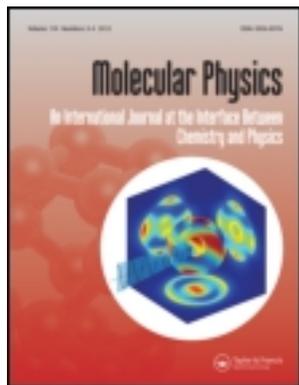


This article was downloaded by: [George Washington University], [Ganhui Lan]

On: 19 December 2012, At: 12:40

Publisher: Taylor & Francis

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Molecular Physics: An International Journal at the Interface Between Chemistry and Physics

Publication details, including instructions for authors and subscription information:

<http://www.tandfonline.com/loi/tmph20>

Mechanochemical models of processive molecular motors

Ganhui Lan^a & Sean X. Sun^b

^a IBM T. J. Watson Research Center, Yorktown Heights, New Jersey 22122, USA

^b Department of Mechanical Engineering and Whitaker Institute of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland 21218, USA

Accepted author version posted online: 10 Apr 2012. Version of record first published: 30 Apr 2012.

To cite this article: Ganhui Lan & Sean X. Sun (2012): Mechanochemical models of processive molecular motors, *Molecular Physics: An International Journal at the Interface Between Chemistry and Physics*, 110:9-10, 1017-1034

To link to this article: <http://dx.doi.org/10.1080/00268976.2012.677863>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.tandfonline.com/page/terms-and-conditions>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

INVITED ARTICLE

Mechanochemical models of processive molecular motors

Ganhui Lan^a and Sean X. Sun^{b*}

^aIBM T. J. Watson Research Center, Yorktown Heights, New Jersey 22122, USA; ^bDepartment of Mechanical Engineering and Whitaker Institute of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland 21218, USA

(Received 8 February 2012; final version received 9 March 2012)

Motor proteins are the molecular engines powering the living cell. These nanometre-sized molecules convert chemical energy, both enthalpic and entropic, into useful mechanical work. High resolution single molecule experiments can now observe motor protein movement with increasing precision. The emerging data must be combined with structural and kinetic measurements to develop a quantitative mechanism. This article describes a modelling framework where quantitative understanding of motor behaviour can be developed based on the protein structure. The framework is applied to myosin motors, with emphasis on how synchrony between motor domains give rise to processive unidirectional movement. The modelling approach shows that the elasticity of protein domains are important in regulating motor function. Simple models of protein domain elasticity are presented. The framework can be generalized to other motor systems, or an ensemble of motors such as muscle contraction. Indeed, for hundreds of myosins, our framework can be reduced to the Huxely–Simmons description of muscle movement in the mean-field limit.

Keywords: molecular motors; myosin movement; stochastic models; coarse-grained modelling

1. Introduction

A basic paradigm of structure biology is that a protein's three-dimensional structure gives rise to its function. Establishing a quantitative relationship between structure and function requires quantitative models of protein conformational dynamics and kinetics. For this purpose, computational modelling based on molecular dynamics (MD) simulations is an essential tool [1–3]. For a class of large protein complexes such as molecular motors, however, operating time scales of these macromolecules are currently beyond the capability of MD methods. At the same time, single molecule techniques are revealing exquisitely detailed information about the motor mechanism [4–7]. Thus, there is a need for models that can quantitatively explain experimental data. In addition, essential motor protein functions have been shown to be insensitive to mutations of amino-acid residues away from the catalytic site [8]. Thus, as long as the hydrolysis chemistry is preserved, the operating principles of molecular motors are emergent properties of the protein structure, and should be independent of the precise atomistic details. An understanding of these emergent principles is the subject of this paper.

Because motor proteins convert chemical energy to mechanical work, analogies to mechanical engines can be invoked [9,10]. Indeed, the F_1 -ATPase roughly

resembles a three-chamber rotary engine [11]. Concepts such as force–velocity relations, duty ratio and operating efficiency are equally applicable to biological motors as well as man-made engines. From a theoretical stand point, descriptions have been often less transparent. Concepts and approaches such as ratchets [12–15], power-strokes [16–18], Markov models [19–22] have been used to explain single molecule data (also summarized in [23]). The foundation of these theoretical approaches, however, must rest with the molecular structure. We will show that these approaches are limits of a structure-based model. Just as when discussing mechanical engines, configurations of cylinders, pistons, and forces between moving parts are quantities of interest, a mechanical theory of molecular motors must focus on the roles of motor subunits and how forces are transmitted between subunits. Quantitative relationships between single motor properties and the dynamics of the motor complex must be explicitly shown. In this article, we discuss a general framework where the free energy of dimeric processive molecular motors is modelled using simple mechanical concepts. Influence of geometrical shapes of the protein subunits is emphasized. From the free energy landscape, conformational dynamics of the motor domain and forces transmitted between motor domains are analysed. Approaches such as Markov

*Corresponding author. Email: ssun@jhu.edu

models, ratchets and power-strokes are various limits of a description based on the free energy landscape.

Processive molecular motors are responsible for cytoplasmic transport of vesicles and organelles in eukaryotic cells [24]. Isoforms of kinesin and dynein are cytoplasmic transporters operating on microtubules [25–29]. Isoforms of myosin are actin-associated molecular motors [30–32]. Exquisitely precise single molecular data have been obtained for both microtubule and actin motors [8,33–42]. The bewildering variety of molecular motors may appear complex, however, we propose that a unified framework exists to explain processive motor dynamics, provided that a set of mechanical and kinetic properties of the protein subunits are understood. Using dimeric myosins as examples, we show how the mechanics of the motor domains can regulate enzymatic activity and lead to processive motion. Quantitative comparisons from the mechanical model with single molecule data are made for the processive myosin system.

Developing a general framework of motor dynamics has important implications for understanding ensembles of motors inside cells. In eukaryotic cells, motors interact with networks of cytoskeletal filaments [43]. We show that by changing a few geometric parameters, our framework can be extended to describe motor motion across cytoskeletal junctions [44], motions of myosin isoforms and mutants [8,33,45–47]. The approach can also be used to describe several motors interacting with the same cargo, or a tug of war between motors travelling in opposite directions [48–52]. Thus, a range of interesting phenomena are contained in one theoretical approach.

In the next section, an energy landscape description of single myosin motors is discussed. The approach of describing single myosins is reminiscent of the framework developed by Eisenberg and Hill [53]. We then describe the motor dimer energy landscape using a simple rod model of the linker domain. Given the energy landscape, we describe how the mechanical structure of the motor gives rise to processive motion. In Section 5, the energy landscape picture is combined with stochastic dynamics to compute the expected motor motion. These results are expanded to understand motor dynamics under forces greater than stall force and motor efficiency. Finally, the same framework is applied to motors interacting with each other and walking on cytoskeletal networks.

2. Energy landscape description of a single protein motor

Motor proteins typically are comprised of several subunits. Each subunit contains a single ATP

hydrolysis site. Coordinated action between the subunits generate mechanical work [54–56]. For instance, myosin V and VI are dimeric actin-based motors with two myosin subunits; kinesin is also a dimeric motor operating on microtubules; ATP synthase is a rotary machine with three subunits. Before discussing how the motor subunits are coupled, it is useful to consider how a single motor subunit must operate. We will use the myosin monomer as an example where there is ample structural and biochemical data.

2.1. Lever-arm movement of myosin

From extensive studies on muscle movement, it is clear that muscle myosin undergoes a conformational transition where the converter domain rotates with respect to the actin-binding domain. This leads to a rotation of the light chain domain (S1) with respect to the actin binding domain (Figure 1) [57]. X-ray structures of scallop myosin II show that the rotation occurs in conjunction with the release of Pi after hydrolysis of ATP [58,59]. The structures are solved without F-actin, although it is believed that a similar set of conformational changes occur while myosin is bound to actin [58]. Other isoforms such as myosin V also show a similar set of structural transitions. An anomaly is the myosin VI isoform which appears to reverse the direction of the power-stroke [60–62]. Structural studies revealed that there is a ~ 50 residue insert between the converter domain and the light chain domain that changes the rotational direction [63,64].

In conjunction with the ATP cycle, myosin also binds F-actin, although actin affinity depends on the identity of the nucleotide in the catalytic site. F-actin binding also influences the reaction rates in the catalytic site. From extensive kinetic studies, de la Cruz *et al.* has investigated the reaction rates of myosin V monomers interacting with actin [65]. A fairly complete picture of the myosin monomer cycle is shown in Figure 1(c). Note that these reaction rates are obtained *in vitro* where myosin can make conformational changes without mechanical hindrance.

2.2. Mechanical description of lever-arm movement

Given the available structural and kinetic information, it is possible to consider a mechanochemical description of the myosin monomer based on an energy landscape picture. We can describe the conformation of the monomer using variables (θ, ϕ) where these angles are defined with respect to the monomer body axis and the F-actin plane, and are shown in Figure 2. (Figure 1(a) shows θ and θ_0 only. To fully specify the

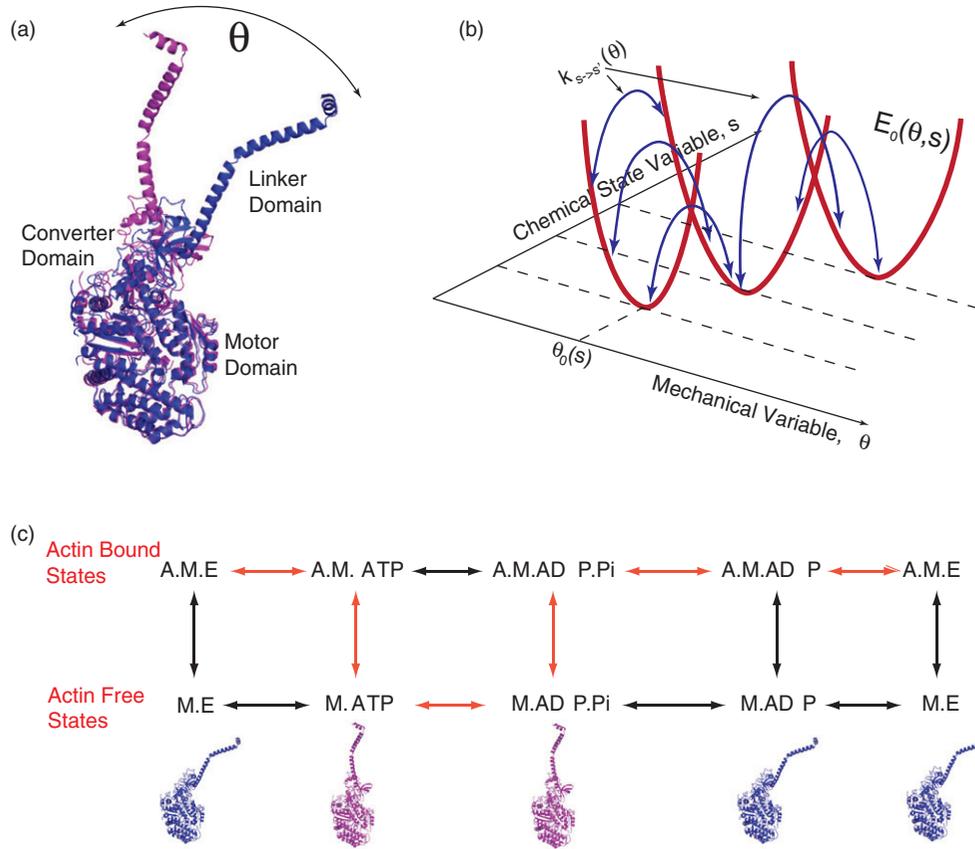


Figure 1. Mechanical model of a single myosin motor. (a) Myosin motors undergo a conformation change (lever-arm movement) upon hydrolysing ATP. The structures shown are the APO and ADP states of scallop myosin II. (b) The free energy of a single motor domain can be visualized as a landscape involving at least two variables, a mechanical variable describing the motor conformation, and a chemical variable describing the hydrolysis and actin-binding states. Chemical transition rate functions, $k_{s \rightarrow s'}$, depend on the motor conformation. If the motor conformation is fixed (θ remains constant), hydrolysis of ATP still proceeds, but with different rates. (c) A typical myosin motor hydrolysing ATP as well as binding and unbinding to actin. The approximate myosin conformation is shown. The top row are actin-bound states and the bottom row are actin-free states. Approximately eight chemical states are considered. The red pathway is the most probably the pathway, although the path probabilities are ultimately determined by the kinetic parameters. The transition rates going from conformational equilibrium have been measured.

orientation of the linker domain, θ and ϕ are needed.) The preferred conformation, (θ_0, ϕ_0) , depends on the nucleotide in the binding pocket and whether myosin is bound to actin. The conformations of X-ray structures are presumed to be the preferred conformation.

