Published in final edited form as:

J Mol Biol. 2008 September 19; 381(5): 1407–1420. doi:10.1016/j.jmb.2008.06.071.

Extensive conformational transitions are required to turn on ATP hydrolysis in myosin

Yang Yang, Haibo Yu, and Qiang Cui¹

Department of Chemistry and Theoretical Chemistry Institute, University of Wisconsin, Madison, 1101 University Avenue, Madison, WI 53706

Abstract

Conventional myosin is representative of biomolecular motors in which the hydrolysis of Adenosine triphosphate (ATP) is coupled to large-scale structural transitions both in and remote from the active site. The mechanism that underlies such “mechanochemical coupling”, especially the causal relationship between the hydrolysis and the allosteric structural changes, has remained elusive despite extensive experimental and computational analyses. In this study, using combined quantum mechanical and molecular mechanical (QM/MM) simulations and different conformations of the myosin motor domain, we provide evidence to support that regulation of ATP hydrolysis activity is not limited to residues in the immediate environment of the γ phosphate. Specifically, we illustrate that efficient hydrolysis of ATP depends not only on the proper orientation of the lytic water but also the structural stability of several nearby residues, especially the Arg238-Glu459 salt-bridge (the numbering of residues follows the myosin II in *Dictyostelium discoideum*) and the water molecule that spans this salt-bridge and the lytic water. More importantly, by comparing the hydrolysis activity in two motor conformations with very similar active site (i.e., Switch-I and II) configurations, which distinguishes this work from our previous study, the results clearly indicate that the ability of these residues to make their crucial electrostatic stabilization relies on the configuration of residues in the nearby N-terminus of the relay helix and the “wedge loop”. Without the structural support from those motifs, residues in a closed active site in the post-rigor motor domain undergo subtle structural variations that lead to consistently higher calculated ATP hydrolysis barriers than in the pre-powerstroke state. In other words, starting from the post-rigor state, turning on the ATPase activity requires not only displacement of Switch II to close the active site but also structural transitions in the N-terminus of the relay helix and the “wedge loop”, which have been proposed previously to be ultimately coupled to the rotation of the converter subdomain 40 Å away.

Keywords

Mechanochemical coupling; ATP hydrolysis; Molecular motors; QM/MM simulations; Potential of Mean Force

Introduction

A tight coupling between chemistry and conformation is observed in many biological systems. [1] Such “mechanochemical coupling” is most notable in biomolecular motors,[2] where nucleotide hydrolysis activity needs to be regulated to avoid futile chemistry.[3] Therefore,

¹Communication should be addressed to: cui@chem.wisc.edu.

Supporting Information

Results for perturbative analysis of the minimum energy path results and additional structural comparisons between the pre-powerstroke and closed post-rigor simulations are also included.

revealing the mechanistic basis for how nucleotide hydrolysis is controlled by the conformational properties of biomolecular motors is a key step towards understanding how these “nanomachines” achieve their high energy-transduction efficiency.[4] In particular, in some motor systems, hydrolysis (rather than binding) of the nucleotide is believed to be coupled to not only local changes but also large-scale conformational transitions of remote domain(s). For example, FRET studies of Shih et al.[5] found that the motor domain of conventional myosin adopts a single conformation with non-hydrolyzable ATP analogue and that large-scale lever rotation is observed when ATP is used as the ligand. Bulk fluorescence experiments[6] also suggest that structural changes near the interface between the converter and the relay helix precede the hydrolysis of ATP. In previous structural[7] and molecular simulation studies, [8–12] others and we have shown that conformational transitions in the converter subdomain are coupled to changes 40 Å away in the position and flexibility of active site amino acids and water molecules. However, the precise identity of the structural changes that directly control the ATP hydrolysis activity and the mechanism through which these structural changes are coupled to the striking converter (and lever arm) rotation remain elusive.

