doi:10.1006/jmbi.2001.5103 available online at http://www.idealibrary.com on IDEAL® *J. Mol. Biol.* (2001) **314**, 589–605

JMB

Forces and Energetics of Hapten-Antibody Dissociation: A Biased Molecular Dynamics Simulation Study

Emanuele Paci[1,2], Amedeo Caflisch[3], Andreas Plückthun[3] and Martin Karplus[1,4]*

[1]*Laboratoire de Chimie Biophysique Institut Le Bel Université Louis Pasteur 4 rue Blaise Pascal 67000, Strasbourg, France*

[2]*Oxford Centre for Molecular Sciences New Chemistry Laboratory, University of Oxford, South Parks Road Oxford OX1 3QT, UK*

[3]*Biochemisches Institut Universität Zürich, Winterthurerstrasse 190 CH-8057 Zürich, Switzerland*

[4]*Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street Cambridge, MA 02138, USA*

**Corresponding author*

The unbinding of fluorescein from the single-chain Fv fragment of the 4D5Flu antibody is investigated by biased molecular dynamics with an implicit solvation model. To obtain statistically meaningful results, a large number of unbinding trajectories are calculated; they involve a total simulation time of more than 200 ns. Simulations are carried out with a time-dependent perturbation and in the presence of a constant force. The two techniques, which provide complementary information, induce unbinding by favoring an increase in the distance between the ligand and the antibody. This distance is an appropriate progress variable for the dissociation reaction and permits direct comparison of the unbinding forces in the simulations with data from atomic force microscopy (AFM). The time-dependent perturbation generates unfolding pathways that are close to equilibrium and can be used to reconstruct the mean force; i.e. the derivative of the potential of mean force, along the reaction coordinate. This is supported by an analysis of the overall unbinding profile and the magnitude of the mean force, which are similar to those of the unbinding force (i.e. the external force due to the time-dependent perturbation) averaged over several unbinding events.

The multiple simulations show that unbinding proceeds along a rather well-defined pathway for a broad range of effective pulling speeds. Initially, there is a distortion of the protein localized in the C-terminal region followed by the fluorescein exit from the binding site. This occurs in steps that involve breaking of specific electrostatic and van der Waals interactions. It appears that the simulations do not explore the same barriers as those measured in the AFM experiments because of the much higher unfolding speed in the former. The dependence of the force on the logarithm of the loading rate is linear and the slope is higher than in the AFM, in agreement with experiment in other systems, where different slopes were observed for different regimes. Based on the unbinding events, mutations in the 4D5Flu antigen binding site are predicted to result in significant changes in the unbinding force.

Keywords: single-chain antibody; fluorescein; binding; molecular dynamics; atomic force microscopy

Introduction

Understanding the binding of a ligand to its macromolecular receptor is important both for its fundamental interest and as an aid in designing new drug candidates. Recently data from equilibrium and kinetic measurements have been supplemented by single molecule experiments in which an external force is applied to dissociate the ligand-receptor complex.[1–11] Molecular dynamics

Abbreviations used: AFM, atomic force microscopy; BDM, biased molecular dynamics; CDR, complementarity-determining region; CFMD, constant force molecular dynamics.

E-mail address of the corresponding author: marci@tammy.harvard.edu

0022-2836/01/030589–17 $35.00/0

simulations[12] which complemented the original atomic force microscopic analysis of the streptavidin-biotin complex[1,2] gave some insights into the interactions along the dissociation path.

One binding phenomenon of particular interest concerns the interactions between an antibody and its strongly binding hapten. Plückthun and coworkers[6,13] have recently used atomic force microscopy (AFM) to measure the binding force between recombinant antibody single-chain Fv (scFv) fragment[14] and fluorescein and compared it to the results measured in solution binding; an scFv fragment is the minimal part of an antibody molecule that includes the complete antigen-binding site.[15–17] Fluorescein is ideal for such a study because it is a rigid molecule, which avoids complications arising from more than one conformation of the ligand.

Here, we report the results of multiple molecular dynamics simulations of the unbinding process. We use a molecular-mechanics potential[18] in the CHARMM program,[43] an implicit solvent model, EEF1[19] and a time-dependent or constant force perturbation applied to a suitable reaction coordinate. In analogy with the AFM experiments, the reaction coordinate was defined as a function of the distance between the C terminus of the scFv fragment and the C2 atom of fluorescein (see Figure 1(a)). The methods employed are the same as those used to induce unfolding of β-sandwich and α-helical proteins on a computationally accessible timescale.[20,21]

Computational limitations require that the simulations separating the antigen from the antibody be performed about seven to nine orders of magnitude faster than the AFM experiments; e.g. in 1 ns instead of 10 ms to one second. The biased molecular dynamics (BMD) method is well suited for such a study because the time-dependent perturbation promotes unbinding by adding a minimal perturbation that has little effect on the short-time dynamics of the molecule and allows a uniform sampling of the whole range of values of the reaction coordinate. The application of a constant force, while more easily understandable in experimental terms, generally requires a stronger perturbation and induces a stepwise unbinding, so that mostly metastable states (i.e. states that are stable on the timescale of the simulation) are sampled at different force magnitudes. The use of an implicit solvent model means that the solvent responds adiabatically to a change of the reaction coordinate, as it would on the millisecond timescale. Because the implicit solvent simulations are relatively fast, multiple simulations can be performed to obtain statistically meaningful results.

The present simulation methodology differs somewhat from the original study of the biotin-streptavidin unbinding[12] and subsequent work on the same system,[22,23] as well as that on another antibody-hapten complex.[24] Either no solvent or explicit solvent were used in these studies. In the unbinding simulations, one atom of the receptor was kept fixed and one atom of the ligand was attached to a spring that was pulled at constant velocity, instead of the variable or constant biasing force. As in the present work, the simulation time

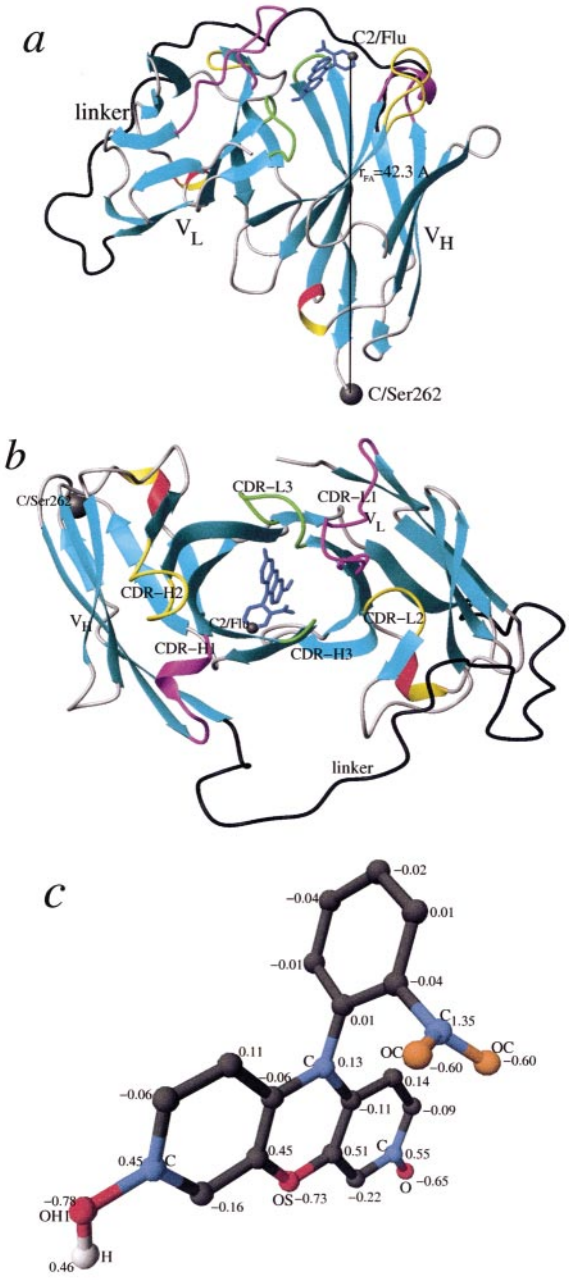

Figure 1. Model[14] of the scFv (ribbons) bound to fluorescein (thick line). (a) The vector $\mathbf{r}_{FA}$ joining C-2 of fluorescein to the C terminus of the antibody, which defines the reaction coordinate, is in the plane of the Figure. (b) View of the scFv fragment from the hapten binding site. (c) Fluorescein structure with atom types and atomic partial charges; H (hydrogen), O (carbonyl oxygen), OS (ester oxygen), OH1 (hydroxy oxygen), OC (carboxy oxygen), C (carbonyl carbon); all the others are CR1E (extended carbon in aromatic ring with one hydrogen atom). This Figure was made with the program MOLMOL.[53]

was many orders of magnitude shorter than the timescale of the experiment.

The present simulations examine the response of both the ligand and the receptor to different external perturbations that have the same macroscopic effect, that of inducing unbinding over a broad range of timescales. The main goal is to determine the magnitudes of the forces and the energetics involved, and to probe the atomic interactions that contribute along the unbinding pathway(s). By carrying out multiple simulations (77 out of 85 result in full unbinding) we are able to show that the average unbinding pathway and the force along that pathway have rather well-defined average properties. Moreover, the mean force (i.e. the derivative of the potential of mean force) along the average unbinding pathway is very similar to the average external force obtained from the time-dependent biasing perturbation. This indicates that the perturbation is such that unbinding pathways are sampled under quasi-equilibrium conditions, which makes it possible to estimate free energy profiles. Thus, the present equilibrium approach, which corresponds to the so-called blue moon method,[25] is complementary to the non-equilibrium free energy simulations proposed recently.[26,27]

The energetics during unbinding are found to be dominated by the intermolecular energy with the van der Waals and electrostatic contribution changing in a corresponding fashion. Further, there is a partial compensation between total electrostatic and solvation energy, as expected, and the variation of the intra-receptor energy is small. The energetics and mechanism described here for the dissociation of fluorescein from an antibody scFv fragment provide an example of mechanically induced hapten-antibody dissociation.

Results

The control simulation

The scFv-fluorescein complex is stable over the 2.4 ns time of the control run (Figure 2). The overall conformation and secondary structural elements are preserved, as is the hapten-binding site. As expected, the largest fluctuations (not shown) are localized in the linker atoms (see Materials and Methods). This is consistent with an analysis of the ^{15}N NOE data for the complex between the scFv fragment of the antibody McPC603 and phosphorylcholine.[28] During the last 800 ps of the control run, the average non-linker RMSD from the original model structure (with superposition of the non-linker C^{α} atoms) is 2.7 Å for C^{α} atoms and 3.3 Å for all heavy atoms. The RMSD averaged over the final 800 ps for the C^{α} atoms in the framework region is 2.2 Å (maximum of 2.9 Å) and for the C^{α} atoms in the complementarity-determining region (CDR) loops (residues 24-39, 55-61, 94-102, 175-179, 194-212 and 245-251) is 3.3 Å (maximum of 4 Å). The CDRs of the V_L domain deviate by 3.2 Å on average, while the CDRs of the V_H domain deviate by only 1.6 Å. These values are comparable to the 3.2 Å deviation from the X-ray structure for the backbone heavy atoms reported by Heymann & Grubmüller[24] in a recent 1.3 ns control simulation of the monoclonal antibody AN02 Fab fragment in a stochastic boundary droplet containing 13,461 water molecules and 75 ions.

