Biophysical Chemistry

Myosin dynamics on the millisecond time scale

Thomas P. Burghardt [a,b,*], Jimmy Yan Hu [a], Katalin Ajtai [a]

[a] *Department of Biochemistry and Molecular Biology, Mayo Clinic College of Medicine, Rochester, MN 55905, United States*
[b] *Department of Physiology and Biomedical Engineering, Mayo Clinic College of Medicine, Rochester, MN 55905, United States*

Abstract

Myosin is a motor protein associating with actin and ATP. It translates along actin filaments against a force by transduction of free energy liberated with ATP hydrolysis. Various myosin crystal structures define time points during ATPase showing the protein undergoes large conformation change during transduction over a cycle with ~10 ms periodicity. The protein conformation trajectory between two intermediates in the cycle is surmised by non-equilibrium Monte Carlo simulation utilizing free-energy minimization. The trajectory shows myosin transduction of free energy to mechanical work giving evidence for: (i) a causal relationship between product release and work production in the native isoform that is correctly disrupted in a chemically modified protein, (ii) the molecular basis of ATP-sensitive tryptophan fluorescence enhancement and acrylamide quenching, (iii) an actin-binding site peptide containing the free-energy barrier to ATPase product release defining the rate limiting step and, (iv) a scenario for actin-activation of myosin ATPase.
© 2007 Elsevier B.V. All rights reserved.

1. Introduction

Myosin is a molecular motor binding ATP and actin to produce work by causing relative translation of the two proteins. Work production is from bound ATP free-energy transduction at an active site in myosin. It is often proposed that myosin

conformation and ATP catalysis steps are strongly coupled in one-to-one correspondence [1–4] summarized by the reaction,

$$M + ATP \leftrightarrow M^*.ATP \leftrightarrow M^{**}.ADP.Pi \leftrightarrow M + ADP + Pi$$

Scheme1

where M, M*, and M** are distinct myosin conformations. Alternatively, weak coupling rationalizes observations suggesting myosin can maintain different structures for identical substrate [5]. Scheme 1 transients converted into quasi-static conformations by active site bound nucleotide or trapped nucleotide analog [6–11] were crystallized for structure determination. Crystal structures of several isoforms representing Scheme 1 intermediates indicate myosin consists of a catalytic domain containing active- and actin-binding sites and a lever arm that swings in the course of ATP hydrolysis [8,12–15]. Crystal structures were obtained without actin, nevertheless, a lever-arm swing taking place with myosin attached to actin is a compelling image for the power stroke translating the two proteins against resistive force. The M** → M transition corresponds to a lever-arm swing that could mimic an active muscle power stroke.

The large functional protein conformation change observed in myosin with the lever-arm swing occurs on the millisecond time

domain [16]. Similar spatial scale conformation change occurs on the micro/millisecond time domain in other systems such as cell membrane receptors actuating ion pores [17]. Both "slowly" cycling systems involve many atoms and perform multi-step tasks like ligand recognition and binding coupled to larger conformation change to perform energy transduction (myosin) or alter ion pore permeability (membrane channel). When transient intermediate structures are known, as they are for myosin, they form time points in the cycle that we will link by a conformation trajectory. Anticipating conformation trajectory facilitates hypothesis formation and experiment design in a structure-function investigation of these systems, however, time scale and system complexity excludes use of equilibrium molecular dynamics (MD) techniques.

Myosin MD simulation in solvent water with bound ATP or ATP analogs could provide the conformation trajectory that we seek, however, practical computing limitations confine its application to protein fragments or to time spans on the order of nanoseconds. MD calculations have already addressed: how water participates in the hydrolytic reaction [18], by what pathway inorganic phosphate leaves the active site [19], and which residues participate in the actomyosin bond [20,21]. Other MD calculations have looked at how local heating at the active site propagates throughout myosin [22]. Constrained MD has been done in the past for a variety of applications [23] some similar to the myosin dynamics problem undertaken here [24–27] and newer work addressing myosin conformation during the power stroke [28]. The difficulty of applying MD to slow conformation change in proteins was addressed with computation of minimum energy paths (MEPs) [29,30]. The MEP for the lever-arm swing transition in myosin provided a detailed sequence of structural states accomplishing lever-arm swing [31]. Normal mode analysis is another approach useful for looking at the large scale protein conformation changes in a system like myosin [32–34].

The power stroke, envisioned by comparison of the M** and M crystal structures, rotates the lever-arm through ∼ 70° and involves substantial myosin secondary structure change. Brant, Miller, and Flory [35] showed that a dipeptide potential energy computed as the sum of dihedral torsion, Coulomb, and van der Waals potentials then expressed as functions of the (ϕ,ψ) Ramachandran angles accurately reproduced the stability of the α-helix and β-sheet protein secondary structures. This simple Ramachandran angle parameterized potential reproduced observation in highly complex protein structures suggesting that substantial secondary structure change could be parameterized as a sequence of Ramachandran angle moves. Backbone dihedral angles are included in the move set for the Monte Carlo (MC) simulation module in CHARMM [36]. There it is pointed out that the change of a single Ramachandran angle can translate many atoms by large distances thereby instigating an avalanche of atomic clashes. We propose that the instigation of clashes makes the backbone dihedral angle a highly constrained degree of freedom (DOF) valuable for efficiently driving large spatial and time scale changes in protein conformation. In real life, impulses from ATP hydrolysis might induce torques on the backbone causing myosin shape change like the Ramachandran DOF.

A MC simulation of the myosin conformation trajectory connecting the M** and M structures was attempted. At the heart of the algorithm is creation of trial structures by small changes in the backbone dihedral angles. The myosin system is still too large to do all of the backbone conformation change via small steps in its backbone dihedral angles hence we also introduced block conformation changes to selected parts of the molecule that change conformation in an all or nothing manner. Thus the molecule is modeled by highly flexible regions with a multitude of DOFs determining global shape and with other localized regions undergoing their dynamics with a single DOF. With each trial structure generated we either accept the trial into the conformation trajectory or reject it to wait for a more favorable future possibility. At equilibrium, we would select structures with criteria obeying detailed balance [36,37]. We find that we must employ a biased selection criterion, consistent with minimal free energy, but favoring the endpoint over the immediate structure by an amount that makes the simulation convergent. The thoughtful selection of flexible and block domains in myosin, together with the biased selection criterion, leads to a tractable MC simulation of the power stroke that imitates key features of the system.

We have two motivations for attempting to visualize the myosin power stroke. First, we often use spectroscopic probes in myosin to detect and characterize local and global conformation changes in the protein [38,39]. Probe circular dichroism (CD) detects side chain and backbone conformation [40,41] while fluorescence lifetime and intensity detects the same as well as accessible surface area (ASA) to an extrinsic solution borne quencher [42,43]. We can surmise the signals expected from the probes given a structural model of the probe neighborhood. The latter requires modeling of the myosin dynamics accompanying energy transduction. We have applied improvised lever-arm dynamics for model development [39,44] but herein describe a systematic approach to the complex modeling problem and discuss what it implies about two spectroscopic signals that we use.

Second, as mentioned above, MC simulation involves a free-energy criterion for trial structure acceptance/rejection. Gibbs free-energy potential change (ΔG) combines enthalpic (ΔH) and constant temperature times entropic change $(T\Delta S)$ in $\Delta G = \Delta H - T\Delta S$. These distinct aspects of ΔG are probably linked during protein conformation change [45]. Nonetheless, enthalpic or entropic intensive conformation change are useful concepts because they might partition into separated protein domains that when identified give insight into how the motor works. Particular steps undertaken by the myosin motor have often been described as entropically or enthalpically driven precisely for the purpose of assigning structure/function relationships within the protein. We will compute the free energy from the separate ΔH and $T\Delta S$ components and assign the changes in these quantities to regions in myosin.

The MC simulation was carried out on native and reactive lysine residue (K84) trinitrophenylated (TNP-) skeletal myosin during the M** → M half-reaction simulating the power stroke in muscle contraction. Simulations cover ∼ 10–100 ms. Free-energy increase, due to an entropy decreasing conformation transition in the actin-binding domain of myosin, opens the active site back

door to release ATPase products and initiates lever-arm movement in native myosin simulations. The actin-binding domain transition is the rate limiting free-energy barrier to product release in the simulation. This result mimics the widely accepted experimental finding that product release is rate limiting and precedes force production with the lever-arm swing. The myosin domain contributing the rate limiting entropic barrier suggests a scenario for the role of actin binding in actin-activated myosin ATPase.

The electron transfer rate (k_{ET}) and ASA for the nucleotide sensitive tryptophan (W510) evaluates how static quenching groups in the protein and solution borne quenchers in a Stern–Volmer experiment might influence W510 fluorescence lifetime and intensity [43,46,47]. Calculation of k_{ET} identifies the groups responsible for modulating W510 fluorescence intensity by the electron transfer mechanism. We find k_{ET} and ASA decrease and increase, respectively, during the trajectory in agreement with experiment.

Statistical causality among time-dependent economic indicators is frequently investigated using a test devised by Granger [48]. Granger and other causality tests are used to study causality among neurophysiological signals [49,50]. In the Granger-test, time-dependent variables $X(t)$ and $Y(t)$ have a relationship where Y is Granger-causal for X when prediction of X is statistically improved by inclusion of past information from both X and Y compared to a prediction based solely on past values of X. We introduce the Granger-test here among time-dependent structural variables characterizing protein dynamics to identify direction sensitive pathways linking distant functional sites in a protein. In particular, we compared time-dependent variables from myosin indicating hydrolysis product release, work producing lever-arm movement, and ASA for W510. In the native protein, hydrolysis product release is Granger-causal for lever-arm rotation accurately restating what is clear by inspection that product release initiates work producing lever-arm movement. Additionally, work producing lever-arm movement is causal for W510 ASA suggesting Stern–Volmer quenching results reflect lever-arm movement rather than direct communication with the active site. No relationship among these parameters is detected in the probe modified myosin. Asymmetrical causality between native and probe modified proteins shows modification alters the protein conformation trajectory and mirrors previous experimental work using *in vitro* motility of modified skeletal myosin [44] and site directed mutagenesis of Dictyostelium myosin II [51]. Graphical depiction of the probe modified myosin trajectory shows collisions between the extrinsic probe and lever arm anticipating the experimental finding that trinitrophenylated myosin does not support in vitro motility.

