On the interpretation of kinetics and thermodynamics probed by single-molecule experiments

Stefano Angioletti-Uberti¹

Received: 20 January 2020 / Revised: 16 April 2020 / Accepted: 17 April 2020 / Published online: 15 May 2020
© The Author(s) 2020

Abstract

Single-molecule pulling experiments are widely used to extract both thermodynamic and kinetic data on ligand-receptor pairs, typically by fitting different models to the probability distribution of rupture forces of the corresponding bond. Here, a theoretical model is presented that shows how a measurement of the number of binding and unbinding events as a function of the observation time can also give access to both the binding (k_{on}) and the unbinding (k_{off}) rates of bonds, which combined provide a well-defined bond free-energy ΔG^{bond} . The connection between ΔG^{bond} and the ligand-receptor binding constant measured by typical binding essays is critically discussed. The role played by the molecular construct used to tether ligands and receptors to a surface is considered, highlighting the various approximations necessary to derive general expressions that connect its structure to its contribution, termed ΔG^{cnf} , to the bond free-energy. In this way, the validity and the assumptions underpinning widely employed formulas and experimental protocols used to extract binding constants from single-molecule experiments are assessed. Finally, the role of ΔG^{cnf} in processes mediated by ligand-receptor binding is briefly considered, and an experiment to unambiguously measure this quantity proposed.

Keywords Ligand-receptor interactions · Single-molecule experiments · Statistical mechanics · Kinetics · Modelling

Since the seminal papers of Evans and Ritchie about two decades ago [1, 2], based on previous theoretical models of bond rupture by Bell [3], dynamic force spectroscopy (DFS) has been widely used as a tool to measure the thermodynamics and kinetics of bonds between a wide variety of ligand-receptor pairs, providing very detailed information regarding important biological processes [4]. The power of this experiment stems from the fact that it allows to prove binding at the level of a single ligand-receptor pair, without the averaging implicit in bulk binding essays as ELISA. In DFS, the kinetics of binding is inferred by measuring the probability distribution of rupture forces at breakage, and fitting this quantity to model distribution functions, such as the Dudko-Hummer-Szabo's [5, 6] or Bell's [3] models. In this way, microscopic quantities such as the free-energy barrier for bond rupture and the equilibrium bond lifetime can be recovered.

DFS experiments can be carried out in various ways, but the general procedure can be exemplified in the following way: In a typical experiment, a ligand (receptor) is tethered to a substrate via a linking construct, typically another small molecule, a protein or a polymer. Colloids controlled via optical tweezers are a prominent example of a substrate [7, 8], although the tip of an AFM cantilever, or just a simple planar surface, has also often been used in this type of experiments [9]. Starting from afar, ligand and receptor are brought to a distance where bonding can occur at time $t = 0$, and retracted after a certain time $t = t_{\text{obs}}$, which will be referred to as the observation time. If a bond is present within the ligand-receptor pair, a resisting force is measured, and none otherwise. In a typical DFS experiment, the ligand bound is retracted from the receptor and the exact value of the force at rupture is recorded and analysed. Here, instead of looking at the distribution of rupture forces, we will look at the probability to observe a certain number of binding and unbinding events as a function of time during an experiment where the ligand-receptor distance is kept fixed. We will derive a mathematical model to describe this probability and see that this quantity provides us with another root to extract k_{on} and k_{off} from a DFS experiment. As we shall see, it turns out that such probability is connected to the

✉ Stefano Angioletti-Uberti
sangiole@imperial.ac.uk

¹ Department of Materials, Imperial College London, London, UK

bond energy ΔG^{bond} for a *tethered* ligand-receptor pair. The statistical mechanical interpretation of such bond energy, its dependence on microscopic details of the ligand-receptor pair and its connection to the binding constant of such pair will be the other main concern of this paper.

Let us first imagine a thought experiment, the practical realisation of which will be discussed later. The system of concern is schematically depicted in Fig. 1. We have a ligand and a receptor, both tethered via a linker to a substrate, the relative position of which is fixed. At time $t = 0$, as the ligand-receptor pair was brought to binding distance from afar, the system is in an unbound state. What is the number of binding and unbinding events measured within a time t_{obs} ? Using the same setup typically employed in DFS, this is equivalent to measuring the number of times that a non-zero attractive force (bond formation) and its subsequent disappearance (bond rupture) is recorded within t_{obs} . Unlike in normal DFS, however, where after formation the bond is stretched until rupture, in this case, the ligand-receptor pair must be kept at a *fixed* distance during the measurement. In this sense, this setup restricts this potential experiment to ligand-receptor pairs where a dynamic equilibrium between bound and unbound states

Fig. 1 Schematic representation of the system and the proposed experiment. Ligand and receptor (white and red circle) are grafted via a linker to a substrate (e.g. a colloid). (1) At time $t = 0$, the unbound ligand-receptor pair is brought to a distance where binding can occur. In this state, binding from an unbound state can occur at each time interval with probability $k_{\text{on}}dt$. The opposite process starting from a bound state occurs with probability $k_{\text{off}}dt$. Upon bond formation (breakage), a force on the substrate is measured, signalling a binding event

exist, or in other words to reversible bonds. In the following, I present a general but simple microscopic model to connect the outcome of this measurement to the binding, k_{on} , and unbinding, k_{off} , rate of the bond between the ligand-receptor pair. Before doing any calculation, let us think about this problem for a second to see what we should expect. In fact, instead of concentrating on the number of binding and unbinding events observed, let us consider the probability P_B for the system to be found in a bound state at a certain time t_{obs} . If the system was initially unbound at $t = 0$, it means the system must have gone through N binding and $N - 1$ unbinding events. This probability is thus given by:

$$P_B(t_{\text{obs}}) = \sum_{N=1}^{\infty} P_{N,N-1}(t_{\text{obs}}), \quad (1)$$

where $P_{N,N-1}(t_{\text{obs}})$ is the probability to observe *exactly* N binding and $N - 1$ unbinding events. Similarly, the probability to find the system in an unbound state $P_{\bar{B}} = 1 - P_B = \sum_{N=1}^{\infty} P_{N,N}(t_{\text{obs}})$, $P_{N,N}(t_{\text{obs}})$ now being the probability to observe an equal number N of binding and unbinding events. If one waits long enough for the system to equilibrate before making an observation, i.e. formally for $t_{\text{obs}} \rightarrow \infty$, the probability for the ligand-receptor pair to be found in a bound state must reach its equilibrium value. The system we describe can be regarded as a two-state system, where the bond is formed, or not. In a two-state system, the equilibrium value for the bonding probability is trivially given by:

$$P_B^{\text{eq}} = \frac{\exp(-\beta \Delta G^{\text{bond}})}{1 + \exp(-\beta \Delta G^{\text{bond}})} \quad (2)$$

$\beta = 1/k_B T$ being the inverse of the thermal energy and ΔG^{bond} the difference in free-energy between the bound and unbound state, i.e. the bond free-energy. Whereas this result is only valid at equilibrium, or in other words at very long times, we should expect for finite times this probability to depend on both k_{on} and k_{off} , the rates for bond formation and bond breakage, respectively. Qualitatively, for example, since the system starts in an unbound state, it will typically be found in a bound state after a time $t \approx k_{\text{on}}^{-1}$. If we wait a bit more, the bond can dissociate again, in a time dictated by k_{off} . In general, however, we should expect the long-time limit of any kinetic model to converge to the equilibrium value. This will allow us to connect k_{on} , k_{off} and ΔG^{bond} .

In order to calculate a general expression for $P_{N,N}(t)$ and $P_{N,N-1}(t)$, and thus $P_B(t)$, valid for all timescales, one should consider the stochastic nature of the binding and unbinding process. The key ingredients are the conditional binding(unbinding) probabilities. Given the over-damped nature of the bond formation and breaking in a typical

biological environment, i.e. in solution, we can consider its description using a simple Markov model, with a constant rate for these two processes. Thus, if $k_{\text{on}}dt$ is the probability that an unbound system becomes bound in a time interval dt , and $k_{\text{off}}dt$ the probability for the reverse process, the probabilities that given that one start with an unbound(bound) state at time $t = 0$ this state will persist up to a later time t' can be given as:

$$P_{\text{on}} = \exp(-k_{\text{on}}t') \quad (3)$$

$$P_{\text{off}} = \exp(-k_{\text{off}}t') \quad (4)$$

Using Eqs. 3 and 4, $P_{N,N}(t_{\text{obs}})$ and $P_{N,N-1}(t_{\text{obs}})$ can be written in terms of a convolution as:

$$P_{N,N}(t_{\text{obs}}) = \int_0^{t_{\text{obs}}} dt_1 P_{N,N-1}(t_1) k_{\text{off}} P_{\text{on}}(t_{\text{obs}} - t_1) \quad (5)$$

$$P_{N,N-1}(t_{\text{obs}}) = \int_0^{t_{\text{obs}}} dt_1 P_{N-1,N-1}(t_1) k_{\text{on}} P_{\text{off}}(t_{\text{obs}} - t_1). \quad (6)$$

In other words, the probability that the system has made exactly N binding and N unbinding transitions within an observation time t_{obs} is equal to the sum over all possible times $t_1 < t_{\text{obs}}$ that the system has made exactly N binding transitions and $N - 1$ unbinding transitions within t_1 , makes an unbinding transition at time $t = t_1$ and then remains in this state for the remaining time $t_{\text{obs}} - t_1$, when the state of the system is observed via a measurement. A similar interpretation can be given to the expression for $P_{N,N-1}(t)$. Because Eqs. 5 and 6 have the structure of a convolution, one can take their Laplace transform $f(s) \equiv \mathcal{L}(f(t)) = \int_0^\infty f(t) \exp(-st) dt$ to obtain the following, equivalent representation in terms of two recurrence relations:

$$P_{N,N}(s) = P_{N,N-1}(s) k_{\text{off}} \frac{1}{k_{\text{on}} + s} \quad (7)$$

$$P_{N,N-1}(s) = P_{N-1,N-1}(s) k_{\text{on}} \frac{1}{k_{\text{off}} + s} \quad (8)$$

where we used the fact that $\mathcal{L}\{\exp(-kt)\} = \frac{1}{k+s}$ and $\mathcal{L}(f * g) = f(s)g(s)$. Starting from $P_{0,0} = P_{\text{on}} \Rightarrow P_{0,0}(s) = \frac{1}{k_{\text{on}} + s}$, Eqs. 7 and 8 can be used to recursively calculate $P_{N,N}(s)$ and $P_{N,N-1}(s)$ for every N , leading to the closed form expressions:

$$P_{N,N}(s) = k_{\text{on}}^N k_{\text{off}}^N \left(\frac{1}{k_{\text{on}} + s} \right)^{N+1} \left(\frac{1}{k_{\text{off}} + s} \right)^N \quad (9)$$

$$P_{N,N-1}(s) = k_{\text{on}}^N k_{\text{off}}^{N-1} \left(\frac{1}{k_{\text{on}} + s} \right)^N \left(\frac{1}{k_{\text{off}} + s} \right)^N, \quad (10)$$

