# Balls and Chains—A Mesoscopic Approach to Tethered Protein Domains

Bernhard Windisch,* Dennis Bray,† and Thomas Duke*
*Department of Physics, Cavendish Laboratory, †Physiology, Development and Neuroscience, University of Cambridge, Cambridge, United Kingdom

ABSTRACT Many proteins contain regions of unstructured polypeptide chain that appear to be flexible and to undergo random thermal motion. In some cases the unfolded sequence acts as a flexible tether that restricts the diffusion of a globular protein domain for the purpose of catalysis or self-assembly. In this article, we present a stochastic model for tethered protein domains under various conditions and solve it numerically to deduce the general and dynamic properties of these systems. A critical domain size dependent on the length of the tether is presented, above which a spherical domain tethered to an impenetrable wall by a flexible chain displays a restricted localization between two concentric half-shells. Results suggest that the diffusion of such a spherical domain is effectively reduced in its dimensionality and able to explore the available space with high efficiency. It also becomes clear that the orientation of the ball is not independent of the distance from the tethering point but becomes more constrained as the linking tether is extended. The possible biological significance of these and other results is discussed.

## INTRODUCTION

With the development of high-throughput crystallographic techniques, the structures of more and more proteins have been elucidated (1). However, it should be kept in mind that only the parts of proteins that have a rigid, reproducible structure show up in x-ray pictures of crystallized proteins. Only by comparing the DNA-predicted amino-acid sequence with the crystallographic structure and images produced by other imaging techniques such as NMR is it possible to identify the existence of flexible chains of amino acids in otherwise structured proteins. The number of proteins in which flexible regions have been detected has increased dramatically in recent years (2–5). (A survey of proteins with disordered sequences can be found in (6).) In some cases, it appears that flexible tethers serve to increase the local concentration of catalytic or inhibitory domains (7–10). Elsewhere, flexible loops or tethers seem to provide proteins with a versatile allosteric mechanism for modulating their activity (11) or promoting self-assembly of large structures (12). Another completely different use of unstructured polypeptide chains has been found in polymer brushes as a means to mediate forces on the surface of microtubules (13) and neurofilaments (14). While freely diffusing unstructured polypeptide chains in the cellular medium are well described by the Flory description of random coils, only very few studies have attempted to examine more complicated models of unstructured polypeptide chains (14,15) and in particular the effects of attaching relatively big diffusing molecules to the tethering chains (16).

One example of a well-characterized system employing unstructured sequences of amino acids that display a variety of the aforementioned features is the bacterial chemotaxis pathway in *Escherichia coli*. Proteins with unstructured sequences include the kinase CheA (17,18), the CheZ phosphatase (5), and the Tar and Tsr receptors (19). In the case of Tar and Tsr, the flexible C-terminus binds, with its terminal pentapeptide, to the enzymes CheR and CheB, which modify the methylation state of the receptors during adaptation (see Fig. 1). These flexible linkers are relatively short, comprising not more than 35 residues. At one end, they are usually attached to a relatively large immobile structure, the receptor dimer—which is itself part of a large subcellular complex, while the other end is either freely diffusing or linked to globular protein. They therefore fit the description of a ball-and-chain.

## METHODS

### Balls and chains

Our model consists of a flexible chain composed of 32 subunits equivalent to a flexible polypeptide of 32 amino acids. The chain was attached at one end to an impenetrable wall (tethering wall), which might correspond to a biological membrane or a large protein complex. The other end was attached to a freely diffusing ball, representing a structured globular domain. To examine how properties vary with the ratio of ball size to tether length, the radius of the ball was varied while the tether was maintained at 32 subunits. Fig. 2 shows two screenshots to illustrate the model system.

The mathematical analysis of flexible polymers is already highly developed (20–22). Some of these methods have already been applied successfully to the dynamics of protein folding (23) and to unstructured charged polypeptides under confinement (15). In our case, Brownian dynamics provides the most suitable mathematical description of the system. Here, instead of solving the Newton equations of motion for all particles in the system including the solvent, the interaction between solvent and dissolved particles is subsumed into a random force acting on the dissolved particle (for a more detailed account, see (24)). The ensuing stochastic equation is called a Langevin equation, which we then solve numerically using a slightly modified time-step adjusting Euler integration method.

*Submitted November 30, 2005, and accepted for publication May 16, 2006.*

Address reprint requests to B. Windisch, Tel.: 0-1-223-50-5581; E-mail: bw247@cam.aca.uk.



0006-3495/06/10/2383/10 $2.00

doi: 10.1529/biophysj.105.078543

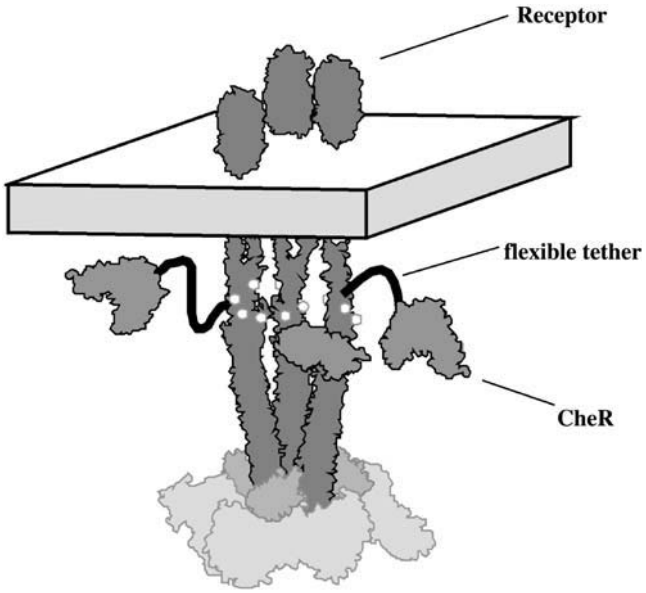


FIGURE 1 Example of a ball-and-chain configuration: CheR is tethered to the C-terminus of a Tar-receptor in the *E. coli* chemotaxis receptor cluster (34).

## Advantages of Brownian dynamics

In general, the Langevin equation has the form

$$\nu \frac{\partial \mathbf{r}}{\partial t} = \mathbf{f}^{\text{deterministic}} + \xi^{\text{random}}, \tag{1}$$

where $\nu$ represents the friction coefficient. The right-hand side is a sum of deterministic forces acting on the simulation units plus the random force through the solvent ($\boldsymbol{\xi}^{\text{random}}$). This force-based approach has a number of advantages for our system. Firstly, every discrete object that is rigidly structured, as, for example, an amino acid or a folded protein domain, can be treated as one simulation unit regardless of its absolute size. This reduces the number of simulation objects significantly compared for example to molecular dynamics. Secondly, additional forces can be included easily, such as a restriction of the angle between successive chain links. Furthermore, using a force-based approach instead of random sampling lets the system exhibit the correct dynamic development over time and allows its diffusion dynamics to be interrogated. Together, these three features make it possible to simulate rather large cellular systems in comparatively complex environments while keeping their dynamic evolution intact. The Langevin equation for the whole system in the proposed model has the form

$$\begin{aligned}\nu_i \frac{\partial \mathbf{r}_i}{\partial t} = {} & \mathbf{f}_{i+1}^{\text{tether}} - \mathbf{f}_i^{\text{tether}} + \mathbf{f}_i^{\text{angle}} \\ & + \mathbf{F}_i^{\text{hardCore}} + \mathbf{F}_i^{\text{confining walls}} \\ & + \xi_i^{\text{random}} \quad i = 1 \ldots n,\end{aligned} \tag{2}$$

and includes four types of forces: the coupling forces between successive links $\mathbf{f}^{\text{tether}}$, an optional force to restrict the angle between successive links $\mathbf{f}^{\text{angle}}$, the hardcore repulsion forces $\mathbf{F}^{\text{hardCore}}$ and $\mathbf{F}^{\text{confining walls}}$, and the random force $\boldsymbol{\xi}^{\text{random}}$. On the left-hand side, $\nu_i$ represents the friction coefficient of the monomer and is calculated as the Stokes friction of a sphere of radius $a$,

$$\nu = 6\pi a \eta, \tag{3}$$

with $\eta$ being the viscosity of the cellular medium chosen to be the same as water. Fig. 3 illustrates the forces within the polypeptide chain.