2. Methods

2.1. MC simulation degrees of freedom

Brant, Miller, and Flory [35] computed potential energy for a dipeptide as the sum of dihedral torsion, Coulomb, and van der Waals potentials, expressed as functions of Ramachandran angles (ϕ,ψ). They predicted the stability of the α-helix and β-sheet protein secondary structures and showed that the dominant contribution to the potential energy comes from steric repulsion in the van der Waals term leading to chain self-avoidance. It follows that substantial secondary structure conformation change could be parameterized as an ordered sequence of Ramachandran angle changes. In the myosin motor, substantial changes in secondary structure accompanying energy transduction produces the lever-arm rotation through ∼70° suggesting the Ramachandran angles are appropriate parameters for the dynamic molecular model. The heart of our algorithm representing myosin transduction is dynamics of a subset of the myosin heavy chain Ramachandran angles. Backbone dihedrals are recognized components of protein dynamics simulation as part of the move set implemented in the Monte Carlo simulation module for CHARMM [36].

Dihedral angles for each residue backbone (ϕ,ψ,ζ) determines protein secondary structure. Ramachandran angles (ϕ,ψ) are wide ranging while ζ, specifying the *cis-* or *trans-*peptide bond, deviates marginally from equilibrium values. Bond lengths and angles likewise affect global conformation and they deviate only slightly from equilibrium values. Side chain and carbonyl oxygen conformation usually does not affect global conformation, except in circumstances where constituent atoms clash, however, it is highly significant in deciding the protein optical spectroscopy. Dihedral angle Ω specifies the side chain C_β position relative to the backbone. Collectively these parameters for a single residue i, separated into the five degrees of freedom $(\phi,\psi,\zeta,\Omega,\theta)_i$ where θ_i stands for the bond lengths, angles, and dihedral angles not included in the other parameters, fully specify protein conformation one residue at a time. A conformation trajectory between initial and final structures is the chronological specification of the single residue degrees of freedom (*DOF*s) where one *DOF* is updated at each step. We generate elementary steps in the conformation trajectory by choosing a *DOF* and the direction for change.

Simulating myosin trajectory during force production is complicated by the large number of *DOF*s. Simplifications that remove *DOF*s without affecting the native contractile mechanism improve prospects for success. Identifying the minimal myosin motor is the first step. Myosin subfragment 1 (S1), containing the motor domain and lever arm (myosin heavy chain or MHC) and essential light chain (ELC) stabilizing lever-arm structure, was isolated by proteolytic cleavage and shown to function normally as an enzyme [52,53] and motor capable of doing mechanical work [54]. Subdivision of S1 into rigid and flexible domains further simplifies simulation. Inspection of S1 crystallographic structures [8,12–15] led to identification of heavy chain domains that remain structurally intact but move relatively during hydrolysis [14,15,55]. They were identified as N-terminal (N-term, AA1–144), upper 50k (U50k, AA145–361), lower 50k (L50k, AA362–462), converter (Conv, AA721–772), and lever (Lever, AA773–812). Intervening flexible regions consist of either single glycine hinges dividing Nterm from U50k, and U50k from L50k, or, longer segments including switch II (AA463–525), and the SH2–SH1 hinge (AA696–720). The peptide intermediate to switch II and SH2–SH1 hinge (AA526–695) contains a large unstructured loop (Loop 2, AA626–647) but otherwise behaves like the other intact domains by maintaining its

shape during transitions represented in the crystal structures. Flexible peptides are called the gear peptides after the analogy to a cog and gear mechanism suggested by Geeves and Holmes [55]. The partitioned S1 is depicted in Fig. 1.

We model the six intact domains (N-term, U50k, L50k, Loop 2, Conv, and Lever) as block-peptides moving rigidly without alteration in a conformational trajectory until their block transition occurs wherein all *DOF*s in the block transit instantaneously from initial to final values thereby reducing a multi-*DOF* peptide to a single *DOF*. Alternatively, the gear peptides make one-residue transitions in one *DOF* from the set $(\phi, \psi, \zeta, \Omega, \theta)_i$. Single dihedral angle transitions in gear peptides move through small amplitudes requiring multiple transitions to exercise the full range of motion. Transitions in θ_i are two state (initial/final). A step in the protein trajectory occurs for any single block-, dihedral-, or θ-transition.

2.2. Sampling

Structures are generated with a biased random walk in a selected subset of the *DOF*s. Biasing the walk is equivalent to applying an external force like that due to the applied electric field in the Repton model for DNA electrophoresis [56]. Bias was adjusted to accomplish protein movement through the M** → M trajectory in simulated time expected for ATP hydrolysis. *DOF* change taking the present towards the target conformation is taken 90% of the time (50% would be unbiased). The exception is when the direction for change in a dihedral angle *DOF* is ambiguous (e.g., when the dihedral is poised 180° from its final value). Then bias is set to 50% until the clear choice emerges. Rather than generating unbiased structures then rejecting 80% of those unfavorably directed to maintain the desired 90% favored

Fig. 1. Myosin S1 subdivisions identified in the simulation. Block peptides: N-terminal (AA1–144, blue cartoon, bottom right), Upper 50 kDa (AA145–361, green ribbon), Lower 50 kDa (AA362–462, red ribbon, top), Loop 2 (AA526–695, blue tube, top right), Converter (AA721–772, red tube, bottom middle), Lever arm (AA773–812, yellow ribbon, left), and Essential Light Chain (orange ribbon, left surrounding the lever arm). Gear peptides: (AA463–525 and AA696–720, white stick backbone bonds). Single residues with side chains shown are Lys84 (yellow within N-terminal block) and W510 (yellow within the gear peptides).

structures, we bias structure generation and decrease the simulation time increment reflecting the higher likelihood that a favorably oriented *DOF* is selected (see discussion for the Repton model [57]).

Time passes for simulated myosin in intervals taken by the protein to sample a trial structure [57,58]. Elapsed time from the last-accepted transition,

$$\Delta t = \frac{N \times t_0}{\cosh\left[\frac{1}{2} \, \text{Ln}\left(\frac{\text{bias}}{1-\text{bias}}\right)\right]} \quad (1)$$

where $N - 1$ is the number of rejected trials before trial N is accepted, t_0 is the time taken to sample one trial, and bias=0.9. When bias=0.5 or 0.9, $\Delta t = N \times t_0$ or $0.6 \times N \times t_0$. One trial conformation is generated in $t_0 \sim 10$ ns because S1 backbone dihedrals fluctuate with this relaxation time as detected by polarization anisotropy [59]. These elementary backbone transitions proceed sequentially hence they determine the total transition time. Side chain relaxations occur in parallel to the slower backbone dynamic and do not affect overall transitioning time. Block transitions likewise do not contribute directly to transitioning time because they are assigned to independent and intact myosin subdomains and parallel the slow backbone dynamic.

DOF selection order, more than any other detail about how peptide energy is computed, determines barrier height resisting conformation change during the M** → M transition. Lacking a specific model for how ATP hydrolysis in the active site influences trajectory pathway, we assign *DOF* selection probabilities to be proportional to the number of steps separating the present from the target conformation such that *DOF*s that are further from target are more heavily weighted. Block- and θ-transitions have minimum probability for selection because only 1 step separates initial and target conformations. When a *DOF* reaches its target value it is trapped there by removing it from the *DOF* list.

Sampling choices are drawn from residues 1 through 961 accounting for 812 MHC (AA1–812) and 149 ELC (AA813–961) residues. For the purposes of this calculation, the lever arm and ELC are joined at AA812 such that residues AA773–961 form one block-peptide. In one program cycle 1–5 steps occur somewhere in the structure. For instance, a cycle might include 1 block-peptide or 5 dihedrals in 5 different gear-peptide residues. S1 has segmental mobility equivalent to a backbone dihedral rotation of ∼3° [59]. Dihedral rotations were confined to 1–5° step amplitudes.

Probability for trial conformation acceptance is given by the Metropolis criterion [60],

$$A = \begin{cases} \text{Exp}\left[\frac{\Delta G}{k_B T}\right] & \text{if } \frac{\Delta G}{k_B T} > 0 \\ 1 & \text{otherwise} \end{cases} \quad (2)$$

where the change in the energy function, $\Delta G = G_{\text{new}} - G_{\text{last}}$. The energy function is free energy computed as described in the next section, and, subscripts *new* and *last* refer to the trial- and the last-accepted-conformation, respectively. Trial acceptance occurs when random number, ξ between 0 and 1, is $\leq A$. Hence all

trials with energy lower than the last-accepted-conformation are accepted (the Metropolis criterion).

The detailed balance condition is given by $W_{i \rightarrow j} \mathrm{Exp}[- G_i / k_B T] = W_{j \rightarrow i} \mathrm{Exp}[- G_j / k_B T]$, where $W_{i \rightarrow j}$ is the transition probability from conformation i to j and $\mathrm{Exp}[- G_i / k_B T]$ is the equilibrium probability of being in state i [37]. Transition probability is the product (Metropolis criterion) × (DOF selection probability) × (bias). The Metropolis criterion obeys detailed balance [61] and DOF selection probabilities for $i \rightarrow j$ and $j \rightarrow i$ are equal and cancel from the equation. Bias favors the forward direction violating detailed balance hence the simulation is non-equilibrium Monte Carlo.

2.3. Free energy

Free-energy change, $\Delta G = \Delta H - T \Delta S$, where ΔH is enthalpy change, ΔS entropy change, and T absolute temperature. Enthalpy and entropy are computed with structure-based thermodynamics (STC, described in Supplementary material). STC free-energy change, $\Delta G_{STC} = \Delta H_{STC} - T \Delta S_{STC}$, has enthalpy computation based entirely on solute hydration while configuration enthalpy change is the dominant force shaping trajectory due to the steric barriers inhibiting backbone chain movement as shown by Brant et al. [35,44]. In contrast, configuration entropy change explicitly contributes to ΔS_{STC}. We introduced an additional term to the total enthalpy change to account for configuration enthalpy.