In this study, using combined quantum mechanical and molecular mechanical (QM/MM) simulations,[13–15] we explicitly show that efficient ATP hydrolysis in myosin relies on the structural stability of the active site, especially that of a key salt-bridge between the Switch I (SwI) and Switch II (SwII) motifs and a water molecule between this salt-bridge and the lytic water. Although the role of these motifs on ATP hydrolysis have been discussed in previous studies,[16–18] the unique contribution of this study is the illustration that formation of a hydrolysis competent active site starting from the post-rigor state requires not only the displacement of SwII but also extensive rearrangements in the nearby structural motifs, which are known to ultimately couple to the rotation of the converter.[7,9–11] As a result, structural transitions remote from the active site can play a very *active* role in regulating the hydrolysis of ATP, rather than merely responding to structural changes in the active site in a *passive* fashion. This scenario of coupling sensitively regulated chemical reactivity to cascades of conformational transitions may contribute to mechanochemical coupling in many motor systems.

Results

ATP hydrolysis in myosin

As described extensively in the literature,[17] the hydrolysis of ATP in myosin occurs in kinetic states where the motor domain is detached from the actin. High-resolution x-ray structures [19,20,16] are available for the two detached states (Fig. 1), commonly referred to as the post-rigor and pre-powerstroke state, respectively, among which only the latter is believed to be capable of hydrolyzing ATP (see below). The converter subdomain at the C-terminus undergoes a striking 60° rotation in going from the post-rigor to the pre-powerstroke state (Fig. 1b,d), and SwII that forms part of the ATP binding site displaces to close the active site during the same transition. As shown by the comparison in Fig. 1a,c, closure of the active site is indicted by the formation of hydrogen bonding interactions between SwI and SwII residues (e.g., between Arg238-Glu459 sidechains; the numbering of residues follows the myosin II in *Dictyostelium discoideum*) as well as between the P-loop and SwII (e.g., between Ser181-Phe458 main chain atoms); it is worth noting that many interactions between SwII and the N-terminus of the relay helix (e.g., Glu459-Gln468, Glu459-Asn472 and Ser456-Asn475) are maintained in both sets of x-ray structures (Fig. 1a,c). Despite more than 40 Å in separation, the converter rotation and active site closure are believed to be tightly coupled through a set of conformational changes that implicate the intervening structural motifs such as the relay helix, the relay loop and the “wedge” loop.[9,21]

Since the water molecule near the γ phosphate in the post-rigor x-ray structure was found far from an ideal “in-line-attack” configuration,[20] it is generally believed that only the pre-powerstroke state with a closed active site is capable of efficiently hydrolyzing ATP,[7] which has also been supported by our recent QM/MM studies[8] using the two x-ray structures. The key remaining question is whether closure of SwII by itself, which has been shown to be a low-barrier process in the post-rigor state in our recent MD simulations,[10] is sufficient for turning on ATP hydrolysis or there are structural features beyond SwII displacement that are also vital to the efficient hydrolysis of ATP. Recent fluorescent experiments[22,6] found that conformational transitions at the interface between the converter subdomain and relay helix precede ATP hydrolysis and therefore seemed to support the latter scenario, although the underlying detail is not known.

ATP hydrolysis in the Pre-powerstroke state

The hydrolysis mechanism of ATP/GTP has been discussed extensively in the past.[23,24] In the enzyme environment, it is generally believed that an associative mechanism[25,26] is operational, although the involvement of a dissociative pathway has not been completely ruled out.[27] In this study, we focus on the associative pathway (also see discussions below) and variation in the energetics of this pathway with different conformations of the myosin motor domain. As shown in Fig. 2a, the Ser236 sidechain serves as a “relay group” for the proton transfer from the lytic water to the γ phosphate during the initial step of the hydrolysis; involvement of a proton relay process is generally thought as necessary based on stereochemical considerations[28,29] (however, see discussions below in **Materials and Methods**). A subsequent rotation of a hydroxyl group in the “pentavalent” intermediate helps completely break the covalent interaction between the ADP moiety and the inorganic phosphate, leading to the final hydrolysis product.