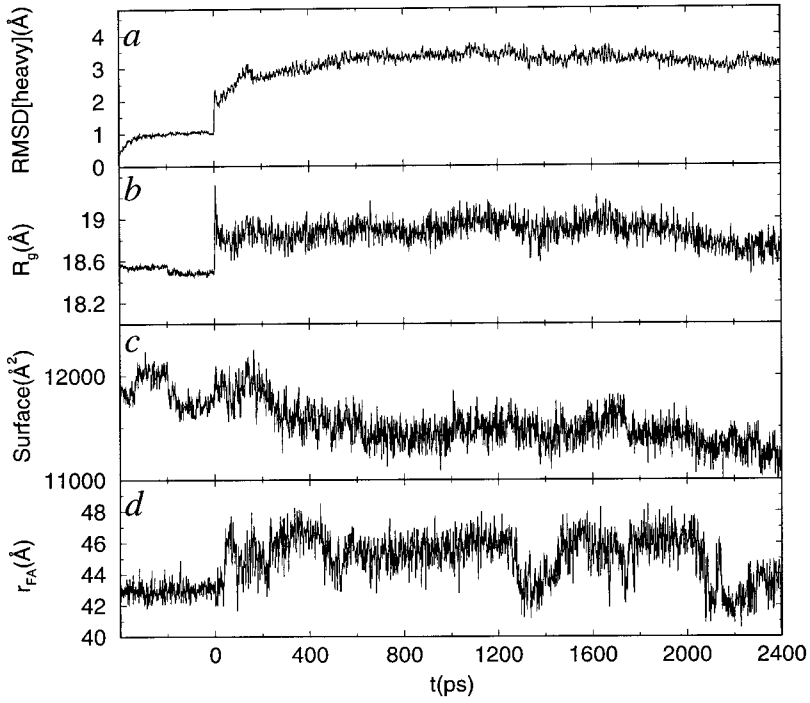

Figure 2. Control simulation. (a) Heavy-atom RMSD of the V_L and V_H domains (excluding the linker). (b) Radius of gyration of scFv. (c) Solvent-accessible surface as defined by Richards.[54] (d) Reaction coordinate: distance between C-2 of the fluorescein and the C terminus of the antibody. The zero of the timescale corresponds to the beginning of the unconstrained control run (i.e. negative times correspond to the 200 ps heating phase and 200 ps equilibration with constrained backbone atoms; see the text).

The relatively large deviation for the light chain is due to the CDR-L1 loop. This is in accord with the fact that the most significant discrepancies between the 1.85 Å structure[29] and the 2.7 Å structure[30] are localized in this loop. However, one cannot exclude the possibility that some of the deviation might originate from the approximations in the model.

Properties of the antibody during the unbiased control simulation shown in Figure 2, together with the behavior of the different energy components (not shown), indicate that the system is equilibrated and stable. The value of the radius of gyration averaged over the last 800 ps is 18.91 Å, which is close to the value of the initial energy-minimized conformation (18.6 Å). The ligand remains bound to the antibody during the control run. The average value of r_{FA} during the last 800 ps of the control simulation is 45.3 Å with a standard deviation of 1.4 Å; the initial value is 43.3 Å. The small difference is due primarily to the fluctuations in the C-terminal segment of the antibody. The solvent-accessible surface decreases slightly from an initial value of 11,840 Å^2 to an average value of 11,480 Å^2 during the final 800 ps. Overall, there is a tendency to approach closer to the crystal structure values in the final stages of the equilibration period, suggesting that the model would be stable on longer timescales.

The unbinding simulations

Unbinding is induced by adding an external perturbation that favors the increase in the distance between antibody and ligand. Two kinds of time-dependent external perturbations have been used. The timescale for the unbinding is modulated by an adjustable parameter α in the BMD approach (see Materials and Methods) or by the constant external force F in an alternative approach where a constant force is applied. In both cases, several values of the parameters α or F were used to investigate the dependence of the unbinding pathways and of the measured properties along the same pathways. For each value of α or F, six simulations were performed from different bound configurations (coordinates and velocities extracted at the end of every 0.4 ns interval of the control run). Table 1 lists all the simulations performed: these were stopped when unbinding occurred, i.e. the value of r_{FA} was greater than 80 Å.

In general, a smaller value of α (see equation (8)) or of F (see equation (10)) is expected to yield more realistic results, since the perturbation has a smaller effect on the system and the unbinding process is closer to the unbiased one. However, this has to be balanced by the need to simulate the transition in a finite time. Although the exact kinetics depends on the initial conditions (i.e. each trajectory is somewhat different), it was found that, in the BMD case, a value $\alpha = 0.0002$ kcal mol^{-1} Å^{-4} (1 cal = 4.184 J; 1 kcal mol^{-1} Å^{-1} = 69.4786 pN) induces complete unbinding in less than 8 ns in all the six simulations. Higher values of the parameter α lead to unbinding on a shorter timescale, and the comparison provides information on the effect of α on the computed quantities. Also, six runs with a lower α value (0.0001 kcal mol^{-1} Å^{-4}) were performed for 10 ns or longer and only two led to complete separation.

Biased molecular dynamics

With a coupling $\alpha = 0.0002$ kcal mol^{-1} Å^{-4}, full unbinding always occurs in less than 8 ns. The results from simulations with larger or lower values of α are mentioned as a basis of comparison and/or to underline important differences. The reaction coordinate r_{FA} as a function of simulation time is shown in Figure 3. Once r_{FA} has increased

Table 1. Overview of the simulations

Simulation method	Force constant	Starting points on the equilibration trajectory					
		$t = 0.4$	$t = 0.8$	$t = 1.2$	$t = 1.6$	$t = 2.0$	$t = 2.4$
BMD	α=0.0001	**30 ns**	14.8 ns	**10 ns**	9.5 ns	**10 ns**	**10 ns**
BMD	α=0.0002	3.2 ns	7.4 ns	3.1 ns	2.5 ns	3.7 ns	2.3 ns
BMD	α=0.0003	0.64 ns	2.6 ns	2.4 ns	1.1 ns	0.80 ns	2.8 ns
BMD	α=0.0005	0.53 ns	0.79 ns	3.5 ns	1 ns	0.89 ns	0.57 ns
BMD	α=0.001	228 ps	179 ps	318 ps	299 ps	289 ps	198 ps
BMD	α=0.002	126 ps	164 ps	200 ps	119 ps	75 ps	133 ps
BMD	α=0.003	54 ps	69 ps	80 ps	71 ps	80 ps	68 ps
BMD	α=0.004	57 ps	60 ps	62 ps	55 ps	35 ps	66 ps
CFMD	F=150 pN	**20 ns**					
CFMD	F=200 pN	5.2 ns	5.1 ns	2.4 ns	**6 ns**	1.6 ns	3.4 ns
CFMD	F=250 pN	0.21 ns	2.6 ns	0.97 ns	0.97 ns	**4 ns**	2.2 ns
CFMD	F=300 pN	63 ps	226 ps	545 ps	788 ps	384 ps	85 ps
CFMD	F=375 pN	269 ps	69 ps	156 ps	**1 ns**	128 ps	88 ps
CFMD	F=500 pN	34 ps	16 ps	41 ps	55 ps	28 ps	26 ps
CFMD	F=625 pN	8.4 ps	4.5 ps	11 ps	12 ps	7.6 ps	7.7 ps

In boldface are those that did not result in full unbinding ($r_{FA} > 80$ Å). BMD; biased molecular dynamics; CFMD, constant force molecular dynamics. Initial conformations where taken after t ns of the control run (see the text).

to about 70 Å from its initial value near 45 Å, the unbinding proceeds without hindrance, since all interaction terms have gone to zero. In one run, started from the t = 0.8 ns conformation of the control simulation, the r_{FA} distance has increased by only 10 Å after 7 ns, and the unbinding is completed only after continuing the simulation to 7.4 ns. It is clear from the trajectories shown in Figure 3, as well as from those for different values of α, that unbinding proceeds by a sequence of relative discrete jumps. This suggests that there are a series of barriers for r_{FA} values between 47 and 61 Å. The intra-receptor and ligand-receptor interactions involved are analyzed below.

The timescale for unbinding has a random component, as expected for a thermally activated process. It is not imposed externally as it would be if the two ends were pulled at constant speed. This is an important feature of BMD, in contrast to other methods. In BMD, the effective unbinding speed in the regions characterized by barriers can be as low as 0.01 Å ns^{-1}; however, even this speed is still $\sim 10^3$ faster than the fastest experimental pulling speed (about 10^{-5} Å ns^{-1}). A longer search time is allowed when the system encounters a barrier, which makes the motion closer to what might be expected under the quasi-equilibrium condition of the AFM experiment. The values of r_{FA} as a function of time in Figure 3 show several bottleneck regions (i.e. regions where intermolecular interactions oppose a further increase of r_{FA}, which then remains almost constant for long time intervals). For different trajectories, the bottlenecks do not occur at the same time, though they occur mainly at corresponding distances.

A twice smaller value of α (α = 0.0001 kcal mol^{-1} $Å^{-4}$) is not enough to always induce unbinding in 10 ns or longer simulations (see Table 1). The magnitude of the RMS fluctuations in the reaction coordinate (relative to the running average over 10 ps intervals) is larger with α = 0.0001 kcal mol^{-1} $Å^{-4}$ (0.48 Å) than with α = 0.0002 (0.44 Å for $r_{FA} < 55$ Å), and somewhat smaller than in the control run (0.73 Å). Hence, the perturbation reduces the spontaneous fluctuations of the reaction coordinate by about 35 % to 40 %, when the value of the coupling α is large enough to observe full unbinding on the nanosecond timescale. However, the overall atomic fluctuations of the protein are not altered significantly.