Total solute enthalpy change at constant pressure and volume, $\Delta H = \Delta H_{solute} + \Delta H_{solute/solvent}$, includes solute configuration (ΔH_{solute}) and hydration ($\Delta H_{solute/solvent}$). Energy computation with the Groningen Molecular Simulation program package (GROMOS, described under potential energy section heading below) force field and energy minimization procedure in the presence of solvent produces the total system energy separated into solute ($\Delta E_{G,solute}$), hydration ($\Delta E_{G,solute/solvent}$), and solvent ($\Delta E_{G,solvent}$) components. Energy and enthalpy are equitable when no external work is done such that GROMOS derived total solute enthalpy, $\Delta H_G = \Delta E_{G,solute} + \Delta E_{G,solute/solvent}$. The GROMOS and STC formalisms are rationalized by assuming that under appropriate circumstances,

$$\Delta H_{STC} = a \Delta H_G \qquad (3)$$

where a is a constant for fixed T, and Δ implies the difference between the trial structure under consideration and a standard structure. Eq. (3) holds for solved protein crystal structures because they are the basis for the STC empirical constants [62–64]. We used crystal structure representations of the skeletal myosin sequence to estimate parameter a including skeletal [12], scallop [15], smooth [14], and myosin V [65] structures. We made GROMOS total energy calculations for S1 and TNP-S1 in water then extracted $\Delta H_G = \Delta E_{G,solute} + \Delta E_{G,solute/solvent}$. The relationship of ΔH_{STC} and ΔH_G is plotted in Fig. 2 where the native M** model (1BR1) is the standard structure. Native (■) and TNP-modified (▲) S1 structures indicated a linear relationship between the GROMOS and STC computations with $a = 0.0028 \pm 0.0004$ (standard error). Linearity implies the STC enthalpy estimates compare favorably with the rigorous GROMOS estimates under
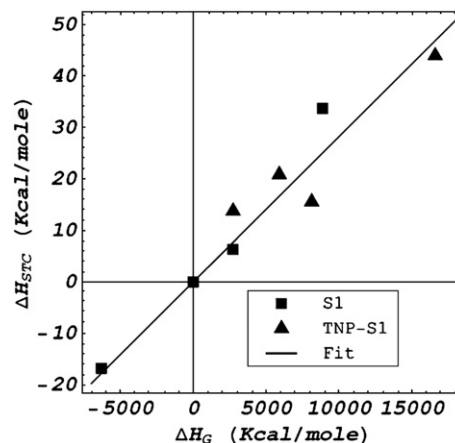

Fig. 2. ΔH_{STC} vs ΔH_G for native and TNP-modified skeletal S1 homology modeled by 2MYS [12], 1B7T [15], 1BR1 [14], and 1OE9 [65] crystal structures with the native M** model (1BR1) the standard enthalpy structure such that ΔH_{STC}(native M**) = ΔH_G(native M**) = 0. Native (■) and TNP-modified (▲) S1 structures indicated a linear relationship between GROMOS and full-STC computations. For $\Delta H_{STC} = a \Delta H_G$ (Eq. (3)) we find $a = 0.0028 \pm 0.0004$ (standard error).

these conditions. Coefficient a renormalizes GROMOS energy for combination with the STC derived quantities. We estimate the total configuration enthalpy,

$$\Delta H = \Delta H_{STC} + a \Delta E_{G,solute} \qquad (4)$$

accounting for both solvent and configuration enthalpy.

Sometimes $b \Delta E_{BMF,solute}$ replaces $\Delta E_{G,solute}$ in Eq. (4) to speed up free-energy computation. $\Delta E_{BMF,solute}$ uses the Brand, Miller, Flory force field (BMF, described under potential energy section heading below) and no structure relaxation. Coefficient b is a proportionality constant accommodating differences in force field parameters. By a method identical to that used to find a, we find $b = 1.0 \pm 0.6$. The large standard error for the 8 structures tested suggested a low significance correlation between the quantities because $\Delta E_{BMF,solute}$ is a lower accuracy energy approximation. Lower accuracy in $\Delta E_{BMF,solute}$ is inconsequential because it is used for trial triaging to eliminate structures that have serious steric backbone atom clashes and/or thermodynamic instabilities as discussed below.

Generated structures are triaged to avoid lengthy computations on trials with no chance for acceptance. The free energy,

$$\Delta G = \Delta H - T \Delta S = \Delta H_{STC} - T \Delta S_{STC} + \Delta E_{solute} \qquad (5)$$

is computed in *quick* or *full* modes. The ΔG_{quick} or ΔG_{full} estimates utilize the quick- or full-STC (quick- and full-STC described in Supplementary material) for $\Delta H_{STC} - T \Delta S_{STC}$, and, $\Delta E_{solute} = a \times b \Delta E_{BMF,solute}$ or $a \Delta E_{G,solute}$, respectively. For every trial, ΔG_{quick} is computed and subjected to the $\Delta G_{quick} \leq \Delta G_{last} + E_0$ test, where ΔG_{last} is free energy for the last structure accepted and E_0 is adjusted to admit a reasonable number of trials throughout the trajectory. Trials failing the test are not considered further although their generation contributes to the elapsed time between accepted transitions like any other rejected trial. Trials

fulfilling the test on ΔG_{quick} have ΔG_{full} computed for acceptance or rejection by Eq. (2).

Occasionally, after generating 800 trial structures, all possibilities are rejected by Eq. (2). The trials with known ΔG_{full} approximate a complete set of possible moves called the "move-set". We perform another round of configuration energy minimization on the move set and reevaluate ΔG_{full}. Position constraints on the solute atoms are gradually relaxed by this procedure to further relieve atomic clashes remaining in the structure and usually reducing ΔE_{solute} while leaving ΔH_{STC} and ΔS_{STC} essentially unchanged. The move-set structures are selected randomly with Boltzmann factor weights and retested for acceptance by Eq. (2). If all cases are again rejected by Eq. (2) the procedure repeats steps starting with configuration energy minimization. Each failed cycle adds the time increment in Eq. (1) for $N=800$ and bias$=0.5$ to total simulation time.

We found that the procedure outlined above will not always locate a free energy allowed transition because the free-energy change from ΔG_{last} is too large. In this case, ΔG_{last} is offset upward incrementally until a trial can be selected from the move set. An arbitrary time penalty is assigned (5ms for each increment offsetting ΔG_{last}) to mark the point in the trajectory where forward movement stopped.

Systematic free-energy computation inaccuracies originate from round-off-error when performing conformation change by a series of DOFs. We observed that visibly identical structures created by change in several DOFs, but where order in which the DOFs undergo transition is permuted, can alter computed free energy by as much as 20kcal/mol. We estimate a running free-energy standard deviation due to round-off-error within which each computed free energy is considered. We assume actual free energy is equivalent to any value within 1 standard deviation of its computed value throughout our trajectory calculation. Free-energy standard deviation never exceeded 20kcal/mol (10–15kcal/mol typical).

2.4. Potential energy

Protein (solute) conformational energy was computed in two ways designated $E_{BMF,solute}$ and $E_{G,solute}$. $E_{BMF,solute}$ used the sum of torsion, Coulomb, and van der Waals potentials as described previously by Brant et al., 1967 [35,44]. Energy evaluation was rapid for triaging trial structures. $E_{G,solute}$ used the GROMOS force field (43A1 and 43B1 for molecules in solution and in *vacuo*, respectively) and software for energy minimization [66]. We amended the GROMOS topology library file for energy minimization of myosin containing TNP-Lys using directions in the documentation. Energy minimization improved trial structures by reducing effects of minor or avoidable steric clashes while restraining all bond lengths and bond angles using the SHAKE procedure.

2.5. Homology modeling

MC simulation requires crystal structures with a common peptide sequence for the intermediates serving as endpoints in the trajectory, in this case M and M**. We choose the skeletal myosin

isoform for which the M conformation structure comes directly from the first myosin crystal structure of chicken skeletal myosin [12] while M** is from smooth muscle myosin [14]. Homology modeling was done with Modeller 8.2 [67] on MHC and ELC. The skeletal myosin target sequence, consisting of the first 812 residues from MHC (sequence identifier gi13432175) and all 149 residues from ELC (gi1942540), was folded into the M and M** templates, 2MYS and 1BR1, respectively. Folding the target with the M template permitted replacement of methylated lysines in the crystal structure sequence with native lysines. Smooth muscle MHC (gi3915778) and ELC (gi5542589) targets were likewise folded into the M** template for assessing modeling accuracy.

Accurate model building depends on target and template sequence identity and alignment. MHC skeletal target and smooth muscle template sequences share 48% identity suggesting Modeller produces a reliable M** structure from the skeletal sequence [68]. Light chain sequence identity is higher at 60%. We did alignment using the NCBI BLAST 2 sequences protocol (http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi). Unstructured loops on the surface of myosin were not resolved in the crystal structures. Modeller added energy minimized surface loops to complete the structures for M and M**. Other templates utilized with the skeletal myosin target sequence in the course of this work include scallop (1B7T) and myosin V (1OE9) atomic structures [15,65]. We amended Modeller library files for the trinitrophenylated lysine residue (TNP-Lys) using directions in the documentation.

Homology model accuracy was evaluated using the Discrete Optimized Protein Energy (DOPE) score as suggested in the Modeller tutorial. The DOPE score is proportional to potential energy per residue, smoothed over a 15 residue window, and normalized by the number of restraints acting on each residue. It indicates problem regions in the homology model when profiles are compared from target and template sequences. Fig. 3A compares DOPE profiles for skeletal and smooth muscle sequences consisting of 961 residues from truncated MHC and ELC used in this study and folded with the 1BR1 template. The DOPE profiles are comparable over the entire peptide suggesting the skeletal template was accurately folded. Fig. 3B compares DOPE profiles for skeletal and TNP-modified skeletal sequences folded with the 2MYS template. TNP modification of K84 causes little perturbation in the DOPE scores indicating a reasonable accommodation of the larger side chain. Similar results were obtained for the TNP-modified skeletal sequence folded with the 1BR1 template (data not shown).