For the coupling forces between successive links ($\mathbf{f}^{\text{tether}}$) (two of which act on every link in the chain apart from the first and the last one) and the force to restrict the angle between successive links, two types of forces have been tested:

1), a conservative Hookean spring force of the form

$$\mathbf{f}^{\text{tether}} = -k(v - v_0), \tag{4}$$

where $k$ is the Hooke constant, $v$ the actual value (either monomer length or angle), and $v_0$ the equilibrium value; and

2), a finitely elastic nonlinear extensible (FENE)-potential of the form

$$\mathbf{f}^{\text{tether}} = -k\frac{(v - v_0)}{\left(1 - \frac{(\mathrm{v} - \mathrm{v}_0)}{\Delta \mathrm{v}_{\max}}\right)^2}, \tag{5}$$

where $\Delta v_{\max}$ is the maximally allowed deviation from the equilibrium value (again either monomer length or angle). Note that the overall difference between Hooke and FENE potential is that the latter has upper and lower limits to its extension.

The hardcore repulsion forces that prevent the chain from crossing itself or any of the restrictive boundaries are calculated by taking the radial gradient of the following hardcore potential:

$$U^{\text{hardCore}} = \begin{cases} C\dfrac{\mathrm{r}_{\text{int}} - \mathrm{r}}{\mathrm{r}^2 - \mathrm{r}_{\text{hard}}^2} & \text{for } r < r_{\text{int}} \\ 0, & \text{for } r > r_{\text{int}} \end{cases}, \tag{6}$$

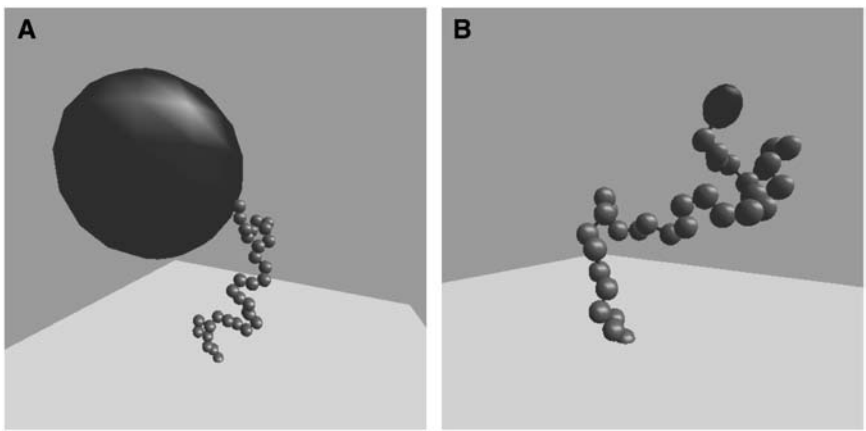


FIGURE 2 Two snapshots of the simulation with a ball of radius 2.7 nm (*left*) and 0.3 nm (*right*).

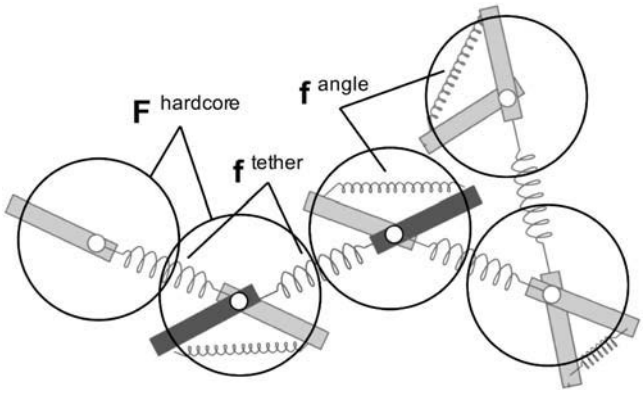


FIGURE 3 Model of the polypeptide chain employed in this study. Shown are the springs linking the monomers in a chain ($\mathbf{f}^{\text{tether}}$), the springs that restrict the angle between successive links ($\mathbf{f}^{\text{angle}}$), and the range of the hardcore-interaction force around each bead ($\mathbf{F}^{\text{hardCore}}$).

$$\mathbf{F}^{\text{hardCore}} = -\nabla \mathrm{U}. \tag{7}$$

Here, $C$ is a constant adjusting the absolute magnitude of the potential, $r$ the distance from the center of the object (in case of the tethering wall from its surface), $r_{\text{int}}$ the range of the potential, and $r_{\text{hard}}$ the actual hardcore radius of the object. The interaction distance between objects $r_{\text{int}}$–$r_{\text{hard}}$ has been chosen to be a quarter of the radius of the object, though maximally 0.1 nm.

The random force $\boldsymbol{\xi}^{\text{random}}$ due to the interaction with the solvent is calculated by drawing random numbers from white noise with a standard deviation of

$$\sigma = \left(\frac{2\nu\, kT}{\Delta t}\right)^{\frac{1}{2}}, \tag{8}$$

where $\nu$ is the friction constant for the monomer, $k$ the Boltzmann constant, $T$ the temperature, and $\Delta t$ the numerical timestep of the simulation.

## Model parameters

The chain in all simulations has 32 amino acids each represented by a sphere with the radius of half the length of the peptide bond, 0.19 nm. The globular domain at the end of the chain is represented by a sphere with a radius varying between 0.3 and 6.0 nm. As can be seen above, the Hookean potential is characterized by two parameters, the equilibrium value and the Hookean spring constant, while the FENE-potential needs an additional parameter—the maximally allowed deviation from the equilibrium value. Of these, the equilibrium value for the monomer-length potential is given by the length of the peptide bond, 0.38 nm, as indicated above.

Apart from the monomer-length restriction, the model has an optional force to limit the angle between two successive monomers in the chain. It seems sensible to include such a restriction into the model since it has been experimentally shown that in random polypeptide chains, the angle between successive links is quite heavily restricted (dependent on the type of amino acid). In a general description of polypeptides, the restriction on the angle between successive monomers is described by the Ramachandran plots. Here, the two degrees of freedom of rotation ($\phi$, $\psi$) at every $\alpha$-carbon in the chain are used to define regions of allowed conformations. Using geometric calculations, it is possible to relate each pair of ($\phi$, $\psi$) to a single angle $\delta$ between successive links loosing information about the actual positions of the side chains in three dimensions. Considering allowed and disallowed regions in the Ramachandran plot, it is then possible to formulate a restriction for $\delta$ as well. The main question to answer is whether the introduction of such a heavily simplified restriction on the angle into the model adds to its explanatory power.

## Testing the model

Atomic force microscopy (AFM) measurements on single molecules (25,26) provide a wealth of experimental data on single polypeptide chains. Generally, a protein is attached to an immobile surface at one end and to the AFM on the other. The latter is then slowly moved away from the surface to exert a pulling force onto the protein. Experiments have been performed on proteins known to have elastic properties such as titin from skeletal muscle (27) or P-selectin/ligand complexes (28) but increasingly also on other proteins like ubiquitin to gain a deeper understanding of protein folding dynamics. In some proteins, the tertiary structure unravels in a series of stages, giving rise to a distinctive sawtooth pattern (29) (see experimental AFM-data on ubiquitin from the Fernandez Laboratory at Columbia University in Fig. 4 (*plus symbols*)).