Constant force molecular dynamics (CFMD)

In the BMD approach, the bias is a time-dependent perturbation that is non-zero only for displacements in the direction that decreases the reaction coordinate. Although the BMD method has important advantages, as described above, complementary information on the unbinding process can be obtained by switching on a constant force on r_{FA} at time zero to pull apart the atoms F and A (see Materials and Methods). This approach has been used in related studies of the stretch-induced unfolding of proteins.[20,31] Such an external force adds a term Fr_{FA} to the potential energy function. This "tilts" the potential energy surface and modifies the shape, magnitude and location of the minima and transition states.[32] In fact, in the presence of a constant force the unbound state becomes energetically the most favorable, although, depending on the magnitude of the force, the bound state may remain a local minimum. Unlike BMD, the magnitude of the force is fixed, so that a force large enough to overcome a barrier situated at a value of r_{FA} close to the bound state might be too large to carefully sample conformations at larger r_{FA} values. However, such a constant force can serve to delineate barriers that cannot be overcome on the simulation timescale, so that comparison of

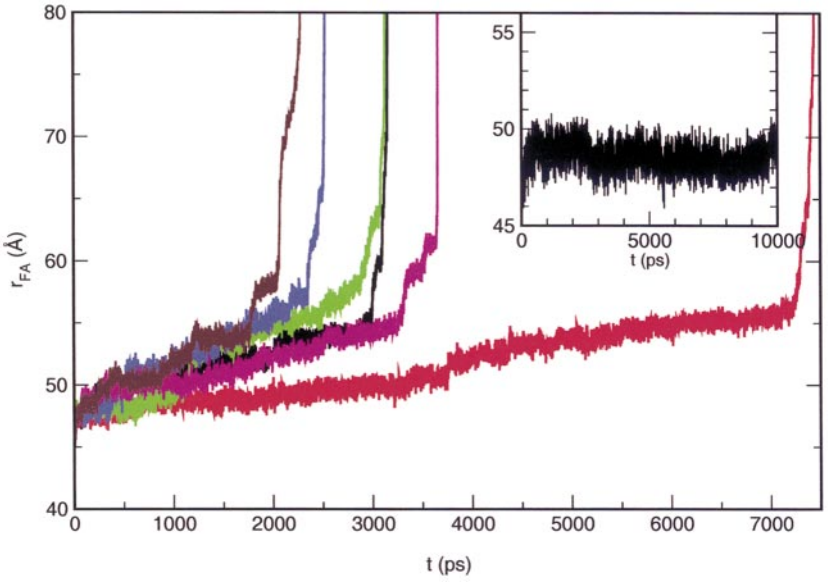

Figure 3. The reaction coordinate r_{FA} as a function of time during six BMD simulations. The coupling parameter α is 0.0002 kcal mol^{-1} $Å^{-4}$, which is always sufficient to dissociate the ligand from its substrate on a 10 ns timescale. The inset shows the behavior of r_{FA} in one of the α = 0.0001 kcal mol^{-1} $Å^{-4}$ simulations that did not yield unbinding.

the results obtained with different values of force are of interest.

The smallest force that induced unbinding was 200 pN (five simulations over six, in less than 6 ns); for this force, Figure 4 shows r_{FA} as a function of the simulation time. In one simulation, the reaction coordinate does not increase above $r_{FA} \sim 53$ Å within 6 ns. From the behavior of the other trajectories, an r_{FA} value of about 53 Å corresponds to one of a series of discrete metastable states for the potential modified by the constant force. Bottlenecks around $r_{FA} = 58$ Å and $r_{FA} = 66$ Å are particularly clear in the Figure. With a 150 pN constant force, a 20 ns simulation did not yield unbinding.

Comparison of results from the two unbinding procedures

To analyze and compare the BMD and CFMD results, it is useful to consider two different coordinates for describing the unbinding process. The first is the reaction coordinate, r_{FA}, used in the simulations and the second is the distance, d_{CM}, between the centers of mass of the fluorescein molecule and the antibody, where the seven residues closest to the C terminus are disregarded. The latter allows us to focus on the relative motion of the fluorescein and the ligand, without the protein distortion that contributes significantly to r_{FA}. In Figure 5, histograms are plotted for the times (in ps) during which the r_{FA} values were in given intervals for the BMD and CFMD simulations; the results correspond to a sum over all the simulations of a given kind (i.e. with a given value of α or F). BMD results for the trajectories with α between 0.0001 and 0.002 kcal mol^{-1} Å^{-4} are shown in Figure 5(a); analogous results for the CFMD trajectories for F between 150 and 500 pN are shown in Figure 5(b). For the number of trajectories used in the averages, see Materials and Methods.

The histograms provide information that amplifies the analysis of the varying slopes seen in the unfolding diagrams in Figures 3 and 4. Choosing r_{FA} as the progress variable, the BMD trajectories (Figure 5(a) show a probability distribution that varies significantly as α increases. At the smallest α value ($\alpha = 0.0001$ kcal mol^{-1} Å^{-4}) there is a single well-defined peak at 52 Å (bottleneck) that moves inwards as α increases to 0.0003 and then disappears. Simultaneously, a second peak appears at $\alpha = 0.0002$, which increases in probability and moves inward at $\alpha = 0.0005$. Finally, at the largest α value (0.002), there is a broad distribution of residence times, which suggests that there is no major barrier; i.e. the BMD perturbation is so large that the intrinsic potential of the system no longer plays a significant role. The constant force result (Figure 5(b)) reveals a similar transition between different regimes. One occurs at forces below 300 pN, where the most populated state is around $r_{FA} = 50$ Å, and another one at forces larger than 300 pN where the most populated state is around $r_{FA} = 60$ Å. At 300 pN, the longest-lived intermediate is present between $r_{FA} = 50$ and 54 Å.

The above results correspond to the different ways in which the effective energy surface is perturbed in the two approaches: BMD introduces a bias that locally ''flattens'' the energy surface so that the reaction coordinate increases rather smoothly as barriers are crossed, while CFMD globally tilts the energy surface, increasing the probability of jumps over the barriers. If the process of unbinding could be described effectively as the crossing of multiple barriers along a one-

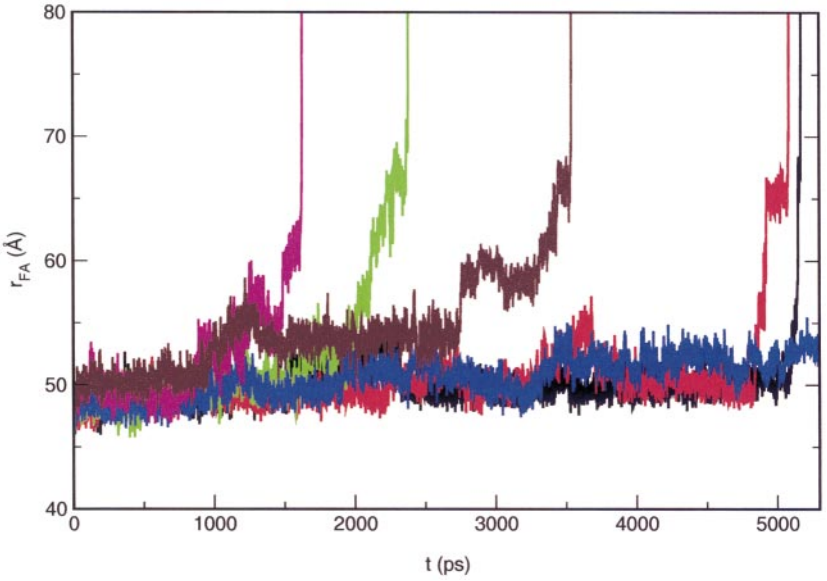

Figure 4. Time behavior of the reaction coordinate during six CFMD simulations. In five cases out of six, a 200 pN force is sufficient to dissociate the ligand from its substrate completely in less than 6 ns.

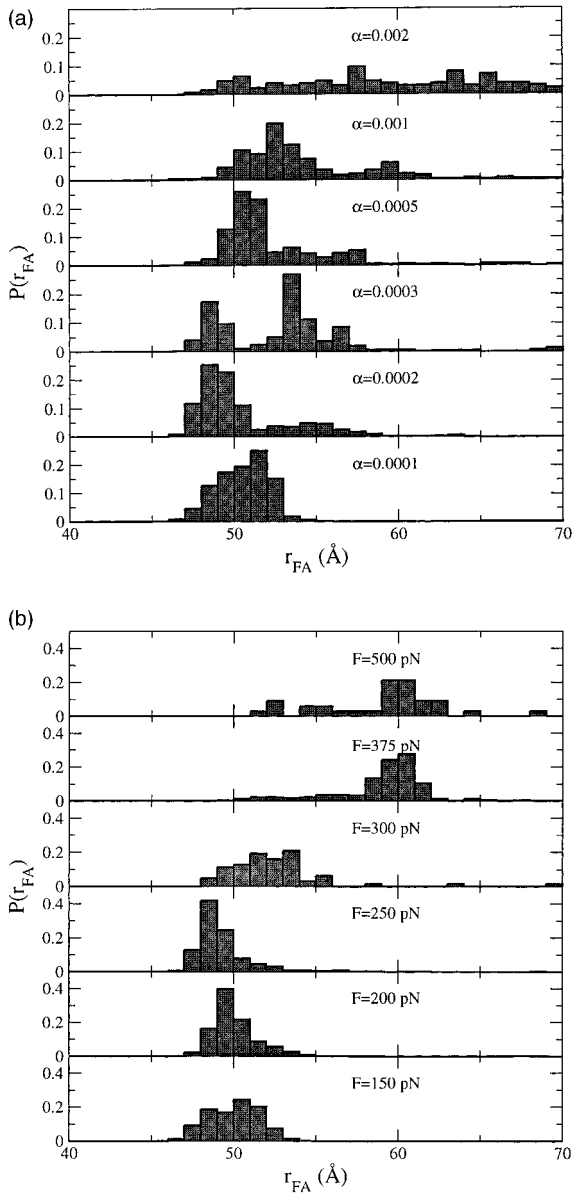

Figure 5. (a) Frequency of occurrence of r_{FA} for the BMD simulations with $0.002 \leqslant \alpha \leqslant 0.0001$ kcal mol^{-1} Å^{-4} and (b) for the constant force trajectories with $150 \leqslant F \leqslant 500$ pN.

dimensional reaction coordinate r_{FA}, a constant force would decrease the outer barriers more than inner barriers (see below), possibly displacing the position of the main transition state and affecting the unbinding pathway. This is the suggested origin of the different regimes observed experimentally by Merkel *et al.*[33] for the forced unbinding of streptavidin from biotin. In our simulations, we observe a more complex behavior. There are several regimes, in each of which the barrier moves inward as the perturbation increases, while in the different regimes the barrier moves outward as the force increases.

Unbinding forces in the BMD trajectories

Figure 6 shows the magnitude of the force as a function of r_{FA} for the six unbinding trajectories with $\alpha = 0.0002$ kcal mol^{-1} Å^{-4}; the force was calculated by block averaging over intervals of 2 Å in r_{FA}, since there are large instantaneous fluctuations that obscure the overall behavior. The force profiles as a function of r_{FA} for the different trajectories are similar, though not identical. In all cases, the force rises quickly as r_{FA} increases by only a few ångström units relative to the initial value. The force tends to have an absolute maximum of about 250 pN for r_{FA} values between 48 and 54 Å. In some cases, the force peak is broader (48 to 64 Å) and certain trajectories have secondary peaks (about 150-200 pN) at r_{FA} values between 63 and 72 Å. The force decreases to zero at r_{FA} between 75 and 80 Å, i.e. at the distance where the hapten and the scFv fragment are separated completely.

Figure 7 shows the unbinding forces averaged over the BMD trajectories (Table 1) at different α values. The overall shapes of the force curves are similar, with a major broad peak between 50 and 60 Å, where the magnitude increases for increasing α, as expected. The results are compared with the mean force below.