As shown by the two-dimensional potential of mean force (PMF) in Fig. 2b, the first step of the reaction has a barrier about 16 kcal/mol; the intermediate is 12 kcal/mol higher than the reactant state, and the P-O bond between the γ and β phosphate groups is largely broken (the $P^\gamma-O^{3\beta}$ distance is 2.60 Å, see Supporting Information). In fact, the $P^\gamma-O^{3\beta}$ bond is already significantly weakened in the transition state (see Fig. 2c). Therefore, the actual mechanism has partial dissociative character. For the second step of the reaction, the two-dimensional PMF (Fig. 2d) indicates a very small barrier from the intermediate followed by an energetically downhill proton transfer from the γ phosphate to $O^{3\beta}$ in ADP. The free energy profile along the $P^\gamma-O^{3\beta}$ distance is rather flat, which is consistent with the observation that the covalent interaction is already broken in the intermediate. Summing the energetics from the two PMFs, the ATP hydrolysis in the pre-powerstroke state is calculated to be nearly thermoneutral and the rate-limiting barrier is 16 kcal/mol. These are consistent with experimentally measured equilibrium constant and rate constant,[30,6] and this is the first time that the experimental data are computationally reproduced for ATP hydrolysis in myosin. Such close agreements indicate that the SCC-DFTBPR/MM approach (see **Materials and Methods** for additional discussions) is at least semi-quantitatively reliable for the current purpose.

For the first step of the hydrolysis, which is shown to be rate-limiting by the PMF results, an ensemble of minimum energy path (MEP) calculations starting from snapshots collected from equilibrium MD simulations for the pre-powerstroke structure are also carried out. These results at the SCC-DFTBPR/MM level are consistent with those from previous QM/MM simulations[8,31] using the minimized x-ray structure[16] and different QM levels. The variation among the MEP ensemble results is fairly small with a standard deviation of merely 1.8 kcal/mol, which suggests that the active site in the pre-powerstroke state does not have significant structural fluctuations, as also indicated by our previous analysis for the active site open/close transition.[10] Interestingly, the average barrier from the MEP calculations is 31.3

kcal/mol, which is substantially higher than the experimental value of 15–17 kcal/mol[30,6] and the PMF result. Superposition of the transition state structures from the PMF simulations and MEP calculations (Fig. 2c) indicates that the key residues, including water molecules in the active site, have very similar positions except for the orientation of Ser181 sidechain and the water that hydrogen bonds to it (Wat3 in Fig. 2c). Evidently, as proton transfer from the lytic water to $O^{2\gamma}$ proceeds, it is energetically more favorable for Ser181 to break hydrogen bonding to $O^{2\gamma}$ and seek alternative interactions with nearby groups. Since such an isomerization also involves changing the orientation of Wat3, the process is difficult to sample in local MEP calculations but readily accessible in PMF calculations; the scenario is similar to our recent analysis of long-range proton transfer reactions using MEP and PMF calculations. [32] Perturbation analysis of the MEP results (see Supporting Information) indicates that removing the hydrogen bonding interaction between Ser181 and ATP lowers the barrier by ~8 kcal/mol, which is a significant portion of the difference between MEP and PMF barriers. Nevertheless, similarity in the average position of most key active site residues suggests that the MEP calculations, if compared carefully, are appropriate for analyzing how the hydrolysis energetics depend on the conformational state of the motor domain.