To our knowledge, no AFM-experiments have been performed on random polypeptides lacking structure, so a direct comparison of simulation results with force-extension data over a full range of forces cannot be made. We can nevertheless compare our model with the AFM data on proteins in the high force regime, where it might be expected that none of the secondary structure remains. This regime is apparent beyond the last peak in the sawtooth pattern in Fig. 4. We find that our model can fit the data with good accuracy, apart from slight deviations at very high forces shortly before the end of the polypeptide chain detaches from the tip of the AFM.

To make this fit, the equilibrium length of the linear spring was set to 0.38 nm, to represent the known length of a peptide bond. The equilibrium angle between successive bonds was set to 109°, consistent with the mean of the probability distribution of $\delta$, calculated from the Ramachandran plots as explained above. We found that the choice of a FENE potential for the spring length improved the quality of the fit (the largest deviation in length was chosen to be 0.3 nm) but that a simple Hookean angular spring was sufficient. The two spring constants were treated as fitting parameters and the

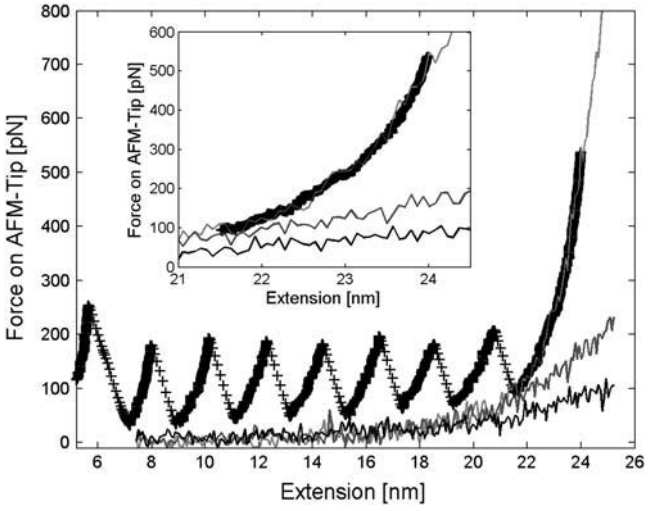


FIGURE 4 MePSim simulation fit for the last sawtooth of a force extension curve for Ubiquitin (with thanks for the data provided by J. Fernandez). The experimental curve (+) shows the characteristic sawtooth pattern, originating from the breaking of bonds in the tertiary structure under increasing tension. None of the MePSim simulations of a random polypeptide coil of the same length show this pattern (since they are random coils). Nevertheless, simulations with strong angular restriction (*light gray line*) fit the last peak in position and gradient, which originates from pulling the completely unfolded polypeptide. This last portion of the curve is shown in more detail in the inset. It can be seen that up to ~21 nm, the weakened parameters (*dark gray line*) are statistically indistinguishable from the strong parameters while keeping the strong linear spring and releasing the angular restriction (*black line*) entirely leads to a difference in force-extension behavior already under 20 nm.

best fit was achieved with a linear spring constant of 35 N/m and an angular spring constant of 45 N/rad.

The angular restriction is essential. As shown in Fig. 4, if the chain is freely jointed, the force rises much less rapidly at large extensions. Indeed, a 76-residue chain (which corresponds to the number of residues in ubiquitin) with a peptide bond length of 0.38 nm could be extended to a full contour length of 28.9 nm, if no angular restriction were imposed. The force would rise significantly only as this value of the extension is approached. We consider that a chain of links with angular restrictions is a more appropriate statistical model of an unstructured polypeptide than the wormlike chain, which is often used to fit elasticity data. The wormlike chain models the polymer as a continuously bending elastic rod. While it provides a good fit to the data using the contour length and a bending modulus as fitting parameters, the value of the contour length obtained is artificial; it does not correspond to the total length of all the peptide bonds along the chain. Recent analytical results also suggest that discrete-chain models such as the freely-rotating chain and the elastically jointed chain can be used to better explain high-force stretching responses in AFM experiments (30).

The comparison with the high-force AFM data establishes that the discrete-link model can describe the elasticity of an unstructured polypeptide chain. In this article, we are interested in the thermal equilibrium regime, where the chain is much less extended. Simulations of a chain of 76 monomers tethered to a ball on one end and to a wall on the other show that extensions $>15$ nm are extremely infrequent. For the range of extensions that typically occur at equilibrium, variation of the values of the two spring constants over an order of magnitude does not significantly alter the quality of the fit. Weaker springs permit greater fluctuation at the scale of individual monomers, but do not change the statistical properties of the molecule as a whole. For simulation purposes, the increased fluctuations are helpful because they greatly reduce the amount of CPU time required to simulate a given interval of real time. We therefore found it advantageous to use Hooke constants of 3 N/m and 12 N/rad, respectively, for the linear and angular springs. As shown in Fig. 4, there is no difference in the force-extension curve using these weaker values of the spring constants for extensions $<21$ nm, so this modified model is appropriate for the equilibrium regime. We emphasize, however, that the angular restriction remains important in the equilibrium regime; without it, the force-extension relation deviates from the best-fit curve when the extension exceeds 18 nm.

## Definition of variables

In this article, emphasis has been laid on the system properties at equilibrium. These comprise the end-to-end chain length of the tether and its radius of gyration, together with the position of the tethered ball and its orientation (Fig. 5). Note that the end-to-end chain length and radius of gyration refer to the tether alone, not the ball attached to its end. The orientation of the ball has been defined generally in relation to the end-to-end length vector of the tethering chain and is given as the angle between this vector and the vector through attachment point and center of the ball (*orientational angle* $\chi$ in Fig. 5). Accordingly, a low $\chi$-value is equivalent to the ball facing away from the tethering point while a high value means that the ball faces toward the tethering point. The orientation is biologically relevant since for a chemical reaction to take place it is not enough for the two reacting partners to come close to each other but also that their active sites come to face each other in close proximity and with the correct orientation. A different definition for the orientation is used to characterize a ball touching the wall. The angle $\phi$ between the orientation vector $d_{\text{orient}}$ and the normal vector characterizing the wall $d_{\text{wall}}$ defines the areas on the surface of the ball that touch the wall most often.

Throughout the article, the subset of conformations where the surface of the ball is closer to the wall than 0.1 nm is referred to as ''touching conformations.'' In addition to the static properties, the diffusion behavior of the ball has been analyzed. In general, diffusion rates have been measured as average square displacement over time. All simulations have been run representing 10 $\mu$s of real-time.

## RESULTS

The properties of freely diffusing chains predicted by our model, including the distributions for end-to-end chain length, radius of gyration, and diffusion behavior conform to the classic predictions of the Flory analysis of homopolymer chains. However, the tethered system exhibits a number of changes (Fig. 6). Tethering the chain to an infinite wall breaks the symmetry of the spatial distribution of chain monomers around the tethering point and exerts an entropic pressure toward more outstretched conformations. The end-to-end chain length distribution is therefore shifted toward longer lengths while the radius of gyration is also slightly shifted toward higher values or more outstretched conformations. These effects are much more pronounced when the ball is present. Because of its size, the ball prohibits certain conformations of the tethering chain, mainly those where the terminal (tethering) monomer is buried within the contours of the current conformation. In this case, there is, in general, insufficient space to accommodate the ball. The effect of eliminating all these chain conformations which are, on average, characterized by a shorter end-to-end chain length is to displace the distribution toward more outstretched conformations. Moreover, due to its interaction with the wall, the ball exerts an additional force on the chain, stretching it away from the wall toward more extended conformations. This can also be seen from the distribution of the radius of gyration, which is also shifted toward higher values.