The chemical state of the monomer is described by s which specifies the nucleotide occupancy as well as whether myosin is bound to actin. s can be any of the eight possible values such as (A.M.ATP) which means that myosin is bound to F-actin with ATP in the binding site. The simplest form of the motor energy as a function of its conformation is then

$$E_0(\theta, \phi, s) = \frac{1}{2}\kappa_1(s)[\theta - \theta_0(s)]^2 + \frac{1}{2}\kappa_2(s)\phi^2 + c(s), \quad (1)$$

where $\kappa_{1,2}(s)$ are mechanical constants of the protein structure which may depend on the chemical state. $\kappa_{1,2}$ describes the stiffness of the myosin monomer and governs how the conformation fluctuates. The numerical values of κ are not known, but estimates can be obtained by fitting to experimental data or from MD simulations. $c(s)$ is a constant that defines the energy difference between the equilibrium structures. $c(s)$ can be obtained from kinetic data as described in the next paragraph. Equation (1) takes the simplest approach and assumes a linear elastic model for the monomer, although nonlinear models are also possible. For example, recent experiments on kinesin seems to suggest a more complex elastic model [66].

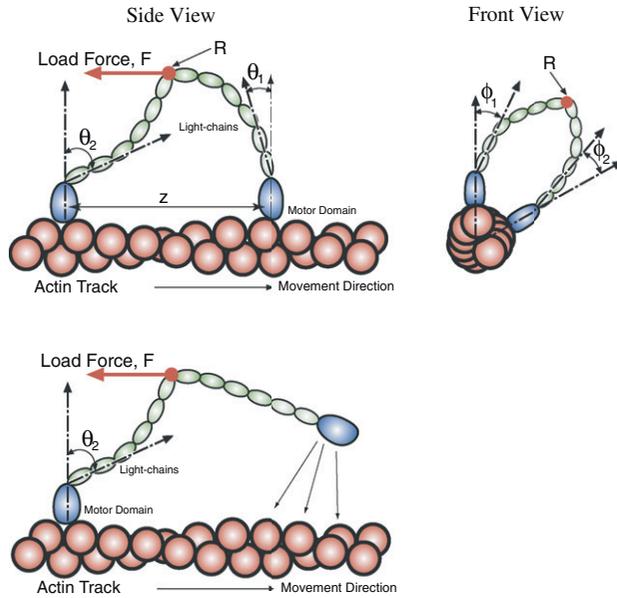


Figure 2. The total free energy landscape of the motor dimer can be obtained using a simple rod-like model for the linker domains. When both domains are bound to the track, the linker domain is uniquely defined by specifying the motor domain conformations $(\theta_1, \phi_1, \theta_2, \phi_2)$, and the distance between bound sites, z , and the force \mathbf{F} applied at position \mathbf{R} . When only one of the motors is bound, only one of the linkers bears the force. The free motor can bind to any available site, but the binding probability depends on the energy difference before and after binding.

2.3. Chemical transition rate functions

Transition rates between different chemical states are denoted as $k_{s \rightarrow s'}$. It is important to note that the transition rate is a function of the conformation of the myosin molecule. Imagine that we constrain the conformation of myosin by holding the light-chain domain fixed. This constraint in fact does not stop the hydrolysis cycle and ATP is still converted to ADP. However, the rates of ATP hydrolysis, and potentially actin-binding will depend on the conformation of the molecule [67,68]. Indeed, enzymatic activities in general are conformationally dependent. Thus, instead of rate constants, we must define rate functions for each of the chemical transitions. In our modelling, we assume that the chemical rate functions only depend on θ , i.e. $k_{s \rightarrow s'}(\theta)$ (Figure 3 shows an example of conformation dependent ADP release rate function). The measured rate constants from de la Cruz *et al.* represent the net rate of going from the preferred conformation $\theta_0(s)$ to $\theta_0(s')$. In those experiments, conformational relaxation should be much faster than chemical conversion in the binding pocket. Thus, those rates are approximately equivalent to $k_{s \rightarrow s'}(\theta_0(s))$.

From the condition of detailed balance, the transition rates must satisfy the constraint:

$$\frac{k_{s \rightarrow s'}(\theta)}{k_{s' \rightarrow s}(\theta)} = \exp[-\beta(E_0(\theta, \phi_0, s) - E_0(\theta, \phi_0, s'))] \quad (2)$$

where $\beta = 1/k_B T$ and k_B is the Boltzmann. This constraint arises from our basic assumption that the underlying free energy landscape of the protein is a

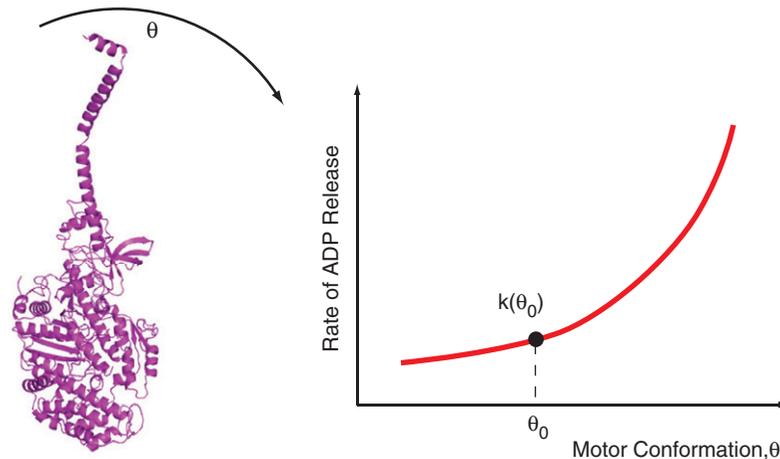


Figure 3. Enzymatic activity of the motor domain is conformation dependent. Here, ADP release rate should increase as the linker domain rotates past the preferred conformation (plot at right). Kinetics measurements using purified proteins measure the reaction rate at conformational equilibrium: $k_{s \rightarrow s'}(\theta_0)$. Conformation of the molecule can be changed by applying a force to the linker domain, and experiments have measured some reaction rate functions.

continuous single valued function. Together with the monomer energies of Equation (1), the transition rate functions therefore specifies a free energy landscape of the form shown in Figure 1(b). In some cases, transition rate functions have been measured for myosin monomers by applying forces to the molecule [67,68]. Using theoretical modelling, it is possible to relate the mechanical properties of the monomer and the conformationally dependent transition rate functions to single molecule data (see below).

2.4. Induced fit model of enzymatic dynamics

Note that our specification of the monomer energy in Equation (1) has a connection with the induced-fit picture of enzyme allostery. Upon binding of ATP and release of hydrolysis products, the enzyme changes shape. This is reflected in the fact that θ_0 is a function of s . For each s , we postulate that there is an unique low energy conformation, (θ_0, ϕ_0) . Within this picture, ‘power-stroke’ of the motor occurs after the nucleotide has changed in the binding pocket, and the protein relaxes to a new lower energy conformation. If there are forces and torques on the motor, then these forces will oppose the conformational relaxation (power-stroke), and when force balance is achieved, a new conformational equilibrium different from θ_0 will occur. Thus, a consequence of forces on the motor is that it alters the conformation and therefore the transition rate $k_{s \rightarrow s'}(\theta)$. This result will explain synchrony between hydrolysis sites and lead to coordinated motion of the motor dimer (see next section).

2.5. Simulations of the energy landscape

Finally, $E_0(\theta, \phi, s)$ can be obtained from molecular dynamics simulations by computing the free energy of the molecule as a function of its conformation. Equation (1) postulates that the motor behaves as torsional springs parameterized by spring constants κ_1 and κ_2 , this postulate can be checked by more careful simulations of the enzyme system to estimate the spring constants.

3. Motor dimers and the elasticity of protein domains

Processive motors are typically dimers where two motor domains are connected by a linker domain [54]. For myosin dimers such as myosin V, VI and X, the motor domains undergo a similar conformational transition, but the linker domains have different structure and composition [31,32]. A fundamental question is: what is the connection between linker

domain structure and the observed processive motion? The developed model should explain not only the processive behaviour, but also the response of the motor to applied forces in single molecule experiments.

3.1. Energy of motor dimer

Within our energy landscape framework, the total energy of the actin-bound dimer can be expressed as

$$E = E_0(\theta_1, \phi_1, s_1) + E_0(\theta_2, \phi_2, s_2) + E_L(\theta_1, \theta_2, \phi_1, \phi_2, z, \mathbf{F}), \quad (3)$$

where $(\theta_1, \phi_1, \theta_2, \phi_2)$ are the conformations of the motor domains 1 and 2; E_L is the free energy of the linker domain which can depend on the conformations of the motor domains and the externally applied force \mathbf{F} . In single molecule experiments, forces from optical traps are typically applied to the trapped bead that is tethered to the linker domain (see Figure 1). Finally, when both motor domains are bound to actin, z is the distance between the actin binding sites. We assume that the conformation of the linker domain is in rapid thermal equilibrium, i.e. given the conformation of the motor domains $(\theta_1, \theta_2, \phi_1, \phi_2)$, the linker domain will reach an equilibrium conformation instantaneously. The equilibrium conformation of the linker domain and its free energy will be computed using a mechanical model below.

3.2. Force transmission between motor domains

In our specification of the linker domain free energy, we have assumed that the linker conformation is not affected by chemical state of the motor domain. The linker domain is a passive mechanical element. However, the linker domain is an elastic element, and can transmit forces between the motor domains. If one of the motor domains makes a conformational change by changing s (and therefore $\theta_0(s)$), E_0 of that domain changes. The new configuration of the dimer is obtained by force balance between the motor domains and the linker domain, which can be computed by minimizing the total energy:

$$\left. \frac{\partial E}{\partial \theta_i} \right|_{\theta_i^*, \phi_i^*} = 0, \quad \left. \frac{\partial E}{\partial \phi_i} \right|_{\theta_i^*, \phi_i^*} = 0, \quad (4)$$

where $(\theta_1^*, \theta_2^*, \phi_1^*, \phi_2^*)$ are conformations of the motor domain that minimizes the total motor energy. These conformations are clearly different from the monomer preferred conformations (θ_0, ϕ_0) due to the presence of the linker domain energy E_L . Note that Equation (4) assumes that the conformation of the linker domain

relaxes to mechanical equilibrium faster than chemical transitions in the motors. We expect this to be the case when ATP hydrolysis rates are relatively slow (> 1 ms). This assumption should be checked when these time scales are comparable.