ATP hydrolysis in a closed post-rigor state

As discussed in both experimental literature[17,7,18] and simulation studies,[8,31] there is consensus that the post-rigor state, which has an open active site in the crystal structure,[20] is unable to catalyze ATP hydrolysis. The precise reason behind this, however, remains debatable and likely involves both water structure[16,28] and electrostatics contributions[8, 24] (Yang et al., work in progress). Here we focus on the issue of whether *locally* closing the active site in the post-rigor state is sufficient for turning on efficient ATP hydrolysis. For this purpose, we generate a closed post-rigor configuration by displacing only SwII in the post-rigor x-ray structure (see **Materials and Methods**); it is meaningful to displace SwII itself because our previous simulations have shown that the open/close transition of SwII in the post-rigor state is a low-barrier and nearly thermoneutral process.[10] This structure is carefully equilibrated with extensive MD simulations, and a large set of QM/MM MEP calculations are then carried out for the first step of the hydrolysis reaction starting from snapshots sampled in the MD trajectories. MEP instead of PMF calculations are employed here because it is easier to correlate structural features of the active site and energetics of the hydrolysis in the MEP framework.

Remarkably, unlike the small variations in the MEP barrier found for configurations sampled in the pre-powerstroke state, the MEP barrier fluctuates significantly (ranging from 32.0 to 53.0 kcal/mol with a standard deviation of 4.4 kcal/mol) and is almost always substantially higher in the closed post-rigor state (Fig. 3a). This is rather unexpected given that the active site remains essentially closed during the simulations of both states (see Supporting Information). To understand the differences in the barriers, we examine the properties of both the lytic water and other active site residues/waters.

As shown in Fig. 3b, the distributions of distance between the lytic water and γ phosphate are very similar in both states throughout the nanosecond scale simulations. There is notable difference between the distributions of the “nucleophilic attack” ($O^{\text{lytic}}\text{-P}^{\gamma}\text{-O}^{3\beta}$) angle (Fig. 3c); for the pre-powerstroke state, the distribution peaks around 165° , while it is $\sim 150^\circ$ for the closed post-rigor state. Since the average nucleophilic attack angle in the transition state is $\sim 169^\circ$ according to both MEP and PMF calculations (see Supporting Information), the free energy penalty associated with properly aligning the lytic water in the closed post-rigor state (i.e., the difference in the angular potential of mean force between 150° and 169° in the closed post-rigor state, see Fig. 3c) is on the order of 2–3 kcal/mol. Although this is *not* a small contribution in kinetic terms (corresponds to 30–150 fold change in the rate constant at 300 K

according to transition state theory), it is clear that the nucleophilic attack angle is *not* as dominating a factor as commonly suggested[16,28] for dictating the hydrolysis activity.

With further analysis of properties of other active site residues and their electrostatic contributions to the hydrolysis barrier (see Supporting Information), it is found that the water molecule (labeled as Wat2 in Figures 2 and 4) hydrogen-bonded to the lytic water plays an important role. Importantly, the contribution from Wat2 to the hydrolysis barrier is observed to be influenced by the configuration of other residues in the active site, especially the Arg238-Glu459 salt-bridge.[18] In the pre-powerstroke state and early segment (<150 ps) of the closed post-rigor state MD simulations, this critical salt-bridge is well-maintained and Wat2 is stabilized by a stable hydrogen-bonding network involving the lytic water, Glu459 and carbonyl of Gly457 or another water molecule. As a result, Wat2 maintains interaction with the lytic water throughout the nucleophilic attack, which is consistent with the finding of perturbation analysis (see Fig. 3a of Supporting Information) that the electrostatic contribution from Wat2 to the hydrolysis barrier is rather modest (~3 kcal/mol).