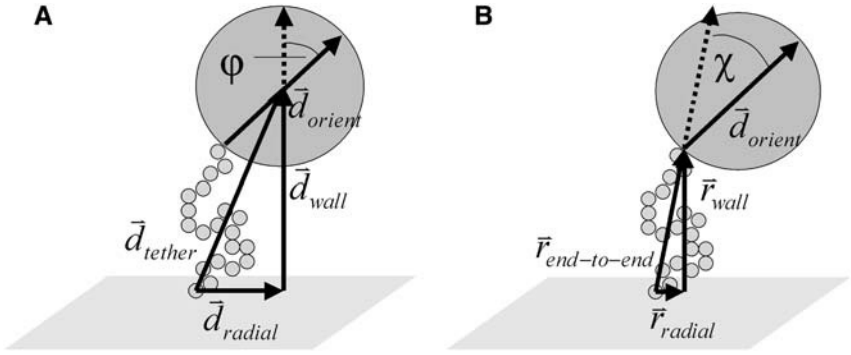


FIGURE 5 Variables used in the analysis, characterizing the properties of the tether alone (identifier *r*) and taking into account the domain attached to the end of the tether (identifier *d*). The orientational angle $\chi$ between the $\vec{r}_{\text{end-to-end}}$ and $\vec{d}_{\text{orient}}$ is large when the ball faces toward the tethering point and small when it faces away from it. This convention is always used when talking about orientation. The orientational angle $\phi$ between $\vec{d}_{\text{orient}}$ and $\vec{d}_{\text{wall}}$ is important when analyzing the orientation of the ball in relation to the wall.

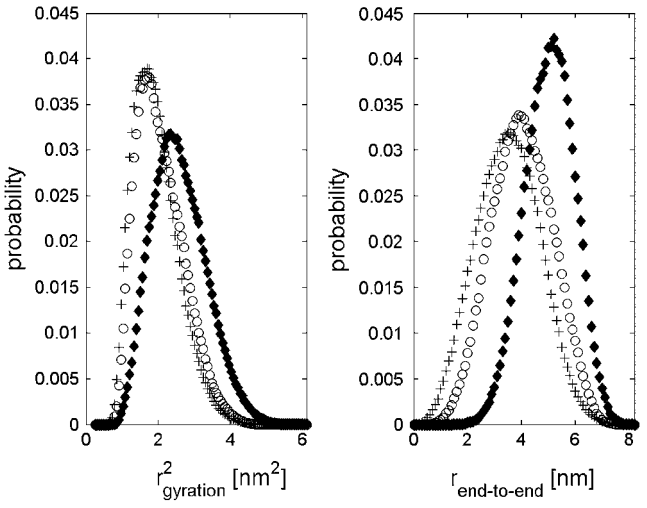


FIGURE 6 Distributions of radius of gyration (*left*) and end-to-end chain length (*right*). Results for a freely diffusing chain of 32 residues (+), for the same chain tethered to a wall (○) and with a ball of $R_{ball} = 2.7$ nm attached to its end (♦). It can be seen that attachment to the wall leads to slightly longer end-to-end chain lengths and more outstretched conformations while the attachment of the ball considerably increases both end-to-end chain length and radius of gyration.

These changes are expected to be dependent on the actual size of the ball attached to the freely diffusing end of the tethered chain. However, Fig. 7 shows that the end-to-end length remains constant, within statistical fluctuations, once the ball radius exceeds 1.2 nm. This represents a critical ball size above which the ball is excluded from the region occupied by the chain. Comparison with Fig. 6 indicates that the critical ball size is similar to the radius of gyration of the free polypeptide chain. When the radius of the ball exceeds the critical size, the average separation of the chain ends does not vary, but the average direction in which the chain is stretched changes as the radius increases. As shown in Fig. 7, the mean normal component $r_{wall}$ of the end-to-end vector increases with the ball radius, while the radial component $r_{radial}$ decreases. The wall exerts an entropic force on the ball, pushing it away; consequently the chain gets increasingly oriented normal to the wall as the ball gets bigger. The existence of this entropic force has recently also been established by analytical methods (16).

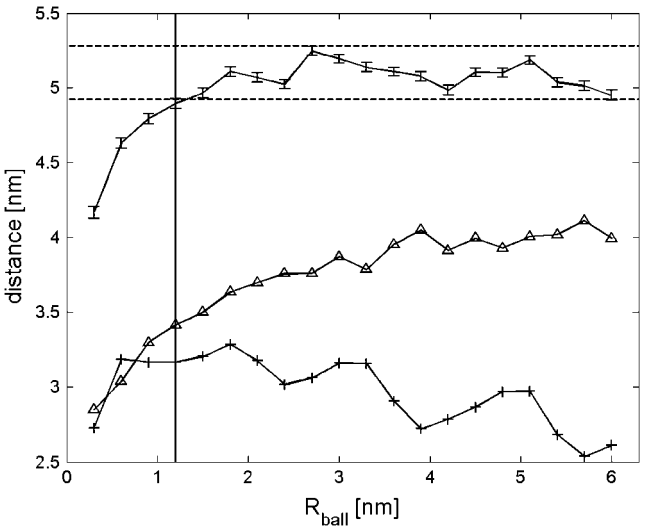


FIGURE 7 Influence of ball size on the end-to-end chain length of the tether. Distribution averages with error are shown for $r_{end\text{-}to\text{-}end}$ for a chain of 32 residues with balls of different radii attached to one end (I). The area confined by the dashed lines indicates where the end-to-end chain length becomes independent of the ball size. This area is entered once the ball radius has reached 1.2 nm, as indicated by the solid line. Also shown are distribution averages for $r_{wall}$ (+) and $r_{radial}$ (△). The development of $r_{wall}$ and $r_{radial}$ demonstrates that the chain becomes successively stretched further away from the wall.

Because a ball that exceeds the critical size is excluded from the region of space occupied by both the chain and the wall, it tends itself to occupy a localized region of space. Fig. 8 displays the probability distribution of the center of a ball of radius 2.7 nm, in cylindrical coordinates. Evidently, the ball

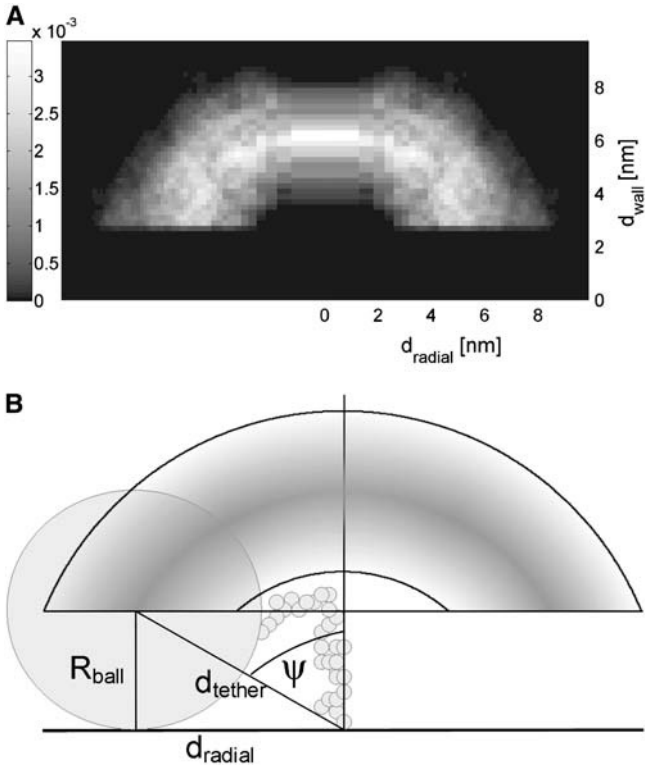


FIGURE 8 Spatial density distribution for a tethered ball. Simulation results and schematic sketch of the ball diffusion around the tethering point in cylindrical coordinates $\pi dr dz$. On the *x*-axis, the distance from the center axis is $d_{radial} = \sqrt{x^2+y^2}$ and on the *y* axis, the height above the wall is $d_{wall} = z$. The step size in $d_{radial}$ is chosen so as to keep the volume between successive values of $d_{radial}$ constant to avoid geometrical distortions in the depicted densities. In the schematic sketch, one possible ball-and-chain conformation has been depicted. Two effects can be seen: Firstly, the wall prevents the center of the ball from coming closer to the wall than its radius, in this case 2.7 nm. Secondly, most of the time the ball is localized between two half-spheres around the tethering point and its center hardly ever comes closer than ~3.7 nm to the tethering point.