3.3. Rod-like model of the linker domain

The conformational free energy of the linker domain, E_L , is not known, but simple models may be able to capture the essential physics. For myosin V, the link domain is a single long α -helix decorated by calmodulins [8]. The domain is roughly 30 nm in length, and is structurally similar to a rod. Thus, we postulate that the conformational energy can be modelled by considering bending and twisting motions of the linker domain. The mechanical energy of a continuous rod within a linear elastic model is

$$E_L = k_B T \int_0^L \frac{1}{2} \sum_{i=1}^3 l_p \omega_i(s)^2 ds + k_B T \int_0^L \frac{1}{2} \sum_{i=1}^3 l'_p \omega'_i(s)^2 ds - \mathbf{F} \cdot \mathbf{R}, \quad (5)$$

where \mathbf{R} is the position at the middle of the linker domain where an external force \mathbf{F} such as from an optical trap, is applied. $\omega_i(s)$ and $\omega'_i(s)$ are the bending and twist densities of the two linker domain segments, which are in units of radians per unit length. L is the length of the linker. l_p and l'_p are the bending and twist persistence lengths of the rod, respectively. These quantities are related to the bending and twisting modulus, which are the *effective* mechanical constants that describe the energies associated with mechanical deformations. For isotropic and homogeneous elastic bodies, bending and twisting moduli are further related to the Young's modulus of the material. Here, we resist the temptation of using a Young's modulus to describe protein domains, which are neither homogeneous or isotropic.

The configuration of the rod cannot be solved unless we specify the forces at the boundaries (ends) of the rod. These boundary conditions are described by variables (θ_1, ϕ_1) and (θ_2, ϕ_2) , and the positions of the motor domains on F-actin. (The structure of F-actin is known [69], and the separation distance z , is sufficient to specify the positions of the bound domains.) Thus, within the linear elastic model, given boundary conditions and the external force, it is possible to solve for the mechanical equilibrium rod configuration by minimizing the elastic energy of Equation (5). Techniques for solving the configuration are found in standard mechanics textbooks. It is clear that there is a unique rod configuration for each set of boundary conditions

and external force. The elastic energy of the mechanical equilibrium configuration is used for $E_L(\theta_1, \theta_2, \phi_1, \phi_2, z, \mathbf{F})$ in Equation (3).

A case of importance is when only one of motor domain is bound to actin; the other domain is unbound. Here, the external force acts on the bound motor domain only and bends its linker domain. The other domain is presumably free to diffuse. Therefore, the total energy of the linker domains involves only the bound linker:

$$E'_L = k_B T \int_0^L \frac{1}{2} \sum_{i=1}^3 l_p \omega_i(s)^2 ds - \mathbf{F} \cdot \mathbf{R}. \quad (6)$$

The rod model requires stiffness parameters l_p and l'_p , which are ultimately results of linker domain structure. In wild-type myosin V, the linker domain is a single α -helix decorated by six calmodulin subunits. In myosin VI, the structure of the linker domain is still being debated [70,71], but it is clear that at least regions of the linker domain is disordered and unfolding. Thus, we expect the myosin V linker domain to be stiffer, with larger l_p and l'_p values.

Note that the rod-like behaviour of the linker domain is an emergent property of any structure where one of its dimensions is much larger than the other two. Rod models have been used for polymers, DNA, and even small organic molecules [72]. The mechanical behaviour of a rod not only depends on its stiffness, but also on its overall length, L . Therefore, even though Equation (6) is linear in ω_i , the overall motor responds nonlinearly to the applied force \mathbf{F} .

4. Origin of motor processivity

Having specified the total energy of the motor dimer, and the transition rate functions for ATP hydrolysis and actin binding, it is possible to mathematically solve for the motor processive motion using standard methods of stochastic dynamics. Before doing so, it is useful to examine the physical factors leading to coordinated stepping of motor dimers. Two crucial steps are needed for processivity: (1) preferential detachment of the trailing motor domain, and (2) binding of the free motor domain in the forward direction. The overall mechanics of the motor dimer regulate both of these steps.

When both motor domains are bound to actin, the most common configuration is where both motor catalytic sites are occupied by ADP [73,74]. This is because ADP release is kinetically the slowest step in the hydrolysis cycle [65,68]. Since the catalytic site occupancy determines the preferred conformation of the motor domain [58–61], (θ_0, ϕ_0) in Equation (1),

both light-chain domains would like to point in the same direction. However, because the light-chain domains are physically connected, it implies that the mechanically favourable configuration is obtained from Equation (4) by minimizing the total dimer energy. The mechanical equilibrium configurations, $(\theta_1^*, \phi_1^*, \theta_2^*, \phi_2^*)$, are different from the preferred configuration, and $\theta_1^* \neq \theta_2^*$, $\phi_1^* \neq \phi_2^*$. Since all kinetic steps are described by rate functions, ADP release rate in the leading motor domain must be different from the trailing motor domain. To achieve processivity, the trailing motor domain, most of the time, must release ADP first.

Once ADP is released from the trailing motor domain, ATP binding is fast and the trailing head detaches quickly from actin [65]. To make a forward step, the unbound motor domain must preferentially bind to the available actin site *in front* of the bound motor domain. The relative probability of binding to the different binding sites are determined again by the total energy of the motor dimer. We assume that the joint between the light-chain domains is free to rotate, then the unbound monomer can diffusively sample the 30 or so available actin binding sites around the bound motor. The spatial positions of the available binding sites are uniquely determined by z . The rates of binding to the available sites, however, are determined by the energy difference before and after binding:

$$\Delta E(z) = \underbrace{[E_0(\theta_1, \phi_1, s_1) + E_0(\theta_2, \phi_2, s_2) + E_L(\theta_1, \phi_1, \theta_2, \phi_2, \mathbf{F}, z)]}_{\text{energy after binding}} - \underbrace{[E_0(\theta_1, \phi_1, s_1) + E_0(\theta_2, \phi_2, s_2') + E_L'(\theta_1, \phi_1, \mathbf{F})]}_{\text{energy before binding}}, \quad (7)$$

where s_1 is typically A.M.D and s_1 is usually M.DP. This energy difference is also the free energy change for the binding reaction, and we see that the elasticity of the light-chains affect the binding energy; depending on z , binding to some sites are favourable ($\Delta E < 0$) and other sites are unfavourable ($\Delta E > 0$). External forces can also affect this binding energy change.

The rate of binding to the available sites will depend on the energy of the transition state of the reaction. A precise understanding of the binding reaction transition state is presently unavailable. However, a phenomenological approach is to write the binding rate as a function of $\Delta E(z)$:

$$k_{s_2 \rightarrow s_2'}(z) = k_{s_2 \rightarrow s_2'}^0 e^{-\lambda \Delta E(z)} \quad (8)$$

where λ is a phenomenological parameter. λ is a measure of how close the transition state is, along the reaction coordinate, to the product and presently has to be guessed. Several typical $\Delta E(z)$ are shown in

Figure 2 for myosin V and myosin VI. We see that the binding rate is highest to the +36 nm binding site for myosin V. The -36 nm site is also favourable, however it is higher in energy when compared to the +36 nm site. This energy difference controls the relative probability of the forward and backward step. Interestingly, when load forces are applied, $\Delta E(-36 \text{ nm})$ becomes similar to $\Delta E(+36 \text{ nm})$. The force at which these two energies become equal is approximately the stall force, where the probability of taking a forward step becomes equal to the probability of taking a backward step.

After analysing the two critical steps that lead to processivity, trailing motor domain detachment and binding of the free motor domain, it is clear that the elasticity of the light-chain domains regulate the processive movement. Changing the light-chain domain stiffness will affect the mechanical equilibrium conformation of the motor domains through Equation (4) and the rate of binding to actin through Equations (7) and (8). A hand-over-hand mechanism is directly the result of mechanical properties of protein domains. Thus, the emergent properties of the molecular structure ultimately controls motor protein behaviour. Quantitative details of the stepping behaviour will depend on quantitative values of elastic constants and geometrical parameters such as the size of the domains. These effects can be investigated from computational studies outlined in the next section.

5. Stochastic modelling approaches based on the energy landscape framework

Starting from the total free energy of the motor dimer in Equation (3), it is possible to completely solve the dynamics of the system using standard approaches from stochastic dynamics [75,76]. Two essentially equivalent approaches are the Fokker–Planck equation where one solves for a time-dependent probability distribution, and the Langevin equation where one solves for a stochastic trajectory. We denote the mechanical variables by a vector \mathbf{x} , e.g. $\mathbf{x} = (\theta_1, \phi_1, \theta_2, \phi_2)$ for dimeric myosin motors. We denote the combined chemical state of the dimer as $\sigma = (s_1, s_2)$. In the high friction limit where inertia is negligible and momenta are in thermal equilibrium, we may use the Smoluchowski limit of the Fokker–Planck equation to describe probability distribution of the motor configuration $P_\sigma(\mathbf{x}, t)$:

$$\frac{\partial P_\sigma(\mathbf{x}, t)}{\partial t} = \nabla_{\mathbf{x}} \cdot [\mathbf{D} \nabla_{\mathbf{x}} P_\sigma(\mathbf{x}, t) + P_\sigma(\mathbf{x}, t) \nabla_{\mathbf{x}} E(\sigma, \mathbf{x})] - \sum_{\sigma'} k_{\sigma \rightarrow \sigma'}(\mathbf{x}) P_\sigma(\mathbf{x}, t) + \sum_{\sigma'} k_{\sigma' \rightarrow \sigma}(\mathbf{x}) P_{\sigma'}(\mathbf{x}, t), \quad (9)$$

where the first line is the ordinary Smoluchowski equation describing probability flow on an energy landscape with fixed σ and \mathbf{D} is the diffusion tensor. $E(\mathbf{x}, \sigma)$ is the same free energy appearing in Equation (3). The second line describes probability flow between different σ 's with rate functions $k_{\sigma \rightarrow \sigma'}(\mathbf{x})$. This equation is equivalent to the high friction limit of the Langevin equation, which describes stochastic trajectories as:

$$\gamma \frac{\partial \mathbf{x}}{\partial t} = -\nabla_{\mathbf{x}} E(\mathbf{x}, \sigma(t)) + \mathbf{R}(t), \quad (10)$$

where $\sigma(t)$ is a stochastic trajectory that jumps between different states and γ is the friction tensor: $\gamma = k_B T \mathbf{D}^{-1}$; $\mathbf{R}(t)$ is the random force vector. Single molecule experiments have collected both probability distribution data as well as trajectory data [33,47,77,78]. Therefore, both of these approaches are useful for direct comparisons with experiments.

5.1. Ratchets and powerstrokes

A number of limits can be examined based on the energy landscape approach. For instance, concepts such as the Brownian Ratchet have been invoked to explain motor activity [12,14]. In the context of an energy landscape, a ratchet is a sudden drop in free energy that is essentially irreversible (Figure 4). Most of the time, the system is undergoing diffusive dynamics on a flat free energy landscape until encountering the free energy drop. On the other hand, a powerstroke is simply a more gradual lowering of a free energy. In the limit of a linear decline in free energy, a constant force is obtained from the free energy change. For biological motors, the shape of the free energy landscape is determined by the motor structure and the enzymatic activity. Both ratchets and power-stroke-like landscapes are possible.