2.6. Molecule alignment and visualization

CCP4 contains subroutine Lsqkab aligning two protein structures by rotation and translation [69]. Lsqkab aligns MHC residues 145 to 361 of each transient structure with the starting myosin crystal structure to keep the protein in an approximately constant relationship with a lab fixed frame. These residues make up the U50k portion of the actin-binding site such that myosin's main actin-binding site remains fixed in space. The resulting trajectory simulates a power stroke where the main actin-binding site does not change relationship to actin.
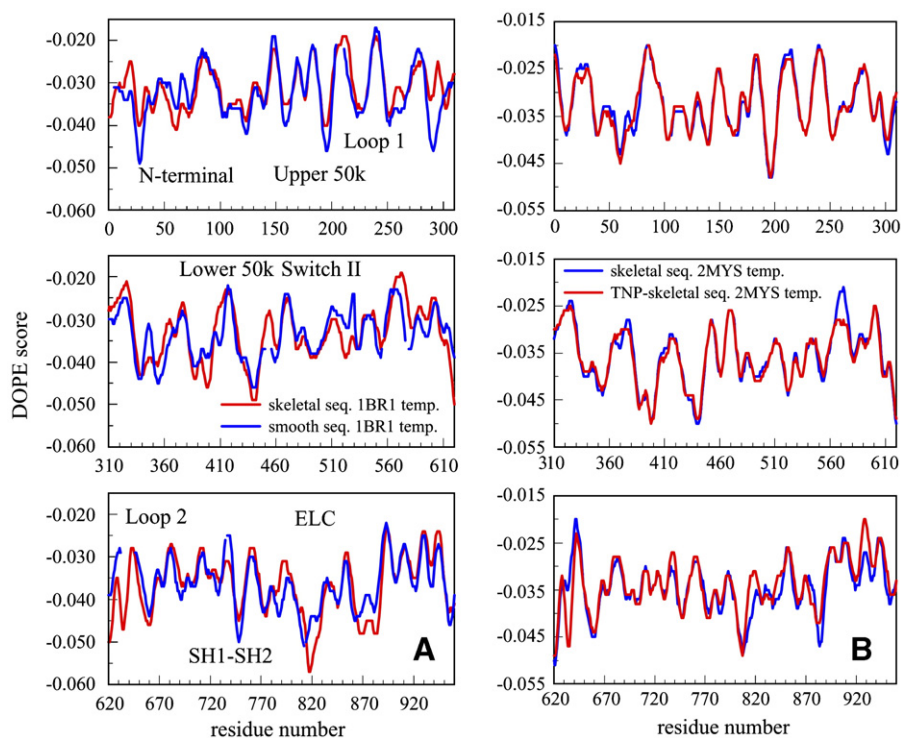
Fig. 3. The Discrete Optimized Protein Energy (DOPE) scores vs residue number for skeletal and smooth muscle myosin sequences. Panel A: the DOPE score profile for sequences folded with the 1BR1 template. The smooth muscle myosin sequence DOPE score (blue) is the best possible case to which the skeletal sequence score (red) is compared. Several portions of the peptides are identified by their common name. The DOPE profiles are comparable over the length of the peptide suggesting the skeletal template was accurately folded. Panel B: the DOPE score profile for native (blue) and TNP-modified (red) skeletal myosin sequences folded with the 2MYS template. The TNP group modifying K84 does not degrade the DOPE profile indicating a reasonable accommodation of the larger side chain into the homology modeled structure. Accommodation of TNP has a mild effect felt throughout the folded structure including ELC.

Visual M** → M trajectories were created in Visual Molecular Dynamics (VMD) [70] then rendered in POV-Ray (http://www.povray.org/) and output as one bitmap file for each frame. Rendered frames were combined with an audio description of the myosin dynamics in Adobe Premiere Pro video editing software. WMV (Windows Media Video) format movies are available for downloading in Supplementary material.

2.7. Optical spectroscopy

Ground and excited state spectroscopic signals from the optically active side chains in native or TNP-modified myosin after interaction with each other and the peptide backbone are computed for each intermediate in the conformation trajectory. We calculate peak absorption energy, dipole strength, and rotatory strength (ground state spectroscopy) for W510 or TNP in TNP-S1 as described previously [39–41,71]. W510 electron transfer rate (excited state spectroscopy) to the backbone or other side chains is derived after the method suggested by Adams et al. and described in the Supplementary material [72]. Acrylamide quenching (excited state spectroscopy) is assesses by tryptophan ASA.

Ground state signals were calculated by the matrix method [73]. The latter calculation combines the (n,π^*) and (π,π^*) transitions of the peptide amide groups, with observable (π,π^*) transitions from tryptophan, tyrosine, phenylalanine, histidine, and TNP-Lys side chains. The Coulomb potential, V, with vacuum

dielectric constant couples the transition monopoles from interacting groups. Positions and values for the ground state and transition monopole charges for the peptide primary, secondary, and tertiary amide groups, and, tryptophan, tyrosine, phenylalanine, and histidine side chains were taken from the literature [74–79]. In addition, transition monopoles amplitudes for $(0,^1L_b)$ and $(0,^1L_a)$ in tryptophan, or $(0,^1L_b)$ in tyrosine, were normalized to reproduce the dipole strength we observed from N-acetyl-L-tryptophanamide (NATpA) or N-acetyl-L-tyrosineamide (NATrA) in aqueous buffer at pH 7 [71]. Positions and values for ground state and transition monopole charges for the TNP-Lys side chain were computed as described previously [41].

2.6. Granger-causality relationships in myosin dynamics

A statistical definition of causality between two time-dependent stationary variables, $X(t)$ and $Y(t)$, was introduced by Granger [48]. $Y(t)$ is Granger-causal for $X(t)$ if prediction of $X(t)$ is statistically improved by inclusion of past information from both $X(t)$ and $Y(t)$ compared to a prediction based solely on past values of $X(t)$ while the converse does not hold. When the converse also holds the relationship is not causal but mutually dependent on a third variable. The lag time in the relationship is the optimal time length extending into the past for use in prediction. We test Granger causality using R statistical software [80] with the MSBVAR package (http://www.R-project.org). For native S1, five time-varying parameters {P,L,A,D,F} were tested

where $P(t)$ is back door displacement signaling commitment to ATP turnover, $L(t)$ is lever-arm displacement linear with work production, $A(t)$ is W510 ASA, $D(t)$ is W510 dipole strength in the indole ^1L_a and ^1L_b transitions, and $F(t)$ is the W510 electron transfer rate. For TNP-S1, four time-varying parameters {P,L,A, C} were tested where C(t) is rotatory strength in TNP modifying K84 in the TNP transition at ∼420 nm.

Tests are applied two variables at a time (bivariate) using a linear least squares model. Lag time, estimated by Akaike information criterion minimization, was the same for all data sets at ∼10 μs. Multivariate causality testing was also done using partial directed coherence (PDC) with results qualitatively confirming those from the bivariate tests [81].

Causality testing is designed for stationary time-varying signals with constant mean and variance over the time domain. This is rarely the case for real applications. With the occasional exception of the electron transfer rate (F), signals tested are nonstationary by inspection and by the Phillips–Perron unit root test (available in R). The proclivity for causality false positives due to nonstationarity or any other artifact was estimated by testing causality on unrelated control signal sets. N myosin simulations were performed producing set {P_i,L_i,A_i,D_i,F_i} for i=1, …, N. For the ith set, we causality tested pairwise relationships between time-varying signals from one myosin simulation. An example is causality testing of the pair {P_i,L_i}. Unrelated signal sets were constructed by combining time-varying signals from different simulations. An example is causality testing of the pair {P_i,L_{i+1}}. There are N independent trials of related pairs and $N(N-1)$ trials of unrelated pairs. The unrelated pairs contain the unwanted features introduced by nonstationarity and any other artifacts without the causal coherence introduced by the myosin structure preserved in a single myosin simulation.

Causality significance testing used the F-statistic with the null hypothesis (that the time-varying signals are not causal) rejected at the 5% significance level. Unrelated pair results produced the probability density for false positives. The false positive probability densities varied noticeably among the different signal pairs. Related pair test results were compared to their corresponding false positive probabilities with the null hypothesis (that there is no difference between related and unrelated pairs) rejected at the 5% significance level.

3. Results

3.1. General trajectory characteristics

Myosin atomic structures in a time series represent the M**→M conformation trajectory. There were 37 trajectories simulated, 20 for native S1 and 17 for reactive lysine residue (K84) trinitrophenylated S1 (TNP-S1). Fig. 4 indicates room temperature energetics where ΔG (♦), ΔH (■), and $T\Delta S$ (▲) are shown for native- and TNP-S1. Block transitions occur at time positions identified in the figure as Nterm, Lever, Conv, U50k, Loop2, and L50k. For native-S1 (Panel A), the U50k transition substantially decreases entropy [$\Delta(\Delta H)=-43.0$, $T\Delta(\Delta S)=-81.4$, $\Delta(\Delta G)=38.4$ kcal/mol, where $\Delta()$ is the difference between
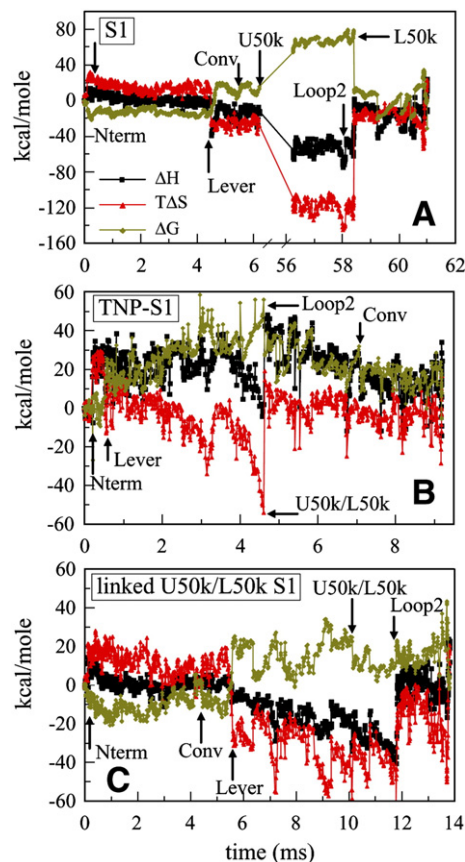

