# Simple Models for Extracting Mechanical Work from the ATP Hydrolysis Cycle

Jonathan L. Eide,* Arup K. Chakraborty,*†§ and George F. Oster‡
*Department of Chemical Engineering, †Department of Chemistry and Biophysics Graduate Group, ‡Departments of Molecular & Cellular Biology and Environmental Science, Policy & Management, University of California, Berkeley, California; and §Physical Biosciences Division and Materials Science Division, Lawrence Berkeley National Laboratory, Berkeley, California

ABSTRACT According to the binding-zipper model, the RecA class of ATPase motors converts chemical energy into mechanical force by the progressive annealing of hydrogen bonds between the nucleotide and the catalytic pocket. The role of hydrolysis is to weaken the binding of products, allowing them to be released so that the cycle can repeat. Molecular dynamics can be used to study the unbinding process, but the binding process is more complex, so that inferences about it are made indirectly from structural, mutation, and biochemical studies. Here we present a series of models of varying complexity that illustrate the basic processes involved in force production during ATP binding. These models reveal the role of solvent and geometry in determining the amount of mechanical work that can be extracted from the binding process.

## INTRODUCTION

Most molecular motors convert chemical energy into mechanical work through a cycle involving nucleotide hydrolysis. The operating principles of these motors are quite different from macroscopic motors because their dynamics is dominated by Brownian motion. For purposes of classification, a nomenclature has evolved to describe two extreme cases of what is actually a continuum of intermediates: Brownian ratchet versus power stroke (1). A Brownian ratchet refers to the situation where the load is driven directly by its thermal fluctuations and its diffusive motion is rectified by a chemical reaction with a large free energy drop involving the motor protein. A typical example is a load driven by polymerization (2). A power stroke refers to a situation where the motion of the load is driven directly by the protein motor, generally through an elastic coupling. Power strokes generate mechanical force using the energy acquired during progressive substrate binding, each step of which is driven by relatively small (Ångstrom-sized) Brownian fluctuations with modest free energy drops. This progressive binding is diffusive, biased by attractive intermolecular forces, a process that has been called the binding-zipper (3).

Power stroke motors generally run in a continuous cycle by using a fuel molecule, usually ATP. The fuel ligand ends up being tightly bound to the protein. To release the ligand so that the cycle can be repeated, it is cleaved into two products in a manner that will allow the products to be released from the catalytic pocket, which can then return to its original binding configuration, whereupon the cycle can repeat. Some motors, such as the $F_0F_1$-ATPase, have characteristics of both types of mechanisms (4,5). This article deals with events in the catalytic site of power stroke molecular motors.

There are two general classes of ATPase motors, grouped according to the structure of their catalytic sites. The RecA class share common structural motifs, including a central $\beta$-sheet whose loops grasp the nucleotide: the Walker A (or P-loop) and Walker B structures (reviewed in (6)). Among these motors are the helicases, AAA motors, proteases, and the $F_1$ ATPase. A second class of motors whose catalytic sites resemble that of the G-proteins include myosin and kinesin. In these motors, the mechanism of force generation is different, involving binding to a polymer track (actin or microtubules), and the energy of the nucleotide hydrolysis cycle is used to regulate the affinity of the track binding site (7). Here we restrict our attention to the RecA class of motor proteins.

Within each class of motors, several other criteria are in use. Motors that move on tracks (e.g., DNA, RNA, or peptide chains) may be classified according to their processivity: how many steps, or hydrolysis cycles, they undergo before dissociating from their track. Tightly coupled motors move, on average, one step per hydrolysis cycle (8,9).

There are two measures of motor protein efficiency. The thermodynamic efficiency can be measured when a motor is processive, tightly coupled, and operating against a conservative load close to its stall force. Frequently, the only load that can be imposed experimentally on a motor is viscous drag. In this situation, the Stokes efficiency measures the ratio of energy dissipation rate to the hydrolysis rate (9). This quantity measures how close the motor force is to a constant force; it is related to motor performance since the dissipation increases with the square of the velocity. Many motors operate near 100% thermodynamic efficiency near stall, and some molecular motors approach a Stokes efficiency of 100% because they output a nearly constant force or torque (3,9). Recently, quantitative models have been used to explain the high Stokes efficiency of ATP synthase (3).

*Submitted August 30, 2005, and accepted for publication February 23, 2006.*

Address reprint requests to George F. Oster, E-mail: goster@nature.berkeley.edu.

Arup K. Chakraborty's present address is Departments of Chemical Engineering, Chemistry, and Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA.



0006-3495/06/06/4281/14 $2.00

doi: 10.1529/biophysj.105.073320

Although constrained and forced molecular dynamics studies have been used to elucidate some features of nucleotide binding (10,11), the size of molecular motors makes unconstrained, atomistic molecular dynamics studies unfeasible. Here we present a coarse-grained model that bridges the gap between detailed atomistic molecular dynamic simulations and mesoscopic Fokker-Planck models. The model provides insight into the general principles that govern how nucleotide binding is efficiently translated into mechanical work.

RecA type motors hold the nucleotide by loops emanating from the central $\beta$-sheet. The actual force is generated principally by the sliding of the P-loop over the nucleotide. This motion is levered up to larger, nanometer size motions in different ways. Generally, the catalytic site is located asymmetrically in the interface between two subunits of the protein, or in the faces of a protein fold, with the P-loop located on the major side. Arginine residues on the minor side frequently assist in holding the nucleotide and transmit the binding stress to the adjacent subunit (12). Here we investigate a cartoon version of the ATP binding process by coarse-graining the P-loop as a polymer chain and modeling ATP as a static surface with reactive sites; we do not treat the catch action of the arginine residues. This molecular motor model converts binding energy into mechanical work, and can operate in a cycle by hydrolyzing the bound nucleotide into two subunits that can be released.

This model illuminates how binding depends on geometric and chemical matching between the catalytic site and the ligand, the proper binding progression for an efficient power stroke, the requirements for the release of spent substrates, and the role of solvent effects on the free energy changes during the binding process. Our results are in general agreement with the binding-zipper model and provide insight into how the P-loop binding to ATP generates a mechanical force. We anticipate that these insights into how a power stroke functions in biological motors can steer coarse-grained models for the design of biomimetic synthetic motors.

The article is organized as follows. First, we describe the model and simulation methods. Then we discuss our primary results and offer concluding remarks. Appendices provide further details relevant to the discussion and details on how the computations were carried out.

## MODELS AND METHODS

To mimic nucleotide binding we use a coarse-grained model where the P-loop (Walker A motif) is represented as a polymer, both with and without side chains, in both two-dimensional and three-dimensional representations. The level of coarse graining is such that one polymer segment represents a single amino acid in the two-dimensional model. The dynamics is computed using the Rouse and modified Binder models for the P-loop, and Brownian and Monte Carlo algorithms. The model is summarized in Figs. 1 and 2. The monomers in the P-loop are connected via harmonic springs,

$$U(\vec{R}) = k_{\text{spring}}(\vec{R} - \vec{R}_{\text{o}})^2, \tag{1}$$

where $\vec{R}$ is the radial distance between adjacent monomers, and $k_{\text{spring}} = 0.34\ k_{\text{B}}T/\text{Å}^2$ for all two-dimensional runs. The value $\vec{R}_{\text{o}}$ is the rest length of 3.8 Å, roughly that of a peptide bond, in the two-dimensional model, and half-this in the three-dimensional model, because each amino acid is represented by two beads as opposed to one. (Throughout our calculations, $k_{\text{B}}$ is Boltzmann's constant and $T = 300$ K is the absolute temperature, so that $k_{\text{B}}T = 4.1$ pN nm.) There is also a potential energy penalty associated with bending the chain given by

$$U(\vec{X}) = \kappa(\Delta\vec{X})^2, \tag{2}$$

where $\Delta\vec{X}$ is the distance from the $n^{\text{th}}$ monomer to the vector between the $n-1$ and $n+1$ monomer (see Fig. 1) and $\kappa$ is the bending modulus ($= 10.4\ k_{\text{B}}T/\text{Å}^2$ for all two-dimensional simulations unless otherwise noted). Excluded volume interactions between segments representing the P-loop are modeled by a repulsive Lennard-Jones (LJ) potential,

$$U_{\text{ev}}(R) = \begin{cases} 4\varepsilon_{\text{ev}}\left(\left(\frac{\sigma_{\text{ev}}}{R}\right)^{12} - \left(\frac{\sigma_{\text{ev}}}{R}\right)^{6}\right) & \text{if} \quad R < \sigma_{\text{ev}} \\ 0 \quad \text{if} \quad R \geq \sigma_{\text{ev}} \end{cases}, \tag{3}$$

where $\sigma_{\text{ev}} = 3.8$ Å in the two-dimensional model (1.9 Å in the three-dimensional) is the monomer radius cutoff and $\varepsilon_{\text{ev}} = 5\ k_{\text{B}}T$ is the LJ energy. We have excluded loop-loop affinities because of the short length and stiffness of the chains representing the P-loop.

The P-loop polymer interacts with a surface representing ATP with evenly spaced attractive sites while otherwise being impenetrable. The shape of the substrate is slightly curved with radius of curvature 234 Å. All surface sites are static and have an LJ radial potential and an angle-dependent potential representing the hydrogen bonds. The potential for the chain segments representing the P-loop is

$$U_{\text{LJ}}(R) = \begin{cases} 4\varepsilon_{\text{PS}}\left(\left(\frac{\sigma_{\text{PS}}}{R}\right)^{12} - \left(\frac{\sigma_{\text{PS}}}{R}\right)^{6}\right) - 4\varepsilon_{\text{PS}}\left(\left(\frac{\sigma_{\text{PS}}}{R_{\text{Cut}}}\right)^{12} - \left(\frac{\sigma_{\text{PS}}}{R_{\text{Cut}}}\right)^{6}\right) & \text{if} \quad R < R_{\text{Cut}} \\ 0 \quad \text{if} \quad R \geq R_{\text{Cut}} \end{cases}, \tag{4}$$

$$U_{\angle}(\angle) = \begin{cases} \dfrac{e^{-\angle^2} - e^{-\angle_{\text{HB}}^2}}{1 - e^{-\angle_{\text{HB}}^2}} & \text{if} \quad \angle < \angle_{\text{HB}} \\ 0 \quad \angle > \angle_{\text{HB}} \end{cases}, \tag{5}$$

where $U_{\text{LJ}}$ and $U_{\angle}$ are the LJ and angle-dependent potentials. The value $\varepsilon_{\text{PS}}$ is the energy of a surface contact ($= 5\ k_{\text{B}}T$), and $\sigma_{\text{PS}} = 3.8$ Å in the

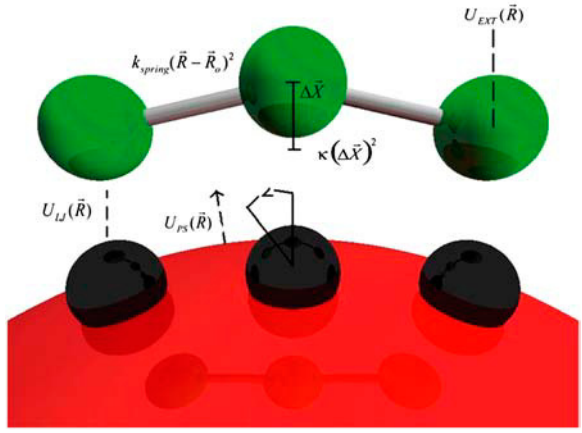


FIGURE 1 Structure of the two-dimensional model, showing the interaction potentials.