which can then be used to obtain the equivalent formula in time domain by applying the inverse Laplace transform, given for every $N > 0$:

$$P_{N,N}(t) = \frac{\sqrt{\pi t |k_{\text{off}} - k_{\text{on}}|}}{2N!} e^{-\frac{1}{2}(k_{\text{on}} + k_{\text{off}})t} \left(\frac{k_{\text{off}} k_{\text{on}} t}{|k_{\text{off}} - k_{\text{on}}|} \right)^N \times \left[I_{N-\frac{1}{2}} \left(\frac{|k_{\text{off}} - k_{\text{on}}|t}{2} \right) \pm I_{N+\frac{1}{2}} \left(\frac{|k_{\text{off}} - k_{\text{on}}|t}{2} \right) \right] \quad (11)$$

$$P_{N,N-1}(t) = \frac{\sqrt{\pi}}{(N-1)!} e^{-\frac{1}{2}(k_{\text{on}} + k_{\text{off}})t} \left(\frac{k_{\text{off}} k_{\text{on}} t}{|k_{\text{off}} - k_{\text{on}}|} \right)^N \times \left(\frac{|k_{\text{off}} - k_{\text{on}}|}{k_{\text{off}}^2 t} \right)^{1/2} I_{N-\frac{1}{2}} \left(\frac{|k_{\text{off}} - k_{\text{on}}|t}{2} \right), \quad (12)$$

where $I_\nu(x)$ is the modified Bessel function of the first kind [10] and $P_{0,0} = P_{\text{on}}$. The $+$ sign in Eq. 11 corresponds to the case $k_{\text{off}} > k_{\text{on}}$, whereas the $-$ sign should be used in the opposite case. Notice that these formulas can be simplified to somewhat simpler but less compact expressions containing only exponentials and powers of k_{on} , k_{off} and t , which can be derived by inverting (9) and (10) for a specific value of N .

Equations 11 and 12 are the most general result and can be used to fit the number of observed binding and unbinding events during the previously described experiments to provide an estimate of k_{on} and k_{off} . It should be noted that all the various $P_{N,N}(t)$ and $P_{N,N-1}(t)$ are connected and to fully use all the information recorded during an experiment one should fit all of them concurrently to get the best statistical estimate for k_{on} and k_{off} . Notably, more specialised models appeared in the literature, where a given number of binding and unbinding events is assumed based on physical constraints of the system, can be derived from the previous expressions. For example, Fritz et al. [11] in their experiments on P-selectin bindings used the fact that for their system $k_{\text{on}} \gg k_{\text{off}}$ and that a single unbinding event can be measured at most due to their settings during the observation time. This is equivalent to using $P_{1,0} = k_{\text{on}} \frac{\exp(-k_{\text{off}}t) - \exp(-k_{\text{on}}t)}{k_{\text{on}} - k_{\text{off}}}$, which indeed for $k_{\text{on}} \gg k_{\text{off}}$ reduces to $P_B \approx \exp(-k_{\text{off}}t)$ (but all information regarding k_{on} is then lost). Combined with Eq. 1, use of Eqd. 11 and 12 allows to calculate the probability to find the system in a bound state after a certain observation time t_{obs} , regardless of the number of potential binding and unbinding events. In this case, however, it is easier to first take the Laplace transform of the right hand side of Eq. 1 and sum

the resulting infinite series:

$$P_B(s) = \sum_{N=1}^{\infty} \frac{1}{k_{\text{off}}} \left(\frac{k_{\text{on}} k_{\text{off}}}{(k_{\text{on}} + s)(k_{\text{off}} + s)} \right)^N$$

$$= \frac{k_{\text{on}}}{(k_{\text{on}} + k_{\text{off}})s + s^2}, \quad (13)$$

where we have used the fact that $\sum_{i=1}^{\infty} x^i = \frac{x}{1-x}$ (for $|x| < 1$), which after performing the inverse Laplace transform, reads:

$$P_B = \frac{(1 - e^{-(k_{\text{off}} + k_{\text{on}})t}) k_{\text{on}}}{k_{\text{off}} + k_{\text{on}}} \quad (14)$$

and thus for $P_{\tilde{B}}$:

$$P_{\tilde{B}} = \frac{k_{\text{off}} + e^{-(k_{\text{off}} + k_{\text{on}})t} k_{\text{on}}}{k_{\text{off}} + k_{\text{on}}}, \quad (15)$$

which could have also been derived in a simpler way by solving the Master equation for this two-state model:

$$\frac{dP_B}{dt} = -k_{\text{off}} P_B + k_{\text{on}}(1 - P_B) \quad (16)$$

with the boundary conditions $P_B(0) = 0$.

Before we delve into the connection between k_{on} , k_{off} and the ligand-receptor bond energy, let us now discuss how these could be determined from experiments applying either Eqs. 14 or 11, 12, as well as the associated challenges. First of all, let us discuss how the measurement of P_B . In practice, this can be extracted using the typical setting employed in single-molecule pulling experiment, see, e.g. [11]. Starting with an unbound ligand at $t' = 0$, the ligand is subsequently pulled away from the receptor at a later time $t' = t$. The fraction of times that a non-zero retraction force is measured is our estimate of $P_B(t)$. In this regard, it should be noticed that pulling the ligand away from the receptor takes a finite time τ_{pull} . In order to have a correct interpretation of $P_B(t)$, the system must not be able to spontaneously change its state from bound to unbound (or vice versa) during this time, hence one needs $\tau_{\text{pull}} \ll \min(k_{\text{on}}^{-1}, k_{\text{off}}^{-1})$. This requirement can be restated in terms of the minimum retraction speed necessary, v_{pull} , which should obey $v_{\text{pull}} \gg \delta/\tau_{\text{pull}}$, δ being the distance at which a ligand-receptor pair cannot rebind anymore, which to first approximation will be of the order of the size of the ligand-receptor interaction range. Whether this speed is achievable depends on the specifics of the experimental apparatus used and could be challenging when the on and off rates are high, as typical for reversible bonds. Assuming this retraction speed can be achieved, determining k_{on} and k_{off} from Eq. 14 can be done by measuring both the short- and long-time limit of P_B . This is easy to see because in the short-time limit one has:

$$P_B(t \rightarrow 0) \approx k_{\text{on}} t + \mathcal{O}(t^2) \quad (17)$$

whereas at long times:

$$P_B(t \rightarrow \infty) = k_{\text{on}} / (k_{\text{on}} + k_{\text{off}}). \quad (18)$$

In this way, Eqs. 17 and 18 provide two equations for the two unknown. The need to measure P_B at this two time limits tells us about other requirements of the experimental apparatus. First, because measuring k_{on} means measuring the short-time behaviour of P_B , the time resolution of the force τ_f , i.e. the inverse of the frequency ν_f at which the force is experimentally sampled, needs to be smaller than k_{on}^{-1} . Note that this is a separate requirement from that on τ_{pull} . Finally, an accurate measurement of the plateau value of P_B can also be challenging, especially in the case of strong bonds, i.e. when $(k_{\text{on}} \gg k_{\text{off}})$. This is because in this regime the plateau value $P_B(t \rightarrow \infty)$ will be always very close to 1 and measuring its deviation from this number with good statistical accuracy requires taking a number of measurements $N_{\text{samples}} \gg \frac{1}{1-P_B} \approx k_{\text{on}}/k_{\text{off}}$. A smaller number of samples will likely result in estimating $P_B = 1$, which will translate in an estimate of zero for k_{off} , or an infinite value for k_{on} , if k_{off} was somehow known and k_{on} could not be estimated from the slope at short timescales.

Estimating k_{on} and k_{off} from Eqs. 11 and 12 requires a slightly different experiment, which can be summarised in the following steps: (1) As for measuring $P_B(t)$, we start with an unbound ligand at time $t' = 0$, which at all times should be kept at a fixed position with respect to the receptor. If at some point a bond is formed, a non-zero force will appear. In practice, this is the thermodynamic average of the instantaneous force over all microscopic configurations of the bound ligand-receptor pair. A force of opposite value will thus be required to keep the ligand fixed, which can be done via a feedback mechanism using the same apparatus used to maintain a fixed retraction velocity in a ligand during a standard pulling experiment. This feedback force is measured and recorded. The appearance of a force (its disappearance) at a certain time thus signals a binding (unbinding) event. (2) The time series of the force is collected for analysis for a certain time t_{max} . (3) t_{max} is divided in sub-intervals of time t_1 and used to make an histogram of the number of observed binding and unbinding events within it. In this way, one has experimentally measured the values of $P_{N,N}(t_1)$ and $P_{N,N-1}(t_1)$. (4) Using any preferred optimisation algorithm to minimise the difference between the experimentally measured $P_{N,N}$, $P_{N,N-1}$ and those in Eqs. 11 and 12 in the paper for a certain value of k_{on} and k_{off} , one can find those that best fit the measured distribution, which will represent our estimate. Using this second approach comes with its own experimental requirements. One is that the total observation time must be $t_{\text{max}} \gg t_1 \gg \max(k_{\text{on}}^{-1}, k_{\text{off}}^{-1})$ to obtain good statistics. Also, similarly to what occurs for the measurement of P_B , the frequency at which the

force is recorded, v_f , should be much higher than both k_{on} and k_{off} to be able not to miss any binding or unbinding event. However, use of Eqs. 11 and 12 might be easier because no retraction of the ligand is required; hence, there is no minimum pulling velocity to deal with. This could be an advantage in particular when reversible ligand-receptor pairs, for which both k_{on} and k_{off} can be quite high, are involved. At the same time, however, it is important to recognise that in this second type of experiment one would require an apparatus able to accurately measure relatively lower forces compared with those involved in measuring P_B . The reason is that in a normal pulling experiment the force becomes higher, and thus more easily detectable, once the ligand-receptor pair is sufficiently stretched. This does not necessarily occur instead in the case where the ligand is kept fixed, because in this case the thermodynamic force can be arbitrarily small, in fact, even zero if no stretching of the ligand is at all required to form a bond with the receptor. Eventually, whether measuring the distributions P_{NN} , $P_{N,N-1}$ or P_B provides a more accurate estimate of k_{on} and k_{off} for a given ligand-receptor pair should be evaluated on a system by system basis, as well as on the specifics of the experimental apparatus available, in light of the above requirements and limitations. In any case, however, it is important to highlight that the measurement of $P_{N,N}$, $P_{N,N-1}$ provides a second, independent route towards the estimation of the binding and unbinding rates, a fact that can be used to check for consistency of other experiments.