Mean force along the reaction coordinate

Since BMD is expected to yield pathways that are close to equilibrium when a low value of α is used, it is of interest to determine the mean force along the reaction coordinate (equation (11)). The mean force profile was obtained by using the structures calculated during the simulation as starting points for runs with r_{FA} constrained to different values. Specifically, the initial phase points (positions and velocities) for the constrained trajectories were taken either from the control run (for $r_{FA} = 45.8$ Å) or from the BMD unbinding runs. Simulations with r_{FA} constrained to 49.9, 55.2, 59.4, 70.2, 75.6 and 78.8 Å, each for 1 ns, were performed to calculate the first term on the right of equation (11); in general, it converges within the first 200 ps. The second term in equation (11), which is always less than ~2 pN, was neglected. The calculated mean force is shown in Figure 7 and is compared to the average external forces obtained during BMD unbinding simulations with different α values. Both the magnitudes of the forces and their dependence on r_{FA} are similar. Moreover, for $r_{FA} < 65$ Å the BMD force for the lowest α that leads to unbinding ($\alpha = 0.0002$ kcal mol^{-1} Å^{-4}) gives results closest to the mean force, while slightly larger forces are obtained for larger α. This supports the conclusion that the initial conformations for the mean force calculations, although extracted from BMD trajectories, are not far from the equilibrium unfolding pathways. The standard deviation for the mean force computed by averaging over 2 ps blocks ranges from 1.1 to 5.4 pN. Use of larger blocks suggests that the error

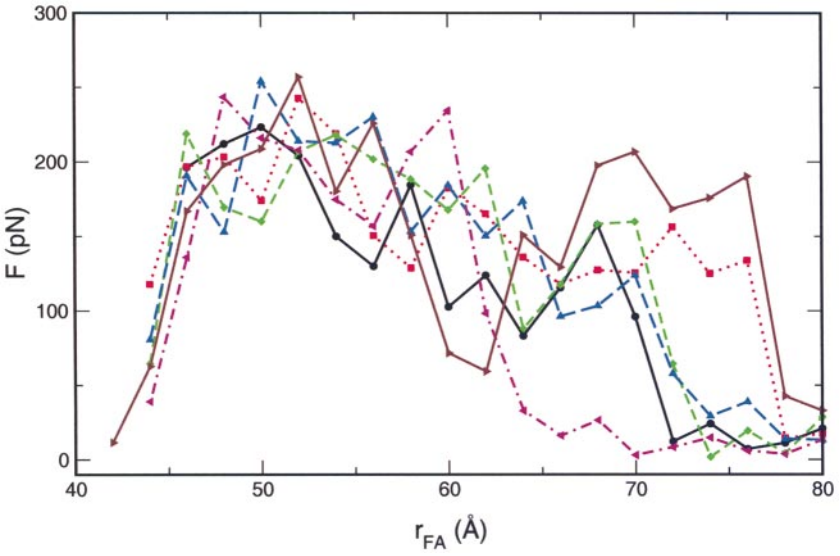

Figure 6. Force due to the external perturbation as a function of the distance for the six trajectories with $\alpha = 0.0002$ kcal mol^{-1} Å^{-4}.

might be underestimated in this way, as is often the case in this type of analysis; e.g. for the trajectory with r_{FA} constrained at 55.2 Å, the difference between the averages over the first and the second half nanosecond is about 19 pN, compared to the mean force of 352 pN. Although this indicates a larger error, it is still small relative to the total mean force.

The mean force shows smooth behavior as a function of r_{FA}, except for the point at $r_{FA} = 59.4$ Å. This is lower than might be expected from the other values, though not in disagreement with the double peak observed in most BMD results with $\alpha \geqslant 0.001$. Analysis showed that for r_{FA} constrained to 59.4 Å, the force has an average value of 409 pN during the first 0.7 ns and 47.3 pN during the last 0.7 ns (from 1.3 to 2 ns). An examination of the constrained trajectory revealed that during the first 0.7 ns the fluorescein ligand is buried in the antigen-binding site, while the C-terminal portion of the scFv fragment (from Gly255 to Ser262) is completely stretched and separated from the rest of the antibody (see also below). During 0.7 ns to 1.3 ns, the fluorescein ligand leaves the binding site, and after 1.3 ns it is almost completely unbound except for the oxygen atoms of the three-

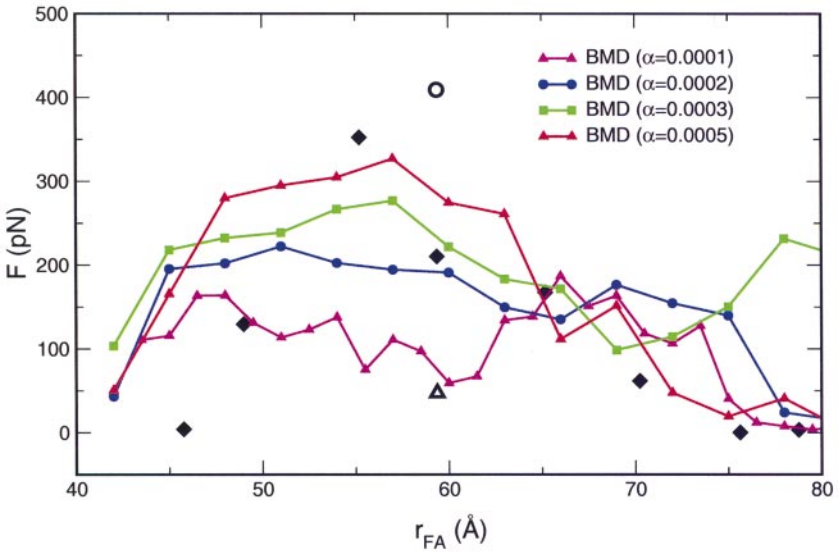

Figure 7. Mean force (diamonds) obtained from eight MD simulations using equation (11), started from conformations from BMD unbinding runs (see the text). Simulations were 1 ns long except for $r_{FA} = 59.4$ Å, which was run for 2 ns. At $r_{FA} = 59.4$ Å, the three values of the mean force correspond to the average over the whole 2 ns trajectory (diamond), the first 700 ps (open circle), and the last 700 ps (open triangle). The latter two averages were computed because the mean force jumps from a high to a low value during the 2 ns simulation (see the text). The curves represent the external force averaged over the BMD trajectories at the four smallest values of α used (see Table 1).

ring system, which still accept hydrogen bonds from the side-chains of His31, Gln33, and the main chain NH of Asn35. The departure of fluorescein from the binding site is concomitant with re-association of the C-terminal residues with the rest of the scFv fragment. A parallel β-sheet involving residues 150-154 and 257-261 is formed; it is shifted by two residues with respect to the native parallel β-arrangement. This indicates that the effective barrier to breaking the contacts between fluorescein and the binding site residues is high, in accord with the large force required. For r_{FA} values between 45 and 55 Å, it is apparently easier to increase the value of the reaction coordinate by the displacement and partial stretching of the C-terminal residues than by unbinding of the fluorescein. The time-dependence of distance d_{CM} between the centers of mass of the antibody and the fluorescein clearly reflects this result. For the first 0.7 ns, d_{CM} is about 16 Å, and between 0.7 and 1.3 ns d_{CM} increases to about 27 Å in three steps. The constrained simulation at $r_{FA} = 59.4$ Å indicates that the initial conformation is not an equilibrium one, even on the nanosecond timescale. This emphasizes the fact that the unbinding takes place in a many-dimensional space and that use of a single reaction coordinate involves a projection. For most of the unbinding trajectories, r_{FA} is a meaningful choice. However, at $r_{FA} = 59.4$ Å, a second dimension is required. While the reaction coordinate r_{FA} is kept constant, d_{CM} changes significantly and results in a large variation of the mean force over different segments of the simulation with r_{FA} fixed. The choice of d_{CM} as reaction coordinate would possibly improve the convergence because the system has less freedom to relax when d_{CM}, and not r_{FA}, is constrained to a fixed value. However, we have chosen r_{FA} as the reaction coordinate, because it corresponds to that used in the experiment for the anchoring sites of the antibody and fluorescein to the AFM device. It is clear that the slower the forced unbinding is, the more critical is the role of the relaxation in a direction orthogonal to the chosen reaction coordinate; i.e. in the present case, that relaxation is on the nanosecond timescale of the simulation. Comparing Figures 6 and 7, it is evident that corresponding events in the different simulations happen for slightly different values of r_{FA} even at a given α value. Thus, in the average over a set of trajectories, the barriers at different positions appear as a single broad barrier (Figure 7). Although the AFM experiments are on a much slower timescale (millisecond to second), corresponding relaxation has been observed in single-molecule experiments,[34] as well as in time-resolved photodissociation experiments.[35] It is clear that protein relaxation covers many orders of magnitude in time that extend into the second range. Thus, processes along directions orthogonal to the pulling coordinate are likely to complicate the experimental result.

Except for the trajectory with r_{FA} constrained at 59.4 Å, convergence is achieved on the nanosecond timescale, as monitored by relevant parameters. The total effective energy is essentially constant during the constrained simulations, and the RMSD of the protein from the starting configurations for each r_{FA} converges quickly to low values, between 2 and 4 Å for heavy atoms (without the linker). This is in accord with the conclusion that conformations extracted from the biased unbinding are, in most cases, similar to equilibrium conformations obtained with a constraint on the reaction coordinate. In other words, the non-equilibrium sampling provides an estimation of the conditional probability of a conformation, given a value of r_{FA}.

Energetics of unbinding

Figure 8 shows the total effective energy and the non-bonded energy terms (van der Waals, electrostatic, and solvation, see Materials and Methods) as a function of r_{FA} for the six runs with $\alpha = 0.0005$ kcal mol^{-1} Å^{-4}; the values correspond to averages over intervals of 3 Å. Although the fluctuations are large, the total energy can be seen to increase by approximately 100 kcal mol^{-1}. This is due to increases in both the van der Waals and electrostatic interactions, which are not compensated completely by the decrease in the solvation energy. The incomplete compensation is due primarily to the exposure to solvent of the apolar sur-

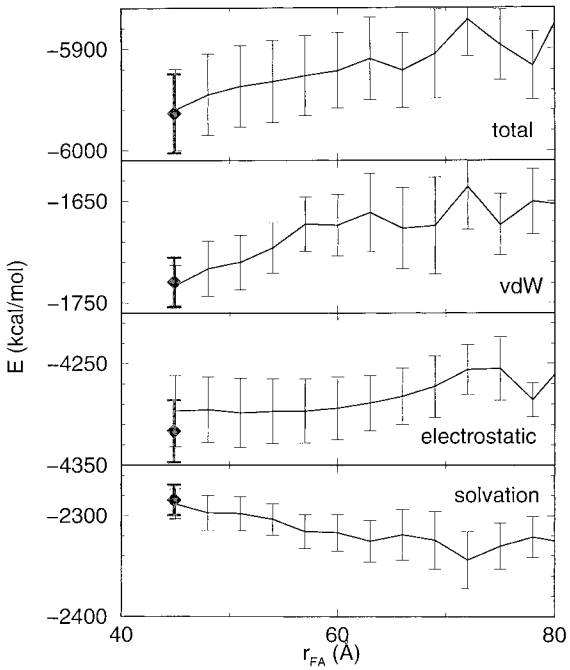

Figure 8. Components of the non-bonded potential energy of the whole system. The diamond represents the corresponding value over the last 800 ps on the control run, while the error bar is the RMS fluctuation of the corresponding quantity within the 3 Å interval of r_{FA} over which it is averaged. (To make comparison simple, the scale is the same in all the windows of the plot).

face of fluorescein during unbinding. The bonded terms are almost constant along the unbinding pathway (not shown). We note that this analysis is easy to do only with the implicit solvation model, which provides a value, albeit approximate, for the free energy of solvation. By contrast, the estimation of solvation free energies from explicit water simulations would require very costly evaluations of the potential of mean force along the unbinding pathways.