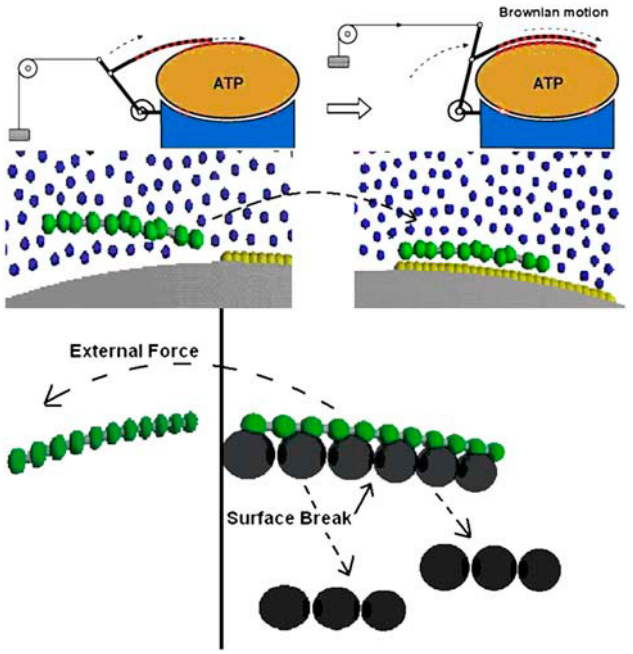


FIGURE 2 (*Top panel*) Cartoon of the binding process used in the model. The top panel is a mechanical analog and the bottom panel shows the simulation setup. The polymer begins with one segment on the surface, while the opposite end is subject to a load force pulling it back to the starting point (*solid circle*). The dashed arrow depicts a typical trajectory. The work done during the binding process is the product of this distance and the applied force. (*Bottom panel*) Model for diffusion of the substrate out of the catalytic pocket. The substrate is broken (hydrolyzed) into two fragments and allowed to diffuse away from the chain. The chain is subject to an external force as before. The surface fragments cannot cross the *y* axis (*vertical line*). The initial configuration has the chain completely zipped to the surface.

two-dimensional model is the normal LJ radius scalar (= 3.3 Å in the three-dimensional model). In the two-dimensional model, 3.8 Å is used to simplify the model because all the lengths are 3.8 Å, whereas, in the three-dimensional model, 3.3 Å is closer to what a real hydrogen bond is, including the length of the carbon-hydrogen bond (remember we are using unified atoms). The value $R_{\text{cut}} = 4\ \sigma_{\text{PS}}$ in both the two-dimensional and three-dimensional models. To mimic hydrogen bonds we set the maximum angle over which force is felt to $\angle_{\text{HB}} = 22°$. In the three-dimensional model, the hydrogen-bond angle is measured using the line formed between the P-O bond on ATP and the reacting monomer.

Finally, we include a short-range repulsion at the surface to render it impenetrable:

$$U_{\text{PS}}(R) = \begin{cases} 4\varepsilon_{\text{PS}}\left(\left(\frac{\sigma_{\text{PS}}}{R}\right)^{12} - \left(\frac{\sigma_{\text{PS}}}{R}\right)^{6}\right) & \text{if } R < \sigma_{\text{PS}} \\ 0 \quad \text{if} \quad R \geq \sigma_{\text{PS}} \end{cases}. \qquad (6)$$

To model substrate cleavage and diffusion out of the catalytic pocket, the surface was divided into two fragments, each of which is mobile (see Fig. 2). The polymer chain begins fully bound to the surface and subject to a constant external pulling force. The surface fragments are constrained not to cross the vertical axis that represents the side of the catalytic pocket into which the P-loop coils back. The fragments of the surface interact through Coulomb repulsion similar to the hydrolyzed fragments of ATP (ADP + $P_i$). The Coulomb repulsion between the surface fragments is given by

$$U(\vec{r}) = \frac{\varepsilon_{\text{surf}}}{|\vec{r}|^2}, \qquad (7)$$

where $\varepsilon_{\text{surf}} = 5\ k_{\text{B}}T$ and $\vec{r}$ is the distance between the center of mass of each fragment.

The three-dimensional model has a more detailed representation of the polymer (P-loop) binding to ATP (see Fig. 3). As in the two-dimensional model, segments of the P-loop are connected by springs between segments ($k_{\text{spring}} = 100\ k_{\text{B}}T/\text{Å}^2$), an energetic penalty for bending is imposed ($\kappa = 50\ k_{\text{B}}T/\text{Å}^2$), and a LJ repulsive potential represents intersegment interactions. The distance between each segment on the chain is 1.9 Å, making every two segments equal to one amino acid (the distance between peptide bonds is 3.8 Å). We also include side chains chosen to match closely those found in the P-loop (13,14). The P-loop amino acid sequence varies between organisms but the functions of the variable side groups are similar. The sequence we used was similar to the one used in other simulation studies (13,14): Glycine-Valine-Glycine-Lysine-Threonine-Glutamic Acid.

The bond length of the side chains are 1.54 Å (the length of a C−C bond). We have added several nonreactive segments to each end of the chain to prevent it from making unphysical moves such as moving above the surface or diffusing away. In the actual catalytic pocket, the P-loop is attached to the central $\beta$-sheet at both ends. This is modeled by preventing the tip of the chain (which is not attracted to the surface) from leaving a region above the surface of ATP (this is ~10 Å away from ATP allowing for unhindered movement along the reaction pathway). We have assumed that ATP has already docked into the catalytic pocket so the movement of the P-loop is limited in this manner.

The backbone of an amino acid has an amine portion, which can form hydrogen bonds, and an acid side that will repel electronegative atoms. To mimic this in our model, only alternating segments on the backbone are attracted to surface sites, whereas the adjacent segment feels an LJ repulsion to those same sites (*red*, attractive; *blue*, repulsive, in the figures and cartoons). The magnitudes of the side-chain attractions and repulsions differ.

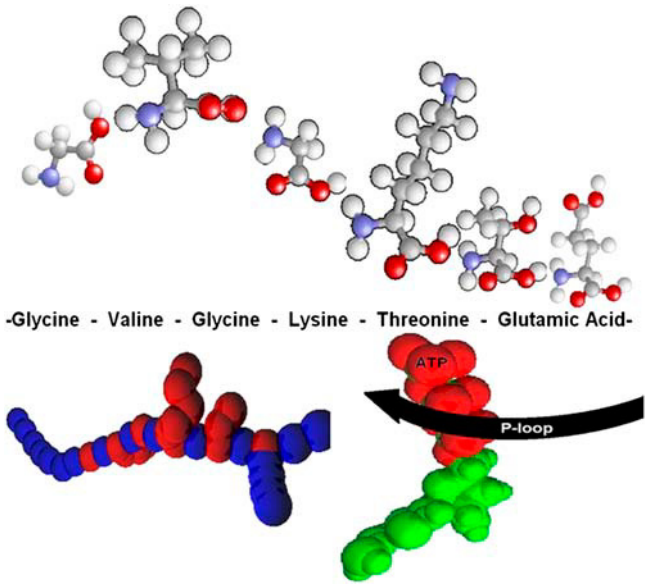


FIGURE 3 United atom model used in the three-dimensional P-loop simulation. The structure and sequence of the P-loop (*top panel*) is approximated in the cartoon on the bottom left. In the top image: red, oxygen; blue, nitrogen; gray, carbon; and white, hydrogen. In our cartoon, red segments are hydrogen-bond-forming (see Table 1 for the strength of these bonds) and blue segments are repulsive to oxygen on ATP. ATP, on the bottom right, is modeled with red as hydrogen-bond-forming oxygen, whereas the rest of the molecule (*green*) is repulsive. The bottom right depicts the approximate path that P-loop takes during the simulated binding to the surface of ATP.

The strength of these interactions is approximated based on other work (15). For example, the side chain for lysine has four carbon segments and then a $NH_2$ tip. This $NH_2$ tip is strongly electropositive and therefore can form a much stronger H-bond than the rest of the side group or backbone. The lysine tip has a LJ energy of just over 8 $k_BT$ while the backbone can only bond with $\sim 3\ k_BT$ (see Fig. 1). The length of the bonds between ATP and the P-loop ($\sigma$ in Eq. 4) was set at 3.3 Å.

We have modeled our surface after ATP as taken from the Protein DataBank (http://www.rcsb.org/pdb/). In the simulations, we have held ATP stationary and allowed only the oxygen on the phosphate groups to form H-bonds with the P-loop (16). The H-bonds are modeled as before with an LJ and angle potential ($\angle_{HB} = 90°$ in Eq. 5). The moment used to measure the angle in the potential points radially out from the axis of the phosphate groups. To reduce computation time we have not modeled all of the atoms in the sugar end of ATP; instead, we have surrounded them with a repulsive LJ sphere.

Explicit solvent was modeled as an LJ fluid with the same attractive potential to the surface as the backbone reactive segments of the chain ($\varepsilon_{\text{solvent-surf}} = 1.5\ k_BT$) with $\sigma = 3.8$ Å (two dimensions) and 3.3 Å (three dimensions). The LJ solvent-solvent parameters were chosen to maintain them in a liquid state (17). The LJ parameters for solvent-chain interactions are the same for solvent-solvent interactions. In RecA type motors, once the P-loop has slid over the nucleotide, water molecules are squeezed out of the catalytic pocket. The analog in our model is once the P-loop slides over the substrate, all of the solvent molecules are displaced away from the interface.