Fig. 4. Conformation trajectory free-energy ΔG (♦), enthalpy ΔH (■), and temperature in Kelvin times entropy $T\Delta S$ (▲) all at room temperature for native S1 (Panel A), TNP-S1 (Panel B), and native S1 with U50k and L50k block transitions constrained to occur simultaneously (Panel C). Values are related at each time point by $\Delta G=\Delta H-T\Delta S$. Block DOFs occur at time positions roughly identified by Nterm (AA1–144), U50k (AA145–361), L50k (AA362–462), Loop2 (AA526–695), Conv (AA721–772), and Lever (AA773–812). For the TNP-S1 simulation (Panel B), block transitions U50k and L50k occur simultaneously.

the first post block-transition structure and the preceding structure], while L50k substantially increases entropy [$\Delta(\Delta H)=29.4$, $T\Delta(\Delta S)=97.5$, $\Delta(\Delta G)=-68.1$ kcal/mol]. Changes in enthalpy and entropy for the U50k and L50k transitions are compensatory. For TNP-S1 (Panel B), simultaneous U50k and L50k transitions (U50k / L50k) make a substantial decrease in ΔG [$\Delta(\Delta H)=44.4$, $T\Delta(\Delta S)=73.3$, $\Delta(\Delta G)=-28.9$ kcal/mole] with entropy dominant in the free-energy change. Fig. 4 (Panel C) shows a native-S1 trajectory where U50k and L50k transitions are constrained to occur simultaneously. It imitates TNP-S1 energetics.

With no exception, native-S1 trajectories have identical block-transition chronology in Nterm, U50k, Loop2, and L50k. Conv and Lever transitions usually occur between the Nterm and U50k transitions but there are exceptions. The S1 trajectories all have a time course interrupted by the large entropic free-energy barrier seen in Fig. 4 (Panel A). TNP-S1 block transitions are similarly ordered but more variable in chronology. TNP-S1 trajectories usually avoid the free-energy barrier by combining U50K and L50k transitions as shown in Fig. 4 (Panel B) although there are

instances (7 out of 17) where these transitions occur separately. With U50k and L50k combined, the time course for the TNP-S1 M∗∗→M trajectory is ∼ 10 ms. Constraining the U50k and L50k transitions to occur simultaneously in native S1 (Fig. 4 **Panel C**) also eliminates the free-energy barrier in native S1 although the time courses are ∼ 50% longer that their TNP − S1 counterpart. It has been observed that modification of S1 with TNP increases MgATPase 20-fold [82,83]. The simulations anticipate this trend by suggesting TNP-S1 efficiently maneuvers around the entropic free-energy barrier inhibiting native-S1 M**→M.

Fig. 4 shows a free-energy barrier of ∼ 75 kcal/mol separating initial and final states of the myosin conformation trajectory due to the U50k transition. This free-energy barrier significantly exceeds energy liberated by ATP hydrolysis theoretically prohibiting myosin ATPase at room temperature. In vitro, myosin has measurable MgATPase in these conditions. The discrepancy is probably not due to inaccuracy in free-energy estimation since our method accurately represents free energy from myosin crystal structures (see Methods under Free energy). The discrepancy is certainly due in part to the selected order of the elementary transitions (see Methods under *sampling*). To rule out poor selection of block-peptide segments within S1 as a source of the discrepancy, block-peptides were tested for structure conserving behavior by running simulations where they were subdivided. We placed three subdivisions in Nterm, four in U50k, and one in L50k. We did not subdivide Conv or Lever domains. Subdivision took place at glycine residues in ∼ 50 residue intervals when minimal backbone swiveling occurred at the division point. For choices tested, domain subdivision had little effect on trajectory pathway because it did not lower S1 free energy. The Nterm made block transitions simultaneously in all subdivisions suggesting it is truly a functional unit. We expect subdivision of established domains will be advantageous to finding more favorable minimum free-energy pathways, however, further testing of sub-domain hinge residues will be needed to identify them.

Fig. 5 shows an average over the 20 native-S1 (top) and 17 TNP-S1 (bottom) simulations of $\Delta(\Delta H)$, $T\Delta(\Delta S)$, and $\Delta(\Delta G)$ for the block and gear peptide transitions. In native-S1, the U50k transition is the free-energy barrier to the overall M**→M transition. L50k has lowest (and negative) $\Delta(\Delta G)$ driving the forward reaction. Nterm, Conv and Lever block transitions have balanced enthalpic and entropic parts thereby minimizing free energy while Loop 2 produces little change in entropy or enthalpy. The Gear peptide is predominantly enthalpic. The free-energy profile for TNP-S1 (Fig. 5, bottom) has enthalpy, entropy, and free-energy distributions all perturbed relative to the native-S1. We find that modification of K84 in the Nterm peptide instigates global structural effects that preserve the ability of the protein to hydrolyze ATP but disturbs its regulation.

3.2. Hydrolysis and lever-arm swing synchronization

Structures in the protein conformation trajectory were mined for information characterizing the energy transduction mechanism. We correlated work producing lever-arm translation with opening of the Arg245/Glu468 active site salt-bridge acting as the back door for phosphate release [84]. Fig. 6 shows the time
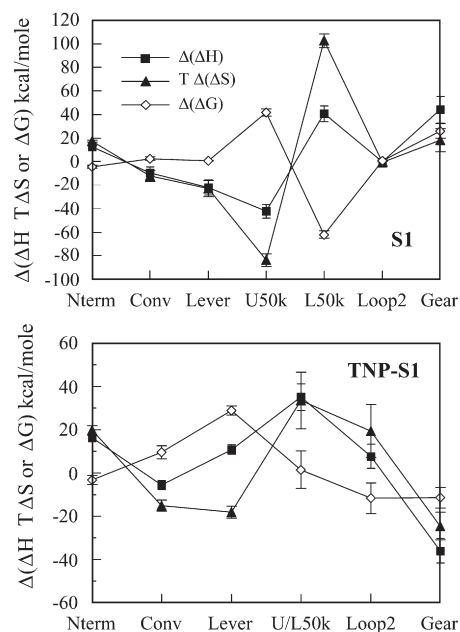

Fig. 5. The average free-energy change, $\Delta(\Delta G)$ (◇), enthalpy, $\Delta(\Delta H)$ (■), and temperature in Kelvin times entropy, $T\Delta(\Delta S)$ (▲), changes for transitions in native S1 (top) and TNP-S1 (bottom). $\Delta()$ is the difference between the first post block-transition structure and the preceding structure except for the gear-peptide where $\Delta()$ is the difference between the total enthalpy, entropy, or free-energy change of the reaction and the sum of these quantities for the block transitions.

dependence for free-energy change, work producing lever-arm displacement, and back door opening for S1 and TNP-S1 during the M**→M transition, from the simulations characterized in Panels A and B in Fig. 4. Normalized displacement rescales the difference between initial (M**) and final (M) positions to 1. Normalized ΔG rescales its peak value to 1. Normalizations facilitate comparison of ΔG time-dependency with mechanical events in the myosin cycle.

In the native protein, the U50k transition lowers both enthalpy and entropy but entropy dominates to increase free energy. The entropy decreasing transition opens the back door and initiates lever-arm movement. All native S1 conformation trajectories confirm that during ATPase the entropy dominated U50k transition is a free-energy barrier whose height makes it rate limiting ensuring that product release is rate limiting and prior to force production by the lever-arm swing in agreement with the conventional myosin mechanism.

Compensating the U50k free energy increasing transition is the L50k free energy decreasing transition. The U50k and L50k peptides form a large part of the actin-binding site. The free-energy barrier presented by U50k, holding the back door closed and preventing lever-arm swing in native S1, suggests a scenario for the role of actin binding in actin-activated ATPase. Actin binding to its binding site at the U50k and L50k junction lowers the free-energy barrier by compensating the U50k free-energy increase with the L50k free-energy decrease. This is accomplished by facilitating their simultaneous transition. A constrained S1 trajectory required the simultaneous transition of the U50k and L50k peptides (Fig. 4 Panel C) and confirmed this mechanism for removing the free-energy barrier to product release.
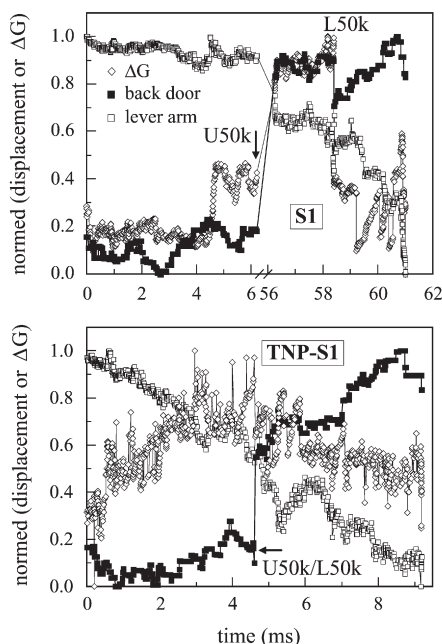
Fig. 6. Synchronization of free-energy change, ΔG ($\diamond$), work producing lever-arm displacement ($\square$), and back door opening ($\blacksquare$), for S1 and TNP-S1 during the M**→M transition. For the TNP-S1 simulation (bottom panel), block transitions U50k and L50k occur simultaneously. Data taken from the simulations in Panels A and B of Fig. 4. Normalized displacement rescales the difference between initial (M**) and final (M) positions to 1. Normalized ΔG rescales the peak value to 1.