Let us now turn back again to Eq. 14 to discuss its connection with physical quantities related to the ligand-receptor pair. In practice, it has the form of an exponential, describing the approach to equilibrium with a typical time $t_{\text{eq}} = 1/(k_{\text{on}} + k_{\text{off}})$. The long-time limit of P_B , Eq. 18, can be compared with the equilibrium behaviour described by Eq. 2 in terms of ΔG^{bond} , which allows us to identify the following relation:

$$k_{\text{on}}/k_{\text{off}} = e^{-\beta\Delta G^{\text{bond}}} \equiv K_{\text{tether}}^{\text{bind}} \quad (19)$$

At this point, one should stop and consider an important point regarding k_{on} , k_{off} and ΔG^{bond} . Given their definition here via Eq. 16, k_{on} and k_{off} have both the dimension of a rate and thus $K_{\text{tether}}^{\text{bind}} \equiv k_{\text{on}}/k_{\text{off}}$ is a non-dimensional quantity. In describing ligand-receptor bonds, however, k_{on} is often provided with the dimension of a rate times an inverse number density (i.e. in units of $1/\text{s M}$), or, equivalently, as a rate times a (molar) volume. To avoid confusion, let us call this second quantity $k_{\text{on}}^{\text{sol}}$. The binding constant of the ligand-receptor pair is then defined as $K_{\text{solution}}^{\text{bind}} = k_{\text{on}}^{\text{sol}}/k_{\text{off}}$, which has the dimension of a molar volume, exactly the dimension in which ligand-receptor binding constants are typically reported in the literature. Clearly, k_{on} and $k_{\text{on}}^{\text{sol}}$, and thus $K_{\text{tether}}^{\text{bind}}$ and

$K_{\text{solution}}^{\text{bind}}$, cannot be the same quantity and a question arises: what is their connection, if any? The difference between the two actually arises from the exact definition of the equilibrium binding constant. For free ligand and receptor in solutions, the equilibrium constant is typically defined as $K_{\text{solution}}^{\text{bind}} = \frac{[LR]}{[L][R]}$, where $[LR]$, $[L]$ and $[R]$ are the equilibrium concentration of bound ligand receptor pairs, free ligands and free receptors, respectively. This $K_{\text{solution}}^{\text{bind}}$ has the familiar dimension of an inverse density. However, $K_{\text{solution}}^{\text{bind}}$ is derived considering *free* ligand and receptors in solution, whereas single-molecule experiments probe binding of *tethered ligands to tethered receptors*. Tethering has important consequences in both the kinetics and thermodynamics of bond formation, and some of its effects in the context of DFS experiments have already been discussed, albeit using simplified models, as we will discuss later.

Generally speaking, to fully quantify and understand the role of tethering, a microscopic description of the linker is required. Models of this sort have already appeared in the soft matter literature, where the concern has been to understand the role of different tethers on colloidal self-assembly mediated by grafted ligand-receptor pairs [12–15]. These works, however, do not seem to have yet reached the full attention of the community of single-molecule experiments. Since they are important for the correct interpretation of ΔG^{bond} , and to shed some light on its microscopic origins, we will briefly discuss them here, focusing on highlighting the correct statistical mechanical definition of the quantities measured when tethered ligands and receptors are involved and highlight the approximations one must use to connect ΔG^{bond} , $K_{\text{tether}}^{\text{bind}}$ and $K_{\text{solution}}^{\text{bind}}$. In the following, we will thus follow the derivation first provided in [12] and later generalised in [13], trying to put their arguments in the context of the single-molecule experiments discussed here. The first step is to look at the binding between ligand-receptor pairs in solution. From a statistical mechanics point of view, the chemical potential for a free ligand(receptor) $L(R)$ in solution in a volume V is given by [13, 15]:

$$\mu_{L(R)} = -k_B T \log \frac{N_{L(R)} \Lambda^3}{V \mathcal{Z}_{L(R)}} = -k_B T \log \frac{[L(R)] \Lambda^3}{\mathcal{Z}_{L(R)}} \quad (20)$$

whereas that for a bound ligand-receptor pair is:

$$\mu_{LR} = -k_B T \log \frac{N_{LR} \Lambda^3}{V \mathcal{Z}_{LR}} = -k_B T \log \frac{[LR] \Lambda^3}{\mathcal{Z}_{LR}}, \quad (21)$$

where $\mathcal{Z}_{L(R)}$ and $\mathcal{Z}_{LR}$ are the so-called molecular partition function of the free ligand(receptor) and of the bound ligand-receptor pair in solution [13], or in other words the partition function of a single ligand(receptor), or that of a bound ligand-receptor pair, at fixed centre of mass. These expressions simply describe an ideal gas of particles *with*

an internal structure, quantified by the molecular partition function $\mathcal{Z}_s$. At equilibrium $\mu_{LR} - (\mu_F + \mu_R) = 0$, leading to $\frac{[LR]}{[L][R]} = \frac{\mathcal{Z}_{LR}}{\mathcal{Z}_L \mathcal{Z}_R}$. This should be compared with the classical thermodynamic expression $\frac{[LR]}{[L][R]} = K^e = \exp(-\frac{\Delta G^\theta}{k_B T})/\rho^\theta$ (where, importantly a reference density ρ^θ appears), leading to:

$$\frac{\mathcal{Z}_{LR}}{\mathcal{Z}_L \mathcal{Z}_R} = \exp(-\frac{\Delta G^\theta}{k_B T})/\rho^\theta = K_{\text{solution}}^{\text{bind}}. \quad (22)$$