The motion of the system, with implicit solvent, was modeled using Brownian dynamics simulations according to the Langevin equation,

$$\zeta \frac{d\vec{X}_n(t)}{dt} = -\frac{dU(\vec{X}_n)}{d\vec{x}} + f(t), \tag{8}$$

where $f(t)$ is the random force characterized by a Gaussian distribution with zero mean, and variance given by

$$\langle f(t) f(t') \rangle = 2k_b T \zeta \delta(t - t'). \tag{9}$$

The friction constant is approximated using the Stokes-Einstein relation,

$$\zeta = D/k_BT = 2\pi\eta a \times \Delta, \tag{10}$$

where $\eta$ is the viscosity (set to 1 centipoise), $\Delta$ is the dimensionality (two or three dimensions), and $a$ is the radius of the monomer (= 1.9 and 4.0 Å for two dimensions and three dimensions, respectively). A cartoon of the model, displaying all the potential functions, is shown in Fig. 3.

To mimic P-loop binding in two dimensions, the polymer begins with only one segment at the leading edge under the influence of the surface potential (see Fig. 2). This first segment (on the *left* in Fig. 2), has a constant load force pulling it back to the original position. A constant force produces the highest Stokes efficiency and is also the consequence of a small motor coupled to a much larger cargo via a weak elastic linkage (5,9,18,19). Varying the constant load force acting on the first segment changes the work output from the binding process. The spontaneous curvature preferred by the polymer representing the P-loop is set to match the slight curvature in the binding surface. The number of surface sites is varied in different simulations.

To measure the free energy of binding, Monte Carlo umbrella sampling was used (20). The reaction coordinate is defined as the radial distance of the first segment from its initial position; this is proportional to the work output. Free energy changes were computed from the probabilities along the reaction coordinate.

For the explicit solvent cases, Monte Carlo simulations were performed with LJ solvent in both the two-dimensional and three-dimensional models. In the two-dimensional model, periodic boundary conditions were imposed on both sides of the P-loop, but a impenetrable surface was placed above the surface with the distance between the surface and opposite boundary large enough that bulk two-dimensional LJ liquid conditions could be established (21). The solvent molecule density and interaction strength were set to ensure that the system was in a liquid state at 300 K (17). The P-loop was initially pinned at both ends to prevent from zipping onto the surface, solvent molecules were added randomly and then the system was allowed to equilibrate for $5 \times 10^6$ MC steps. The reaction coordinate was separated into four regions for umbrella sampling, with each region sampled for at least $5 \times 10^6$ MC steps for both implicit and explicit solvent cases.

The boundary conditions in three-dimensional employ a mean field force by making a mean field approximation of all solvent particles that would be outside an imaginary sphere around the system of interest (a technique used in simulation packages such as CHARMM) (22). This greatly reduces the computational time compared to periodic boundary conditions, at the cost of introducing a reasonably small error. The 750 LJ solvent molecules were used in a 21 Å sphere around ATP and the P-loop. To initialize the system, a random P-loop configuration was taken from an implicit P-loop configuration (positioning varying on where in the reaction coordinate we were performing the umbrella sampling). The solvent molecules were added randomly and then allowed to relax, without P-loop movement for $2 \times 10^6$ MC steps. The reaction coordinate was separated into five regions for umbrella sampling, with each region sampled for at least $5 \times 10^7$ (explicit cases) or $5 \times 10^6$ (implicit cases) MC steps.

## RESULTS

In this section we discuss first the two-dimensional model. The three-dimensional model is discussed insofar as its results deviate from that obtained using the two-dimensional model.

### Two-dimensional model

#### *Bonds form sequentially and smoothly*

To demonstrate that the two-dimensional model forms bonds sequentially with the substrate in a relatively smooth and constant manner, we first measured the binding free energy profile of the system using umbrella-sampling Monte Carlo (20). Results from a load force of 16.5 pN ($\sim 15\ k_BT$ work done upon a complete zip) and a dense surface of 22 sites are shown in Fig. 4 (all other surface densities look quantitatively similar for other cases where the solvent is represented explicitly). See the Appendices for comparison with the implicit case.

Monte Carlo simulations were run in the NVT ensemble where energy transfer between the infinite thermal heat bath and the polymer makes it possible for the external work to be larger than the binding free energy. Fig. 4 shows that the free energy and enthalpy decrease fairly smoothly and nearly linearly, in agreement with the binding-zipper model. The computed bonds form sequentially along the surface in a smooth manner, much as the cartoon depicts in Fig. 2 (see the movie in the Supplementary Material). The bumps that appear near the end of the reaction coordinate are caused by the breaking and creating of bonds between the polymer and surface. For a new segment to bind to the surface, the segments already bound must break their current bonds and create new bonds at one step further along on the surface. This creates an increasing energy barrier that leads to the nonlinear behavior. At the start of the binding process, there are few bound segments, so this behavior is not yet manifest.

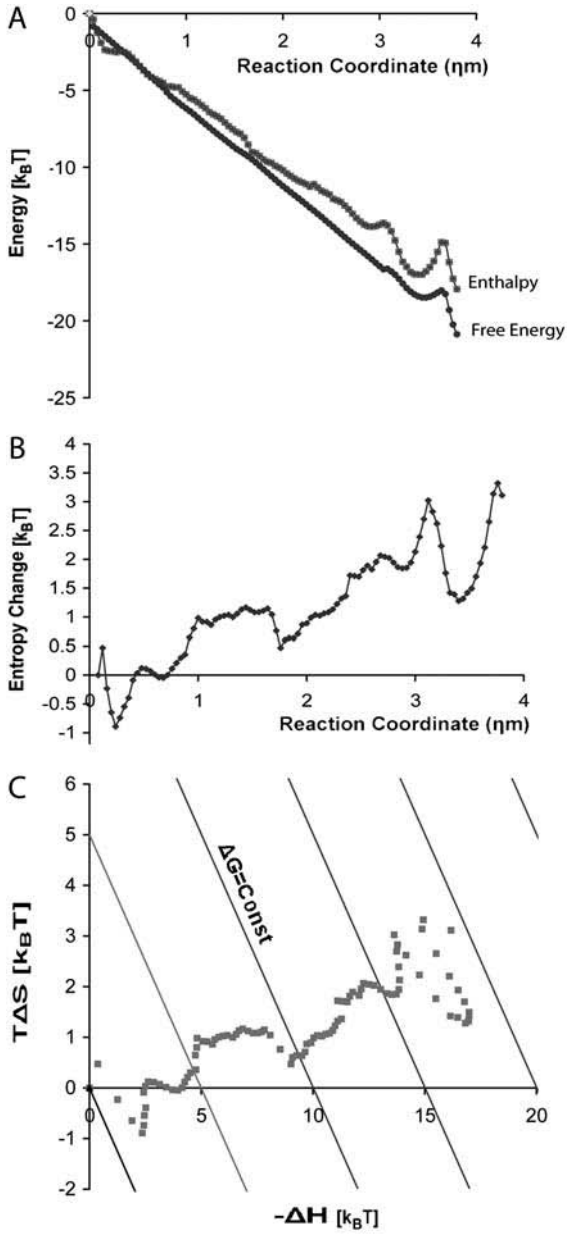


FIGURE 4 Energy profiles. Simulations were created using umbrella-sampling Monte Carlo with a load force of 16.5 pN and 22 surface sites. Error bars not shown, but the standard deviation in the measurement bins (bins were 0.38 Å) is 4–5 $k_BT$. (*A*) Free energy and enthalpy profile. The work is 15 $k_BT$; i.e., the amount of energy stored in the chain as it anneals to the surface against a constant force. See text for discussion on the nonlinear behavior near the end of each trajectory. (*B*) Entropic profile. The entropy is increasing because of the entropic gains due to solvent molecules being displaced from the surface. The solvent-surface LJ interaction energy ($\varepsilon$) was set to 1.5 $k_BT$. (*C*) Entropic versus enthalpic energy. As the polymer moves along the substrate and forms bonds, the explicit solvent model shows a steadily increasing free energy monotonic with enthalpy.

This type of barrier behavior has been observed in other molecular models (23). Another method of characterizing the binding process is the effective driving potential (EDP) (1), which is discussed in the Appendices.

Fig. 4 *B* shows that the entropy does not change nearly as smoothly but the trend is an increase of entropy as the chain zips to the surface. The error bars for the enthalpy in Fig. 4 *A*, and thus Fig. 4 *B*, are ~4–5 $k_BT$. The majority of this uncertainty comes from the large enthalpic fluctuations, characteristic of an explicit solvent (see Appendices for comparison with implicit solvent). The enthalpy due to chain interactions changes smoothly but the hundreds of solvent molecules, responsible for the majority of the total enthalpy in the system, lead to much larger fluctuations. The largest nonlinearities in the entropy profile also take place near the end of the reaction coordinate where we see the bumps in both the free energy and enthalpy.

Generally, there are several solvent molecules bound to both the chain and substrate. The release of these bound molecules by P-loop binding can dramatically increase the entropy of the system. This effect can be seen in the simulations with explicit solvent where there is ~3–4 $k_BT$ entropic gain upon successful zipping; note that this trend is reversed when the solvent is represented implicitly. The size of the entropic gain or loss varies with the interaction strength of the solvent with the surface as well as the stiffness of the chain, $\kappa$, since this affects the size of its configuration space. The P-loop is competing with solvent for surface sites. Therefore, the requirement for the binding process to be favorable is that the gain in solvent entropy must be greater than the enthalpic loss due to a P-loop segment displacing solvent molecules plus the entropic changes due to a chain segment binding to the surface. If we set the surface-solvent LJ energy, $\sum_{\text{surf-solvent}}$, to zero, the entropy change is the same as when the solvent is represented implicitly. If $\sum_{\text{surf-solvent}} \gg \varepsilon_{\text{surf-chain}}$, then the chain will never bind to the surface because the P-loop will be unable to displace solvent molecules. When $\varepsilon \approx 1.5\ k_BT$, we found that the entropic gains were maximized, as shown in Fig. 4 *B*. Note that this is not the maximum free energy drop, which would take place at $\varepsilon_{\text{surf-solvent}} = 0$, or equal to the implicit case. As discussed in the Appendices, this is not the case for the three-dimensional model. The entropy versus enthalpy profile in Fig. 4 *C* shows this graphically: as the polymer moves along the substrate and forms bonds (increasing the enthalpic energy), the chain gains entropy linearly (the slope moving through the lines of free energy is constant). The volatility is similar to that found in Fig. 4 *B*.