5.2. Markovian models

Further simplifications based on the energy landscape is possible by considering only a set of interested regions on the landscape [21,22]. These regions can be called 'Markov states', and the transition probability per unit time between these Markov states are Markov transition probabilities. The Markov transition probabilities can be obtained from a more detailed analysis of probability fluxes between Markov states using the Fokker-Planck equation, similar to the analysis Kramers carried out for the reaction over a barrier. Thus, a Markovian description of the motor dynamics is possible starting from the energy landscape picture,

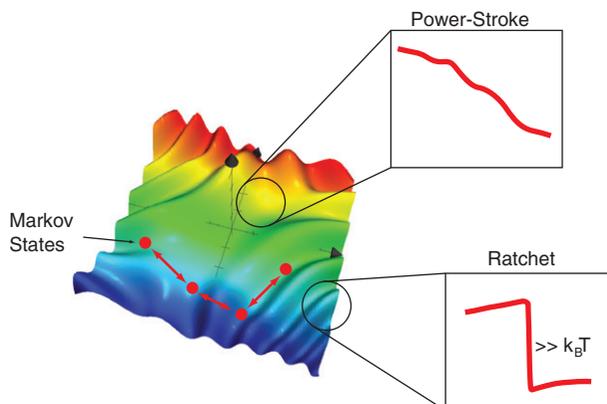


Figure 4. The energy landscape is a general picture that includes concepts such as ratchets and power-strokes. A ratchet-like region in the energy landscape is a steep drop in free energy (much greater than $k_B T$). Once motion goes over the drop, the motion is rectified. A power-stroke is a more gradual change in free energy. The slope of the free energy change is the power-stroke force. Regions on the energy landscape may be of special interest (red dots). These regions are Markov states. Jumps between these Markov states, which is equivalent to stochastic dynamics on the energy landscape, can be described by Markov models.

provided a set of important states are identified. Alternatively, without identifying these Markov states, symmetries of the energy landscape and a minimum number of states can be assumed to explain single molecule data. This approach asks what is the simplest model that can capture the observed phenomena without having to provide more detail of the characters of the assumed states.

The Markov approach can be extended further and thought of as a discrete solution to the Smoluchowski equation in Equation (9). The energy landscape considers a set of continuous mechanical variables, and a set of discrete chemical variables. In practical calculations, the continuous mechanical variables are discretized. Since the definition of a state is flexible in a Markov model, a state α can be defined by a set of mechanical and chemical variables: $\alpha = (\theta_{1,i}, \phi_{1,j}, \theta_{2,k}, \phi_{2,l}, s_1, s_2)$ where i, j, k, l are the indices in the discretized mechanical variables. If the Markov jump is a change in a mechanical variable, the transition rate between states, $k_{\alpha \rightarrow \beta}$ can be determined by solving the Fokker-Planck equation in the local region where the transition takes place. If the Markov jump is a change in the chemical variable (s_1, s_2), the transition rate is completely determined by chemical rate functions described in Section 2. Details of these procedures are summarized in other publications [16,79].

Thus, starting from the energy landscape picture, a continuum of models can arise, depending on the level

of simplifications and assumptions. To explain experimental data, and predict new experiments, however, the energy landscape picture provides the most fundamental description. It makes connections with the protein structure, and asks which emergent properties of the structure contribute to motor motion. In order to develop an energy landscape model, coarse-grained descriptions of protein structures are necessary. In Appendix 1, a coarse-graining approach based on mechanical models of protein secondary structures is explained.

6. Results for myosin V, myosin VI, isoforms and mutants

The modelling framework described here has been applied to myosin V and VI [16,17], which are two prototypical processive motors on actin filaments. Atomic structures of myosin V and VI monomers have been solved [58–62]. Therefore, reasonable parameters such as $(\theta_0(s), \phi_0(s))$ can be obtained from the structures. The chemical transition rate functions have not been measured, but kinetic measurements from purified motor monomers are available [65]. The measured rate constants represent the rate function evaluated at the preferred conformation: $k_{s \rightarrow s'}(\theta_0(s), \phi_0(s))$. Moreover, the influence of applied forces on the rates for motor monomers have been measured as well. Therefore, there is increasing experimental input to constrain the model.

The mechanical stiffness parameters for the light-chain domains, l_{pi} , have not been measured. Given the three-dimensional geometry of the actin filament, these mechanical parameters dictate the conformation and energy of the motor on the F-actin track under external load forces (Figure 5), which directly determine the force–velocity relationship and the step-size distributions. Hence, comparing model predictions with experimental observations through parameter study provides information about the mechanical properties of motor proteins, which can be validated in other experiments. For instance, we simulated the processive motion of myosin V. Force–velocity relationships computed from the trajectories compare favourably with experimental results (Figure 6): myosin V processes along F-actin at $\sim 500 \text{ nm s}^{-1}$ velocity against external load force $|\mathbf{F}|$ up to 1 pN, and then slows down rapidly with increasing force for $|\mathbf{F}| > 1 \text{ pN}$; myosin V stalls at around $|\mathbf{F}| = 2.7 \text{ pN}$.

As observed in single molecule experiments, the computed motor trajectories based on stochastic dynamics simulations of Equation (9) show distinct steps. However, the stepping motion is stochastic, and

myosin V has a broad distribution of step sizes with an average value at 36 nm (Figure 6). The step size distribution changes with the applied load force, but the average step size is always 30–36 nm. A strength of modelling that the stepping trajectories have essentially infinite spatial and temporal resolution. For example, the model easily captures the $\sim 10 \text{ nm}$ sub-steps that have only been recently observed with improved instruments [39,40]. The model also predicts the processes that give rise to these substeps.

The same theoretical framework with a different set of parameters were applied to myosin VI [17]. To properly describe the conformational energy of the motor molecule, we decomposed the linker domain of the myosin VI dimer into three parts: two short elastic rods that extend from the motor domains and one flexible linker that connects the two rods. This description can be still written as Equation (5), but now, l_{pi} is a function of s , representing the non-homogeneous elasticity along the linker extension. In the limit of $l_{pi} \ll L$, the middle flexible linker reduces to a Worm-Like-Chain (WLC) model whose extension property has been studied intensively [80–82]. It can be shown that the WLC model behaves like an entropic spring.

Different mechanical properties of the myosin VI linker give rise to different conformation energies during processive stepping along actin filaments (Figure 5). For forward stepping, the lowest binding energy change for myosin VI occurs at around 30 nm interhead distance. For backward stepping, -10 nm interhead distance is the lowest binding energy change. The computed trajectories using the stochastic simulation shows a broad step-size distribution for myosin VI processivity with average forward step-size 30 nm and backward step-size -10 nm . Both are consistent with experiments (Figure 6).

Since the myosin VI linker is flexible, the Brownian search process prior to binding to the actin track can be time consuming and can not be ignored. We used the conformation energy (Figure 5) to compute the mean-first-passage-time (MFPT) of the free motor head searching for the proper binding site [33]. Adding this searching time to the kinetic transition time when both heads bound, we explained the observed dwell-time between successive steps under different load forces at different ATP levels (Figure 6).

To further examine the role of the elastic linker domains in processive myosins, we studied the myosin X isoform and myosin V mutants with truncated linker domains. Again, the same modelling framework applies, but the length of the linker domain is shorter. Results show that, given the same linker stiffness, the linker length controls the conformation energy of

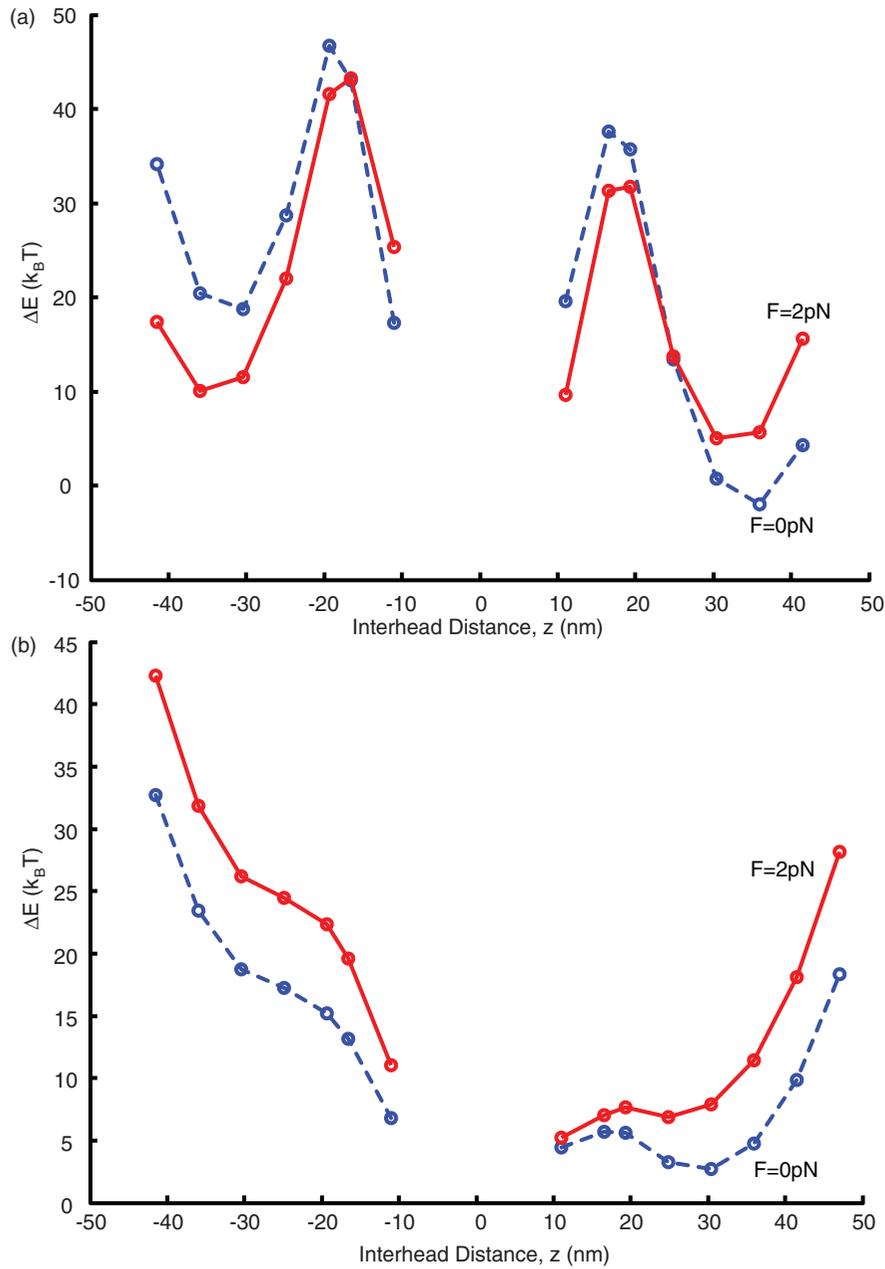


Figure 5. Mechanical model estimates for the energy difference before and after motor binding to the actin track, $\Delta E(z)$ in Equation (7). (a) Myosin V. (b) Myosin VI. Two load forces, $F=0$ pN (blue) and $F=2$ pN (red) are shown.

myosin motors binding with F-actin. Shorter linker domain leads to higher ΔE and more significant barriers for stepping, reducing the duty ratio of the myosin dimer. Interestingly, the mechanical equilibrium conformation for the bound dimer, $(\theta_1^*, \phi_1^*, \theta_2^*, \phi_2^*)$, depends on the linker length. Since the conformation $(\theta_1^*, \phi_1^*, \theta_2^*, \phi_2^*)$ directly affects the conformation of the ATP-catalytic pocket, changing $(\theta_1^*, \phi_1^*, \theta_2^*, \phi_2^*)$ could potentially accelerate or decelerate ATP hydrolysis. Hence, truncating linker domain can

counter-intuitively increase the velocity of the motor. Understanding the dependence of rate constants on conformation can explain this puzzling observation in myosin V mutant experiments [8].