In TNP-S1, coincidence of the U50k transition with back door opening is intact but mutual coincidence with initiation of the lever-arm movement is lost. Fig. 6 bottom shows that when the back door opens, the lever arm has already completed ∼45% of its swing. Other simulations of TNP-S1 showed detachment of back door and lever-arm movement but where lever-arm swing occurred much later than back door opening. We conclude that TNP modification of S1 decouples lever-arm swing from product release. It has been proposed that TNP modification of S1 emulates actin binding to myosin. The modification increases S1 MgATPase ∼20-fold and TNP-S1 cannot be actin-activated [85]. It seems apparent from our simulation that actin binding can not affect MgATPase of TNP-S1 because lever-arm movement is not inhibited by structural transitioning in U50k. Some TNP-S1 simulations (including the one shown in Fig. 6) had lower overall free-energy barrier to the M**→M transition because of the simultaneous transition of the U50k and L50k peptides. If the simultaneous transition of the U50k and L50k peptides is how actin activates ATPase then TNP modification emulates actin binding to myosin. Modification of K84 in the Nterm peptide instigates global structural effects that preserve the ability of the protein to hydrolyze ATP but disturbs its regulation and mechanochemical coupling.

Neighboring U50k and L50k peptides when uncoupled present a free-energy barrier equal to the activation energy of U50k that is eliminated by their coupling possibly by interaction with actin. The latter mechanism may represent a general design for a de-activatable free-energy barrier where in this case the reaction is myosin ATPase and the catalyst is actin.

Downloadable S1 trajectory movies in Supplementary material depict trajectories characterized in Fig. 4 Panels A and B. The movie for native S1 demonstrates synchronization of the U50k block transition and initiation of productive lever-arm movement. The TNP-S1 trajectory illustrates dysfunctional protein dynamics where productive lever-arm movement precedes the U50k transition.

3.3. W510 spectroscopy

ATP-sensitive tryptophan k_{ET} and ASA evaluate how static quenching groups in the protein matrix and solution borne quenchers in a Stern–Volmer experiment influence W510 fluorescence lifetime and intensity. Fig. 7 shows k_{ET} and ASA computed as a function of lever-arm displacement during force production for trajectories characterized in Fig. 4 Panel A but are characteristic to all S1 simulations. We find that k_{ET} uniquely characterizes work producing displacement of the lever arm except at the beginning of the power stroke and that it decreases during force production to a minimum at the end of the power stroke. Consequently, W510 fluorescence lifetime is expected to be longest in the absence of nucleotide (apo S1) and shortest in nucleotide bound S1 if electron transfer is the predominant quenching mechanism. W510 ASA uniquely characterizes all of the work producing displacement of the lever arm and increases during force production to a maximum at the end of the power stroke. A specific example of this phenomenon is demonstrated in the native S1 simulation in the Supplementary material. There the W510 side chain is highlighted in yellow during the simulation and clearly increases ASA during the course of the M**→M trajectory. The Stern–Volmer quenching constant for W510 fluorescence intensity, due to solution borne dynamic tryptophan quenchers, is proportional to the square root of ASA and expected to be largest in apo S1 and smallest in nucleotide bound S1.

Table 1 summarizes experimental lifetimes, τ, and Stern–Volmer acrylamide quenching constants, K_{SV}, for W510 in several S1 conformations. Acrylamide is a solution borne, dynamic quencher of W510 [46]. W510 fluorescence lifetime is longest in apo S1 and shortens significantly in the presence of nucleotide or trapped nucleotide analogs. Simulated k_{ET} (Fig. 7)
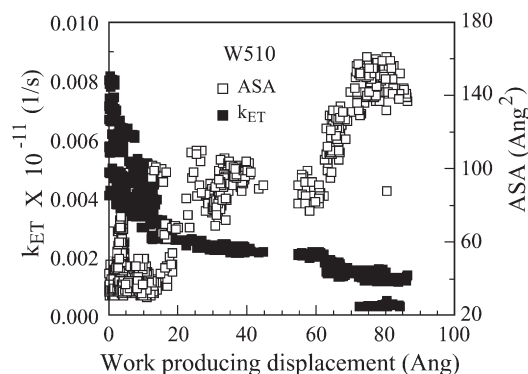

Fig. 7. W510 electron transfer rate, k_{ET} ($\blacksquare$), and accessible surface area, ASA ($\square$), vs work producing displacement during the M**→M transition. The work producing displacement is MHC Ser812 position projected onto the actin filament axis.

and these experimental data shows identical trends suggesting electron exchange with static quenching groups in the protein matrix is the quenching mechanism. The simulation identifies the Y503 phenol side group and backbone amide groups from W510, E511, K709, and G710 as the main electron exchange partners. The influence of the Y503 phenol side group on W510 fluorescence sensitivity to nucleotide or trapped nucleotide analogs was recently confirmed experimentally using site directed mutagenesis of a smooth muscle myosin construct [86]. K_{SV} shows acrylamide quenching is most effective in apo S1 and least effective in the presence of nucleotide or trapped nucleotide analogs reflecting maximal and minimal W510 ASA in the apo and trapped nucleotide analog S1. Simulated W510 ASA (Fig. 7) shows identical results.

3.4. Granger-causal relationships in S1 and TNP-S1

Statistical causality testing was performed on several time-varying parameters extracted from the MC generated trajectories. For S1 they include: (i) $P(t)$, the distance between residues forming the back door salt-bridge (Arg245/Glu468) that opens for product release and commitment to ATP turnover, (ii) $L(t)$, C-terminus translation along the actin filament equaling work producing lever-arm displacement, (iii) $A(t)$, W510 ASA, (iv) $D(t)$, W510 absorption dipole strength in the two lowest energy indole transitions (^1L_a and ^1L_b), and (v) $F(t)$, W510 electron transfer rate. For TNP-S1, $C(t)$, the rotatory strength from the lowest energy transition of bound TNP absorbing light at ~ 420 nm, replaces optical signals $D(t)$ and $F(t)$ from W510.

Parameters (circles containing the one letter codes) and significant causal relationships (arrows) linking the parameters are indicated in Fig. 8 for S1. TNP-S1 parameters are not causally related. In S1, we find that active site commitment to ATP turnover (P) is causal for lever-arm movement leading to work production (L) and for W510 dipole strength (D). The notion that propagating conformation change within the protein matrix facilitates active site control of distant sites is central to the
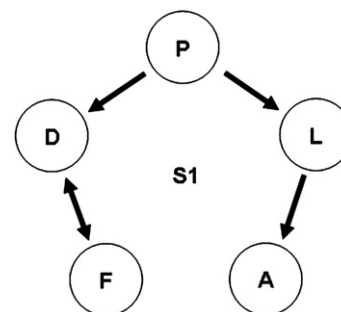

Fig. 8. Statistical causality relationships among the five time-varying parameters {P,L,A,D,F} generated in native S1 with MC simulation where $P(t)$ is back door displacement signaling commitment to ATP turnover, $L(t)$ is lever-arm movement proportional to work production, $A(t)$ is W510 ASA, $D(t)$ is W510 dipole strength in the indole ^1L_a and ^1L_b transitions, and $F(t)$ is the W510 electron transfer rate. No significant causality relationships were detected in the four time-varying parameters {P,L,A,C} generated in TNP-S1 where $C(t)$ is rotatory strength in TNP modifying K84 in the TNP transition at ~ 420 nm (not shown).

modern formulation of the molecular mechanism of muscle contraction [1]. Our results suggest the causal relationships identify propagating conformation change within the protein matrix [87].

Lever-arm movement (L) causal for W510 ASA (A) indicates lever-arm position inhibits W510 acrylamide accessibility in the pre-powerstroke myosin rather than the effect being under active site control. The control mechanism contrasts with the P→D causality suggesting W510 ATP sensitivity is under active site control. Thus the W510 site is structurally perturbed by two distinct mechanisms, one by its proximity to the moving lever arm, the other by its proximity to a propagating conformation change pathway emanating from the active site.

The bi-directional arrow connecting W510 electron transfer rate (F) and dipole strength (D) indicates they are not causal but mutually dependent on an unidentified parameter.

4. Discussion

Dynamic molecular models relating protein structure and function provide a basis for interpreting experimental observation and have a role in stimulating experimental investigation by inexpensively testing ideas for feasibility. Frequently physically interesting temporal and spatial domains are impractical to model with rigorous MD. For example, kinetics might run on a millisecond clock while rigorous MD simulation ends after several nanoseconds. Coarse grained MD might treat individual amino acid residues as identical particles while site directed mutagenesis modifying a single residue side chain demonstrates altered protein function. In this MC application temporal and spatial resolution scales to match the application and provide the modeling tools needed.

The MC simulation utilizes elementary DOFs defining amino acid configuration on the atomic level. The elementary DOFs include Ramachandran angles and other dihedral angles in an amino acid residue that often undergo large change, and, collectively all other parameters such as bond angles and lengths that in comparison remain within tightly constrained limits during

Table 1
ATP-sensitive W510 fluorescence lifetime, τ, and Stern–Volmer quenching constant, K_{SV}[a]

Species	S1 form	τ (ns)[b]	K_{SV} (M^{-1})[c]
Apo S1	M	9.4±0.2(5)	7.6±0.2(5)
S1.MgADP	M^	9.2±0.2(4)	7.1±0.03(4)
S1.MgADP·BeF$_x$	M*	8.0±0.2(2)	6.5±0.3(4)
S1.MgATP	M**	7.2±0.3(4)	6.11±0.04(4)
S1.MgADP·AlF$_4^-$	M**	7.0±0.3(3)	4.8±0.3(4)

[a] All errors are standard error of the mean for (n) measurements. S1 conformation forms are defined in Scheme 1. Nucleotide or nucleotide analogs are bound or trapped in the S1 active site to induce statically the S1 conformation form appearing in Scheme 1.

[b] Lifetimes taken from Park & Burghardt, 2002 [42].

[c] Stern–Volmer constants for acrylamide quenching of W510 defined by,

$$\frac{F_0}{F} = 1 - K_{SV}[Q]$$

for fluorescence intensity F, and acrylamide concentration [Q], are taken from Park et al., 1996 [46].

protein functions. Elementary DOFs can undergo dynamics individually to provide the highest spatial and temporal resolution or are collected into block DOFs that undergo two state isomerization. In functioning proteins, blocked domains undergo little or peripheral collective conformation change. Suitable choices of elementary- and block DOFs provide the appropriate time and spatial resolution for all regions of the system.