### *Load velocity behavior*

Next, we examined the speed and power of the binding-zipper using the two-dimensional implicit solvent model. This addresses the issue of how quickly and efficiently chemical energy of binding is converted into mechanical work. A trajectory is deemed successful when the entire polymer is bound to the surface, which corresponds to the maximum work output (see Fig. 2). At the start of the simulation, only the leading edge segment is bound to the surface. Using

Brownian dynamics simulations, we generated the velocity versus load profile shown in Fig. 5 *A* for binding site densities that varied from 22 to three evenly spaced sites; all data points are the result of 1000+ trajectories. The velocity profile is nearly linear for all loads and surface site densities, except when it asymptotically approaches the stall force. These load-velocity profiles are qualitatively similar to the experiments using conservative force (force clamps) in kinesin (24) and for ATP synthase operating against a viscous load (25,26).

Interestingly, the surface with six sites binds faster than the surface with 11 sites for small forces but the case where there are three sites leads to slower binding than when there are six or 11 sites. Fig. 5 *B* examines the effects of surface site density versus the zipping velocity while holding the load force constant. With high surface density (see *cartoon inset* on the *right* in Fig. 5 *B*), the sites are close enough together so that their potentials overlap (recall: each surface site has an angular dependence as described by Eq. 5). This leads to relatively small barriers as the P-loop moves down the surface; that is, only small fluctuations are required for each segment to move to the next surface potential. With 11 sites (*middle cartoon inset* in Fig. 5 *B*) the potentials no longer overlap, thus requiring larger fluctuations for the segments to move out of the current energy well and into a new well, one step further down. Further frustrating the 11-site binding is the perfect match between the chain and the binding surface, which raises the enthalpic barrier between steps considerably, while lowering the entropy gain when bound. When there are a small number of surface sites (*cartoon inset* on the *left* in Fig. 5 *B*), the barrier to moving down the surface never rises very high, and the binding is not slowed appreciably. Moreover, the chain can take on many more configurations along the surface, thus increasing its entropy relative to the denser site case. This permits faster binding at lower load forces, but the stall force is much smaller compared to the cases with more surface sites. The surface site density determines the slope of the velocity-force curve. This power-stroke type of behavior has been found in other biological models (23). Could natural selection have chosen substrates with fewer binding sites because of this phenomenon?

In Fig. 5 *C*, we have plotted the power-versus-load force. The power output is defined as the amount of work output (the external work against a constant force) divided by the average time to fully bind. Each case exhibits a fairly symmetric profile, passing through a maximum where the power output is optimal. This is the rate of work a motor exerts in a single power stroke. Nature may have evolved molecular motors to work at these optimal load forces to maximize the work they could produce.

Combining the velocity and binding free energy, we computed the difference between the work output and the

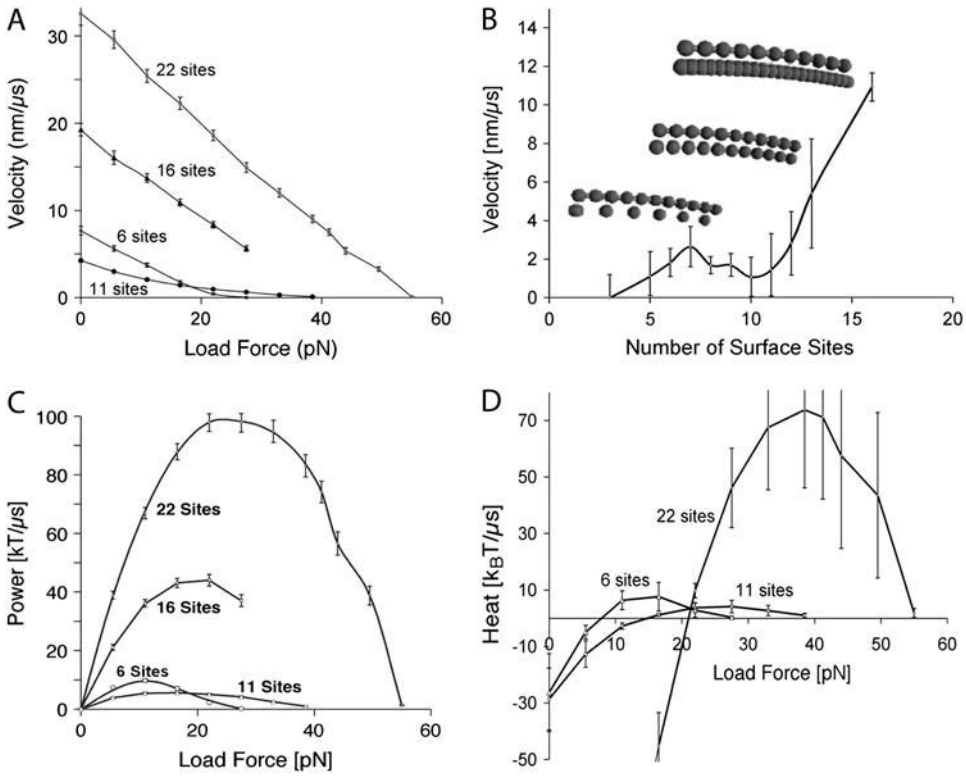


FIGURE 5 Mechanical performance. (*A*) Load-velocity relation at various surface site densities. All surface sites are evenly spaced. All cases are decreasing linearly until they approach their stall force. Error bars are standard deviations of at least 1000 runs. The load-velocity curves are nearly linear. (*B*) Velocity dependence on the surface site density with a load force of 16.5 pN. Cartoon insets depict the energetic frustration when the surfaces match and the overlap of sites when the surface is densely filled with active sites. The local minimum around 10 surface sites arises because of the perfect matching of the chain to the surface, whereas the Vernier effect smoothes out the net energy profile when the surfaces mismatch. (*C*) Power versus force profile for various surface site densities. Because the load-velocity curves are nearly linear, there is an optimal force that produces the maximum power output. (*D*) Heat flux versus force profile for various surface site densities. Positive heat is defined as thermal energy captured by the motor, whereas negative heat is the energy dissipated to the constant temperature bath. There is an optimal force where the maximum energy is captured from the thermal bath. When the heat change is zero (nearly reversible), the system is working adiabatically. This also roughly corresponds to the maximum power in panel *C*.

binding free energy divided by the time for a full zip. This measures how much energy flows into the system from the thermal bath via favorable Brownian fluctuations, or how much energy is dissipated into the bath via viscous dissipation. We adopt the usual thermodynamic convention that treats heat into the system as positive (endothermic), and heat out of the system as negative (exothermic). Fig. 5 *D* shows this for a few cases. As with the power versus load profile, there is an optimal load where the most energy is captured from the thermal bath. When the heat flow is zero, the system is working adiabatically, which roughly corresponds to the maximum power in Fig. 5 *D*. When the heat transfer between the bath and the motor is zero, the energy flowing into the polymer chain (the size of favorable fluctuations required to move down the surface) is equal to the amount of energy that is being dissipated (viscous dissipation). If the heat energy required to move the chain is high then the motor moves very slowly, having to wait a long time for favorable thermal fluctuations. These results could have implications for synthetic motor design.

### *The role of protein elasticity*

Peskin and Elston demonstrated how protein flexibility affects motor performance (18,19). In a similar manner, we investigated the effect of varying the parameters $k_{spring}$ and $\kappa$. Fig. 6, *A* and *B*, show how the velocity of zipping depends on $k_{spring}$ (stretching) and $\kappa$ (bending). Just as Peskin and Elston found, there is an optimal stretching spring constant with respect to motor velocity (this is discussed further in the Appendices for both the two-dimensional and three-dimensional models). When $k_{spring}$ is small, the increased flexibility wastes fluctuations: it is much more likely to make moves that will not lower the free energy, and so it takes a longer time to move the chain against a constant force down the reaction coordinate. At large $k_{spring}$ values, the entire chain is so stiff that it is effectively attached to the constant force so it can only diffuse as fast as the whole chain. The optimal elasticity is found at an intermediate value of $k_{spring}$ as shown in Fig. 6 *A*. This profile would look different if the surface were not uniform and larger fluctuations were required to zip, for the chain would need the added flexibility to access adjacent local energy minima.

Fig. 6 *B* shows that the velocity dependence on $\kappa$ (bending) does not pass through a maximum, but quickly approaches an asymptote at large $\kappa$ that decreases as the number of sites increases. At a small $\kappa$, the velocity is reduced because the chain flexibility leads to wasted fluctuations that do not match the surface. At large $\kappa$ the stiff chain approaches a perfectly matched rod sliding over the surface. If the surface and chain were not matched perfectly, then the chain would require some flexibility to search out the lowest free energy configuration. Optimal flexibility strategies depend on the type and shape of substrate and how specific the binding chain has to be.

## The motor working in a cycle

So far, we have discussed only a single power stroke. However, a motor must operate in a cycle, and so it must rid itself of tightly bound nucleotide so that the cycle can repeat. This is accomplished by hydrolyzing the nucleotide into two parts: ATP $\leftrightarrow$ ADP + $P_i$ (1,16). In many motors (but not all), the phosphate moiety diffuses out of the catalytic pocket spontaneously, but the ADP fragment may require energy from a second binding site to release (16). Here we present a very simplified model to illustrate how the substrate might diffuse out of the catalytic pocket after hydrolysis.

The two fragments (ADP + $P_i$) are negatively charged, and so repel each other. This distorts the highly angle-dependent hydrogen bond network, dramatically decreasing the free energy that binds the substrates into the catalytic site. In the $F_1$ motor, the binding-zipper stores elastic energy that aids the protein in recoiling to its open (empty) configuration. This elastic energy aids in the release of hydrolysis product (27,28). A cartoon of this system is depicted in Fig. 7 *A* where one catalytic site zipping onto its substrate uses some of its mechanical energy to help another site unzip and release a spent substrate. Using this model with our previous results we can determine the optimal conditions for product release.