diffuses between two spherical half-shells centered on the tethering point. The center of the ball is located with highest probability at a distance of 6.3 nm from the tethering point and >97% of all positions lie between the two half-shells with radii 4.3 nm and 8.3 nm. The permitted locations are bounded by a plane due to the interaction of the ball with the wall, as illustrated by the schematic diagram in Fig. 8. It is clear that the high degree of localization of the ball effectively reduces the dimensionality of its diffusion, and that this could serve a useful purpose in the case where the ball is a functional protein domain. Note that in general, the thickness of the shell depends on the length of the polypeptide chain, while the mean distance of the ball from the tethering point depends on both the chain length and the ball size, as will be described later. Only balls that are larger than the critical size are effectively confined in a two-dimensional shell. Smaller balls can be enveloped by the chain, and so are not excluded from the region immediately adjacent to the tethering point.

In many biological situations, tethered protein domains perform a function either directly on the surface to which they are linked, or by moving a substrate close to the surface. It is therefore important to know where and in which orientation (refer to Fig. 5) a tethered ball makes contact with the wall. Fig. 9 compares the distribution of positions (in the *x-y* plane) of the ball in 1), all conformations; and 2), the subset of touching conformations. It is clear that the tethered ball makes contact with the adjacent wall on a ring, defined by the intersection of the two half-spheres with the plane of exclusion due to the wall. The radius of the ring increases with ball radius (Fig. 9).

Having established in which positions the ball makes contact with the wall, Fig. 10 correlates the orientational angle $\chi$ with $d_{\text{tether}}$ for a ball of radius 2.7 nm. The upper plot takes into account all conformations while the lower plot displays the distribution only for touching conformations. The distribution for all conformations shows a clear negative correlation between distance and orientation—the more stretched the chain, the more aligned are the tethering and orientation vectors. It also shows that for the ball center to come closer to the tethering point than ~5.5 nm, the chain has to wrap around the ball (at least partly) and alignment between tethering and orientation vectors becomes virtually impossible. It is interesting to note that the touching conformations are characterized in general by shorter distances (lower $d_{\text{tether}}$) and sharper orientational angles (higher $\chi$). Excluded by both the ball and the wall, the chain has to partly envelop the ball, especially when the ball makes contact with the wall close to the tethering point. These conformations are characterized by a shorter $r_{\text{end-to-end}}$ and therefore $d_{\text{tether}}$ and at the same time by sharper orientation angles (see the *schematic sketch* in Fig. 8 for an illustration of a touching conformation).

How can the influence of the ball size on orientation and the distance between ball center and tethering point be quantified? And how does the ball size affect the touching conformations and their different properties? Looking at the

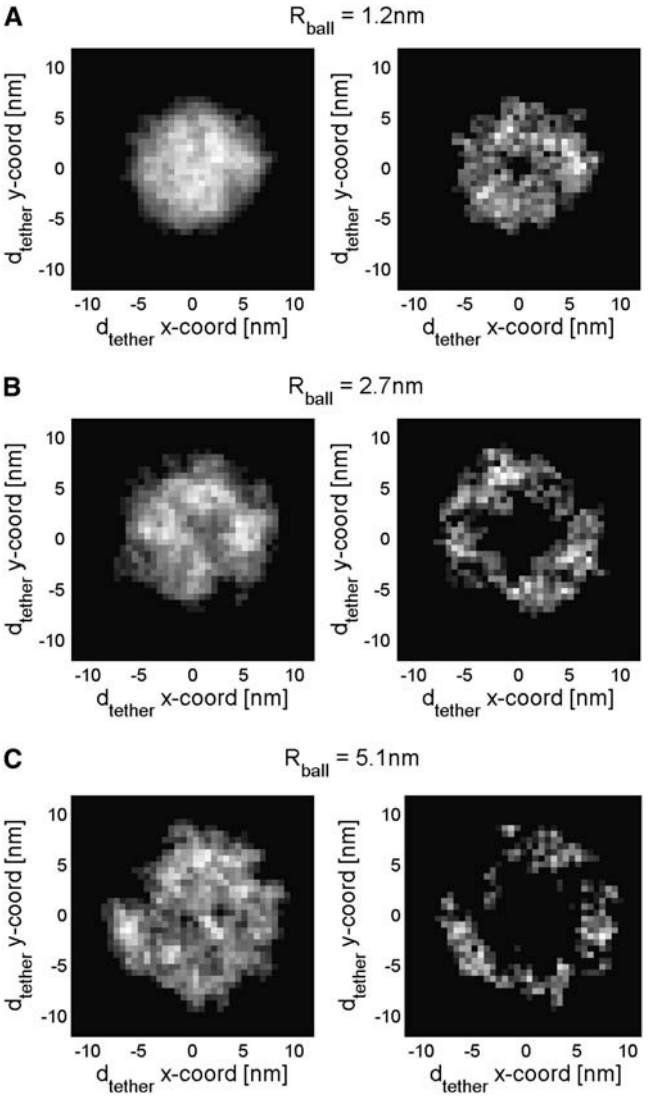


FIGURE 9 Distribution of ball positions over *x-y* plane for three different ball sizes. The position of the ball center over the *x-y* plane is shown for all conformations on the left and for touching conformations on the right. It becomes clear that the points of contact between ball and wall are distributed on a ring around the tethering point. The average radius of this ring increases with ball radius.

schematic sketch in Fig. 8, the ball is excluded by the chain from a hemispherical region surrounding the tethering point of radius $r$. We might expect $r$ to be similar in magnitude to the average end-to-end separation of the tethering chain in absence of the wall. Then the average tethering distance $d_{\text{tether}}$ is given by

$$d_{\text{tether}} = R_{\text{ball}} + r. \tag{9}$$

The average radius of the ring around the tethering point on which the ball touches the wall (see Fig. 9) can then be calculated using simple geometric principles

$$d_{\text{radial}} = (2R_{\text{ball}}r + r^2)^{1/2}. \tag{10}$$

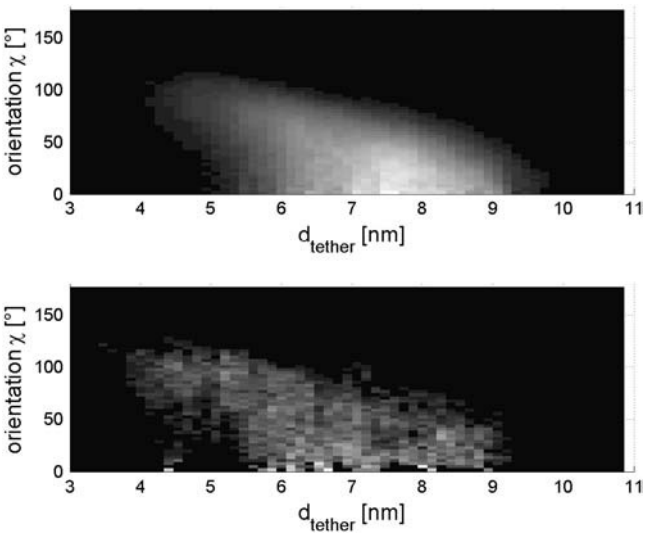


FIGURE 10 Ball orientation $\chi$ versus distance between ball center and tethering point. Above, the dependence is shown for all conformations, below for touching conformations only. It shows clearly that, in general, the ball orientation becomes more aligned with the tethering vector the further away the ball diffuses from the tethering point. Sharp angles, on the other hand, appear only for short distances and are more likely in touching conformations.