One important thing to notice is the correct identification of $K_{\text{solution}}^{\text{bind}}$ as the exponential of the bonding energy scaled by the reference density ρ^θ , at which ΔG^θ is measured. In practice, ΔG^θ measures detail about how upon binding molecular interactions change. Equation 22 has been obtained by treating ligands and receptors as an ideal gas of structured particles floating in solution. In a DFS experiment, however, both ligands and receptors are tethered: What is the equivalent of the bond energy for a tethered ligand-receptor pair? Similar statistical mechanical considerations in this case lead to a definition of the bond energy as [13]:

$$\exp(-\beta \Delta G^{\text{bond}}) = \frac{Q_{LR}}{Q_R Q_L} \quad (23)$$

where now $Q_{L(R)}$ and Q_{LR} are the partition functions of the tethered and unbound bound ligand(receptor) and of the bound ligand-receptor pair, respectively. What is the connection between the bond energy ΔG^{bond} , as defined by Eq. 23, and the bond energy ΔG^θ (or the equilibrium binding constant $K_{\text{solution}}^{\text{bind}}$) for binding between free ligands and receptors in solution? In the general case, although one expects ΔG^{bond} to have contributions from both ligand-receptor binding and from their linkers, its value cannot be expressed simply as the sum of two separate and well defined contributions. In this case, a simple connection between ΔG^{bond} , ΔG^θ and $K_{\text{solution}}^{\text{bind}}$ cannot be established. However, a good approximation in many cases is to assume that the linker does not change appreciably the molecular structure of the ligands and receptors [13]. This allows to factorise the partition functions as $Q_{L(R)} = \Omega_{T_{L(R)}} \mathcal{Z}_{L(R)}$ and $Q_{LR} = \Omega_{T_L T_R} \mathcal{Z}_{LR}$. Hence, with this approximation, and further using Eq. 22, we have that:

$$\begin{aligned} \exp(-\beta \Delta G^{\text{bond}}) &= \frac{\mathcal{Z}_{LR}}{\mathcal{Z}_L \mathcal{Z}_R} \frac{\Omega_{T_L, T_R}}{\Omega_{T_L} \Omega_{T_R}} \\ &= \frac{\exp(-\beta \Delta G^\theta)}{\rho^\theta} \frac{\Omega_{T_L, T_R}}{\Omega_{T_L} \Omega_{T_R}} \\ &= K_{\text{solution}}^{\text{bind}} \frac{\Omega_{T_L, T_R}}{\Omega_{T_L} \Omega_{T_R}} \end{aligned} \quad (24)$$

where Ω_{T_L, T_R} and $\Omega_{T_L(T_R)}$ are now the partition functions of the linker only, when the ligand and receptors are bound to each other, or not, respectively. These partition functions account for the fact that the linker is constrained from being

grafted, and additionally that in the bound state the linkers of the ligand and that of the receptor must assume microscopic configurations compatible with the presence of a bond, or in other words that allow ligand and receptors to be at binding distance. Examples of how to calculate or compute these quantities using specific molecular models for the linkers (for example, treated as rigid rods or gaussian chains) are reported in [13]. For a general Monte Carlo formalism to compute this quantity for realistic polymer models see [13] and [16].

An alternative and revealing form of Eq. 24 can also be given by re-adjusting terms,

$$\begin{aligned} \Delta G^{\text{bond}} &= -k_B T \log(K^e \rho^\theta) - k_B T \log\left(\frac{1}{\rho^\theta} \frac{\Omega_{T_L, T_R}}{\Omega_{T_L} \Omega_{T_R}}\right) \\ &= \Delta G^\theta + \Delta G^{\text{cnf}}. \end{aligned} \quad (25)$$

Equations 24 and 25 suggest a few important points regarding the study of ligand-receptor interactions and their quantification via single-molecule experiments such as DFS, which will now be discussed. The first is that the bond energy measured in any experiment using tethered ligands and receptors has an additional contribution ΔG^{cnf} compared with that of free ligands and receptors in solution. This point has previously been made by others, albeit using specific microscopic models, whereas the generalisation discussed allows a broader applicability of these results. For example, Nguyen-Duong et al. [17] showed that if one treats the linker as a spring, an additional positive contribution to the bond free-energy destabilising the bond arises. Generally speaking, this contribution, usually referred to as “configurational free-energy” [13], can be several $k_B T$ in magnitude depending on the type of tether [13, 14], and should thus always be properly taken into account in comparing experiments with the same ligand-receptor pairs but different linkers. It should be noted that since ΔG^{bond} contains the non-negligible effect of tethering, this augmented bond energy ΔG^{bond} , and not ΔG^θ , is the relevant quantity determining binding in various biological settings, where often the ligands and receptors which mediate binding are attached to the surface of a cell, and not free in solution. The definition of ΔG^{cnf} given by Eq. 25 shows that this quantity is strongly influenced by both the molecular architecture of the linker as well as the specifics of the grafting geometry, e.g. the distance between ligand and receptor grafting points [13], the mobility of the grafting points [18] as well as the rigidity of the grafting surface to which grafting occurs [19], since all these quantities affect the molecular partition functions. Indirectly, this also means that the values of k_{on} and k_{off} depend on the same quantities, suggesting that the proposed experiments to measure them must be performed keeping these quantities fixed. Just to make a relevant example, this means that