In Fig. 7 *B*, the size-dependence of the release time for a fragment is examined with an external force of 16.5 pN and

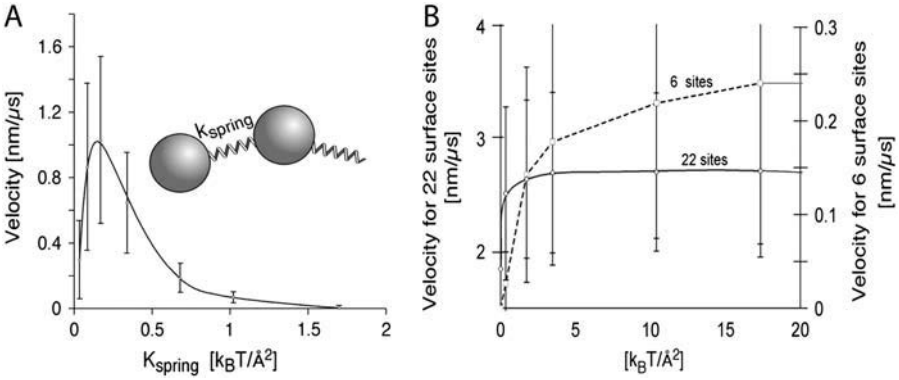


FIGURE 6 The effect of elasticity. (*A*) For nine surface sites, there is a optimal stretching elasticity, $k_{spring}$, that gives the fastest zipping velocity (the results are similar for all cases). (*B*) For the six- and 22-surface-site cases (similar for all cases), the velocity dependence on $\kappa$ (bending) approaches an asymptotic value as the chain approaches a perfectly matched rod to the surface.

nine surface sites. Both substrates diffuse away fastest when the fracture takes place near the middle of the substrate. When the fracture takes place near the end of the substrate away from the impenetrable axis (see Fig. 3), the time for diffusing takes much longer. This is because the smaller outer fragment quickly diffuses away leaving the larger fragment to wait for favorable thermal fluctuations to escape the chain. When the fracture takes place on the side closest to the impenetrable axis, the rattling of the smaller segment against the impenetrable axis helps to push the larger substrate fragment out (both fragments have a columbic repulsion described in Eq. 7), resulting in a quicker release time. This has implications as to why nature has chosen substrates with certain numbers of attractive sites (ATP versus substrates with additional phosphate groups). The requirement for an efficient cycle is such that a substrate must have enough sites to create a strong power stroke but not so many sites that it is unable to fragment in a way as to efficiently release the products. The diffusion of products out of the catalytic is more complicated and involves the interaction of solvent molecules, other protein structures in the catalytic pocket and/or ions—especially $Mg^{+2}$ (29,13,14).

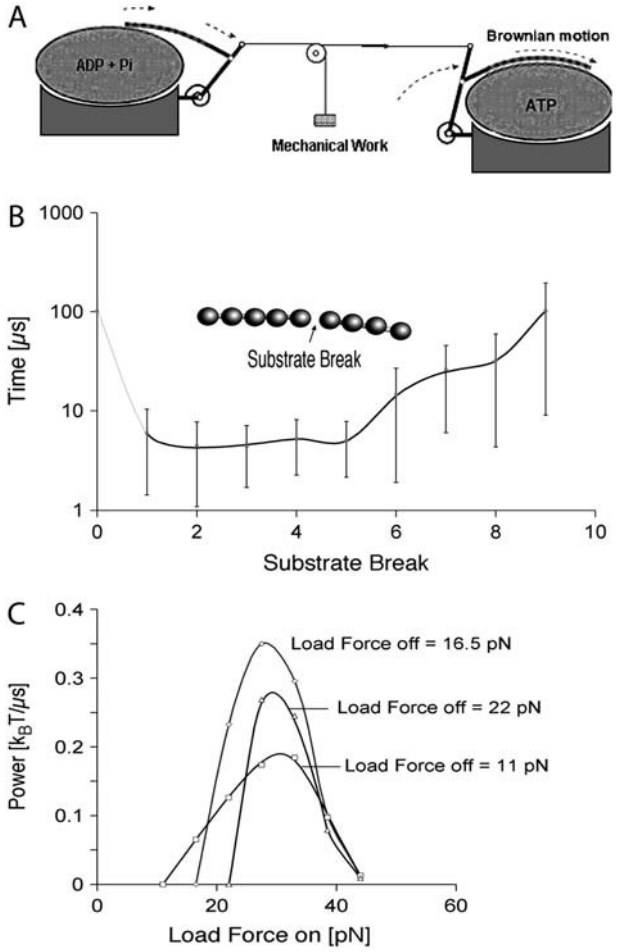


FIGURE 7 Releasing products to make a cycle. (*A*) Cartoon showing how two catalytic sites in a single motor might help each other release spent substrates, and how some of the mechanical work during the power stroke can be stored to aid in product release. As one catalytic site zips (*right*), some of the mechanical work is used to change the conformation of another catalytic site (*left*) such that it aids in the removal of a spent substrate. (*B*) Time for the substrate to diffuse away from the chain depending on where the substrate is fractured. The external force is 16.5 pN with nine surface sites. The substrate break refers to which site is fractured along the surface, where 0 is closest to the protein body while nine is furthest away. (*C*) Cycle power of a substrate with nine surface sites. The load force used is constant when the chain zips to the surface and a different constant load force off is applied when the chain unzips from the fragments. There is an optimal condition for each load force and a global optimum at ~27.5 pN on and 16.5 pN off. This optimum load force on is not the same as found for the power stroke results in Fig. 5 *C*.

With the data accumulated from the unzipping model, we can combine this with the zipping model to determine the optimal conditions for a cycle. This is accomplished by determining how long it takes the chain to zip under a specific external force and then how long it takes to unzip at a different external force. The difference between the external forces determines the total work created. The result for the case where there are nine surface sites is shown in Fig. 7 *C*. The abscissa displays the constant force used when the chain zipped onto the surface and data sequences are for various forces pulling the chain off the fragments. There is an optimal condition for each load force (on and off) and a global optimal at roughly a load-force-on of 27.5 pN and load-force-off of 16.5 pN. Notice that this optimal condition is different from the one found in Fig. 5 *C*. Other surface-site densities had similar results. At lower unbinding forces, although the total work is large, the power becomes small because of the inability of the fragments to diffuse quickly off the chain. This is further evidence that some molecular motors may require stored mechanical energy to remove spent substrate molecules at an appreciable rate (27).

In summary, the two-dimensional model captures the minimum requirements for an efficient power stroke and allowed us to evaluate the effects of chemical, mechanical, and solvent interactions. We have modeled the system after the P-loop of RecA class molecular motors, which may behave in a similar way.

## Three-dimensional model

### *Bonds form sequentially and smoothly in three dimensions*

Next, we investigated the three-dimensional model to further illustrate the basic processes involved in force production during ATP binding. The major differences between the two-dimensional and three-dimensional models are that we include side chains in the three-dimensional P-loop and model the ATP substrate more realistically. The results from a characteristic trajectory of the zipping process are shown in Fig. 8 *A*. As the P-loop first approaches ATP (Fig. 8 *A*-I), the first reactive segment, Glycine, begins to form a hydrogen bond with an Oxygen on ATP. The segments turn black when they

are bound to a surface site(s) (arbitrarily chosen as less than a radius of 4.5 Å and angle of 30°). Once Valine binds to ATP in Fig. 8 *A*-II, the much more reactive and flexible lysine zips soon after (Fig. 8 *A*-III). Lysine will always remain tightly bound to ATP but its length and flexibility allows the rest of the chain to continue snaking along the surface of ATP. Finally, Valine is able to wrap around ATP and the number of H-bond contacts is maximized. In this configuration, the P-loop reaches its free energy minimum (Fig. 8 *A*-IV). Our model shows that when the P-loop zips to ATP, the side groups fluctuate out from the backbone and pull the entire chain along the reaction coordinate, so that the chain walks across the nucleotide like a centipede. The flexibility of the side groups is much greater, and possibly more reactive, than the backbone, and thus well suited to facilitate the progressive annealing process.

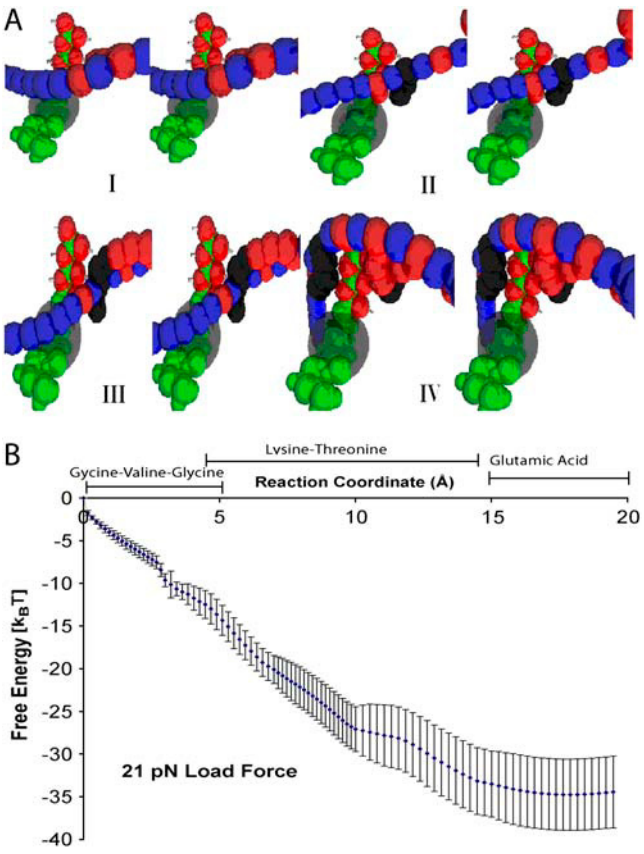


FIGURE 8 Zipping of the P-loop in the three-dimensional model. (*A*) Images from the simulation of the P-loop zipping to ATP. Red segments are LJ-attractive and blue are repulsive. Segments are black when they acquire a large interaction energy with an oxygen on ATP. (*I*) As the P-loop begins to zip onto ATP, no segments have yet completely bound. (*II*) The first two amino acids have bound to the surface, allowing for lysine to fluctuate into a position to do the same. Lysine is the major actor driving the binding process. (*III*) lysine is tightly bound, but the length and flexibility of this side group allows for the rest of the chain to continue along ATP to the other oxygen. (*IV*) The P-loop is near its free energy minimum as the number of surface contacts is maximized. The first few amino acids are now binding to oxygen on the other side of ATP. (*B*) The free energy profile for the zipping of the P-loop onto ATP with a load force of 21 pN. The major enthalpy and free energy drops take place during the lysine binding. Glutamic acid serves as a stopper for the P-loop, keeping it tightly bound to the minimum free energy position.