For balls whose radius greatly exceeds the radius of gyration of the polypeptide chain, the following scaling laws hold:

$$\langle d_{\text{tether}} \rangle \sim R_{\text{ball}}, \tag{11}$$

$$\langle d_{\text{radial}} \rangle \sim (R_{\text{ball}} r)^{1/2}. \tag{12}$$

Similar scaling laws have been derived using different methods (16). That the radial distance increases much more slowly than the average distance of the ball from the tethering point can be understood by considering the angle $\psi$ in the schematic sketch in Fig. 8. This angle becomes more and more restricted as the size of the ball increases, limiting the radial deviation in position. This observation also explains why the chain becomes more oriented in the normal direction with increasing ball size, as remarked in Fig. 7.

The dependence of $d_{\text{tether}}$ and $d_{\text{radial}}$ on the ball size, as measured from the simulations, is shown in Fig. 11. Using Eqs. 9 and 10 to fit the simulation data with a single parameter, $r$, reveals that although $d_{\text{tether}}$ can be fitted well with $r = 4.2$ nm (which corresponds well with the mean end-to-end separation of the chain in Fig. 6), the same value predicts too-high average radii of the ring of touching conformations. Two effects contribute to this shift toward smaller radii. First, it is clear from Fig. 8 that sampling the conformation space along the wall results in a probability distribution that overemphasizes short distances. Secondly, as can be seen in the lower panel from Fig. 11, the touching conformations are characterized by a lower degree of alignment of the ball and the chain. These two effects combine to give a shorter effective length parameter $r = 3.6$ nm with which the data can be fitted reasonably well. Note that the

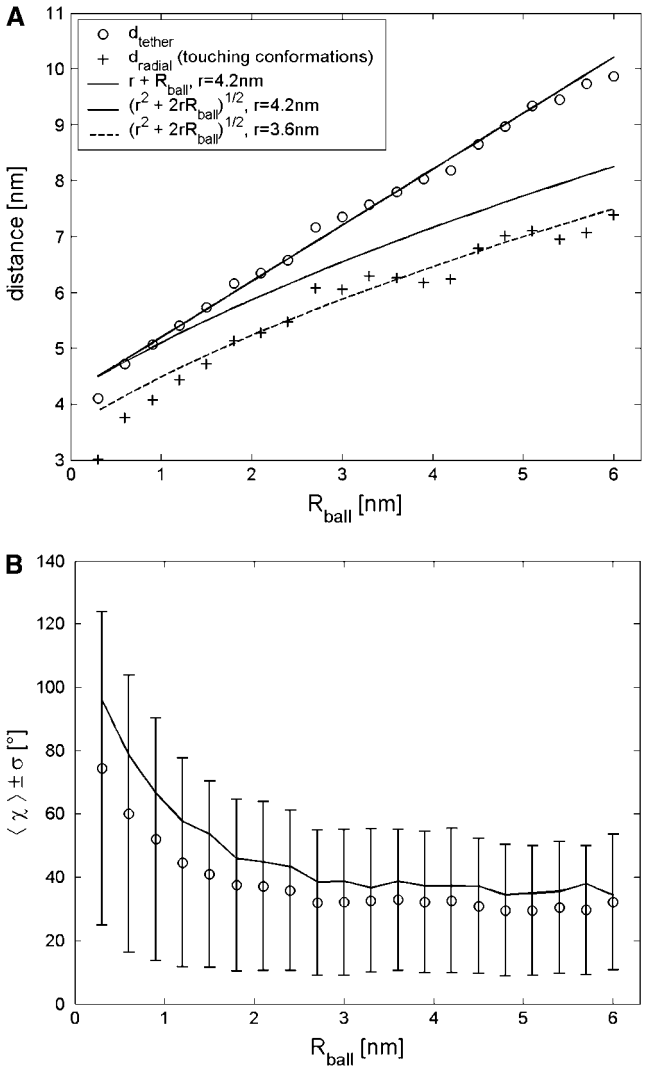


FIGURE 11 Characteristic values for diffusion volume and ball orientation over ball radius. Above, the average radius of the half-shell $d_{\text{tether}}$ and the average radius of the ring of touching conformations $d_{\text{radial}}$. Below, the average orientation angle ($\chi$), defined as in Fig. 5 and its $1\sigma$ environment. It becomes clear that $d_{\text{tether}}$ can be well fitted by $R_{\text{ball}} + r$ with $r = 4.2$ nm. The same parametric value, though, does not fit $d_{\text{radial}}$ very well. Only when taking into account the shorter distances and sharper orientation angles for the touching conformations justifying a lower value for $r = 3.6$ nm, the model produces a good fit.

quality of the fit deteriorates for balls of radius below the critical size, as expected since the ball is not then excluded from the region occupied by the chain.

Considering the biological context in which an active site on a tethered globular domain must bind to a cognate site on the wall, it is clear that the orientation of the ball relative to the wall is important. Fig. 12 shows the distribution of the orientation angle $\phi$ for touching conformations for ball radii between 0.3 and 6 nm. As the size of the ball increases, the angle at which it comes into contact with the wall becomes more narrowly defined. This enhances the probability that certain points on the surface of the ball contact the wall. The

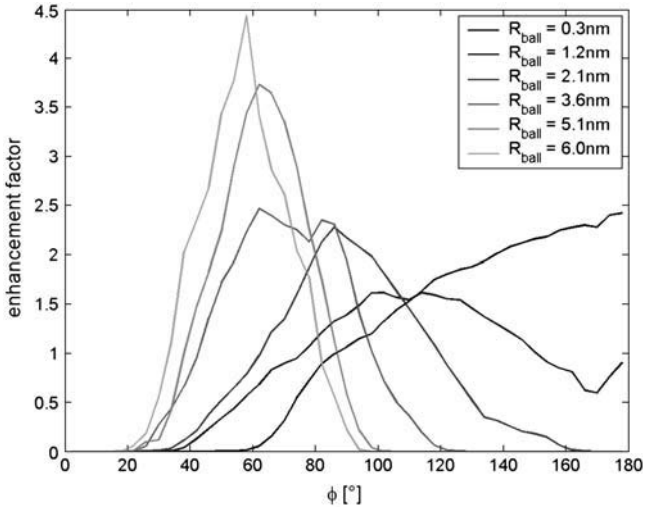


FIGURE 12 $\Phi$-distributions for touching conformations and different ball radii between 0.3 nm (*black line*) and 6 nm (*lightest gray line*). On the *y*-axis is plotted an enhancement factor calculated by dividing the probability of finding a tethered ball of given size with a certain orientation by the probability of finding a freely diffusing ball with the same orientation. It becomes clear that balls of different sizes attached to the same chain touch the wall with different areas in relation to their attachment point to the chain. Moreover, the orientation becomes increasingly narrowly defined the bigger the radius of the ball.

enhancement factor plotted in Fig. 12 is the probability that a tethered ball touches the wall with a certain orientation, compared to the probability that a freely diffusing ball touches the wall with the same orientation. Clearly, the tether would enable an active site that is positioned in the right location on the ball to interact more frequently with its cognate site on the wall. The optimal location varies with the size of the ball: for balls smaller than the critical radius, it is directly opposite the tethering point; for large balls, it is quite close to the tethering point.

Overall, we note that if the size of the ball exceeds a certain critical value in the order of magnitude of the radius of gyration of the tether, the characteristic static properties of the system remain qualitatively unchanged: A diffusion space that is enclosed between two half-shells around the tethering point, an anticorrelation between the orientation $\chi$ and the $d_{\text{tether}}$, and the subset of touching conformations characterized by shorter lengths and sharper orientation angles $\chi$.