The protein energy, which has large fluctuations, is correlated, for r_{FA} values up to about 70 Å, with the RMSD from the initial structure. This suggests that a moderate increase (30 kcal mol^{-1}) in the protein energy arises from the deformation, which precedes the breaking of the ligand-receptor contacts (see above). The ligand energy is not altered during and after unbinding because of its rigidity.

Figure 9 shows the interaction energy between the ligand and the receptor as a function of r_{FA}. The value of r_{FA} (but not the distance between fluorescein and the hapten-binding site) increases by 10 Å up to about 55 Å without any noticeable difference in the intermolecular energy. This is due to the displacement of the extended strand consisting of the seven C-terminal residues of the scFv fragment that was described above, without any significant change in the conformation of the binding site and the binding mode of fluorescein; the distance d_{CM} is essentially at its equilibrium value. The intermolecular energy is about −70 kcal mol^{-1} in the native bound conformation and goes to zero in the unbound state at ~80 Å, the distance at which ligand and antibody no longer interact. The largest change takes place at r_{FA} values between 55 and 57 Å because of the sudden loss of two to four intermolecular hydrogen bonds and favorable van der Waals interactions between fluorescein and a number of aromatic side-chains (see the next section).

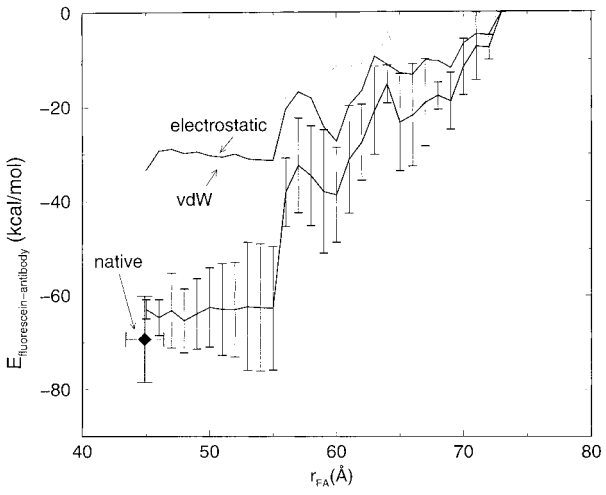

Figure 9. Fluorescein-antibody interaction energy (van der Waals, electrostatic and their sum) as a function of the reaction coordinate. Each point represents an average over all the configurations from the unbinding trajectories for r_{FA} in a 1 Å interval, and error bars represent the magnitude of fluctuations.

Structural analysis of the unbinding

We first describe the sequence of events for the runs with small values of the bias parameters α or F. For r_{FA} values up to about 55 Å the position and orientation of fluorescein in the antigen-binding site varies only slightly. The increase of about 10 Å in the r_{FA} distance (from 45 to 55 Å) is primarily due to a displacement of the extended strand consisting of the seven C-terminal residues of the scFv fragment (Figure 10), without any significant change in the binding mode or interaction energy of fluorescein. The C-terminal strand moves away from the rest of the antibody as a consequence of the force acting on it; this results from the choice of the reaction coordinate, which corresponds to the point of attachment of the antibody in the AFM experiment. Since the scFv fragment consists of only the two variable domains, it is possible that the flexibility of its C-terminal residues is due to the lack of the stabilizing interactions provided by the constant domains in the complete Fab fragment.

To eliminate the complication of the C-terminal end we consider the distance d_{CM} between the center of fluorescein and the protein as the reaction coordinate. As d_{CM} increases from $d = 14$ Å to 17 Å, the fluorescein is slowly moving out of the pocket with certain aromatic residues (Tyr37, Trp177, and Tyr247) and hydrogen bonding groups (Ser96 and Arg196) following the fluorescein. When d_{CM} is between 17 and 19 Å, most of the trajectories undergo a sudden transition, which increases d_{CM} by 2 to 4 Å. For $d < 19$ Å, there is a tendency to unbind rapidly. An analysis of the trajectories reveals that for d_{CM} between 17 and 19 Å the hydrogen bond between the backbone CO of Ser96 and the OH of fluorescein breaks first, followed by that between the guanidinium group of Arg196 and the fluorescein carbonyl group. The breaking of the hydrogen bonds at the transition state is in accord with the interaction energy results (Figure 9), which shows a sudden decrease in magnitude for both the van der Waals and electrostatic contribution at $r_{FA} \sim 55$ Å.

What we have described is a ''consensus'' path, which corresponds to the main features of most unbinding trajectories in the simulations with weak bias parameters, although there are some differences among them. This variation is due to the fact, already mentioned, that the details of the unbinding cannot be described in terms of a single coordinate (either r_{FA} or d_{CM}), but that a more complex progress variable, including other coordinates, is required for a complete description. Such complexity exists even in the ''simple'' case of dihedral angle transitions in the alanine dipeptide, where a broad transition region was found when the transition is expressed in terms of the variables ϕ and

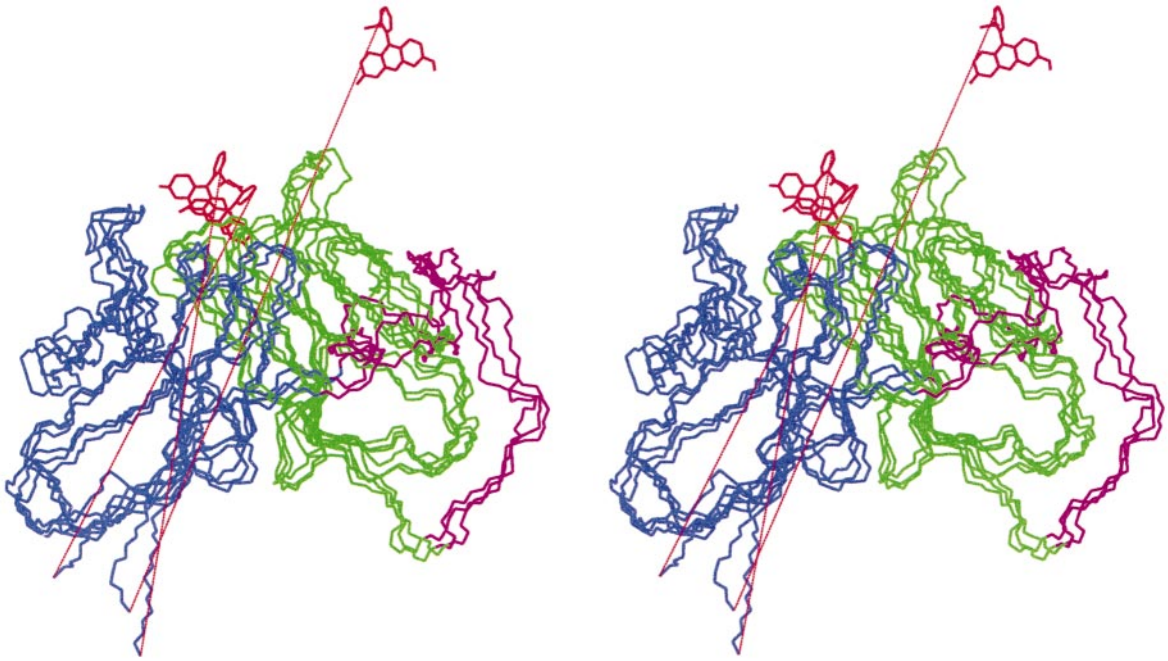

Figure 10. Stereoplot (wall-eye) of the unbinding simulation started at $t = 0.4$ ns with $\alpha = 0.0005$ kcal mol^{-1} Å^{-4}. Snapshots are shown for conformations saved at 1 ps ($r_{FA} = 46.5$ Å), 433 ps (55 Å), and 518 ps (70 Å). The backbone atoms of the scFv fragment are colored in green (V_L), blue (V_H), and black (linker). Fluorescein is shown in red. The r_{FA} distance is shown as a red line. The Figure shows that the most flexible parts of the scFv fragment are the linker and the C-terminal segment (see the text).

ψ.[36] Nevertheless, the important interactions are apparent from the present analysis.

The simulations with $\alpha > 0.0005$ or $F > 300$ pN show a different sequence of events. The value of r_{FA} increases to 55 Å (and that of d_{CM} to about 19 Å) because of the almost complete exit of fluorescein from the binding site without any detachment of the C-terminal strand. In most of these runs, the fluorescein molecule is kept at the entrance of the binding site (i.e. d_{CM} of about 19 Å) by two hydrogen bonds: between its carbonyl oxygen atom and the guanidinium group of Arg196, and between its hydroxyl group and the main-chain carbonyl group of Ser96. During the last steps of unbinding ($r_{FA} > 55$), the seven residues of the C-terminal strand separate from the rest of the scFv either partially (i.e. rupture of only the two hydrogen bonds involving residue 156 on one strand, and the CO of residue 259 and the NH of residue 261 on the other strand) or totally. Hence, the sequence of events under a strong bias is different than at low bias. The rate-limiting step seems to involve, in both cases, the separation of the C-terminal strand from the rest of the scFv. This happens before the exit of fluorescein from the binding site ($d_{CM} < 17$) at low values of the bias parameter, and after fluorescein has undergone a significant displacement from its bound position ($d_{CM} > 19$) at high bias. This partially explains the two different regimes in Figure 5.

Estimation of unbinding rates

A bound state can be idealized as confined by a single energy barrier at a given position along a specific reaction path. Assuming that the kinetics of the forced unbinding can be analyzed in the context of the Kramers theory for activated processes,[32,37] the off-rate for unbinding can be written as:

$$k_{off} = \omega e^{-\beta \Delta E^{\dagger}} \tag{1}$$

where $\beta = 1/k_B T$, k_B is the Boltzmann constant, T is absolute temperature, and ω is a frequency prefactor. The activation energy $\Delta E^{\dagger}$ is equal to $E(r_{FA}^{\dagger}) - E(r_{FA}^{bound})$ and $\dagger$ indicates the transition state.

In the AFM experiments, an external force is applied to a pair of atoms that are pulled apart. Since the timescale of thermal motion and the timescale of the force variation in the AFM experiment are separated by at least eight orders of magnitude, the assumption is usually made that the effect of the external force on the microscopic kinetics is that of a constant force. As first argued by Bell,[38] a constant force $\mathbf{F}$ applied to $\mathbf{r}_{FA}$ modifies the energy (and the free energy) by a factor $\mathbf{F}\mathbf{r}_{FA}$. Thus, if $\mathbf{F}$ and $\mathbf{r}_{FA}$ are parallel:

$$\Delta E^{\dagger}(F) = \Delta E^{\dagger} - F r_{FA}^{\dagger} \tag{2}$$

which gives a rate for unbinding that depends on

the applied force:

$$k_{off}(F) = k_{off} e^{\beta F r^{\dagger}_{FA}} \quad (3)$$

if one assumes (as is commonly done) that k_{off} is given by equation (1) and that the prefactor ω is not affected by the force.