The agreement between modelling and experiments suggests that decomposing motor molecules into multiple mechanical components and combining mechanics with biochemical kinetics is a reasonable and effective approach for studying operating the mechanism of processive motor proteins. The modelling

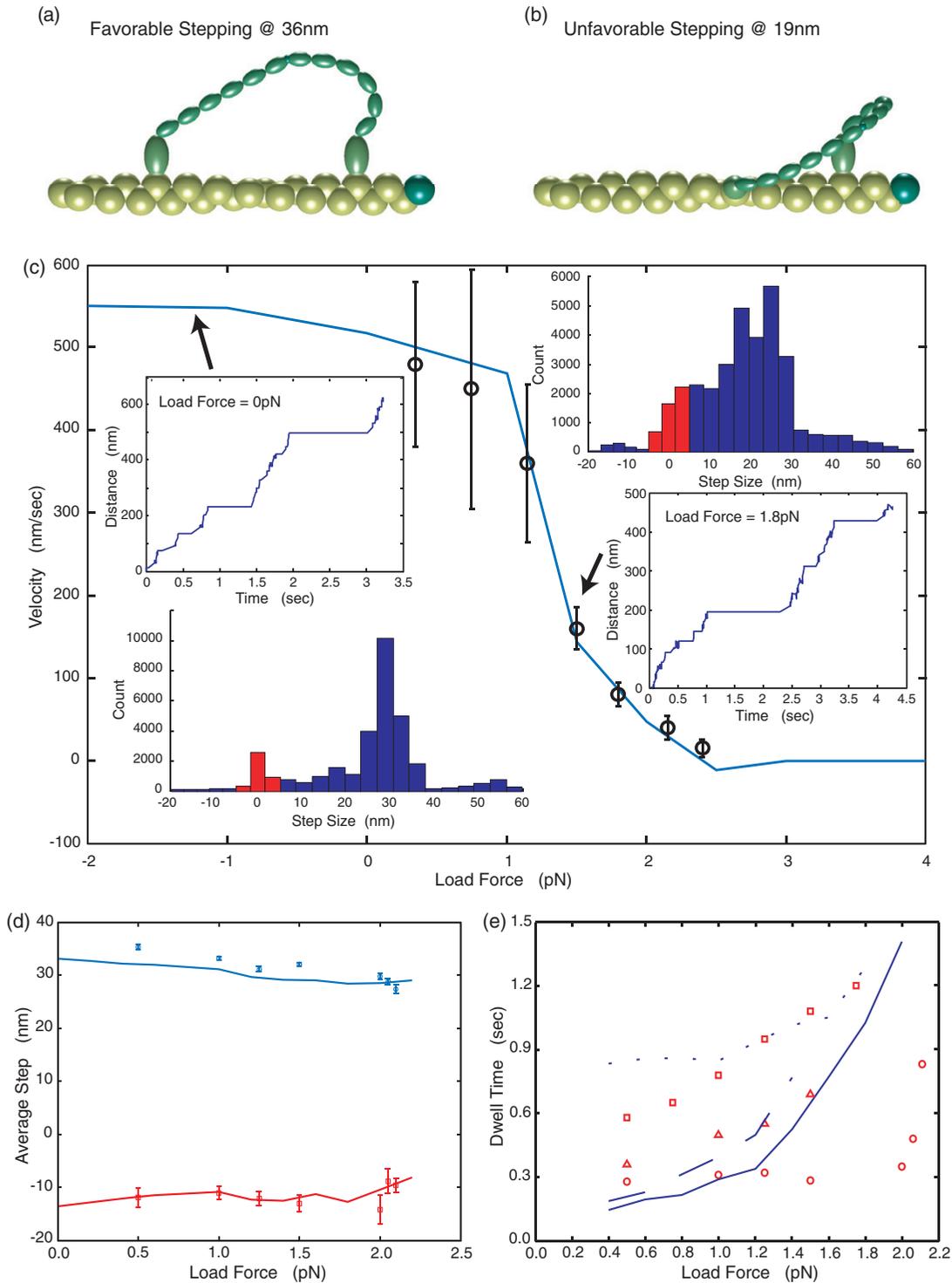


Figure 6. Stepping property of myosin V and myosin VI molecules. Geometries of myosin V dimer at step size of (a) 36 nm and (b) 18 nm. Part (a) shows that the lever-arms of myosin V are moderately bent, which leads to low mechanical energy E_L and defines a favourable stepping geometry. In (b), the lever-arms are heavily bent, which results in high E_L and the stepping at this distance is unfavourable and unlikely to happen. (c) Force-velocity relation of myosin V. Symbols are experimental measurements of Uemura *et al.* [39] and the solid line is the simulation result. Sample trajectories are shown in the insets for $F=0$ pN (upper left) and $F=1.8$ pN (lower right). Forward (blue) and backward (red) step size distributions are shown in the insets for $F=0$ pN (lower left) and $F=1.8$ pN (upper right). (d) The average step size for the forward and backward steps of myosin VI. The experimental results (symbols) are taken from Altman *et al.* [33]. (e) The dwell time before taking a step as a function of the load force for [ATP]=2 mM, [ADP]=0 μ M (solid line and circles); [ATP]=1.5 mM, [ADP]=1 μ M (dashed line and triangles); [ATP]=100 μ M, [ADP]=0 μ M (dotted line and squares). The symbols are measurements from Altman *et al.* [33].

framework not only explains the single molecule experiments but also provides important estimates that are experimentally inaccessible, such as the mechanical properties of the motor domains and the conformation dependence of kinetic rate constants. Our theory reveals that, in general, processive motor proteins utilize their linker domains to transmit forces between monomers (Figure 5), and synchronize the stepping motion of the motor dimer (Figure 6).

7. Motors interacting with cytoskeletal networks

In eukaryotic cells, myosin motors operate within the three-dimensional network of cross-linked F-actin filaments. To understand how transport occurs in the network, and how the motor switches between tracks, we use our model to examine myosin motor stepping on multiple F-actin polymers with varying positions and orientations. Because our theoretical framework is based on a three-dimensional mechanical model, and the motor can at most simultaneously bind to two actin binding sites, investigating how motors interact with two filaments is sufficient to cover all possibilities. A number of additional features needs to be considered: (1) increased number of available binding sites along different F-actin tracks, (2) three-dimensional geometry of available actin sites and the geometry of the actin filament at cross-linked junctions, and (3) exclusive volume effect preventing the access of binding sites that are spatially close. A detailed mathematical description of F-actin geometry at cross-linked junctions is presented in Appendix 2.

For motors stepping on a single actin track, the binding site positions are uniquely determined by the distance z between two binding sites, E_L is only a function of $(\theta_1, \theta_2, \phi_1, \phi_2, z, \mathbf{F})$ as shown in Equation (3). In contrast, for motors at the junction of two intersecting tracks, E_L in Equation (3) now depends on $(\theta_1, \theta_2, \phi_1, \phi_2, \mathbf{p}_1, \mathbf{p}_2, \mathbf{o}_1, \mathbf{o}_2, \mathbf{F})$, where $(\mathbf{p}_1, \mathbf{p}_2)$ are the coordinates of a pair of available binding sites and $(\mathbf{o}_1, \mathbf{o}_2)$ are their orientations. The number of possible combinations of $(\mathbf{p}_1, \mathbf{p}_2, \mathbf{o}_1, \mathbf{o}_2, \mathbf{F})$ is substantial. We explicitly discuss myosin V stepping through a junction connected by the actin branching protein Arp2/3. We will also discuss myosin X processivity along parallel actin filaments cross-linked by actin bundling protein fascin. Calculations are performed without external load force. All geometrical parameters for Arp2/3 and fascin are summarized in Appendix 2.

Arp2/3 is a protein found at the cellular leading edge that binds to the existing (mother) F-actin and initiates a second (daughter) filament at a 70° angle (Figure 7) [83]. Recent structural studies have provided

detailed geometrical information about the Arp2/3 junction [84–86]. We use this information to calculate $(\mathbf{p}_1, \mathbf{p}_2, \mathbf{o}_1, \mathbf{o}_2)$ for all possible pairs of actin binding sites along the mother and daughter filaments. The energy of the motor, $E_{i,j}$ (i and j are the indices of the available binding sites), is evaluated using Equations (3) – (4) with constraints provided by the branch geometry. Figure 7 shows several sample motor geometries based on the evaluated motor energy. When one motor domain is bound to a binding site (i) on the mother filament, the probability for the free motor domain to step to the daughter filament, i.e. switching, can be computed by:

$$P_s(i) = \frac{\sum_{j \in \text{daughter}} e^{-E_{ij}}}{\sum_{j \in \text{daughter}} e^{-E_{ij}} + \sum_{j \in \text{mother}} e^{-E_{ij}}}. \quad (11)$$

Summing over i gives the total switching probability of a motor dimer near the branch. Our model predicts $\sim 24\%$ switching probability, which is in excellent accord with experiments [44].

For myosin X stepping along parallel actin bundles cross-linked by fascin [45,46], we use the same methodology to compute motor dimer energy as a function of pairs of binding sites (Figure 7). *In vitro* single molecule experiments have shown that myosin X has limited processivity along single actin filaments, but has a much improved processivity along fascin-cross-linked actin bundles [46]. Our modelling demonstrates that this is due to the shorter linker domains (15 nm instead of 30 nm): stepping along the same track requires significant bending in the linker and is energetically unfavourable. A more favourable scenario is to bind the motor domains to two separate actin filaments. Fascin bundles parallel actin filaments and positions the filaments sufficiently close so that myosin X can easily step along two neighbouring tracks [87].

With sufficient information about the actin network geometry, our model can explicitly evaluate the energies of motor dimer configurations. When a myosin motor dimer is walking in a cytoskeleton network, the energies of the motor configurations determine the relative probabilities of the available binding sites. The linker domains of the motor dimer again play a regulatory role in determining motor processive dynamics. This raises the possibility of artificially designing motors that behave in different ways in the cell.