during the measurements of the binding or unbinding events suggested here, the distance between the ligand receptor pair must be kept fixed, since Ω_{T_L, T_R} depends on it [13]. On a more general note, the presence of ΔG^{cnf} in the definition of the bond free-energy suggests that, at least potentially, changes in binding energies might sometimes be dominated by changes in ΔG^{cnf} , and hence to the linker, rather than the specific ligand or receptor used. For example, it is tempting to speculate that some adhesion phenomena, e.g. leukocyte rolling on the epithelium, might be controlled by changing simply the length or rigidity of that part of the selectin molecules acting as a linker to the epithelium's surface. In fact, this might represent a competitive evolutionary strategy with respect to changing instead the binding site (i.e. what we call the ligand) for specific receptors. Another field where exploration of the change in ΔG^{cnf} as a function of linker's features is quite important is that of DNA and DNA-mediated self-assembly [20, 21], where knowledge of this quantity would lead to increased designability of these systems, as well as exploitation of geometrical effects to change binding properties. In this regard, whereas normally single-molecule experiments aim at measuring ΔG^θ for specific ligand-receptor pairs, and ΔG^{cnf} is only considered a nuisance to account for, similar experiments can instead be used to measure this latter quantity. In fact, the second point suggested by Eqs. 24 and 25 is that one can combine known values of $K_{\text{solution}}^{\text{bind}}$ for different ligand-receptor pairs determined in solution to measure ΔG^{cnf} for specific linkers. Until now, this has only been measured *indirectly* by fitting melting curves of DNA-coated colloids, or directly but via computer simulations that must necessarily make assumptions built into the molecular force-field employed [13, 16]. It is suggested here that the use of single-molecule experiments can instead provide a *direct* experimental measurement for this quantity, by simply taking the ratio of $K_{\text{tether}}^{\text{bind}}$ measured with the same ligand-receptor pairs, but different linkers. This procedure would provide a systematic way to explore the connection between ΔG^{cnf} and the microscopic parameters it is linked to. However, it should be pointed out that in this way only the relative value of ΔG^{cnf} can be measured, which nevertheless can provide an accurate measurement of trends within a range of linkers. For similar reasons, Eq. 24 also suggests that a direct comparison of the absolute values of the binding constants measured from single-molecule experiments with that measured by essays measuring binding in solution, e.g. ELISA, cannot be done unless a *proper* estimate of ΔG^{cnf} is available. Similarly to the previous case illustrated for ΔG^{cnf} , however, *ratios* between binding constants measured consistently (i.e. with the same type of experiment and keeping the same types of linkers, while only changing the ligand-receptor pair) will cancel out these configurational effects, and can be directly

compared. This is not a minor point, for the following reason. In DFS experiments, $K_{\text{solution}}^{\text{bind}}$ is typically estimated via the empirical expression [22]:

$$K_{\text{solution}}^{\text{bind}} = \frac{k_{\text{on}} n / V_{\text{eff}}}{k_{\text{off}}}, \quad (26)$$

where n is the number of potential binding receptors within the effective volume V_{eff} spanned by the linker. Considering surface densities of receptors low enough so that a single ligand-receptor bond can occur, i.e. using $n = 1$ in Eq. 26, we can directly compare with our model, where only a single binding partner was considered (a generalisation to many receptors, however, is trivial, one simply replaces $k_{\text{on}} \rightarrow k_{\text{on}} n$ in all the previous expressions, and the total partition function for the bound state must also be multiplied by n). Using the general statistical mechanical definition, one obtains:

$$\begin{aligned} K_{\text{solution}}^{\text{bind}} &= \frac{\exp(-\beta \Delta G^{\text{bond}})}{\frac{\Omega_{T_L, T_R}}{\Omega_{T_L} \Omega_{T_R}}} \\ &= \frac{\exp(-\beta \Delta G^{\text{bond}})}{\rho^\theta \exp(-\beta \Delta G^{\text{cnf}})} \equiv \frac{\exp(-\beta \Delta G^{\text{bond}})}{\bar{V}_{\text{eff, model}}}. \end{aligned} \quad (27)$$

Crucially, although $\bar{V}_{\text{eff, model}}$ has indeed the dimensions of a volume, it is not equal to the definition used to calculate V_{eff} when interpreting DFS experiments. In fact, estimating V_{eff} as the volume spanned by the linker is an approximation that one can arrive at only by making two approximations. The first is that the linker confines the ligand to reside at points within a volume V_{eff} , which is quite reasonable. The second approximation is that each point in this volume is equally likely to be occupied by the ligand and does not depend on the actual microscopic configuration the linker must assume given to put the ligand in that position. This is a very rough approximation. In fact, it is not even valid using a very simple model for a tether, i.e. treating it as a spring [17]. For this reason, one should expect this empirical definition of V_{eff} to provide *at best* a correct order of magnitude for its real value, and its use to determine $K_{\text{solution}}^{\text{bind}}$ from $K_{\text{tether}}^{\text{bind}}$ should be taken carefully, with more than a pinch of salt.

In conclusion, we presented a mathematical model for the calculation of the probability to observe a certain number of binding and unbinding transitions in a single-molecule experiment, which can be used in different ways to extract the binding and unbinding rates for tethered ligand-receptor pairs. Furthermore, we discussed the exact connection between these rates and the bond (free-)energy for tethered ligands and receptors, and elucidated the difference with bond energies, and equilibrium binding constants, measured in solution.

We have used a statistical mechanical model to show that, under well-defined assumptions, the bond free-energy for tethered constructs can be split into two parts. The first

is the bond energy measured from binding experiments of free ligand and receptors in solution (or when at least one species is free), ΔG^θ . The other, ΔG^{cnf} , is a configurational penalty term that depends both on molecular characteristics of the linker, as measured by its molecular partition function, and on details of the grafting geometry and substrate, like the relative position of the grafting points, or the shape and rigidity of the grafting surface. By a proper combination of available information regarding binding between ligand-receptor pairs in solution, it is suggested how single-molecule experiments can be used to systematically investigate tethering effects and the role of ΔG^{cnf} . This latter fact seems quite important, considering that ΔG^{cnf} is currently being exploited for the rational design of nanotechnology applications, in particular for self-assembly purposes. It would be peculiar, indeed, if similar effects were not already exploited by nature to tune the large variety of biological interactions also mediated by ligand-receptor binding between tethered species, and their understanding could thus be further exploited for a large variety of applications connected to them, e.g. in drug-design.