Experimentally, the average unbinding force (the height of the peak) as a function of pulling speed is usually measured. With a series of (reasonable) assumptions it can be shown[32,39] that the most probable unbinding force is given by:

$$F = \frac{1}{\beta r^{\dagger}_{FA}} \ln\left(\frac{\beta \rho r^{\dagger}_{FA}}{k_{off}}\right) \quad (4)$$

where ρ is the loading rate, defined as the time derivative of the force applied to the bond. The loading rate ρ is assumed to be equal to the product of the force constant of the AFM cantilever and its retraction speed. It is equal to the true loading rate ρ only if the ''bond'' is fixed in position. This assumption was used to analyze the experimental AFM results[13] for the same ligand-antibody system as studied here.

In the simulations two different methods (BMD, using various α values and CFMD using various F values) were used to study the unbinding. The results obtained for multiple trajectories are summarized in Table 1. As can be seen, for each α or F set that results in complete unbinding, a range of values for the unbinding times is obtained. One difference between BMD and CFMD is that the unfolding time varies in a narrower range at a given value of α than at a given value of F. This difference between the results of the two methods supports the choice of BMD, which is more complicated to interpret but somewhat more efficient. Approximate average unbinding times and rates (the inverse time) can be computed and extrapolated to $\alpha = 0$ or $F = 0$. These limits correspond to no applied force, when the unbinding rate should approach the off-rate in solution, k_{off}, if the simple model just described is applicable. From the results given in the previous section, including the multidimensional character of the ''true'' reaction coordinate, this is likely not to be true. Nevertheless, it is of interest to determine the dependence of the required unbinding force on the unbinding rate.

Equation (3) can be used to estimate k_{off} and $r^{\dagger}_{FA}$ from the CFMD and BMD simulations. In the latter case, the unbinding force is estimated as the maximum of the bias force $F(r_{FA})$ block averaged over intervals of 1.5 Å in r_{FA}. Rewriting equation (3) as:

$$F = \frac{1}{\beta r^{\dagger}_{FA}} \ln\left(\frac{k_{off}(F)}{k_{off}}\right) \quad (5)$$

the off-rate at zero force can be estimated as the intercept at $F = 0$.

Figure 11 shows the unbinding force as a function of the unbinding rate in the BMD and CFMD simulations. The plots of F as a function of ln(k) show a linear behavior (correlation $r^2 \simeq 0.99$). The results in Table 1 give $k_{off} = 5.8(\pm 0.7) \times 10^{-6}$ ps^{-1} and $r^{\dagger}_{FA} = 0.73(\pm 0.05)$ Å (BMD) and $k_{off} = 2.4(0.4) \times 10^{-5}$ ps^{-1} and $r^{\dagger}_{FA} = 0.58(\pm 0.05)$ Å (CFMD). The two sets of results are the same within the uncertainty of the calculations; a source of systematic error could be that k_{off} is estimated at low forces by disregarding the trajectories that do not lead to unbinding.

The above results for the two parameters of the simple model are very different from the experimental estimates, $k_{off}(0) = 0.1$ s^{-1} and $r^{\dagger}_{FA} = 4$ Å.[13] This could be due to the fact that simulated forced unbinding proceeds through a pathway that is different from that explored in AFM unbinding. More likely, the effective unbinding rates of the forced simulations might be determined by inner barriers (i.e. at lower $r^{\dagger}_{FA}$) that become dominant under a large externally applied force. It is interesting to note that Heymann & Grubmüller[24] found for a similar system that the unbinding length (i.e. the displacement of the two anchoring site at the rupture point) varies between 0 and 8 Å, while the value obtained independently from the fit was 2.2 Å. The present results also indicate that, while the two-state model is appropriate to fit the numerical results (i.e. the dependence of the inverse average unfolding time of the external force is logarithmic), the $r^{\dagger}_{FA}$ parameter appears not to be the true ''position'' of the TS in terms of the r_{FA} coordinate. The same simulations from which we derived such a value for the fitting parameter $r^{\dagger}_{FA}$ showed, in fact, that the fluorescein departure from the binding site occurs in steps over a length of several ångström units. This provides a cautionary note concerning the interpretation of the experimental data; i.e. the r_{FA} parameter obtained from

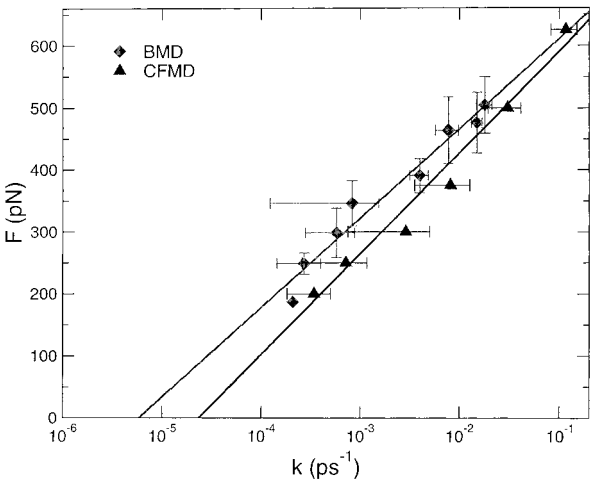

Figure 11. Unbinding force as a function of the unbinding rate in the BMD and CFMD simulations.

the experiment may not have a real physical meaning.

Using the value of $\Delta E^{\dagger} = 95.9$ kcal mol^{-1} reported by Schwesinger *et al.*[13] and the relation:

$$\frac{k_{\mathrm{off}}^{\mathrm{exp}}}{k_{\mathrm{off}}^{\mathrm{sim}}} = \frac{10^{-1}\ \mathrm{s}^{-1}}{10^{7}\ \mathrm{s}^{-1}} = \mathrm{e}^{-\beta(\Delta E_{\mathrm{exp}}^{\dagger} - \Delta E_{\mathrm{sim}}^{\dagger})} \qquad (6)$$

we derive $\Delta E_{\mathrm{sim}}^{\dagger} = \Delta E_{\mathrm{exp}}^{\dagger} + 0.6\ln\ 10^{-8} = 95.9 - 11$ kcal mol^{-1} = 84.9 kcal mol^{-1}; assuming that only two barriers are present, located at $r_{\mathrm{FA}} = 0.7$ and 4 Å, respectively, the crossover between inner and outer barriers happens at a force equal to 11/(4 − 0.7) kcal (Å mol)$^{-1}$ = 230 pN. However, the results reported above suggest that the interpretation of values obtained by assuming that the system is a two-state system following a unique pathway that can be projected onto the variable r_{FA} might be an oversimplification.

Concluding Discussion

The forced unbinding of fluorescein from the 4D5Flu scFv fragment was investigated using two types of external forces and exploring a broad range of timescales by molecular dynamics with an implicit solvent. In the simulations, the force-field was supplemented by either a time-dependent biasing perturbation or a constant pulling force. The former scheme allows one to exploit the spontaneous fluctuations of the system. We find evidence for the importance of comparing different methods for inducing the unbinding on a timescale accessible to the simulation, and for exploring a broad range of timescales.

The present simulation results indicate that the unbinding pathways are close to equilibrium in spite of the very high effective pulling speed. Although the unbinding simulations are performed at a pulling speed at which a linear dependence of the unbinding force on the pulling would be expected due to Stokes' friction,[23] the absence of explicit water is expected to reduce this effect; only the internal friction, which is expected to be small, and the force contribution related to the activated crossing of the free energy barriers to unbinding remain. The derivative of the potential of mean force (i.e. the mean force) computed over the unbinding pathways shows good convergence on a sub-nanosecond timescale. Also, the mean force profile is similar to the externally applied force averaged over the unbinding simulations.

Experimentally, it has been found recently[13] that the unbinding force depends linearly on the logarithm of the loading rate and that by an extrapolation to zero force the off-rate in solution can be determined. This was interpreted as a demonstration of the fact that, in the range of loading rates explored (the loading rate is the product of the force constant of the AFM cantilever and its retraction speed), the unbinding forces of the scFv fragment-fluorescein system are independent of the loading rate and thus determined by the same energy barrier as in solution. In the range of loading rates going from zero to 10^4 pN s^{-1} the same transition state is found for the unbinding. This is similar to what found by Merkel *et al.*[33] for streptavidin unbinding from biotin. However, in that case, at larger loading rates, linear relations with different slopes resulted between force and the logarithm of the loading rate, suggesting that different transition state(s) are explored at larger loading rates.

In the present simulations, number of quantities computed along the unbinding trajectories suggest multiple pathways that depend on the external perturbation and the timescale over which the unbinding is induced. However, of the structural details and the breaking of individual atom-atom interactions during the unbinding show that these pathways follow a well-defined sequence of events. These are shown to involve the whole antibody, not only the binding site. In particular, we observe a distortion of the scFv fragment that is localized mainly at the C-terminal region (residues 255-262) and are partially reversible, while fluorescein is not deformed during unbinding. This ''compliance'' of the antibody, involving a reversible deformation, is often ignored in the analysis of AFM data.

The structural details of the unbinding have revealed that most of the residues involved in the interactions with fluorescein, which are responsible for the peak of the force, are the same in all simulations. The peak of the force precedes the loss of three to five intermolecular hydrogen bonds and the rupture of favorable van der Waals contacts with the aromatic side-chains of the hapten-binding site (Tyr37, Trp177 and Tyr247). It would be of interest to mutate these residues to determine whether any of them modulate the unbinding force.

Materials and Methods

Simulation system

The 4D5Flu scFv fragment consists of the complementary determining regions (CDRs) of the fluorescein-binding antibody 4-4-20 transplanted into the framework of the humanized 4D5Flu antibody.[14] The loop grafting improved the stability and folding yield relative to 4-4-20, an antibody with very high aggregation tendency,[40,41] which yields almost no soluble expression in the periplasm of *Escherichia coli*. The 4D5Flu scFv fragment consists of two variable domains V_L (residues 1-114, or L1-L109 in Kabat[42] numbering) and V_H (residues 145-262, or H1-H113 in Kabat numbering) connected by the flexible linker $(\mathrm{Gly\text{-}Gly\text{-}Gly\text{-}Gly\text{-}Ser})_6$. The scFv 4D5Flu has been characterized by a number of biophysical techniques. Fluorescence quenching data have shown that it has a binding affinity for fluorescein ($K_d = 2.2 \times 10^{-8}$ M) in good agreement with that of scFv 4-4-20 ($K_d = 2.33 \times 10^{-8}$ M);[14] the on and off-rate in solution have been measured.[13] A model of the 4D5Flu scFv-fluorescein complex was constructed by Jung & Plückthun[14] by superposing its V_L and the V_H domains with the crystal structures of the Fv 4D5 (PDB entry 1fvc) and Fab

fragment 4-4-20 (PDB entry 4fab) and checking the CDR-framework contacts. The simulations reported here make use of an improved model (A. Honegger and A. P., unpublished results) based on the 1.85 Å resolution structure of the 4-4-20 Fab bound to fluorescein[29] (PDB entry 1flr). A ribbon representation of the model of the complex is shown in Figure 1.