8. Motor assemblies, tug of war, and muscle contraction

One of the primary functions of myosin motors is to transport cargo inside cells [31]. These cargos are

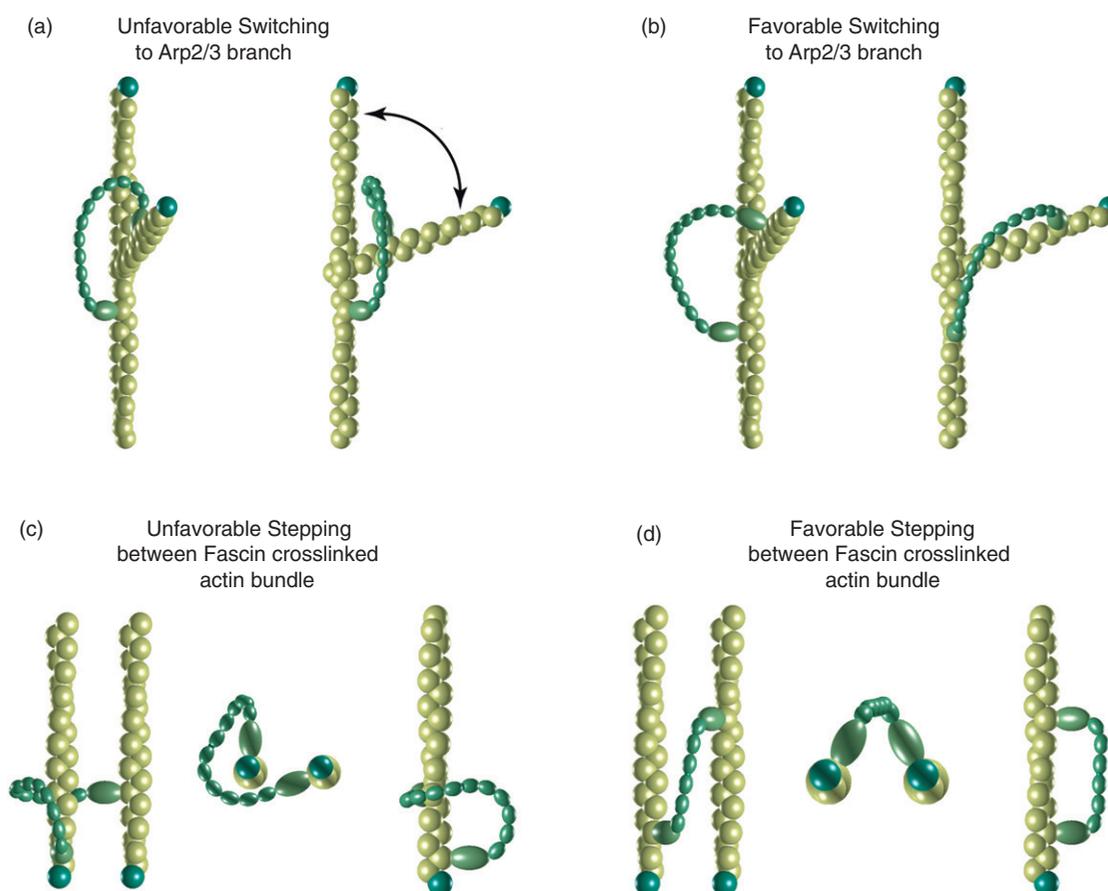


Figure 7. Stepping of dimeric myosin motors on cross-linked F-actin filaments. Parts (a) and (b) show three-dimensional views of the unfavourable and favourable stepping geometries of myosin V at the junction of an arp2/3 branch, respectively. Parts (c) and (d) show three-dimensional views of the unfavourable and favourable stepping geometries of myosin X at a fascin F-actin bundle, respectively. The darker actin monomers indicate the barbed-ends of individual F-actin filaments.

hundreds of nanometres in size; a single cargo can associate with multiple motors, including microtubule motors, to form a large transport assembly. The transport assembly is also likely to interact with multiple cytoskeletal tracks. When the cargo is transported, all the motors within the assembly undergo their own stepping dynamics and the force generated by each motor is felt by all other motors. Thus, motors within the assembly ‘communicate’ with each other. If there are motors travelling towards the opposite ends of the track, such as myosin V and myosin VI, the cargo is transported in a ‘tug of war’ fashion [51]. *A similar approach could also describe recent interesting results in assemblies of kinesin motors* [88]. Even though the *in vivo* process is complex, our theoretical framework is still applicable for describing motors in the cargo transport assembly. The framework also

provides explanations for interesting collective phenomena, including bistability and hysteresis that are observed in motor assemblies [89].

In addition to the motor forces needed for cargo transport, cells also require larger nano-Newton forces for important functions such as cell division and muscle contraction. To this end, pico-Newton force generators, conventional myosin II molecules, form highly ordered assemblies and operate as a single force generating unit. One of the most important force generating units is the sarcomere in the skeletal muscle [90]. Each sarcomere contains about 300 myosin II molecules that bundle together to form the thick filament. The thick filament is polarized, and two opposing units of 150 myosin II motors form a contractile unit. Multiple F-actin filaments (thin filaments) are around the thick filament to interact

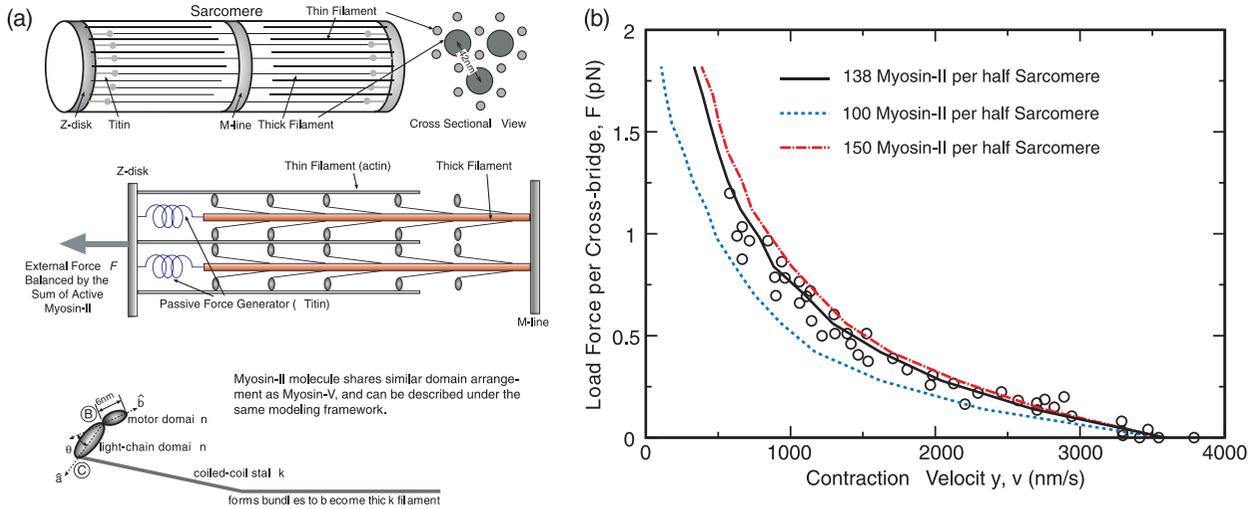


Figure 8. Contraction and force generation of skeleton muscle. (a) Illustration of skeleton muscle structure. More than a hundred myosin II motors bundle to form the thick-filament that collectively interact with the thin filaments (F-actin) to generate force and motion. The mechanochemical dynamics of individual myosin II motors can be described using the presented model. (b) Relation of generated force and the contraction velocity for muscles that are made from different numbers of myosin II motors. The dots are experimental data for 138 motors.

with myosin II. When a myosin II binds to a thin filament (forming an active cross-bridge) and performs a power-stroke, the movement slides the thin filament and causes shortening (contraction) of the sarcomere unit [91–93] (Figure 8).

Even though myosin II is not processive, and unlike the processive motors that use both heads to interact with actin filaments, myosin II uses a single head to bind F-actin and generate forces, the modelling framework for a group of myosins is essentially the same. We successfully modelled the mechanochemistry of the myosin II motor and the collective behaviour of sarcomere contraction. It is possible to incorporate the elasticity of the linker domain in the model [57]. When multiple myosin II motors operate on the same thin filament, the hydrolysis and power-stroke activities are coupled due to the geometrical constraint from the rigid thin filament. The motors are also arranged in opposite directions in the two halves of the sarcomere, leading to a natural tug of war. Thus, when the load force is high, oscillations, hysteresis and other dynamic instabilities naturally arise [89,94].

Within our framework, for a sarcomere with N working myosin II heads (cross-bridges), the total energy of the system can be written as

$$E = \sum_{i=1}^N E_0(\theta_i, \phi_i, s_i) + E_L(\theta_i, \phi_i), \quad (12)$$

where E_0 is motor domain energy. E_L is the conformation energy of the linker domain. If F is the load

force applied in the direction of the thin filament, force balance of the entire system leads to

$$\frac{\partial E}{\partial x} = \sum_{i=1}^N \frac{\partial E_0(\theta_i, \phi_i, s_i)}{\partial x} + \frac{\partial E_L(\theta_i, \phi_i)}{\partial x} = F, \quad (13)$$

where x is the displacement along the axis of the thick filament, which is a function of the motor conformation $x(\theta_i, \phi_i)$. Equations (12) and (13) are not sufficient to describe sarcomere contraction. However, when the chemical transitions are taken into account [Equation (9)], for given initial configurations (θ_i, ϕ_i) and applied force F , hydrolysis and force generation activities of all myosin II motors can be solved. With this approach, isometric, isotonic and jump experiments on muscle can be analysed in detail using our model. Observed collective behaviour of sarcomere contraction is also contained in the model solution. Questions such as the effect of light-chain mechanical properties on contraction dynamics can be answered as well.

Simulation predictions on the maximum contraction velocity ($\sim 3500 \text{ nm s}^{-1}$ under no load) and the maximum load force ($\sim 2 \text{ pN}$ per motor at near-zero velocity) are in quantitative agreement with measurements [95]. The load–velocity and efficiency–velocity relationships both agree well with experimental results as well [95] (Figure 8). Our theory predicted that with increasing load force, more and more myosin II motors interact with the thin filament to generate a balancing force (Figure 8).

For a small motor assembly performing cargo transport within a cytoskeleton network, the geometry of the available binding sites are likely to be complex. Nevertheless, the total motor energy is still in the form of Equation (12), and the dynamics of the motor assembly can be described using the same approach as for muscle contraction.

9. Conclusions

The modelling framework reviewed in this article is a general method to relate the three-dimensional molecular structure with precision single molecule data. The modelling framework is applicable to single motors, to dimers and motor assemblies. By carrying out quantitative studies, a number of general concepts emerge. (1) Proteins are mechanical structures. Geometry and emergent properties of protein domains are important for understanding large scale conformational changes. (2) Enzymatic activity of proteins is related to protein conformation and affected by mechanical forces. A general concept is the enzymatic rate function, which relates the rate of enzymatic reaction with the protein conformation. Applied forces change the protein conformation, and in turn affect reaction rates. (3) By examining the three-dimensional structure, it is possible to use models to describe the overall energy landscape of the protein assembly. Features of the energy landscape can be used to develop concepts such as ratchets and power-strokes. Stochastic methods combined with the energy landscape are able to describe experimental data. We emphasized the general understanding of mechanical performance of molecular motors and their assembly. Using myosin motors as a model system, we showed that the motor function is intimately related to the geometry and mechanical properties of the linker domain. Even though the linker domain does not bind nucleotides, it serves an important role in regulating motor function and the observed stepping behaviour.

To perform computations, the model requires a number of parameters. These parameters include the mechanical constants of the protein domains and the functional forms of the rate functions. Many of these parameters are not known, therefore systematic studies are required. In earlier work, unknown parameters are fitted to experimental data. By manipulating these parameters through mutations and other engineering approaches, it is possible to design molecular motors that function in desired ways. By understanding how these parameters relate to molecular structure, more precise control of motor systems is possible.

Parameters of the coarse-grained model can also be obtained from molecular dynamics studies and atomistic simulations [96,97]. Indeed, the proposed framework can be naturally connected with MD results as described in Appendix 1. Furthermore, MD allows detailed studies of mutations and functional roles of critical residues in the motor, which could be important in some cases.