As the simulation for the closed post-rigor state proceeds further, however, it is apparent that SwII is substantially more flexible than that in the pre-powerstroke state; this difference in flexibility has a direct impact on the hydrolysis barrier and contribution from Wat2 to the barrier. For example, during 150–800 ps of the simulation (green dots in Fig. 3a), Glu459 rotates out of the salt-bridge plane. As a result, Wat2 is tightly associated with the sidechain of Glu459 and forms a hydrogen-bonding interaction with the lytic water in the reactant state but not the transition state of hydrolysis (compare Fig. 4b with Fig. 4a). Accordingly, Wat2 makes an *unfavorable* contribution (~6 kcal/mol, see perturbation analysis in Fig. 3c of Supporting Information) to the hydrolysis barrier, which is consistent with the higher hydrolysis barriers for configurations collected from this part of the trajectory. Similarly, in the last segment (1800–2250 ps) of the simulation, the salt-bridge is further weakened and the active site becomes better solvated (see variations in the active site water distribution shown in Fig. 4c). Once again, Wat2 forms a hydrogen bond with the lytic water only in the reactant but not in the transition state because Wat2 is well stabilized by other water molecules that have diffused into the active site; as a result, the hydrolysis barrier is high. Interestingly, in the middle (800–1800 ps) of the trajectory, Wat2 is observed to diffuse out of the active site (see Fig. 4c) and therefore doesn't form a hydrogen bond with the lytic water in either the reactant or transition state; still, the hydrolysis barriers (red dots in Fig. 3a) for configurations collected from this segment of the trajectory are higher than those sampled in the pre-powerstroke state by a few kcal/mol.

To further understand the higher flexibility of SwII in the closed post-rigor state than in the pre-powerstroke state, we analyze the active site configurations from different sets of simulations. Comparing Fig. 4d with Fig. 1a–c, we note that displacing SwII *alone* to close the active site in the post-rigor state leaves many crucial interactions unformed or even breaks existing interactions (see Supporting Information). For example, although Gly457 forms a stable hydrogen bonding interaction with the γ phosphate of ATP upon SwII displacement, the interaction between Ser456 main chain and Asn475 in the relay helix is broken; the interaction between Gly457 and γ phosphate also becomes weaker in the later segment of the simulations. Similarly, although Glu459 forms a salt-bridge interaction with Arg238 when SwII is displaced, the interactions between the main chain of Glu459 and Asn472 in the relay helix are lost; instead, the carbonyl of Glu459 forms a hydrogen bond with the sidechain of Gln468. As a result, the extensive hydrogen-bonding network that involves Arg238, Glu459, Glu264 and Gln468 observed in the pre-powerstroke state (Fig. 1c) is not present in the closed post-rigor state (Fig. 4d), which explains the higher rotational flexibility of the Glu459 sidechain in the latter structure. Finally, since Tyr573 in the “wedge loop” remains far from SwII in the closed post-rigor state, there is ample space for Phe458 in SwII to sample multiple rotameric

states and its main chain interaction with Ser181 remains unformed, in contrast to the situation in the pre-powerstroke state. Evidently, the emerging picture is that the transition from the post-rigor state to a structurally stable closed active site (which apparently is critical to efficient ATP hydrolysis) relies on not only the displacement of SwII but also more extensive structural rearrangements in the nearby region, such as the N-terminus of the relay helix and the “wedge loop”; changes in these latter motifs have been shown previously[7,9,11] to be ultimately coupled to the rotation of the remote converter subdomain.

Concluding Discussions

The tight coupling between ATPase activity and large-scale conformational transition of remote domain(s) is the hallmark feature of many biomolecular motors[2] and crucial for their high thermodynamic efficiency.[3] While the process of large-scale structural transition induced by ATP *binding* has been visualized and discussed in several systems,[33,4,34] the underlying mechanism for the allosteric coordination of ATP *hydrolysis* and large-scale conformational rearrangements remains elusive; indeed, the stereochemical difference between ATP and ADP·P_i is much more subtle than that brought about by the binding of the highly charged Mg²⁺·ATP. It is common to assume that hydrolysis is strictly controlled by the immediate environment and the large-scale structural changes are the *consequence* of the local structural changes associated with the hydrolysis process; i.e., the hydrolysis product preferentially stabilizes a specific conformation of the active site, which in turn propagates into major structural changes in remote regions.[9,10]