Compliance with ethical standards

Conflict of interest The author declares that he has no conflict of interest.

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/>.

References

1. Evans Evan, Ritchie Ken (1997) Dynamic strength of molecular adhesion bonds. *Biophys J* 72(4):1541–1555
2. Merkel R, Nassoy P, Leung A, Ritchie K, Evans E (1999) Energy landscapes of receptor–ligand bonds explored with dynamic force spectroscopy. *Nature* 397(6714):50–53
3. Bell GI (1978) Models for the specific adhesion of cells to cells. *Science* 200(4342):618–627
4. Evans E (2001) Probing the relation between force-lifetime-and chemistry in single molecular bonds. *Ann Rev Biophys Biomolec Struct* 30(1):105–128
5. Dudko OK, Hummer G, Szabo A (2006) Intrinsic rates and activation free energies from single-molecule pulling experiments. *Phys Rev Lett* 96(10):108101
6. Dudko OK, Hummer G, Szabo A (2008) Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *Proc Natl Acad Sci* 105(41):15755–15760
7. Stangner T, Wagner C, Singer D, Angioletti-Uberti S, Gutsche C, Dzubiella J, Hoffmann R, Kremer F (2013) Determining the specificity of monoclonal antibody HPT-101 to tau-peptides with optical tweezers. *ACS Nano* 7(12):11388–11396
8. Stangner T, Angioletti-Uberti S, Knappe D, Singer D, Wagner C, Hoffmann R, Kremer F (2015) Epitope mapping of monoclonal antibody HPT-101: a study combining dynamic force spectroscopy, ELISA and molecular dynamics simulations. *Phys Biol* 12(6):066018
9. Neuman KC, Nagy A (2008) Single-molecule force spectroscopy: optical tweezers, magnetic tweezers and atomic force microscopy. *Nat Methods* 5(6):491–505
10. Olver FWJ, Lozier DW, Boisvert RF, Clark CW (2010) NIST handbook of mathematical functions hardback and CD-ROM. Cambridge University Press
11. Fritz J, Katopodis AG, Kolbinger F, Anselmetti D (1998) Force-mediated kinetics of single P-selectin/ligand complexes observed by atomic force microscopy. *Proc Natl Acad Sci* 95(21):12283–12288
12. Leunissen ME, Dreyfus R, Sha R, Seeman NC, Chaikin PM (2010) Quantitative study of the association thermodynamics and kinetics of DNA-coated particles for different functionalization schemes. *J Am Chem Soc* 132(6):1903–1913
13. Varilly P, Angioletti-Uberti S, Moggetti BM, Frenkel D (2012) A general theory of DNA-mediated and other valence-limited colloidal interactions. *J Chem Phys* 137:094108–094122
14. Dreyfus R, Leunissen ME, Sha R, Tkachenko AV, Seeman NC, Pine DJ, Chaikin PM (2009) Simple quantitative model for the reversible association of DNA coated colloids. *Phys Rev Lett* 102:048301
15. Leunissen ME, Frenkel D (2011) Numerical study of DNA-functionalized microparticles and nanoparticles: explicit pair potentials and their implications for phase behavior. *J Chem Phys* 134(8):084702
16. De Gernier R, Curk T, Dubacheva GV, Richter RP, Moggetti BM (2014) A new configurational bias scheme for sampling supramolecular structures. *J Chem Phys* 141(24):244909
17. Nguyen-Duong M, Koch KW, Merkel R (2003) Surface anchoring reduces the lifetime of single specific bonds. *Europhys Lett* 61(6):845
18. Angioletti-Uberti S, Varilly P, Moggetti BM, Frenkel D (2014) Mobile linkers on DNA-coated colloids: valency without patches. *Phys Rev Lett* 113:128303–128306
19. Hu J, Lipowsky R, Weikl TR (2013) Binding constants of membrane-anchored receptors and ligands depend strongly on the nanoscale roughness of membranes. *Proc Natl Acad Sci* 110(38):15283–15288
20. Jones MR, Seeman NC, Mirkin CA (2015) Programmable materials and the nature of the DNA bond. *Science* 347(6224):1260901
21. Angioletti-Uberti S, Moggetti BM, Frenkel D (2016) Theory and simulation of DNA-coated colloids: a guide for rational design. *Physical Chemistry Chemical Physics*
22. Hinterdorfer P, Baumgartner W, Gruber HJ, Schilcher K, Schindler H (1996) Detection and localization of individual antibody-antigen recognition events by atomic force microscopy. *Proc Natl Acad Sci* 93(8):3477–3481

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

Stefano Angioletti-Uberti

obtained his PhD in Materials Science at Imperial College London in 2010, working on solid-liquid interfaces. He then moved to the University of Cambridge with Prof. Daan Frenkel, shifting his research to Soft Matter. In 2013, he was awarded an Alexander von Humboldt Fellowship and moved to Berlin, where he worked on protein adsorption and nanoreactors with Prof. Joachim Dzubiella and Prof. Matthias Ballauff. After a

short stint as Professor of Soft Matter at Beijing University of Chemical Technology, since October 2016, he came back as a Lecturer at Imperial College London, applying Soft Matter theory and modelling to smart materials. Prof. Ballauff has been an inspiring influence, making him truly appreciate the importance of collaborating with great experimentalists and always providing plenty of interesting intellectual challenges. Prof. Ballauff also gave him a strong support as a young academic and he will always be grateful to him: “Buon Compleanno Professore!”