Returning to the biological context, a globular domain must often repeatedly and rapidly catalyze a reaction on a surface. The system should therefore not only ensure a high local concentration but also allow sufficiently fast diffusion. In general, using a Stokes friction coefficient for the ball, the diffusion velocity is clearly dependent on the ball size with bigger balls diffusing much more slowly. Analyzing the diffusion behavior of the tethered systems in comparison to the freely diffusing ones shows two main differences.

Fig. 13 compares a freely diffusing ball of radius 2.7 nm with a tethered ball and shows that the tethering affects the diffusion rate of the ball only very slightly at short times (visible from the *right shift* in the lines in the double-log plot). At long times, the mean-square displacement of the tethered ball saturates at a maximum value dependent on the length of the tethering chain. The time at which the ball reaches this plateau, which may be interpreted as the time it takes to explore the available volume indicated in Fig. 8, is dependent on its diffusion rate and therefore its size. The ball of radius 2.7 nm reaches the plateau on a timescale of order 0.5 $\mu$s. Other balls above critical size reach their respective plateau between 0.2 $\mu$s (radius 1.2 nm) and 1.5 $\mu$s (radius 6 nm).

## BIOLOGICAL IMPLICATIONS

Considering the very simple nature of our model, what implications does it have for naturally occurring proteins? As an illustration, we considered the methylation system of CheR tethered to the C-terminus of a Tar receptor within the chemotactic receptor cluster of *E. coli*. X-ray crystallography shows that the receptor has a high degree of $\alpha$-helicity and that its cytoplasmic domain is predominantly coiled-coil (31). However, a region of 32 residues between the coiled-coil structure and the C-terminus is believed to be flexible, since it does not show in the crystallographic structure and is highly variable in sequence between similar receptors and organisms. CheR binds to the C-terminal pentapeptide which has been shown to be essential for methylation. The methylation sites, which are the substrate of CheR, are located on the coiled-coil body of the receptor at various distances from the tethering point. CheR itself is a globular protein, which to a first approximation can be modeled as a ball of 2.7 nm

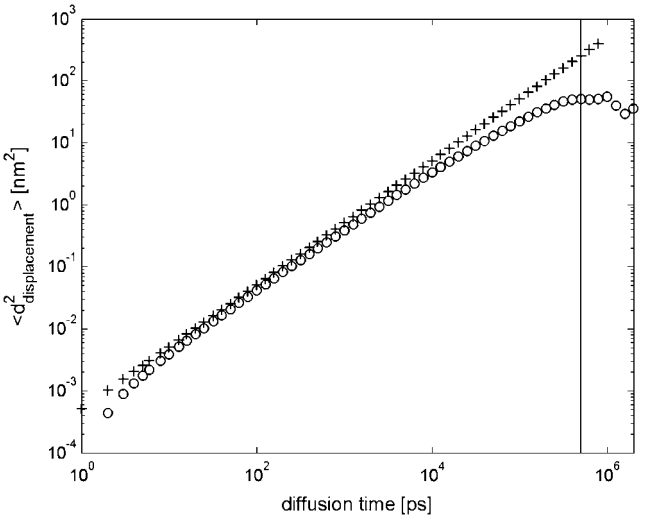


FIGURE 13 Comparative diffusion behavior in terms of mean-square displacement over time. The diffusion of a freely diffusing ball of radius 2.7 nm (+) is compared to the diffusion of the same ball tethered to the wall by a chain of 32 residues (○). Evidently, the tethering impedes the diffusion of the ball on very short timescales and limits it at a maximal value. Above the line at $t = 5 \times 10^5$ ps (0.5 $\mu$s), the diffusion of the tethered ball has reached its maximal level of displacement.

radius. These parameters (ball of radius 2.7 nm on a chain of 32 monomers) indeed guided our choice of model system. The molecular configuration when CheR binds to substrate therefore resembles one of the touching conformations of the ball and chain as discussed above.

Other proteins in the chemotactic pathway of *E. coli* also possess flexible tethers of similar lengths. The methylation system of the Tsr-receptor which also binds CheR has ~28 amino acids. The C-terminus of the phosphatase CheZ has ~32 residues in its flexible region and tethers CheY, which in first approximation can be modeled as a ball of 1.5 nm radius. Note that all of the above proteins are actually dimeric, so that each carries two tethered balls, a further complication. Other examples in prokaryotic and eukaryotic systems suggest that chains of ~25–50 residues and balls varying between 1.5 and 3.5 nm in radius are widespread.

Returning to the CheR-model system, we have shown that in comparison to a freely diffusing protein domain, the tethered domain stays in the region where it is needed, close to the substrate (Fig. 8). Since the distribution given in the figure is stationary over time, it can be translated into a local concentration (probability of finding the tethered ball in a given volume divided by the volume). Calculating the effective local concentration in this way gives concentration values of the ball in touching conformations (1.8 nm $< d_{\text{radial}} <$ 8.5 nm) in the order of 0.1–5.0 molar. This may be compared to a concentration of ~0.2 $\mu$M estimated from the number of CheR molecules per cell, which is much too low to guarantee working methylation in the freely diffusing regime (32). Thus, provided that the binding site of CheR on the C-terminus of the receptor is within reach of the methylation sites, tethering provides an extremely effective way of enhancing the rate at which it encounters its substrate.

The question remains therefore to compare the distances between beginning of the flexible part of the C-terminus and the four different methylation sites on the Tar-receptor with the reach of CheR tethered on the C-terminus.

Fig. 14 shows the methylation sites in relation to the distance where the tethered ball touches the wall. This figure aggregates the data from Fig. 8 in a different way and therefore shows at what distance our model predicts the tethered CheR is most likely to be found. It can be seen that all four methylation sites are clearly in reach of CheR tethered on the C-terminus. It also shows in comparison with experimentally measured probabilities of methylation on the four sites (33) that our model overemphasizes the probability of methylating the sites close to the tethering point. Experimental data suggests that, on the contrary, the sites furthest from the tethering point (3 and 2) are methylated more strongly than sites 4 and 1.

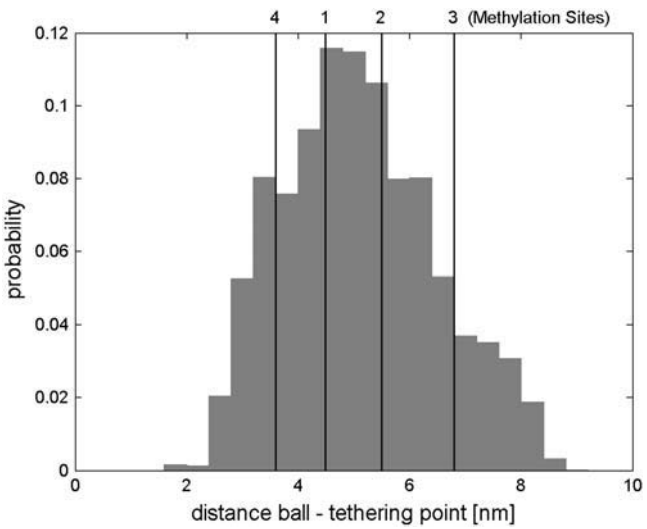


FIGURE 14 Probability distribution of touches on the wall versus the distance from the tethering point. Since the system of the ball tethered to the wall is rotationally symmetric, the probability has been averaged and normalized over the planar angle. Shown as lines are the positions of the methylation sites on the receptor body. It can be seen that all sites are in easy reach of the tethered ball. This supports the notion that tethering CheR on the receptor C-terminus makes it much more readily available to catalyze methylation on the receptor body.

## CONCLUSION

In this article, we have presented a simple model to assess the general properties of globular protein domains (balls) on polypeptide tethers (chains). Although simple in concept, our model serves to highlight several features likely to be important in a biological context. Most markedly, we found that the tethered ball diffuses in a well-defined restricted region, or shell, around its tethering point. The orientation of the ball becomes increasingly constrained the further the ball is from its tethering point. Furthermore, we demonstrated that using this simple model can explain certain properties of the system of CheR tethered on the C-terminus of the Tar-receptor in the chemoreceptor-cluster of *E. coli* qualitatively, if not quantitatively.