The free energy profile for the binding process for the three-dimensional model with explicit solvent molecules and a 21 pN external force is shown in Fig. 8 *B* was calculated using MC umbrella sampling (see the Appendices for comparison to the implicit solvent three-dimensional case). The results in Fig. 8 *B* show that the free energy profile is again smooth and linear, in a manner similar to our two-dimensional model. The process in which the P-loop zips to the surface differs because the chain is not uniform in the three-dimensional model. The three-dimensional model has side chains varying in length and reaction strength. The $NH_2$ tip very strongly binds to the surface (often modeled as $NH_{3+}$ as in (14)), while the rest of the chain is flexible and much less reactive (see Table 1). In this regard, the two-dimensional and three-dimensional simulations differ. Lysine acts as a tether, or hook, fluctuating out to bind ATP allowing for the rest of the chain to diffuse into place. The flexibility and length of Lysine also allows the huge enthalpic gain, due to $NH_2$ binding, to be spread out along the reaction coordinate and thus gently decrease the free energy. Mutation studies show that the Lysine is extremely important for the binding and release of ATP (30,31). The rest of the P-loop, although it does not bind as strongly as $NH_2$, displaces a greater number of solvent molecules than the lysine $NH_2$ tip. This increases the solvent entropy and drives down the free energy. This and the smoothing out of the reactive $NH_2$ tip results in the smooth free energy drop seen in Fig. 8 *B*. The last side group on the P-loop is the electronegative Glutamic Acid residue that repels the oxygen on ATP. This group serves as a stopper so the P-loop remains tightly bound to ATP in the configuration shown in Fig. 8 *A*-IV. Glutamic Acid comes into contact with ATP near the free energy minimum in Fig. 8 *B*. In most RecA ATPases, Arginines from the adjacent subunit also act as a latch onto the $\gamma$-phosphate; although important, we have not included them in our simulations.

As in the two-dimensional case, there is a solvent entropy gain, but instead of $\sim 1\ k_BT$ gain for each solvent molecule (approximated by ideal gas) displaced from the surface in two dimensions, it is $\sim 1.5\ k_BT$ in the three-dimensional

**TABLE 1 Energy used for hydrogen bonds during simulations ($\varepsilon$ in Eq. 4)**

| Unified atom | LJ energy ($k_BT$) |
|---|---|
| $NH_2$ (lysine) | 8.3 |
| Backbone H on lysine | 5.0 |
| $CH_2$ (such as valine) | 0.8 |
| Glycine H | 4.2 |
| Other backbone H | 3.3 |
| Acid groups | −5.0 |

model. The exact solvent entropic gain is difficult to measure, because the three-dimensional enthalpic fluctuations of the LJ solvent are several times larger than the two-dimensional case. Our free energy change for the explicit solvent case will also depend on the choice of our model liquid (LJ versus TIP3, etc.) and the interaction strength of the liquid with ATP and the P-loop. Unlike our two-dimensional model, the free energy drop in the three-dimensional model is actually larger with explicit rather than implicit solvent (see the Appendices for comparison and discussion). Again, our intention is not to determine the exact free energy change of the actual binding of P-loop to ATP, but instead to elucidate the qualitative characteristics of such a binding process. The size of the free energy drop should not be compared to the experimental binding energy of $\sim 20\ k_BT$ (32) because we have only considered the binding stage of the reaction, leaving out the docking and release phase, the $Mg^{+2}$ interactions, solvent interactions, and ATP interactions with other protein structures in the catalytic pocket.

These qualitative results are similar to the two-dimensional case and what others have calculated for the free energy drop of ATP binding to the P-loop. (In addition to the $\sim 35\ k_BT$ free energy minimum, there is $\sim 9\ k_BT$ stored in the constant force spring.) Antes et al. (29) used an atomistic model with TIP3P solvent, and found the free energy difference upon binding (unbinding) between $F_1$-ATPase, solvent, and ATP to be $\sim 80\ k_BT$ (their simulation included $Mg^{2+}$ and other loops in the binding pocket). A direct comparison should not be made between our coarse-grained model and the actual RecA class P-loops. Our model is intended only to illustrate the basic processes involved in force production during ATP binding. Nevertheless, we have shown that the formation and breaking of bonds along the zipping path of the P-loop acts very similar in our model to the atomistic simulation. Massova et al. (33) estimated the free energy including solvent entropic effects, and Karplus et al. (34) used a combination of molecular dynamic simulations and experimental binding constant measurements to determine the free energy change.

## DISCUSSION

Here we have used simplified models to investigate the general principles that govern how nucleotide binding is efficiently translated into mechanical work. The coarse-grained models we present bridge the gap between detailed atomistic molecular dynamics simulations and mean field models that capture the power stroke generated by the binding process via a phenomenological potential function (1,16,35). The results described here give a simple but convincing description of the binding-zipper model for energy transduction (6,29). The model illuminates how binding depends on geometric and chemical matching between the catalytic site and the ligand, the role of binding progression in developing an efficient power stroke, the requirements for the release of spent substrates, and the role of solvent effects on free energy changes during the binding process.

We first examined the P-loop binding process with a simplified two-dimensional model and then increased the complexity toward an atomistic model, showing that the basic mechanisms can be captured in the simple two-dimensional representation. The three-dimensional model elucidates some of the requirements for the P-loop binding to ATP. Our three-dimensional results also show that a more detailed model acts in a similar manner as our much simpler two-dimensional model. A molecular motor is vastly more complicated than our three-dimensional model, but our results indicate much of this detail can be coarse-grained into a simpler model. The details required for an accurate representation are not obvious but depend significantly on the system of interest. Even with this disclaimer, perhaps similar coarse-grained models can be used to elucidate general principles of other protein binding interactions as well as aid in the design of biomimetic devices.

Future work should approximate the free energy changes accompanying P-loop annealing using methods from polymer physics (see, for example, (36)). A similar approach could be applied to the class of proteins exemplified by kinesin and myosin where the free energy of nucleotide binding is used to alter the affinity of the motor for its track—microtubule or actin—via an escapement involving the Switch 1 and 2 motifs. In these motors, the power stroke draws its energy from binding to the track, so that even models such as we have presented here would be considerably more complicated.

## APPENDIX A: IMPLICIT SOLVENT COMPARISON

Fig. 9 *A* shows that the free energy profile for the two-dimensional implicit solvent model is qualitatively similar to the model that treats the solvent explicitly. The total free energy drop is smaller for the explicit solvent case compared to the implicit case because the total enthalpy decreases when the chain zips to the surface and expels solvent molecules. There is an enthalpic energy loss, not present in the implicit case, due to the breaking of bonds between the solvent molecules and surface. This loss effectively decreases the surface sites strength by the enthalpic energy of the surface-solvent interactions (i.e., the enthalpic energy gained by P-loop binding is offset by the enthalpic energy loss due to the breaking of the solvent surface bond). Some, but not all, of the enthalpic energy will be recaptured in the free energy due to the increased entropy of the displaced solvent molecules.

The entropy profile for the implicit and explicit two-dimensional cases (Fig. 9 *B*) differs dramatically due to the solvent entropic effects. In two dimensions, the free energy gain from releasing an ideal solvent molecule is $\sim 1\ k_BT$ and there are $\sim 10$ solvent molecules that must be displaced during binding; this is in good agreement with the $\sim 8$–$9\ k_BT$ difference between the implicit and explicit case shown in Fig. 9 *B*. As we have discussed before, the enthalpy versus entropy profile displayed in Fig. 9 *C*, also emphasizes the importance of solvent entropy. As the binding process moves along the reaction coordinate, the implicit simulation loses entropy while the explicit simulation gains entropy.

The free energy profiles for the implicit and explicit three-dimensional models are compared in Fig. 9 *D*, with a load force of 21 pN. The free energy drop is larger and smoother with explicit solvent. The free energy drop for the implicit case changes slopes between 5 and 8 Å where lysine binds (protein flexibility might smooth this out as discussed later in the Appendices). The