The modelling framework reviewed here is not limited to myosin motors. For other types of processive microtubule motors such as kinesin and dynein, this approach is also applicable. The necessary inputs are single monomer motor energies in Equation (1) and the linker domain energies Equation (5), and the enzymatic rate functions. For kinesin and dynein, these conformational energies are likely to be more complex. Accurate descriptions of these energies will require a more detailed understanding of the molecular structure. For dynein, the number of domains and ATP hydrolysis sites is also significantly larger [98], presumably leading to more complex regulation of motor stepping behaviour. Therefore, more information is needed for the kinetics of the motor system as well. Nevertheless, the essential elements that lead to processivity described in Section 4 should be the same. A unified description of motor systems ranging from single molecules to assemblies in cells appears to be possible.

References

- [1] J.A. McCammon, B.R. Gelin and M. Karplus, *Nature* **267**, 585 (1977).
- [2] J.A. McCammon and S.C. Harvey, *Dynamics of Proteins and Nucleic Acids* (Cambridge University Press, Cambridge, 1988).
- [3] O.M. Becker, *Computational Biochemistry and Biophysics* (CRC, New York, 2001).
- [4] S.M. Block, *Mod. Cell Biol.* **9**, 375 (1990).
- [5] J.T. Finer, R.M. Simmons and J.A. Spudich, *Nature* **368** (6467), 113 (1994).
- [6] M.J. Schnitzer and S.M. Block, *Nature* **388** (6640), 386 (1997).
- [7] K. Svoboda and S.M. Block, *Annu. Rev. Biophys. Biomol. Struct.* **23** (1), 247 (1994).
- [8] T.J. Purcell, C. Morris, J.A. Spudich and H.L. Sweeney, *Proc. Natl. Acad. Sci. U.S.A.* **99** (22), 14159 (2002).
- [9] S.M. Block, *Nature* **386** (6622), 217 (1997).
- [10] G. Oster and H. Wang, *Trends Cell Biol.* **13** (3), 114 (2003).
- [11] S.X. Sun, H. Wang and G. Oster, *Biophys. J.* **86** (3), 1373 (2004).
- [12] R.D. Astumian and M. Bier, *Phys. Rev. Lett.* **72** (11), 1766 (1994).
- [13] T.C. Elston and G. Oster, *Biophys. J.* **73** (2), 703 (1997).

- [14] A.B. Kolomeisky and B. Widom, *J. Stat. Phys.* **93** (3), 633 (1998).
- [15] C.S. Peskin and G. Oster, *Biophys. J.* **68** (4 Suppl.), 202S (1995).
- [16] G. Lan and S.X. Sun, *Biophys. J.* **88** (2), 999 (2005).
- [17] G. Lan and S.X. Sun, *Biophys. J.* **91** (11), 4002 (2006).
- [18] A. Vilfran, *Biophys. J.* **88** (6), 3792 (2005).
- [19] M.E. Fisher and A.B. Kolomeisky, *Physica A* **274** (1), 241 (1999).
- [20] M.E. Fisher and A.B. Kolomeisky, *Proc. Natl. Acad. Sci. U.S.A.* **96** (12), 6597 (1999).
- [21] M.E. Fisher and A.B. Kolomeisky, *Proc. Natl. Acad. Sci. U.S.A.* **98** (14), 7748 (2001).
- [22] A.B. Kolomeisky and M.E. Fisher, *Biophys. J.* **84** (3), 1642 (2003).
- [23] A.B. Kolomeisky and M.E. Fisher, *Annu. Rev. Phys. Chem.* **58**, 675 (2007).
- [24] G.M. Langford, *Traffic* **3** (12), 859 (2002).
- [25] R.P. Brenda, L.R. Serbus, J.B. Duffy and W.M. Saxton, *Science* **289** (5487), 2120 (2000).
- [26] N. Hirokawa, *Science* **279** (5350), 519 (1998).
- [27] E.L.F. Holzbaur and R.B. Vallee, *Annu. Rev. Cell Biol.* **10** (1), 339 (1994).
- [28] V. Prahlad, M. Yoon, R.D. Moir, R.D. Vale and R.D. Goldman, *J. Cell Biol.* **143** (1), 159 (1998).
- [29] R.D. Vale, *Cell* **112** (4), 467 (2003).
- [30] E.M. De La Cruz and E.M. Ostap, *Curr. Opin. Cell Biol.* **16** (1), 61 (2004).
- [31] J.R. Sellers, *Biochim. Biophys. Acta, Mol. Cell Res.* **1496** (1), 3 (2000).
- [32] R.F. Thompson and G.M. Langford, *Anat. Rec., Part A* **268** (3), 276 (2002).
- [33] D. Altman, H.L. Sweeney and J.A. Spudich, *Cell* **116** (5), 737 (2004).
- [34] A. Gennerich, A.P. Carter, S.L. Reck-Peterson and R.D. Vale, *Cell* **131** (5), 952 (2007).
- [35] S.L. Reck-Peterson, A. Yildiz, A.P. Carter, A. Gennerich, N. Zhang and R.D. Vale, *Cell* **126** (2), 335 (2006).
- [36] M.J. Schnitzer, K. Visscher and S.M. Block, *Nat. Cell Biol.* **2** (10), 718 (2000).
- [37] K. Svoboda and S.M. Block, *Cell* **77** (5), 773 (1994).
- [38] H. Tanaka, K. Homma, A.H. Iwane, E. Katayama, R. Ikebe, J. Saito, T. Yanagida and M. Ikebe, *Nature* **415** (6868), 192 (2002).
- [39] S. Uemura, H. Higuchi, A.O. Olivares, E.M. De La Cruz and S. Ishiwata, *Nat. Struct. Mol. Biol.* **11** (9), 877 (2004).
- [40] C. Veigel, F. Wang, M.L. Bartoo, J.R. Sellers and J.E. Molloy, *Nat. Cell Biol.* **4** (1), 59 (2001).
- [41] K. Visscher, M.J. Schnitzer and S.M. Block, *Nature* **400** (6740), 184 (1999).
- [42] A. Yildiz, J.N. Forkey, S.A. McKinney, T. Ha, Y.E. Goldman and P.R. Selvin, *Science* **300** (5628), 2061 (2003).
- [43] J. Howard, *Annu. Rev. Biophys.* **38**, 217 (2009).
- [44] M.Y. Ali, E.B. Kremtsova, G.G. Kennedy, R. Mahaffy, T.D. Pollard, K.M. Trybus and D.M. Warshaw, *Proc. Natl. Acad. Sci. U.S.A.* **104** (11), 4332 (2007).
- [45] A.B. Bohil, B.W. Robertson and R.E. Cheney, *Proc. Natl. Acad. Sci. U.S.A.* **103** (33), 12411 (2006).
- [46] S. Nagy, B.L. Ricca, M.F. Norstrom, D.S. Courson, C.M. Brawley, P.A. Smithback and R.S. Rock, *Proc. Natl. Acad. Sci. U.S.A.* **105** (28), 9616 (2008).
- [47] R.S. Rock, S.E. Rice, A.L. Wells, T.J. Purcell, J.A. Spudich and H.L. Sweeney, *Proc. Natl. Acad. Sci. U.S.A.* **98** (24), 13655 (2001).
- [48] M.Y. Ali, H. Lu, C.S. Bookwalter, D.M. Warshaw and K.M. Trybus, *Proc. Natl. Acad. Sci. U.S.A.* **105** (12), 4691 (2008).
- [49] A. Kunwar, M. Vershinin, J. Xu and S.P. Gross, *Curr. Biol.* **18** (16), 1173 (2008).
- [50] C. Kural, H. Kim, S. Syed, G. Goshima, V.I. Gelfand and P.R. Selvin, *Science* **308** (5727), 1469 (2005).
- [51] J.L. Ross, M.Y. Ali and D.M. Warshaw, *Curr. Opin. Cell Biol.* **20** (1), 41 (2008).
- [52] G.T. Shubeita, S.L. Tran, J. Xu, M. Vershinin, S. Cermelli, S.L. Cotton, M.A. Welte and S.P. Gross, *Cell* **135** (6), 1098 (2008).
- [53] E. Eisenberg and T.L. Hill, *Science* **227** (4690), 999 (1985).
- [54] J. Howard, *Mechanics of Motor Proteins and the Cytoskeleton* (Sinauer Associates, Sunderland, MA, 2001).
- [55] M. Schliwa, *Molecular Motors* (Wiley-VCH, Weinheim, 2002).
- [56] M. Schliwa and G. Woehlke, *Nature* **422** (6933), 759 (2003).
- [57] G. Lan and S.X. Sun, *Biophys. J.* **88** (6), 4107 (2005).
- [58] P.D. Coureux, H.L. Sweeney and A. Houdusse, *EMBO J.* **23** (23), 4527 (2004).
- [59] A. Houdusse, A.G. Szent-Györgyi and C. Cohen, *Proc. Natl. Acad. Sci. U.S.A.* **97** (21), 11238 (2000).
- [60] J. Ménétrey, A. Bahloul, A.L. Wells, C.M. Yengo, C.A. Morris, H.L. Sweeney and A. Houdusse, *Nature* **435** (7043), 779 (2005).
- [61] J. Ménétrey, P. Llinas, J. Cicolari, G. Squires, X. Liu, A. Li, H.L. Sweeney and A. Houdusse, *EMBO J.* **27** (1), 244 (2008).
- [62] J. Ménétrey, P. Llinas, M. Mukherjea, H.L. Sweeney and A. Houdusse, *Cell* **131** (2), 300 (2007).
- [63] F. Buss, G. Spudich and J. Kendrick-Jones, *Annu. Rev. Cell Dev. Biol.* **20**, 649 (2004).
- [64] R. Roberts, I. Lister, S. Schmitz, M. Walker, C. Veigel, J. Trinick, F. Buss and J. Kendrick-Jones, *Philos. Trans. R. Soc. B* **359** (1452), 1931 (2004).
- [65] E.M. De La Cruz, A.L. Wells, S.S. Rosenfeld, E.M. Ostap and H.L. Sweeney, *Proc. Natl. Acad. Sci. U.S.A.* **96** (24), 13726 (1999).
- [66] D.K. Jamison, J.W. Driver, A.R. Rogers, P.E. Constantinou and M.R. Diehl, *Biophys. J.* **99**, 2967 (2010).
- [67] C. Veigel, J.E. Molloy, S. Schmitz and J. Kendrick-Jones, *Nat. Cell Biol.* **5** (11), 980 (2003).

