 1781–1802 (2005).
47. Best, R. B. *et al.* Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone  $\phi$ ,  $\psi$  and side-chain  $\chi(1)$  and  $\chi(2)$  dihedral angles. *J. Chem. Theory Comput.* **8**, 3257–3273 (2012).
48. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graphics* **14**, 33–38 (1996).
49. Weeks, J. D., Chandler, D. & Andersen, H. C. Role of Repulsive Forces in Determining the Equilibrium Structure of Simple Liquids. *J. Chem. Phys.* **54**, 5237–5247 (1971).
50. Lee, H., Venable, R. M., MacKerell, A. D. Jr. & Pastor, R. W. Molecular dynamics studies of polyethylene oxide and polyethylene glycol: hydrodynamic radius and shape anisotropy. *Biophys. J.* **95**, 1590–599 (2008).
51. Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H. & Kollman, P. A. The weighted histogram analysis method for free-energy calculations on biomolecules. I. The method. *J. Comput. Chem.* **13**(8), 1011–1021 (1992).

## Acknowledgements

The authors acknowledge the French ministry of higher education under the PIA2 project BIO-S for funding.

## Author contributions

F.P. conceived the study. S.S. performed the simulations. S.S. and F.P. analyzed the data and wrote the paper.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-64646-5>.

**Correspondence** and requests for materials should be addressed to S.S.

**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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020