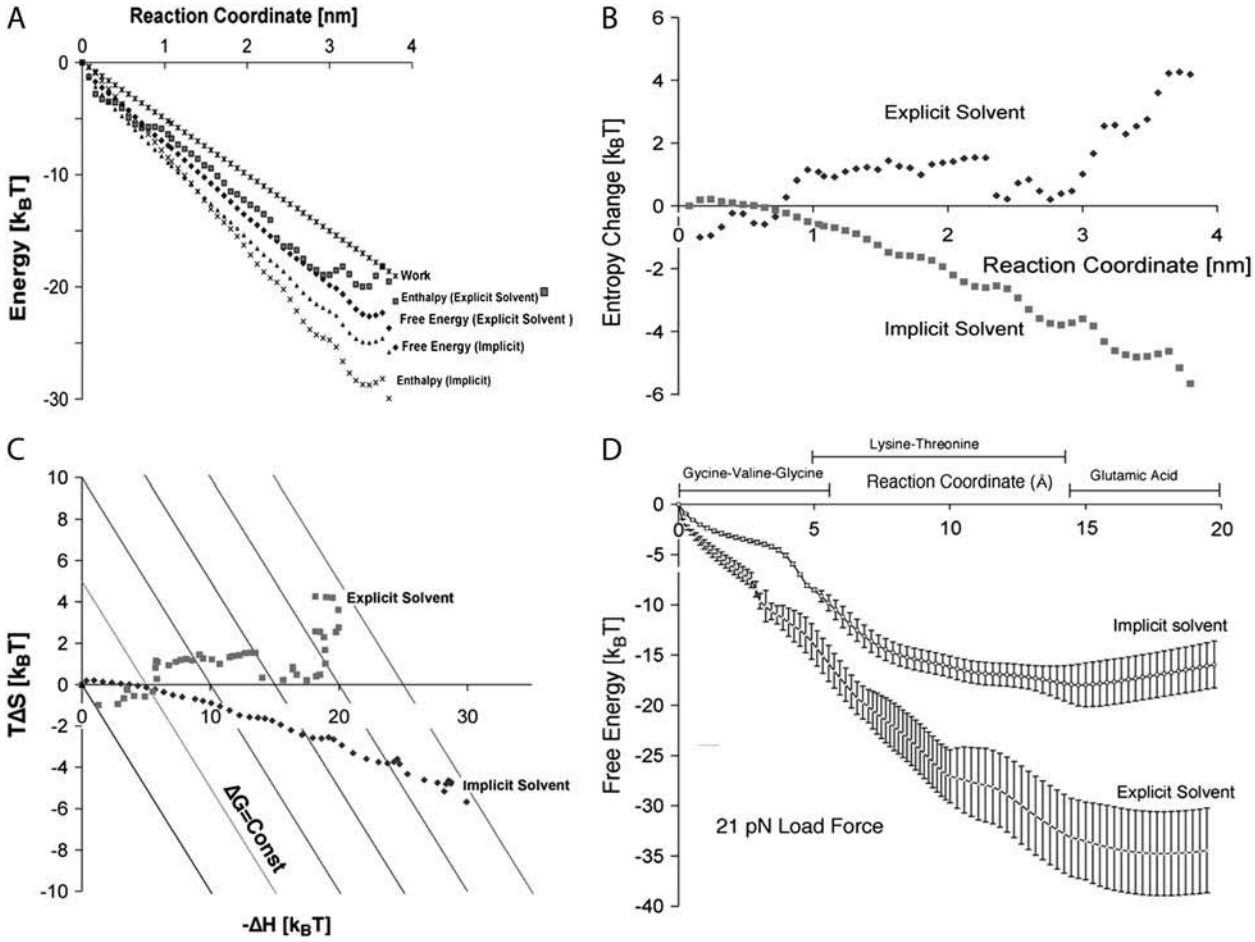


FIGURE 9 Energy profiles. Two-dimensional simulations were created using umbrella-sampling Monte Carlo with a load force of 11 pN and 22 surface sites. Error bars not shown but are the standard deviation in the measurement bins (bins were .38 Å) with a value of 4–5 $k_BT$ for the explicit solvent case and ~1–2 $k_BT$ for the implicit solvent case. (*A*) Two-dimensional free energy and enthalpy profile. The work is the amount of energy stored in the chain as it anneals to the surface against a constant force. (*B*) Two-dimensional entropic profile. The implicit and explicit solvent entropic changes are dramatically different due to the increase in solvent entropy when the water is treated explicitly. (*C*) Two-dimensional entropic versus enthalpic energy. As the polymer moves along the substrate and forms bonds, the explicit solvent model shows a steadily decreasing free energy. (*D*) Three-dimensional implicit and explicit solvent models with a load force of 21 pN. The free energy drop is much greater and smoother with explicit solvent. The two regions before and after lysine have lower free energies because of the entropic gains associated with displacing solvent molecules. Although lysine creates a strong bond with ATP, it does not displace as many solvent molecules, so its entropy change is not affected as much as the other groups. The slope in the lysine region is similar for both explicit and implicit solvent cases.

slope of the lysine region in both the implicit and explicit case is nearly equal because lysine creates a strong bond with ATP, but does not displace as many solvent molecules (small entropic gain). The two regions before and after lysine, have larger free energy drops in the explicit case because of the entropic gains associated with displacing solvent molecules. Due to these interactions, the slopes of these regions are similar to that of lysine leading to a constant slope. The total free energy drop is larger for the explicit case, because, in three dimensions, we are displacing many more solvent molecules and the entropy gained upon displacement is larger compared to the two-dimensional case.

## APPENDIX B: MOTOR FLEXIBILITY

In the soft spring limit $k_{spring} \rightarrow 0$, the cargo trails far behind the motor and it is assumed that the net effect of cargo on the motor is to produce a constant load force, $F$, in the direction opposing the motor. Newton's Third Law of Motion requires that $F$ be sufficient to keep the cargo moving at the same speed as the motor. Thus, the cargo feels a constant force $F$ and moves with a constant velocity,

$$\nu_{kspring \rightarrow 0} = \frac{D}{k_B T} F, \tag{B1}$$

where $\nu_{kspring} \rightarrow 0$ is the velocity of the system as $k_{spring} \rightarrow 0$, $D$ is the diffusion coefficient for the cargo, and $F$ is the force felt by the cargo (or motor in the reverse direction). In our model, we have only a constant force pulling on one side of our chain. By measuring the velocity at a low value of $k_{spring}$, we can solve for $D$ in the above equation to get some effective diffusion if there were a cargo. Our chain is a series of 10 springs so we can approximate it as

$$\frac{1}{k_{spring}^{effective}} = \sum \frac{1}{k_{spring}}. \tag{B2}$$

Converting our units into those used by Peskin and Elston ($L$ now refers to the total distance of the zip) we get the similar quantitative results shown in Fig. 10 *A* for our two-dimensional model (19). To make a more direct comparison we also removed the constant force from our chain and replaced it by a slow moving cargo such as Peskin and Elston used. This noninteracting cargo was attached by a spring to the same segment as the external force was and the diffusion constant was varied (see Fig. 10 *C*). The results are shown in Fig. 10 *B*, again qualitatively very similar to our constant force results and that of Peskin and Elston.

We extended the idea of protein flexibility to our three-dimensional system to determine if flexibility will lead to a more linear free energy drop for the implicit solvent model. To model the rest of the protein body (outside the P-loop) as an elastic cargo, we added a slow diffusing segment attached to the end of the P-loop (analogous to the two-dimensional model in Fig. 10 *C*). The external force now acts on this cargo attempting to pull it back to its starting position. The cargo is attached to the P-loop via a harmonic spring but otherwise can diffuse freely. Our reaction coordinate is now the distance that the cargo has moved against the external force. All other parameters in the system are unchanged.

This arrangement is intended to mimic the relationship between the small catalytic site of a motor and escapement that levers up the force generated at the catalytic site. This larger, and thus slower diffusing, domain outside the catalytic site dramatically increases the zipping time of the P-loop. The elastic properties of the escapement are approximated as a single spring. Some molecular motor systems store energy in the bending of their secondary structures such as the $\beta$-sheet whose loops grasp the nucleotide (27). In this system, the spring corresponds to the $\beta$-sheet. Without the escapement and cargo, the model P-loop completes its power stroke in $\sim$5 $\mu$s, depending on the load force. In a real system, the power stroke takes place in $\sim$100 $\mu$s (37). With a cargo friction constant $\sim$100 times larger than a single backbone segment on the P-loop, the zipping time increased by a factor of $\sim$10. The spring constant between the P-loop and cargo was set to 1.75 $k_BT/Å^2$. For example, with a 21 pN load force without cargo, the P-loop zipped in 4.5 $\pm$ 2.1 $\mu$s; with the cargo the power stroke took 52 $\pm$ 31 $\mu$s. Further increasing the cargo friction constant by a factor of 1000, the zipping time increased $\sim$100 (from 4.5 $\mu$s to 319 $\pm$ 173 $\mu$s), which is within an order of magnitude of the actual time measured in the $F_1$ experiments (37).

Another factor affecting the speed of the power stroke and the size of the free energy change is the strength of the spring between the chain and cargo, $\kappa_{cargo}$. This follows closely what took place in our two-dimensional model (18,19). With a stiff spring, the chain can only zip as quickly as the cargo diffuses, while with a very compliant spring the cargo will take a long time to move against an external force. There is an intermediate between these extremes that produces an optimal zipping speed, but also may smooth out the free energy profile.

In Fig. 10 *D*, we examine the three-dimensional implicit solvent model and show how the free energy depends on $\kappa_{cargo}$. In the implicit solvent model, the free energy profile is not nearly as smooth as the explicit model. We show that the addition and optimization of the cargo and spring constant improves the linearity, and thus efficiency of the zipping process. In the extremely compliant case ($\kappa_{cargo} \ll 1$), the free energy drop is smooth but small as the chain gently pulls against the cargo. As $\kappa_{cargo}$ approaches zero, the free energy drop also approaches zero. As $\kappa_{cargo}$ increases, the free energy drop approaches a maximum, $\kappa_{cargo} > 2\ k_BT/Å^2$. Thus, at large $\kappa_{cargo}$ the free energy drop remains large but will also lead to nonlinearity, while having a small $\kappa_{cargo}$ will lead to a smooth but smaller free energy drop. In the intermediate region, $0.5\ k_BT/Å^2 < \kappa_{cargo} < 1.75\ k_BT/Å^2$, there is tradeoff such that the cargo spring is soft enough to create a relatively smooth free energy profile but stiff enough to maintain a large free energy drop. Also, note that with the cargo, the free energy drop maximum is greater than without the cargo. This is because there is no bending modulus between the cargo and P-loop. We are measuring the reaction coordinate from the cargo and the cargo now has more entropy. If we add a bending modulus to the cargo, we get nearly the same results as without the cargo (slight differences because the effective length of the chain has changed).

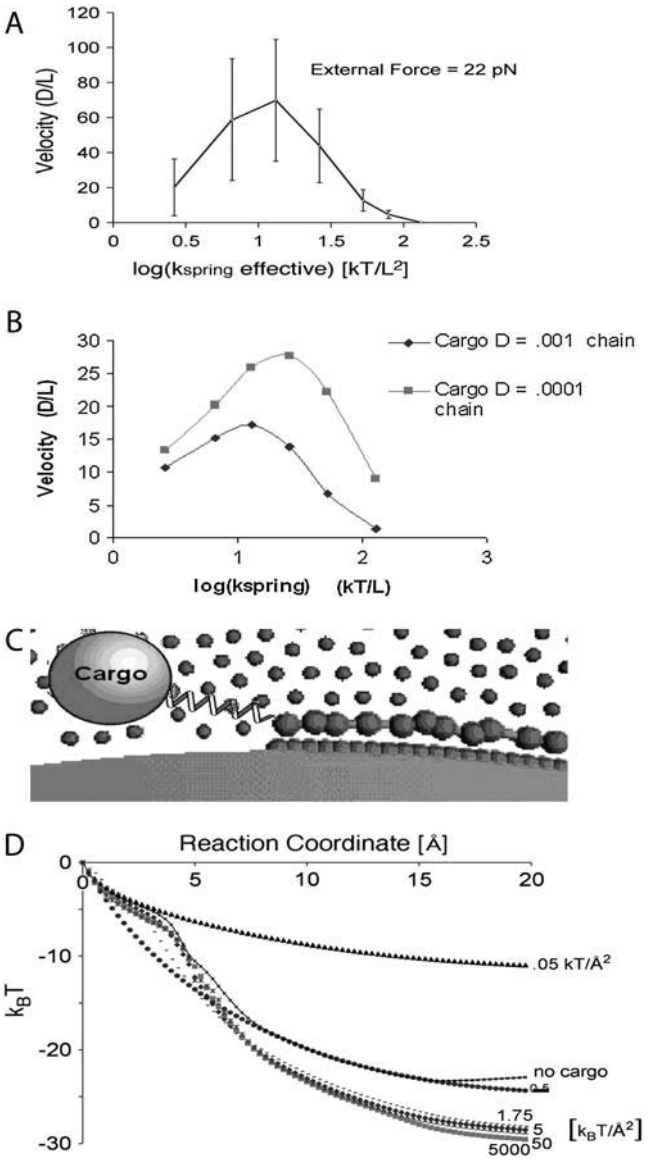


FIGURE 10 (*A*) Velocity dependence on $k_{spring\ effective}$ now in units used by Peskin and Elston in the two-dimensional model. Notice a maximum velocity. The potential we used differs from that of Peskin and Elston; therefore, quantitatively, the results differ, but qualitatively they are similar. The value $k_{spring\ effective}$ is calculated as an effective $k_{spring}$ using springs in series as described in Eq. A2. (*B*) The constant force in our original two-dimensional model is replaced by a slow moving cargo similar to Peskin and Elston. The potential we used and that of Peskin and Elston differ so the velocity differs, but they are qualitatively similar. The value $k_{spring}$ is now an effective spring constant using springs in series as described by Eq. B2. (*C*) Cartoon of model with cargo attached to P-loop via an harmonic spring. All other parameters are unchanged in the system. Cargo has a diffusion constant of $D$ and the distance the cargo travels is now the reaction coordinate. (*D*) Free energy with and without a load (three-dimensional implicit solvent with no external force). The spring constant $k_{cargo}$, connects the P-loop and cargo. Error bars are not shown, but they are $\sim$12 $k_BT$.