Force-field and implicit solvent model

Molecular dynamics simulations were performed with a version of the CHARMM program[43] modified to include the biasing perturbation that accelerates the unbinding process (see below). A polar-hydrogen model was used for the protein[18] and a continuum Gaussian model for the solvent.[19] A cut-off on long-range interactions (group-based switch between 7 and 9 Å) consistent with the solvent model parameters was used. The implicit solvent model leads to a simulation time similar to that required for an *in vacuo* calculation, so that enough simulations can be carried out rather easily to obtain statistically meaningful results. Moreover, it provides a potential-of-mean-force description of the solvent[44] that is appropriate in view of the experimental pulling speed, which is many orders of magnitude slower than the solvent relaxation time.

CHARMM[43] atom types were used to describe the fluorescein molecule; they are indicated in Figure 1(c), which also gives the partial charges of the fluorescein atoms. The latter were derived as follows. A formal charge of −1 was assigned to the carboxyl group, which has a pK_a of 2.2; the pH of the experiment is 7.4.[7] A preliminary set of partial atomic charges was determined by a method based on modified partial equalization of orbital electronegativity (MPEOE).[45,46] The MPEOE approach yields partial charges that depend only on the molecular connectivity and not on the conformation. These were modified by using the implementation of MPEOE in the program WITNOTP (A. Widmer, Novartis Pharma, Basel, Switzerland, unpublished), which introduces constraints on the partial charges based on proteinaceous fragments in the MSI CHARMm all-hydrogen force-field (Molecular Simulations Inc., Burlington, MA, USA). The conformation of fluorescein was obtained from the Cambridge Crystal Data Bank[47] and minimized with the MSI CHARMm all-hydrogen force-field and MPEOE partial charges. The heavy-atom root-mean-square deviation (RMSD) between the minimized structure of fluorescein and the conformation in the complex with the antibody (PDB file 1flr) is 0.24 Å. The only significant difference is the orientation of the phenyl ring which is almost perpendicular to the plane of the three-ring system in the minimized conformation (angle of 80°), while in the fluorescein complexed to the antibody the angle is 65°. The minimized conformation was used for a single-point *ab initio* calculation at the HF/31-G* level of theory with Gaussian94,[48] and the final partial charges were obtained from a Mulliken population analysis of the HF/31-G* wave function. In the polar-hydrogen model the aryl hydrogen atoms are not considered explicitly; hence, for each CH group the partial charge of the hydrogen atom was added to that of the bonded carbon atom. The formal charge on the carboxyl group was neutralized by increasing the polarity ($q_C = 1.35$ and $q_{OC} = -0.60$, see Figure 1(c)) to be consistent with the solvation model for the ionic side-chains of the protein.[19] All parameters for bonds, angles and dihedrals of fluorescein are equal to those of the corresponding CHARMM atom types.

Setup and control run

The model structure was first minimized with 1000 steps of the steepest-descent algorithm. The local minimum was very close to the initial structure; the RMSD is 0.31 Å for all heavy atoms and 0.22 Å for C^α atoms. The system was then heated from 0 to 300 K in 15 K steps by scaling up the velocities by a single factor every 4 ps (80 ps in total). The system was then equilibrated for 120 ps by rescaling the velocities when the temperature deviated from 300 K by more than 30 K. During the 200 ps, all the backbone atoms were constrained to remain near to their initial positions by a harmonic force. The simulation was continued in the canonical ensemble using the Nosé-Hoover thermostat[49,50] for 200 ps; all the backbone atoms except those of the linker (residues 115-144) were restrained harmonically. This was followed by a 2.4 ns unconstrained simulation at 300 K (control run) without rescaling of the velocities. Initial configurations for unbinding simulations were selected each 400 ps from the control run to insure statistical independence. A time-step of 2 fs was used and coordinates were saved every 500 steps (1 ps).

Biased molecular dynamics (BMD)

The biasing force used to unbind fluorescein from the antigen-binding site of the 4D5Flu scFv fragment is a time-dependent perturbation proposed by Ballone & Rubini (unpublished) for studying the crystallization kinetics in model systems. A particular implementation of the technique, inspired by an analogy with the AFM experiments, has been used to simulate the stretch-induced unfolding of a number of proteins.[20,21] In the present case, in accord with the AFM experiments by Plückthun and collaborators,[7,13] the reaction coordinate leading from the initial to the final state was chosen to be:

$$\rho(t) = r_{FA}^2 \tag{7}$$

where r_{FA} is the distance between C-2 of fluorescein (F) and the carboxyl carbon atom of the C-terminal residue (SerH113) of the scFv antibody fragment (A); see Figure 1. Use of the latter as one end of the reaction coordinate was dictated by the experimental method,[7,13] in which the scFv protein was immobilized on a gold surface through a spacer and a single free thiol group, which extends the C terminus. This choice may be important, since it influences the strain in the interior of the scFv fragment during unbinding. By contrast, Heymann & Grubmüller,[24] who pulled a hapten out of an antibody by a somewhat different molecular dynamics simulation procedure, used a coordinate corresponding to the distance between the center of mass of the protein and the fluorescein.

To force the system to sample regions of the configuration space that are separated by either thermodynamic or kinetic (on the simulation timescale) barriers, the molecular effective energy function was supplemented by a perturbation of the form:

$$W(\mathbf{r}, t) = \begin{cases} \frac{\alpha}{2}(\rho - \rho_a)^2 & \text{if} \quad \rho(t) < \rho_a \\ 0 & \text{if} \quad \rho(t) \geqslant \rho_a \end{cases} \tag{8}$$

where $\rho_a(t)$ is the maximum value reached by ρ during time 0 to t and α is the force constant for the biasing perturbation. The choice of r_{FA}^2 rather than r_{FA} as the reac-

tion coordinate means that the force is less for small r_{FA}, the region of most interest.

The perturbation is asymmetric in the sense that is quadratic or zero according to the sign of $\rho(t) - \rho_a$ and depends on the time through the reaction coordinate. The simulation is started at $t = 0$ and the value of $\rho_a(0)$ is set equal to $\rho(0)$, the square of the C2/Flu-C/Ser H113 distance ($r_{FA} = 42.3$ Å) of the equilibrated starting configuration. If this distance increases spontaneously in the simulation step from t to $t + \Delta t$, i.e. $\rho(t + \Delta t) > \rho_a(t)$, the external perturbation is zero and has no effect on the dynamics. In such a case, $\rho_a(t)$ is updated and $W(\mathbf{r}, t)$ is modified accordingly; i.e. $\rho_a(t)$ is set equal to $\rho(t + \Delta t)$. If $\rho(t)$ is smaller than ρ_a, the harmonic force (see equation (8)) acts on ρ to prevent the reaction coordinate from decreasing significantly; the value of α determines the magnitude of the backward fluctuation of the reaction coordinate that can occur. We refer to a molecular dynamics simulation with this type of time-dependent perturbation added to the Hamiltonian as biased molecular dynamics (BMD).

In BMD, the macroscopic state of the system is never changed, since the perturbation is numerically zero when it is added to the Hamiltonian of the unperturbed system. Nevertheless, the perturbation affects the evolution of system; i.e. it acts like a Maxwell demon in that it drives the system in a certain direction by ''selecting'' the sign of the spontaneous fluctuations of a given function of the internal coordinates. If the effective energy surface is such that the motion of the reaction coordinate is diffusive in the absence of a barrier, the temperature of the system is not expected to change during the conformational transition. However, if there is a barrier along the reaction path, the effect of the directed motion induced by the perturbation is to transform some of the kinetic energy associated with the reaction coordinate into potential energy. This would lower the temperature in an isolated system. To avoid possible artifacts from temperature variations of this type, the simulations were performed at constant temperature using a Nosé-Hoover thermostat. In studies of protein unfolding,[21] it was determined that no significant local heating is introduced by this method.

In the BMD method, the force on the system varies with time. The external perturbation corresponds to a pair of forces (identical in magnitude and pointing in opposite directions) applied to the C terminus of the scFv fragment and C2 of fluorescein and directed along the vector $\mathbf{r}_{FA}$ joining the A and F atoms; that is:

$$\mathbf{F}_F^W = -\mathbf{F}_A^W = \begin{cases} -2\alpha(\rho - \rho_a)\mathbf{r}_{FA} & \text{if} \quad \rho(t) < \rho_a \\ 0 & \text{if} \quad \rho(t) \geq \rho_a \end{cases} \tag{9}$$

Whenever a spontaneous fluctuation of the system tends to decrease r_{FA}, the external force acts to pull the F and A atoms apart.

The average force for given intervals of the reaction coordinate is larger if the system is crossing barriers and smaller (or equal to zero) if the system is moving freely in the direction of the reaction coordinate. For comparison, a series of simulations were performed with a force of constant magnitude.

In this case, the potential energy term:

$$W = Fr_{FA} \tag{10}$$

was added to the Hamiltonian. This introduces a pair of constant forces acting on the F and A atoms; the forces are identical in magnitude and point in opposite directions along the $\mathbf{r}_{FA}$ vector.

The use of the implicit representation of the solvent[51] made it possible to perform many simulations of 1 ns or longer (50 hours of CPU time on a Pentium II 400 MHz processor are required for a 1 ns run).

Calculation of mean force

To compute the mean (equilibrium) force along the pathway determined by BMD, molecular dynamics simulations with the reaction coordinate constrained to a given value were performed. The mean force along the reaction coordinate r_{FA} can be written[25,52] as:

$$F(\tilde{r}_{FA}) = \left\langle -\frac{\partial V}{\partial r_{FA}} \right\rangle_{\tilde{r}_{FA}} + \frac{2}{\beta \tilde{r}_{FA}} \tag{11}$$

where $\beta = 1/k_B T$. The first term is the average of the derivative of the total effective energy of the system with respect to the reaction coordinate in an ensemble where r_{FA} is constrained to the value $\tilde{r}_{FA}$. The second term is proportional to the apparent (centrifugal) force on r_{FA} at $\tilde{r}_{FA}$ and is negligible in the present case.

Acknowledgments

We thank A. Honegger for the model of the fluorescein-4D5Flu scFv fragment, and S. Jung and F. Schwesinger for useful discussions. E.P. was supported by a Marie Curie fellowship of the European Union and A.C. by the Swiss National Science Foundation (grant 31-53604.98) the NCCR Structural Biology and the Cancer Society of the Canton Zürich. The work in Strasbourg was supported by the CNRS (ESA 7006), by the Ministère de l'Education Nationale, de la Recherche et de la Technologie and by a grant from the Association pour la Recherche contre le Cancer; that done in Harvard was supported, in part, by a grant from NIH.