Simulated dynamics of driven protein conformation, as opposed to equilibrium dynamics where the structure fluctuates near a free-energy minimum, has conformation change along a free-energy minimum pathway driven by impulse. Our driving force is implemented as a biased random walk in the DOFs. External force amplitude is adjusted by deviation of the bias from random. The specific mechanism for driving force implementation resides in the DOF selection probabilities. It is generally recognized that nucleotide hydrolysis supplies the driving impulse but we made no attempt to implement a specific mechanism. Instead we apply a logical, but ad hoc, rule for deciding DOF selection probabilities governing how conformations are sampled. We demonstrated that our free energies (computed using structure-based thermodynamics calculations and a standard force field) are accurate because they agree with rigorous total energies of solute (myosin) and solvent (water) computed for myosin crystal structures (see Methods).

Results for our non-equilibrium MC application to motor protein dynamics were compared with known structure-function relationships summarized in Scheme 1. In the absence of actin, M**·ADP·Pi is the predominant intermediate. During contraction, actin binds weakly to M**·ADP·Pi accelerating Pi release. Force production occurs when Pi release is followed by strong actin binding and lever-arm swing. The MC simulation of the M**→M transition shows that product release precedes lever-arm swing and that both processes are inhibited by the large free-energy barrier to conformation change in the U50k actin-binding domain of myosin. Thus the MC simulation captures the essence of motor function by providing appropriate coupling of the predominant intermediate to product release and work producing lever-arm movement. The simulation couples lever-arm performance to the U50k transition and shows that a second transition in an actin-binding domain (L50 k) favorably compensates free energy suggesting that actin binding activates ATPase by causing simultaneous U50k and L50k transitions thus lowering the free-energy barrier to product release. MC simulation of trinitrophenylated myosin predicts acceleration of its basal MgATPase by U50k and L50k transition synchronization. It was suggested that myosin trinitrophenylation imitates actin-activation [85] and the simulation supports this suggestion and gives it a molecular rational. Coupling of neighboring free-energy-complementary peptides like U50k and L50k may represent a general design for a de-activatable free-energy barrier. It is unclear for now whether the entropy dominance of free-energy change in U50k and L50k has significance.

We investigated detectable spectroscopic characteristics of the work producing transition. Tryptophan k_{ET} and ASA evaluated how static quenching groups within the protein matrix and solution borne quenchers in a Stern–Volmer experiment influenced W510 fluorescence lifetime and intensity. We demonstrated functional relationships between lever-arm displacement and k_{ET} or ASA for W510 that are consistent with experimental measurements from skeletal myosin (Fig. 7 and Table 1). The simulation identifies the Y503 phenol side group and backbone amide groups from W510, E511, K709, and G710 as the main electron exchange partners. We have recently directly confirmed the W510 and Y503 phenol side group interaction leading to W510 fluorescence intensity perturbation using site directed mutagenesis [86].

Rotatory strength of TNP in TNP-S1 (a measure of CD signal sign and amplitude) senses conformation change in this location consistent with its interaction with the lever arm [39,40]. K84 lies at the interface of the N-terminal domain and lever arm. We hypothesized that TNP attachment to K84 introduces steric clashes between the regions. This hypothesis was confirmed by simulation and observed in experiment [44]. Visual inspection of the TNP-S1 M**→M trajectory (Supplementary material) indicates collision of the TNP group and lever arm during the powerstroke.

Myosin structure-function characteristics in the M**→M transition just outlined are reinforced by statistical causality results summarized in Fig. 8. P→L causality is expected if ATP hydrolysis drives conformation change in myosin leading to work production. Similarly, P→D↔F causality/mutual dependency is expected given the observation that W510 absorption and fluorescence senses active site conformation during hydrolysis. Unexpectedly, statistical causality separates mechanisms for W510 absorption/fluorescence enhancement from its quenching sensitivity. The former is due to W510 proximity to a propagating conformation change pathway emanating from the active site. The latter is due to W510 proximity to the moving lever arm affecting solvent accessibility. Statistical causality is a new tool to mine information content from protein dynamics simulations.

The M**→M transition corresponds to large remodeling of protein structure central to the energy transduction mechanism. Remodeling mechanics have been described most often in the broad terms of helices and sheets melting, rotating, translating, or bending and based on judgments made by comparing M and M** structure representations from different myosin isoforms [4,55]. The MC simulation characterizes the protein remodeling as a function of time producing a visual record of single molecule and single isoform dynamics on the millisecond time scale. It accurately accounts for structure-function relationships in myosin suggesting predictive aspects of the simulation should be exploited in hypothesis driven investigation.

Future applications for myosin MC simulation will include investigation of functional changes induced by mutation. The search for the molecular basis of heart disease using gene sequencing has shown that many myosin motor domain mutations occur in people with heart disease [88–90]. Regions in myosin associated with actin binding, energy transduction, and ATP binding/hydrolysis hosted disease implicated mutations suggesting their location identifies important functional residues. Experimental results have demonstrated interesting and contradictory relationships between myosin functional alteration and mutation [91–94]. Expressed proteins are typically model myosin isoforms [95] not including cardiac because it is difficult to

express [91]. Isoform specific investigation of the functional role of sensitive residues can be done in silico with MC simulation.

5. Conclusion

Free energy minimized energy transduction conformation trajectory for S1 undergoing the $M^{**} \rightarrow M$ state transition was visualized for the native and TNP-modified skeletal myosin using a non-equilibrium MC approach. Simulated results were compared with known structure-function relationships in the protein. MC simulation of native-S1 indicates product release precedes and initiates lever-arm swing thus capturing the essence of the motor function. The rate limiting step for product release is defined by an entropic barrier localized in the actin-binding site. MC simulation of TNP-S1 correctly anticipates its higher MgATPase and lack of in vitro motility due to destabilization of the rate limiting step, the decoupling of product release from lever-arm swing, and collision of TNP with the lever arm. The location of the entropic barrier to product release and the mechanism for its destabilization suggests a general method for lowering activation energy in an enzyme catalyzed reaction. MC simulation applied to tryptophan k_{ET} and ASA correctly evaluated how static quenching groups within the protein matrix and solution borne quenchers in a Stern–Volmer experiment influenced W510 fluorescence lifetime and intensity. Statistical causality applied to several time-dependent signals generated in the simulations identifies two distinct mechanisms at work in sensitizing native W510 emission to the protein conformation state. Statistical causality is a promising new data mining tool for discovery of propagating conformation change pathways in proteins.

Acknowledgements

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bpc.2007.08.008.