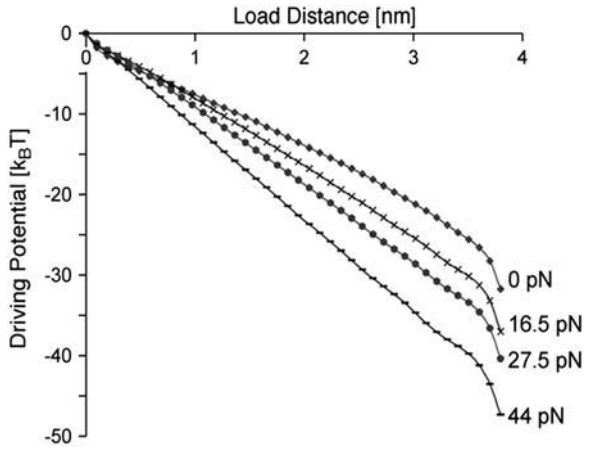


FIGURE 11 Effective driving potential (EDP) for the 22-surface-site case. Notice that the potential becomes more linear as the load force is increased.

## APPENDIX C: EFFECTIVE DRIVING POTENTIAL

Another method of characterizing the binding process is the effective driving potential (EDP) (1). The more linear the EDP is along the reaction coordinate, the higher the efficiency and power-stroke characteristics of the motor. Following Wang and Oster's method (1) for constructing the EDP, we have determined the EDP for various loads with the 22-surface-site case in Fig. 11 (other surface-site case results were similar). Again, we see that our model is making a smooth zip along the surface characteristic of a power stroke. At large forces, the chain is being pulled from both ends, tending to dampen out fluctuations and leading to a lower and more linear potential (similar to increasing $\kappa$). At smaller force, the chain is more flexible, so many fluctuations are wasted, leading to a lower and nonlinear potential. Protein loops in catalytic sites may have evolved to be stiffer to improve the efficiency of the enzyme.

## SUPPLEMENTARY MATERIAL

An online supplement to this article can be found by visiting BJ Online at http://www.biophysj.org.

The authors thank Jung-Chi Liao for valuable input during the preparation of the manuscript.

J.E. and A.C. were supported by the National Science Foundation and Defense Advanced Research Projects Agency. G.O. was supported by National Institutes of Health grant No. GM59875-02.

## REFERENCES

1. Wang, H., and G. Oster. 2001. Ratchets, power strokes, and molecular motors. *Appl. Phys. A.* 75:315–323.
2. Mogilner, A., and G. Oster. 2003. Polymer motors: pushing out the front and pulling out the back. *Curr. Biol.* 13:R721–R733.
3. Oster, G., and H. Wang. 2000. Why is the efficiency of the $F_1$ ATPase so high? *J. Bioenerg. Biomembr.* 332:459–469.
4. Dimroth, P., H. Wang, M. Grabe, and G. Oster. 1999. Energy transduction in the sodium $F_1$-ATPase of *Propionigenium modestum*. *Proc. Natl. Acad. Sci. USA.* 96:4924–4929.
5. Xing, J., H. Wang, P. Dimroth, C. von Balmoos, and G. Oster. 2004. Torque generation by the sodium $F_0$-ATPase. *Biophys. J.* 87:2148–2163.
6. Ye, J., A. R. Osborne, M. Groll, and T. A. Rapoport. 2004. RecA-like motor ATPases—lessons from structures. *Biochim. Biophys. Acta.* 1659:1–18.
7. Vale, R., and R. Milligan. 2000. The way things move: looking under the hood of molecular motor proteins. *Science.* 288:88–95.
8. Howard, J. 2001. Mechanics of Motor Proteins and the Cytoskeleton. Sinauer, Sunderland, MA.
9. Wang, H., and G. Oster. 2002. The Stokes efficiency for molecular motors and its applications. *Europhys. Lett.* 57:134–140.
10. Bockmann, R. A., and H. Grubmuller. 2003. Conformational dynamics of the $F_1$-ATPase $\beta$-subunit: a molecular dynamics study. *Biophys. J.* 85:1482–1491.
11. Bockmann, R., and H. Grubmuller. 2002. Nanoseconds molecular dynamics simulation of primary mechanical energy transfer steps in $F_1$-ATP synthase. *Nat. Struct. Biol.* 9:198–202.
12. Ogura, T., S. Whiteheart, and A. Wilkinson. 2004. Conserved arginine residues implicated in ATP hydrolysis, nucleotide-sensing, and inter-subunit interactions in AAA and $AAA^+$ ATPases. *J. Struct. Biol.* 146: 106–112.
13. Yang, W., Y. Q. Gao, Q. Cui, J. Ma, and M. Karplus. 2003. The missing link between thermodynamics and structure in $F_1$-ATPase. *Proc. Natl. Acad. Sci. USA.* 100:874–879.
14. Minehardt, T., N. Marzari, R. Cooke, P. Edward, P. A. Kollman, and R. Car. 2002. A classical and ab initio study of the interaction of the myosin triphosphate binding domain with ATP. *Biophys. J.* 82:660–675.
15. Scheiner, S., T. Kar, and Y. Gu. 2001. Strength of the CaH-O Hydrogen bond of amino acid residues. *J. Biol. Chem.* 276:9832–9837.
16. Oster, G., and H. Wang. 2000. Reverse engineering a protein: the mechanochemistry of ATP synthase. *Biochim. Biophys. Acta.* 1458:482–510.
17. Johnson, J. K., J. A. Zollweg, and K. E. Gubbins. 1993. The Lennard-Jones equation revisited. *Mol. Phys.* 78:591–618.
18. Peskin, C., D. You, and T. Elston. 2000. Protein flexibility and the correlation ratchet. *J. Appl. Math.* 61:776–791.
19. Peskin, C., and T. Elston. 2000. The role of protein flexibility in molecular motor function: coupled diffusion in a tilted periodic potential. *J. Appl. Math.* 60:842–867.
20. Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. 1953. Equation of state calculations by fast computing machines. *J. Chem. Phys.* 21:1087–1092.
21. Koch, S., R. Desai, and F. Abraham. 1983. Dynamics of phase separation in two-dimensional fluids: spinodal decomposition. *Phys. Rev. A.* 27:2152–2167.
22. Brooks, C. L., and M. Karplus. 1983. Deformable stochastic boundaries in molecular dynamics. *J. Chem. Phys.* 79:6312–6325.
23. Hunt, A. J., and A. P. Joglekar. 2002. A simple, mechanistic model for directional instability during mitotic chromosome movements. *Biophys. J.* 83:42–58.
24. Visscher, K., M. Schnizter, and S. Block. 1999. Single kinesin molecules studied with a molecular force clamp. *Nature.* 400:184–189.
25. Kato-Yamada, Y., H. Noji, R. Yasuda, K. J. Kinosita, and M. Yoshida. 1998. Direct observation of the rotation of $\varepsilon$-subunit in $F_1$-ATPase. *J. Biol. Chem.* 273:19375–19377.
26. Noji, H., R. Yasuda, M. Yoshida, and K. Kinosita. 1997. Direct observation of the rotation of $F_1$-ATPase. *Nature.* 386:299–302.
27. Sun, S., D. Chandler, A. R. Dinner, and G. Oster. 2003. Elastic energy storage in $\beta$-sheets with application to $F_1$-ATPase. *Eur. J. Biophys.* 32:676–683.
28. Wang, H., and G. Oster. 1998. Energy transduction in the $F_1$ motor of ATP synthase. *Nature.* 396:279–282.
29. Antes, I., D. Chandler, H. Wang, and G. Oster. 2003. The unbinding of ATP from $F_1$ ATPase. *Biophys. J.* 85:695–706.
30. Senior, A. E., S. Nadanaciva, and J. Weber. 2002. The molecular mechanism of ATP synthesis by $F_1F_0$-ATP synthase. *Biochim. Biophys. Acta.* 1553:188–211.

31. Futai, M., H. Omote, Y. Sambongi, and Y. Wada. 2000. ATP synthase ($H^+$ ATPase): coupling between catalysis, mechanical work, and proton translocation. *Biochim. Biophys. Acta.* 1458:276–288.

32. Senior, A. E. 1992. Catalytic sites of *Escherichia coli* $F_1$-ATPase. *J. Bioenerg. Biomembr.* 24:479–484.

33. Massova, M., and P. Kollman. 1999. Computational alanine scanning to probe protein-protein interactions: a novel approach to evaluate binding free energies. *J. Am. Chem. Soc.* 121:8133–8143.

34. Gao, Y. Q., W. Yang, R. A. Marcus, and M. Karplus. 2003. A model for the cooperative free energy transduction and kinetics of ATP hydrolysis by $F_1$-ATPase. *Proc. Natl. Acad. Sci. USA.* 100:11339–11344.

35. Sun, S., H. Wang, and G. Oster. 2004. Asymmetry in the $F_1$-ATPase and its implications for the rotational cycle. *Biophys. J.* 86:1373–1384.

36. Chakraborty, A. K. 2000. Disordered heteropolymers: models for biomimetic polymers and polymers with frustrating quenched disorder. *Phys. Rep.* 342:1–61.

37. Yasuda, R., H. Noji, M. Yoshida, K. Kinosita, and H. Itoh. 2001. Resolution of distinct rotational substeps by submillisecond kinetic analysis of $F_1$-ATPase. *Nature.* 410:898–904.