The general analysis presented in this article indicates the range of different effects that can occur in proteins containing tethered domains. The exact ways in which these effects are employed are expected to be highly dependent on the system in question. Therefore, the next step should be a closer examination of specific biological systems along the lines presented in this study. The CheR-methylation system should continue to prove an excellent study object because of the wealth of data from ongoing experiments on how the length and sequence of the tether affects the functioning of the methylation system. Further work along these lines should make an important contribution to our understanding of this and other similar systems employing tethered protein domains.

We are most grateful to J. Brujic and J. Fernandez for providing us with experimental AFM data on Ubiquitin. We also thank Mathew Levin and Ian Graham for advice.

This work was financially supported by Trinity College, Cambridge, and National Institute of General Medical Sciences (grant No. G64713).

## REFERENCES


1. Kundrot, C. D. 2004. Which strategy for a protein crystallization project? *Cell. Mol. Life Sci.* 61:525–536.
2. Wootton, J. C., and M. H. Drummond. 1989. The Q-linker: a class of interdomain sequences found in bacterial multidomain regulatory proteins. *Protein Eng.* 2:535–543.
3. Hoshi, T., W. N. Zagotta, and R. W. Aldrich. 1990. Biophysical and molecular mechanisms of *Shaker* potassium channel inactivation. *Science.* 250:533–538.
4. Wu, J., J. Li, G. Li, D. G. Long, and R. M. Weis. 1996. The receptor binding site for the methyltransferase of bacterial chemotaxis is distinct from the sites of methylation. *Biochemistry.* 35:4984–4993.
5. Zhao, R., E. J. Collins, R. B. Bourret, and R. E. Silversmith. 2002. Structure and catalytic mechanism of the *E-coli* chemotaxis phosphatase CheZ. *Nat. Struct. Biol.* 9:570–575.
6. Dunker, A. K., C. J. Brown, and Z. Obradovic. 2002. Identification and functions of usefully disordered proteins. *Adv. Protein Chem.* 62:25–49.
7. Le Moual, H., T. Quang, and D. E. Koshland, Jr. 1997. Methylation of the *Escherichia coli* chemotaxis receptors:intra- and interdimer mechanisms. *Biochemistry.* 36:13441–13448.
8. Li, J., G. Li, and R. M. Weis. 1997. The serine chemoreceptor from *Escherichia coli* is methylated through an inter-dimer process. *Biochemistry.* 36:11851–11857.
9. Perham, R. N. 2000. Swinging arms and swinging domains in multifunctional enzymes: catalytic machines for multistep reactions. *Annu. Rev. Biochem.* 69:961–1004.
10. Timpe, L. C., and L. Peller. 1995. A random flight chain model for the tether of the *Shaker* $K^+$ channel inactivation domain. *Biophys. J.* 69: 2415–2418.
11. Jordan, A., and P. Reichard. 1998. Ribonucleotide reductases. *Annu. Rev. Biochem.* 67:71–98.
12. Namba, K. 2001. Roles of partly unfolded conformations in macromolecules self-assembly. *Genes Cells.* 6:1–12.
13. Mukhopadhyay, R., S. Kumar, and J. H. Hoh. 2004. Molecular mechanisms for organizing the neuronal cytoskeleton. *Bioessays.* 26:1017–1025.
14. Kumar, S., X. Yin, B. D. Trapp, J. H. Hoh, and M. E. Paulaitis. 2002. Relating interactions between neurofilaments to the structure of axonal neurofilament distributions through polymer brush models. *Biophys. J.* 82:2360–2372.
15. Bright, J. N., M. J. Stevens, J. H. Hoh, and T. B. Woolf. 2001. Characterizing the function of unstructured proteins: simulations of charged polymers under confinement. *J. Chem. Phys.* 115:4909–4918.
16. Segall, D. E., P. C. Nelson, and R. Phillips. 2006. Excluded-volume effects in tethered-particle experiments: bead size matters. *Phys. Rev. Lett.* 96:088306.
17. Jahreis, K., T. B. Morrison, A. Garzon, and J. S. Parkinson. 2004. Chemotactic signaling by an *Escherichia coli* CheA mutant that lacks the binding domain for phosphoacceptor partners. *J. Bacteriol.* 186: 2664–2672.
18. Swanson, R. V., S. C. Schuster, and M. I. Simon. 1993. Expression of CheA fragments which define domains encoding kinase, phosphotransfer, and CheY binding activities. *Biochemistry.* 32:7623–7629.
19. Djordjevic, S., and A. M. Stock. 1998. Chemotaxis receptor recognition by protein methyltransferase CheR. *Nat. Struct. Biol.* 5:445–450.
20. Kremer, K. 2003. Computer simulations for macromolecular science. *Macromol. Chem. Phys.* 204:257–264.
21. Glotzer, S. C., and W. Paul. 2002. Molecular and mesoscale simulation methods for polymer materials. *Annu. Rev. Mater. Res.* 32:401–436.
22. Baschnagel, J., K. Binder, P. Doruker, A. A. Gusev, O. Hahn, K. Kremer, W. L. Mattice, F. Mueller-Plathe, M. Murat, W. Paul, S. Santos, U. W. Suter, and V. Tries. 2000. Bridging the gap between atomistic and coarse-grained models of polymers: status and perspectives. *Adv. Polym. Sci.* 152:41–156.
23. Pande, V. S., A. Y. Grosberg, and T. Tanaka. 2000. Heteropolymer freezing and design: towards physical models of protein folding. *Rev. Mod. Phys.* 72:259–314.
24. Ottinger, H. C. 1996. Stochastic Processes in Polymeric Fluids. Springer, New York.
25. Merkel, R. 2001. Force spectroscopy on single passive biomolecules and single biomolecular bonds. *Phys. Rep.* 346:343–385.
26. Janshoff, A., M. Neitzert, Y. Oberdörfer, and H. Fuchs. 2000. Force spectroscopy. *Angew. Chem. Int. Ed. Engl.* 39:3212–3237.
27. Oberhauser, A. F., P. K. Hansma, M. Carrion-Vazquez, and J. M. Fernandez. 2001. Stepwise unfolding of titin under force-clamp AFM. *Proc. Natl. Acad. Sci. USA.* 98:468–472.
28. Fritz, J., A. Katopodis, F. Kolbinger, and D. Anselmetti. 1998. Force-mediated kinetics of single P-selectin ligand complexes observed by atomic force microscopy. *Proc. Natl. Acad. Sci. USA.* 95:12283–12288.
29. Fernandez, J. M., and H. Li. 2004. Force-clamp spectroscopy monitors the folding trajectory of a single protein. *Science.* 303:1674–1678.
30. Livadaru, L., R. R. Netz, and H. J. Kreuzer. 2003. Stretching response of discrete semiflexible polymers. *Macromolecules.* 36:3732–3744.
31. Kim, K. K., H. Yokota, and S. Kim. 1999. Four-helical-bundle structure of the cytoplasmic domain of a serine chemotaxis receptor. *Nature.* 400:787–792.
32. Li, M., and G. L. Hazelbauer. 2004. Cellular stoichiometry of the components of the chemotaxis signaling complex. *J. Bacteriol.* 186: 3687–3694.
33. Shapiro, M. J., D. Panomitros, and D. E. Koshland, Jr. 1995. Interactions between the methylation sites of the *Escherichia coli* aspartate receptor mediated by the methyltransferase. *J. Biol. Chem.* 2:751–755.
34. Bray, D. 2005. Flexible peptides and cytoplasmic gels. *Genome Biol.* 6:1–4.