References

[1] M.F. Morales, J. Botts, Proc. Natl. Acad. Sci. USA 76 (1979) 3857–3859.
[2] H. Onishi, N. Mochizuki, M.F. Morales, Biochemistry 43 (2004) 3757–3763.
[3] T. Ohki, S.V. Mikhailenko, M.F. Morales, H. Onishi, N. Mochizuki, Biochemistry 43 (2004) 13707–13714.
[4] M.A. Geeves, R. Fedorov, D.J. Manstein, Cell. Mol. Life Sci. 62 (2005) 1462–1477.
[5] T. Yanagida, S. Esaki, A.H. Iwane, Y. Inoue, A. Ishijima, K. Kitaura, H. Tanaka, M. Tokunaga, Philos. Trans. R. Soc. Lond., B 355 (2000) 441–447.
[6] R.S. Goody, W. Hofmann, J. Muscle Res. Cell Motil. 1 (1980) 101–115.
[7] Y.M. Peyser, K. Ajtai, T.P. Burghardt, A. Muhlrad, Biophys. J. 81 (2001) 1101–1114.
[8] A.J. Fisher, C.A. Smith, J.B. Thoden, R. Smith, K. Sutoh, H.M. Holden, I. Rayment, Biochemistry 34 (1995) 8960–8972.
[9] M.M. Werber, Y.M. Peyser, A. Muhlrad, Biochemistry 31 (1992) 7190–7197.
[10] Y.M. Peyser, K. Ajtai, M.M. Werber, T.P. Burghardt, A. Muhlrad, Biochemistry 36 (1997) 5170–5178.
[11] S. Maruta, G.D. Henry, B.D. Sykes, M. Ikebe, J. Biol. Chem. 268 (1993) 7093–7100.
[12] I. Rayment, W.R. Rypniewski, K. Schmidt-Base, R. Smith, D.R. Tomchick, M.M. Benning, D.A. Winkelmann, G. Wesenberg, H.M. Holden, Science 261 (1993) 50–58.
[13] C.A. Smith, I. Rayment, Biochemistry 35 (1996) 5404–5417.
[14] R. Dominguez, Y. Freyzon, K.M. Trybus, C. Cohen, Cell 94 (1998) 559–571.
[15] A. Houdusse, V.N. Kalabokis, D. Himmel, A.G. Szent-Gyorgyi, C. Cohen, Cell 97 (1999) 459–470.
[16] C.R. Bagshaw, D.R. Trentham, Biochem. J. 141 (1974) 331–349.
[17] S.M. Sine, J. Neurobiol. 53 (2002) 431–446.
[18] T.J. Minehardt, N. Marzari, R. Cooke, E. Pate, P.A. Kollman, R. Car, Biophys. J. 82 (2002) 660–675.
[19] J.D. Lawson, E. Pate, I. Rayment, R.G. Yount, Biophys. J. 86 (2004) 3794–3803.
[20] D.D. Root, Cell Biochem. Biophys. 37 (2002) 97–110.
[21] Y. Liu, M. Scolari, W. Im, H.-J. Woo, Proteins: Struct., Funct., and Bioinformatics 64 (2006) 156–166.
[22] T. Kawakubo, O. Okada, T. Minami, Biophys. Chemist. 115 (2005) 77–85.
[23] P. Gros, W.F. van Gunsteren, W.G.J. Hol, Science 249 (1990) 1149–1152.
[24] W. Wriggers, K. Schulten, Proteins 35 (1999) 262–273.
[25] W. Wriggers, K. Schulten, Biophys. J. 75 (1998) 646–661.
[26] J.R. Collins, S.K. Burt, J.W. Erickson, Nat. Struct. Biol. 2 (1995) 334–338.
[27] J. Xing, W. Wriggers, G.M. Jefferson, R. Stein, H.C. Cheung, S.S. Rosenfeld, J. Biol. Chem. 275 (2000) 35413–35423.
[28] H.J. Woo, Biophys. Chemist. 125 (2007) 127–137.
[29] S. Bell, J.S. Crighton, Chem. Phys. Lett. 82 (1981) 122–126.
[30] S. Fischer, M. Karplus, Chem.Phys.Lett. 194 (1992) 252–261.
[31] S. Fischer, B. Windshigel, D. Horak, K.C. Holmes, J.C. Smith, Proc. Natl. Acad. Sci. USA 102 (2005) 6873–6878.
[32] P. Chacón, F. Tama, W. Wriggers, J. Mol. Biol. 326 (2003) 485–492.
[33] W. Zheng, S. Doniach, Proc. Natl. Acad. Sci. USA 100 (2003) 13253–13258.
[34] W.J. Zheng, B. Brooks, J. Mol. Biol. 346 (2005) 745–759.
[35] D.A. Brant, W.G. Miller, P.J. Flory, J. Mol. Biol. 23 (1967) 47–65.
[36] J. Hi, A. Ma, A.R. Dinner, J. Comput. Chem. 27 (2006) 203–216.
[37] M.E.J. Newman, G.T. Barkema, The principles of equilibrium thermal Monte Carlo simulation, Monte Carlo Methods in Statistical Physics, Clarendon Press, Oxford, 1999, pp. 31–44.
[38] T.P. Burghardt, A.R. Cruz-Walker, S. Park, K. Ajtai, Biochemistry 40 (2001) 4821–4833.
[39] T.P. Burghardt, S. Park, K. Ajtai, Biochemistry 40 (2001) 4834–4843.
[40] K. Ajtai, Y.M. Peyser, S. Park, T.P. Burghardt, A. Muhlrad, Biochemistry 38 (1999) 6428–6440.
[41] T.P. Burghardt, N. Juranic, S. Macura, A. Muhlrad, K. Ajtai, J. Am. Chem. Soc. 121 (1999) 10373–10378.
[42] S. Park, T.P. Burghardt, Biochemistry 41 (2002) 1436–1444.
[43] T.P. Burghardt, S. Park, W.-J. Dong, J. Xing, H.C. Cheung, K. Ajtai, Biochemistry 42 (2003) 5877–5884.
[44] A. Muhlrad, Y.M. Peyser, M. Nili, K. Ajtai, E. Reisler, T.P. Burghardt, Biophys. J. 84 (2003) 1047–1056.
[45] G. Weber, The chemical potentials of proteins, Protein Interactions, Chapman and Hall, New York, 1991, pp. 20–30.
[46] S. Park, K. Ajtai, T.P. Burghardt, Biochim. Biophys. Acta 1296 (1996) 1–4.
[47] S. Park, T.P. Burghardt, Biochemistry 39 (2000) 11732–11741.
[48] C.W.J. Granger, Econometrica 37 (1969) 424–438.
[49] E. Pereda, R.Q. Quiroga, J. Bhattacharya, Prog. Neurobiol. 77 (2005) 1–37.
[50] K. Sameshima, L.A. Baccalá, J. Neurosci. Methods 94 (1999) 93–103.

[51] M. Kovacs, J. Toth, A. Malnasi-Csizmadia, C.R. Bagshaw, L. Nyitray, J. Muscle Res. Cell Motil. 25 (2004) 95–102.

[52] S. Lowey, H.S. Slayter, A.G. Weeds, H. Baker, J. Mol. Biol. 42 (1969) 1–29.

[53] A.G. Weeds, R.S. Taylor, Nature 257 (1975) 54–56.

[54] Y.Y. Toyoshima, S.J. Kron, E.M. McNally, K.R. Niebling, C. Toyoshima, J.A. Spudich, Nature 328 (1987) 536–539.

[55] M.A. Geeves, K.C. Holmes, Annu. Rev. Biochem. Allied Res. India 68 (1999) 687–728.

[56] M. Rubinstein, Phys. Rev. Lett. 59 (1987) 1946–1949.

[57] M.E.J. Newman, G.T. Barkema, The repton model, Monte Carlo Methods in Statistical Physics, Claredon Press, Oxford, 1999, pp. 307–328.

[58] A.B. Bortz, M.H. Kalos, J.L. Lebowitz, J. Comput. Phys. 17 (1975) 10–18.

[59] T.P. Burghardt, K. Ajtai, Proc. Natl. Acad. Sci. USA 82 (1985) 8478–8482.

[60] N.C. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, E. Teller, J. Chem. Phys. 21 (1953) 1087–1092.

[61] I. Andricioaei, J.E. Straub, Phys. Rev. E 53 (1996) r3055–r3058.

[62] J.A. D'Aquino, J. Gómez, V.J. Hilser, K.H. Lee, L.M. Amzel, E. Freire, Proteins: Struct. Funct. and Genet. 25 (1996) 143–156.

[63] J. Gomez, V.J. Hilser, D. Xie, E. Freire, Proteins: Struct., Funct., and Genet. 22 (1995) 404–412.

[64] B.M. Baker, K.P. Murphy, Methods Enzymol. 295 (1998) 294–315.

[65] P.-D. Coureux, A.L. Wells, J. Menetrey, C.M. Yengo, C.A. Morris, H.L. Sweeney, A. Houdusse, Nature 425 (2003) 419–422.

[66] W.R.P. Scott, P.H. Hunenberger, I.G. Tironi, A.E. Mark, S.R. Billeter, J. Fennen, A.E. Torda, T. Huber, P. Kruger, W.F. van Gunsteren, J. Phys. Chem. 103 (1999) 3596–3607.

[67] A. Šali, T.L. Blundell, J. Mol. Biol. 234 (1993) 779–815.

[68] M.A. Martí-Renom, A.C. Stuart, A. Fiser, R. Sánchez, F. Melo, A. Šali, Annu. Rev. Biophys. Biomol. Struct. 29 (2000) 291–325.

[69] N. 4. Collaborative Computational Project, Acta Crystallogr. D50 (1994) 760–763.

[70] W. Humphrey, A. Dalke, K. Schulten, J. Mol. Graph. 14 (1999) 33–38.

[71] T.P. Burghardt, N. Juranic, S. Macura, K. Ajtai, Biopolymers 63 (2002) 261–272.

[72] P.D. Adams, Y. Chen, K. Ma, M.G. Zagorski, F.D. Sonnichsen, M.L. McLaughlin, M.D. Barkley, J. Am. Chem. Soc. 124 (2002) 9278–9286.

[73] P.M. Bayley, E.B. Nielsen, J.A. Schellman, J. Phys. Chem. 73 (1969) 228–243.

[74] E.B. Nielsen, J.A. Schellman, Biopolymers 10 (1971) 1559–1581.

[75] W.J. Goux, T.R. Radesch, T.M. Hooker, Biopolymers 15 (1976) 977–997.

[76] T.M. Hooker, J.A. Schellman, Biopolymers 9 (1970) 1319–1348.

[77] M.-C. Hsu, M.-C. Hsu, University of Illinois at Urbana-Champaign, 1970.

[78] P.E. Grebow, T.M. Hooker, Biopolymers 14 (1975) 871–881.

[79] T.M. Hooker, P.M. Bayley, W. Radding, J.A. Schellman, Biopolymers 13 (1974) 549–566.

[80] R Foundation for Statistical Computing, R: A language and environment for statistical computing, Vienna, Austria, 2006.

[81] L.A. Baccalá, K. Sameshima, Biol. Cybern. 84 (2001) 463–474.

[82] S. Kitagawa, J. Yoshimura, Y. Tonomura, J. Biol. Chem. 236 (1961) 902–906.

[83] F. Fabian, A. Muhlrad, Biochim. Biophys. Acta 162 (1968) 596–603.

[84] R.G. Yount, D. Lawson, I. Rayment, Biophys. J. 68 (1995) 44s–49s.

[85] H. Tokuyama, Y. Tonomura, J. Biochem.(Tokyo) 62 (1967) 456–463.

[86] M.F. Halstead, K. Ajtai, A.R. Penheiter, J.D. Spencer, Y. Zheng, E.A. Morrison, T.P. Burghardt, An unusual transduction pathway in human tonic smooth muscle myosin, Biophys. J. (in press).

[87] E.W. Yu, D.E. Koshland, Proc. Natl. Acad. Sci. USA 98 (2001) 9517–9520.

[88] I. Rayment, H.M. Holden, J.R. Sellers, L. Fananapazir, N.D. Epstein, Proc. Natl. Acad. Sci. USA 92 (1995) 3864–3868.

[89] K.L. Vikstrom, L.A. Leinwand, Curr. Opin. Cell Biol. 8 (1996) 97–105.

[90] S.L. Van Driest, M.A. Jaeger, S.R. Ommen, M.L. Will, B.J. Gersh, A.J. Tajik, M.J. Ackerman, J. Am. Coll. Cardiol. 44 (2004) 602–610.

[91] M. Sata, M. Ikebe, J. Clin. Invest. 98 (1996) 2866–2873.

[92] H. Fujita, S. Sugiura, S. Momomura, M. Omata, H. Sugi, K. Sutoh, J. Clin. Invest. 99 (1997) 1010–1015.

[93] K. Poetter, H. Jiang, S. Hassanzadeh, S.R. Master, A. Chang, M.C. Dalakas, I. Rayment, J.R. Sellers, L. Fananapazir, N.D. Epstein, Nat. Genet. 13 (1996) 63–69.

[94] S. Lowey, Trends Cardiovasc. Med. 12 (2002) 348–354.

[95] H. Yamashita, M.J. Tyska, D.M. Warshaw, S. Lowey, K.M. Trybus, J. Biol. Chem. 275 (2000) 28045–28052.