References

1. Florin, E. L., Moy, V. T. & Gaub, H. E. (1994). Adhesion forces between individual ligand-receptor pairs. *Science,* **264**, 415-417.
2. Moy, V. T., Florin, E. L. & Gaub, H. E. (1994). Intermolecular forces and energies between ligands and receptors. *Science,* **266**, 257-259.
3. Hinterdorfer, P., Baumgartner, W., Gruber, H. J., Schilcher, K. & Schindler, H. (1996). Detection and localization of individual antibody-antigen recognition events by atomic force microscopy. *Proc. Natl Acad. Sci. USA,* **93**, 3477-3481.
4. Dammer, U., Hegner, M., Anselmetti, D., Wagner, P., Dreier, M., Huber, W. & Guntherodt, H. J. (1996). Specific antigen/antibody interactions measured by force microscopy. *Biophys. J.* **70**, 2437-2441.
5. Wong, S. S., Joselevich, E., Woolley, A. T., Cheung, C. L. & Lieber, C. M. (1998). Covalently functionalized nanotubes as nanometre-sized probes in chemistry and biology. *Nature,* **394**, 52-55.
6. Ros, R., Schwesinger, F., Anselmetti, D., Kubon, M., Schäfer, R., Plückthun, A. & Tiefenauer, L. (1998). Antigen binding forces of individually addressed single-chain Fv antibody molecules. *Proc. Natl Acad. Sci. USA,* **95**, 7402-7405.

7. Yip, C. M., Yip, C. C. & Ward, M. D. (1998). Direct force measurements of insulin monomer-monomer interactions. *Biochemistry,* **37**, 5439-49.
8. Willemsen, O. H., Snel, M. M., Cambi, A., Greve, J., DeGrooth, B. G. & Figdor, C. G. (2000). Biomolecular interactions measured by atomic force microscopy. *Biophys. J.* **79**, 3267-3281.
9. Gergely, C., Voegel, J., Schaaf, P., Senger, B., Maaloum, M., Horber, J. K. & Hemmerle, J. (2000). Unbinding process of adsorbed proteins under external stress studied by atomic force microscopy spectroscopy. *Proc. Natl Acad. Sci. USA,* **97**, 10802-10807.
10. Yuan, C., Chen, A., Kolb, P. & Moy, V. T. (2000). Energy landscape of streptavidin-biotin complexes measured by atomic force microscopy. *Biochemistry,* **39**, 10219-10223.
11. Baumgartner, W., Hinterdorfer, P., Ness, W., Raab, A., Vestweber, D., Schindler, H. & Drenckhahn, D. (2000). Cadherin interaction probed by atomic force microscopy. *Proc. Natl Acad. Sci. USA,* **97**, 4005-4010.
12. Grubmüller, H., Heymann, B. & Tavan, P. (1996). Ligand binding: molecular mechanics calculation of the streptavidin-biotin rupture force. *Science,* **271**, 997-999.
13. Schwesinger, F., Ros, R., Strunz, T., Anselmetti, D., Güntherodt, H. J., Honegger, A. *et al*. (2000). Unbinding forces of single antibody-antigen complexes correlate with their rates. *Proc. Natl Acad. Sci. USA,* **97**, 9972-9977.
14. Jung, S. & Plückthun, A. (1997). Improving *in vivo* folding and stability of a single-chain Fv antibody fragment by loop grafting. *Protein Eng.* **10**, 959-966.
15. Chaudhary, V. K., Queen, C., Junghans, R. P., Waldmann, T. A., FitzGerald, D. J. & Pastan, I. (1989). A recombinant immunotoxin consisting of two antibody variable domains fused to Pseudomonas exotoxin. *Nature,* **339**, 394-397.
16. Bird, R. E. & Walker, B. W. (1991). Single chain antibody variable regions. *Trends Biotechnol.* **9**, 132-137.
17. Plückthun, A. (1992). Mono- and bivalent antibody fragments produced in *Escherichia coli*: engineering, folding and antigen binding. *Immunol. Rev.* **130**, 151-188.
18. Neria, E., Fischer, S. & Karplus, M. (1996). Simulation of activation free energies in molecular dynamics system. *J. Chem. Phys.* **105**, 1902-1921.
19. Lazaridis, T. & Karplus, M. (1999). Effective energy function for protein dynamics and thermodynamics. *Proteins: Struct. Funct. Genet.* **35**, 133-152.
20. Paci, E. & Karplus, M. (1999). Forced unfolding of fibronectin type 3 modules: an analysis by biased molecular dynamics simulations. *J. Mol. Biol.* **288**, 441-459.
21. Paci, E. & Karplus, M. (2000). Unfolding proteins by external forces and high temperatures: the importance of topology and energetics. *Proc. Natl Acad. Sci. USA,* **97**, 6521-6526.
22. Isralewitz, B., Izrailev, S. & Schulten, K. (1997). Binding pathway of retinal to bacterio-opsin: a prediction by molecular dynamics simulations. *Biophys. J.* **73**, 2972-2979.
23. Izrailev, S., Stepaniatis, S., Balsera, M., Oono, Y. & Schulten, K. (1997). Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophys. J.* **72**, 1568-1581.
24. Heymann, B. & Grubmüller, H. (1999). AN02/DNP-hapten unbinding forces studied by molecular dynamics atomic force microscopy simulations. *Chem. Phys. Letters,* **303**, 1-9.
25. Carter, E. A., Ciccotti, G., Hynes, J. T. & Kapral, R. (1989). Constrained reaction coordinate dynamics for the simulation of rare events. *Chem. Phys. Letters,* **156**, 472-477.
26. Jarzynski, C. (1997). Equilibrium free-energy differences from nonequilibrium measurements: a master-equation approach. *Phys. Rev. ser. E,* **56**, 5018-5035.
27. Hummer, G. & Szabo, A. (2001). Free energy reconstruction from nonequilibrium single-molecule pulling experiments. *Proc. Natl Acad. Sci. USA,* **98**, 3658-3661.
28. Freund, C., Ross, A., Plückthun, A. & Holak, T. A. (1994). Structural and dynamic properties of the Fv fragment and the single-chain Fv fragment of an antibody in solution investigated by heteronuclear three-dimensional NMR spectroscopy. *Biochemistry,* **33**, 3296-3303.
29. Whitlow, M., Howard, A. J., Wood, J. F., Voss, E. W., Jr & Hardman, K. D. (1995). 1.85 Å structure of anti-fluorescein 4-4-20 Fab. *Protein Eng.* **8**, 749-761.
30. Herron, J. N., He, X. M., Mason, M. L., Voss, E. W., Jr & Edmundson, A. B. (1989). Three-dimensional structure of a fluorescein-Fab complex crystallized in 2-methyl-2,4-pentanediol. *Proteins: Struct. Funct. Genet.* **5**, 271-280.
31. Lu, H. & Schulten, K. (1999). Steered molecular dynamics simulations of force-induced protein domain unfolding. *Proteins: Struct. Funct. Genet.* **35**, 453-463.
32. Evans, E. & Ritchie, K. (1997). Dynamic strength of molecular adhesion bonds. *Biophys. J.* **72**, 1541-1555.
33. Merkel, R., Nassoy, P., Leung, A., Ritchie, K. & Evans, E. (1999). Energy landscapes of receptor-ligand bonds explored with dynamic force spectroscopy. *Nature,* **397**, 50-53.
34. Xie, X. S. & Lu, H. P. (1999). Single-molecule enzymology. *J. Biol. Chem.* **274**, 15967-15970.
35. Slayton, R. M. & Anfinrud, P. A. (1997). Time-resolved mid-infrared spectroscopy: methods and biological applications. *Curr. Opin. Struct. Biol.* **7**, 717-721.
36. Bolhuis, P. G., Dellago, C. & Chandler, D. (2000). Reaction coordinates of biomolecular isomerization. *Proc. Natl Acad. Sci. USA,* **97**, 5877-5882.
37. Kramers, H. A. (1940). Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica (Amsterdam),* **7**, 284.
38. Bell, G. I. (1978). Models for the specific adhesion of cells to cells. *Science,* **200**, 618-627.
39. Evans, E. & Ritchie, K. (1999). Strength of a weak bond connecting flexible polymer chains. *Biophys. J.* **76**, 2439-2447.
40. Mallender, W. D., Carrero, J. & Voss, E. W., Jr (1996). Comparative properties of the single chain antibody and Fv derivatives of mAb 4-4-20. Relationship between interdomain interactions and the high affinity for fluorescein ligand. *J. Biol. Chem.* **271**, 5338-5346.
41. Nieba, L., Honegger, A., Krebber, C. & Plückthun, A. (1997). Disrupting the hydrophobic patches at the antibody variable/constant domain interface: improved in vivo folding and physical characterization of an engineered scFv fragment. *Protein Eng.* **10**, 435-444.
42. Kabat, E. A., Wu, T. T., Reid-Miller, M., Perry, H. M. & Gottesman, K. S. (1987). *Sequences of Proteins of Immunological Interest*, Technical Report US Depart-

ment of Health and Human Services, National Institute of Health, Bethesda, USA, MD.
43. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S. & Karplus, M. (1983). CHARMM: a program for macromolecular energy, minimization and dynamics calculations. *J. Comput. Chem.* **4**, 187-217.
44. Roux, B. & Simonson, T. (1999). Implicit solvent models. *Biophys. Chem.* **78**, 1-20.
45. No, K. T., Grant, J. A. & Scheraga, H. A. (1990). Determination of net atomic charges using a modified partial equalization of orbital electronegativity method. 1. Application to neutral molecules as models for polypeptides. *J. Phys. Chem.* **94**, 4732-4739.
46. No, K. T., Grant, J. A., John, M. S. & Scheraga, H. A. (1990). Determination of net atomic charges using a modified partial equalization of orbital electronegativity method. 2. Application to ionic and aromatic molecules as models for polypeptides. *J. Phys. Chem.* **94**, 4740-4746.
47. Allen, F. H., Bellard, S., Brice, M. D., Cartwright, B. A., Doubleday, A., Higgs, H. *et al*. (1979). The Cambridge Crystallographic Data Centre: computer-based search, retrieval, analysis and display of information. *Acta Crystallog. sect. B,* **35**, 2331-2339.
48. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Gill, P. M. W., Johnson, B. G., Robb, M. A. *et al*. (1995). *Gaussian 94, Rev. B.3*, Gaussian, Inc. Pittsburgh, USA, PA.
49. Nosé, S. (1984). A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **52**, 255-268.
50. Hoover, W. G. (1985). Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev. ser. A,* **31**, 1695-1697.
51. Lazaridis, T. & Karplus, M. (1997). ''New View'' of protein folding reconciled with the old through multiple unfolding simulations. *Science,* **278**, 1928-1931.
52. Paci, E., Ciccotti, G., Ferrario, M. & Kapral, R. (1991). Activation energies by molecular dynamics with constraints. *Chem. Phys. Letters,* **176**, 581-587.
53. Koradi, R., Billeter, M. & Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14**, 51-55.
54. Richards, F. M. (1977). Areas, volumes, packing and protein structure. *Annu. Rev. Biophys. Bioeng.* **6**, 151-176.

Edited by B. Honig

(Received 3 August 2001; received in revised form 21 September 2001; accepted 21 September 2001)