- [68] C. Veigel, S. Schmitz, F. Wang and J.R. Sellers, *Nat. Cell Biol.* **7** (9), 861 (2005).
- [69] R.A. Milligan, M. Whittaker and D. Safer, *Nature* **348**, 217 (1990).
- [70] R.S. Rock, B. Ramamurthy, A.R. Dunn, S. Beccafico, B.R. Rami, C. Morris, B.J. Spink, C. Franzini-Armstrong, J.A. Spudich and H.L. Sweeney, *Mol. Cell* **17** (4), 603 (2005).
- [71] B.J. Spink, S. Sivaramkrishnan, J. Lipfert, S. Doniach and J.A. Spudich, *Nat. Struct. Mol. Biol.* **15** (6), 591 (2008).
- [72] A. Balaeff, L. Mahadevan and K. Schulten, *Phys. Rev. Lett.* **83** (23), 4900 (1999).
- [73] J.E. Baker, E.B. Krementsova, G.G. Kennedy, A. Armstrong, K.M. Trybus and D.M. Warshaw, *Proc. Natl. Acad. Sci. U.S.A.* **101** (15), 5542 (2004).
- [74] M.L. Walker, S.A. Burgess, J.R. Sellers, F. Wang, J.A. Hammer, J. Trinick and P.J. Knight, *Nature* **405** (6788), 804 (2000).
- [75] C.P. Fall, *Computational Cell Biology* (Springer Verlag, Berlin, 2002).
- [76] D.T. Gillespie, *J. Phys. Chem.* **81** (25), 2340 (1977).
- [77] A.D. Mehta, R.S. Rock, M. Rief, J.A. Spudich, M.S. Mooseker and R.E. Cheney, *Nature* **400** (6744), 590 (1999).
- [78] M. Rief, R.S. Rock, A.D. Mehta, M.S. Mooseker, R.E. Cheney and J.A. Spudich, *Proc. Natl. Acad. Sci. U.S.A.* **97** (17), 9482 (2000).
- [79] J.H. Xing, H. Wang and G. Oster, *Biophys. J.* **89**, 1551 (2005).
- [80] C. Bustamante, J.F. Marko, E.D. Siggia and S. Smith, *Science* **265** (5178), 1599 (1994).
- [81] J.F. Marko and E.D. Siggia, *Macromolecules* **28** (26), 8759 (1995).
- [82] I. Schwaiger, A. Kardinal, M. Schleicher, A.A. Noegel and M. Rief, *Nat. Struct. Mol. Biol.* **11** (1), 81 (2003).
- [83] H.N. Higgs and T.D. Pollard, *Ann. Rev. Biochem.* **70** (1), 649 (2001).
- [84] R.C. Robinson, K. Turbedsky, D.A. Kaiser, J.B. Marchand, H.N. Higgs, S. Choe and T.D. Pollard, *Science* **294** (5547), 1679 (2001).
- [85] I. Rouiller, X.P. Xu, K.J. Amann, C. Egile, S. Nickell, D. Nicastro, R. Li, T.D. Pollard, N. Volkmann and D. Hanein, *J. Cell Biol.* **180**, 887 (2008).
- [86] K.J. Amann, S. Stoilova-McPhie, C. Egile, D.C. Winter, L. Hazelwood, J.E. Heuser, R. Li, T.D. Pollard, N. Volkmann and D. Hanein, *Science* **293** (5539), 2456 (2001).
- [87] R.A. Edwards and J. Bryan, *Cell Motil. Cytoskeleton* **32** (1), 1 (2005).
- [88] J.W. Driver, D.K. Jamison, K. Uppulury, A.R. Rogers, A.B. Kolomeisky and M.R. Diehl, *Biophys. J.* **101**, 386 (2011).
- [89] S. Walcott and S.X. Sun, *Phys. Chem. Chem. Phys.* **11**, 4871 (2009).
- [90] C.R. Bagshaw, *Muscle Contraction* (Springer, Berlin, 1993).
- [91] H. Huxley and J. Hanson, *Nature* **173** (4412), 973 (1954).
- [92] A.F. Huxley and R. Niedergerke, *Nature* **173** (4412), 971 (1954).
- [93] A.F. Huxley and R.M. Simmons, *Nature* **233**, 533 (1971).
- [94] M. Badoual and F. Jülicher, J. Prost, *Proc. Natl. Acad. Sci. U.S.A.* **99** (10), 6696 (2002).
- [95] R. Cooke and K.C. Holmes, *Crit. Rev. Biochem. Mol. Biol.* **21** (1), 53 (1986).
- [96] H. Yu, L. Ma, Y. Yang and Q. Cui, *PLoS Comp. Biol.* **3**, 0214 (2007).
- [97] H. Yu, L. Ma, Y. Yang and Q. Cui, *PLoS Comp. Biol.* **3**, 0199 (2007).
- [98] A.P. Carter and R.D. Vale, *Biochem. Cell Biol.* **88** (1), 15 (2010).
- [99] S. Choe and S.X. Sun, *J. Chem. Phys.* **122**, 244912 (2005).
- [100] S. Choe and S.X. Sun, *Biophys. J.* **92**, 1204 (2007).
- [101] C.W. Wolgemuth and S.X. Sun, *Phys. Rev. Lett.* **97**, 248101 (2006).

Appendix 1. Mechanical models of protein domains

The question we would like to address in this appendix is how we can arrive at mechanical models such as Equations (1) and (5) starting from the molecular structure. The connection is found through statistical field theory where the geometrical shape of the molecule can be thought of as a function of atomic coordinates: $\mathbf{u}(\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n)$, where \mathbf{q}_i are the positions of the i th atom. The definition of this function is not unique. But since we are interested in the emergent properties of the molecular structure, the results should not depend on the details of this functional relationship. Having defined the geometrical shape of the molecule, the free energy, E , as a function of change in shape is then defined as

$$\exp[-E(\mathbf{u})/k_B T] \propto \int d\mathbf{q}_1 \dots d\mathbf{q}_n \delta(\mathbf{u} - \hat{\mathbf{u}}(\mathbf{q}_1 \dots \mathbf{q}_n)) \exp[-V(\mathbf{q}_1 \dots \mathbf{q}_n)/k_B T], \quad (14)$$

where V is the atomic interaction potential between atoms. The assumption made here is that degrees of freedom orthogonal to \mathbf{u} are in rapid thermal equilibrium.

Geometrical variables such as \mathbf{u} represent the kinematics of protein conformational change. To obtain dynamics, we must know about $E(\mathbf{u})$. One way to obtain E is from direct MD simulations using Equation (14). This was performed for α -helices and β -sheets to obtain the elastic energy and mechanical constants for bending and twisting these structures [99,100]. It was found that α -helices behave quite similar to an elastic rod, and β -sheets can be described as an elastic surface. In these studies, proteins are assumed to remain in a folded state mechanical and conformational changes are variations in structure in the folded state. By viewing protein secondary structures as mechanical elements, properties of tertiary structures such as coiled coils and α -helical bundles can be explained as well [101].

In general, a functional form for E is assumed (the constitutive law). These choices are modelling assumptions that must be tested in experiments. The modelling

assumptions also cannot violate basic requirements such as coordinate frame invariances. Thus, simulations, experiments and some intuition can be used to obtain information about E .

Appendix 2. Mathematical description of actin filament geometry

In this appendix, we provide the general mathematical description of the binding sites on a pair of cross-linked F-actin filaments in the three-dimensional space. The more complicated situation with three or more F-actin filaments can be extended from this presented description which in turn provides a framework to describe the 'track'-geometry in the *in vivo* F-actin cortex.

Geometrically, F-actin is a 7 nm thick filament (radius $r = 3.5$ nm) made from actin monomers and has a right-handed double-stranded helical structure. Each strand is a helix that every 13 monomers (monomer size 5.5 nm) form a full turn and the pitch is 71.5 nm. Therefore, one monomer rotates $2\pi/13$ (the twisting angle) around the filament centre line with respect to the previous monomer. From a pure mathematical point of view, however, we simplify this double-stranded structure to a right-handed single helix with a pitch of 35.75 nm, twisting angle $\alpha = 14\pi/13$ and neighbouring monomer distance $d = 2.75$ nm. We will use this single-stranded description in our study.

We first use two unit vectors, \mathbf{f}_1 and \mathbf{f}_2 , to denote the centre line orientations of the two F-actin filaments towards their barbed-ends. For convenience, we use the following cartesian coordinate system:

$$\mathbf{x} : \mathbf{f}_2; \quad \mathbf{z} : \mathbf{f}_1 \times \mathbf{f}_2 / \sin(\gamma); \quad \mathbf{y} : \mathbf{z} \times \mathbf{x},$$

where \mathbf{z} overlaps with the shortest segment connecting \mathbf{f}_1 and \mathbf{f}_2 ; $\gamma \in [0, \pi]$ is the orientational difference between \mathbf{f}_1 and \mathbf{f}_2 ; $\cos(\gamma) = \mathbf{f}_1 \cdot \mathbf{f}_2$; and let S be the shortest distance between the two filaments.

We then index the binding sites on each filament. We define the zeroth binding site on one filament to be the first binding site arising from the \mathbf{z} axis towards the filament's barbed-end. Incremental positive or negative integer *indexes* are assigned to the successive binding sites towards the barbed- or pointed-ends, respectively. The position and orientation of the i th binding site on the j th filament are denoted as $\mathbf{p}_{j,i}$ and $\mathbf{o}_{j,i}$.

To further identify $\mathbf{p}_{j,i}$ and $\mathbf{o}_{j,i}$, we need to know the *mismatch* (m_j) and the *phase* (β_j) of the j th filament.

The *mismatch* is defined as the distance between the \mathbf{z} axis and the zeroth binding site; and the *phase* is defined as the off-plane angle between the orientation of the zeroth binding site and the \mathbf{x} - \mathbf{y} plane. The cross-linking parameters ($S, \gamma, m_1, m_2, \beta_1, \beta_2$), together with the F-actin intrinsic parameters (r, α, d), completely determine $\mathbf{p}_{j,i}$ and $\mathbf{o}_{j,i}$.

For the binding sites on the first F-actin filament, because it is aligned with the \mathbf{x} axis, it is straightforward to derive $\mathbf{p}_{1,i}$ and $\mathbf{o}_{1,i}$:

$$\mathbf{o}_{1,i} = \begin{bmatrix} 0 \\ \cos(\beta_1 + i \cdot \alpha) \\ \sin(\beta_1 + i \cdot \alpha) \end{bmatrix}, \quad (15)$$

$$\mathbf{p}_{1,i} = (m_1 + i \cdot d) \cdot \mathbf{x} + r \cdot \mathbf{o}_{1,i}, \quad (16)$$

where the first term in $\mathbf{p}_{1,i}$ is the distance along the filament and the second term is due to the incrementing helical rotation ($\mathbf{o}_{1,i}$) of the actin monomers.

For the binding sites on the second F-actin filament, we obtain the rotational matrix from \mathbf{f}_1 to \mathbf{f}_2 :

$$\mathbf{R} = \begin{bmatrix} \cos(\gamma) & -\sin(\gamma) & 0 \\ \sin(\gamma) & \cos(\gamma) & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

$\mathbf{p}_{2,i}$ & $\mathbf{o}_{2,i}$ are then:

$$\mathbf{o}_{2,i} = \mathbf{R} \cdot \begin{bmatrix} 0 \\ \cos(\beta_2 + i \cdot \alpha) \\ \sin(\beta_2 + i \cdot \alpha) \end{bmatrix}, \quad (17)$$

$$\mathbf{p}_{2,i} = (m_2 + i \cdot d) \cdot \mathbf{R} \cdot \mathbf{x} + r \cdot \mathbf{o}_{2,i} + S \cdot \mathbf{z}, \quad (18)$$

where the additional third term is due to the spacing between the two F-actin filaments.

Equations (15–18) mathematically describe any binding site along the two cross-linked F-actin filaments. This information can then be used as the boundary conditions to determine the elastic energy of the myosin motor E_L .

It is worth noting that parameters ($S, \gamma, m_1, m_2, \beta_1, \beta_2$) define the property of the cross-linking protein. In particular, if two F-actin filaments are cross-linked by branching protein Arp2/3, the parameters are (0 nm, 70° , 0 nm, 0 nm, 0° , 90°), as illustrated in Figures 7(a) and (b); whereas if two F-actin filaments are cross-linked by bundling protein Fascin (12 nm, 0° , 0 nm, 0 nm, 0° , 0°), as illustrated in Figures 7(c) and (d).