Combined with results from previous structural[7] and classical molecular simulations,[9,8, 10,11] QM/MM simulations in this study demonstrate that turning on the ATP hydrolysis activity in myosin requires rather extensive conformational rearrangements beyond the immediate environment of ATP. Specifically, we illustrate that efficient ATP hydrolysis relies directly on the proper positioning of not only the first coordination sphere of the γ -phosphate, such as the commonly discussed lytic water,[28] but also second coordination sphere motifs, especially the Arg238-Glu459 salt-bridge between SwI/II and the water molecule that spans this salt-bridge and the lytic water. In fact, the second-sphere electrostatic stabilization is found to be more significant than aligning the lytic water into the in-line attack orientation.[28] Importantly, we demonstrate that the ability of these second coordination sphere residues to make their crucial electrostatic stabilization relies on the configuration of residues in the nearby N-terminus of the relay helix and the “wedge loop”. Without the structural support from those motifs, residues in a closed active site in the post-rigor motor domain undergo subtle structural variations that lead to consistently higher calculated ATP hydrolysis barriers than in the pre-powerstroke state. In other words, starting from the post-rigor state, turning on the ATPase activity requires not only displacement of Switch II to close the active site but also structural transitions in the N-terminus of the relay helix and the “wedge loop”, which have been proposed previously to be ultimately coupled to the rotation of the converter subdomain 40 Å away.[7, 9,11]

It is important to emphasize that the current findings do not suggest that *all* the conformational cascades linking the active site and the converter subdomain discussed in earlier structurally motivated studies[7,9,11] are required to turn on efficient ATP hydrolysis. Indeed, the fact that ATPase activity in many decoupling mutants[35,36] is very close to be normal indicates that the lack of converter/lever-arm rotation doesn't significantly impair ATP hydrolysis. Therefore, a more likely scenario for the recovery stroke in myosin consists of two phases. [9,11,12] In the first phase, structural transitions near the N-terminus of the relay helix are coupled to the SwII displacement to establish a stable active site and turn on the ATP hydrolysis; this is consistent with the finding from bulk fluorescent study[22,6] that conformational transitions propagated to the interface between the converter subdomain and

relay helix precede ATP hydrolysis. In the second “relaxation phase”, the rest of the conformational cascades propagate into the rotations of the converter and lever-arm; the “relaxation” is expected to occur only with the active-site stabilized to the closed conformation by the hydrolysis products,[10] which is compatible with the FRET observation that lever-arm rotation occurs only following ATP hydrolysis.[5] This two-phase description is reminiscent of coupled tertiary and quaternary structural changes for oxygen binding induced allostery in hemoglobin, where it has been shown that tertiary structural changes induced by oxygen binding proceeds quaternary structural changes.[37,38] For myosin, it remains difficult to precisely determine the microscopic details for the two phases (e.g., a structural model for the kinetically significant intermediate immediately prior to ATP hydrolysis[22,6]), although sensible models exist based on computational studies.[9,11,12]

From the functional perspective, an important feature of molecular motors is that futile ATP hydrolysis is minimized to maintain a high thermodynamic efficiency. In the case of myosin, the simplest explanation that has been well received in the literature[17] is that the SwII configuration is tightly coupled to the orientation of the converter. Since SwII closure is accepted to be a requirement for efficient ATP hydrolysis, a tight structural coupling between SwII and converter seems sufficient for avoiding futile ATP hydrolysis (i.e., hydrolysis of ATP without converter rotation). The finding of our current study that ATP hydrolysis is sensitive to more extensive structural changes reveals another layer of subtlety to mechanochemical coupling and adds to the mechanism that Nature uses to avoid futile hydrolysis in myosin and possibly other molecular motors. In particular, the QM/MM analysis here further suggests that highly conserved residues near the N-terminus of the relay helix (e.g., Gln468, Asn472, Tyr573) play a role not only in relaying structural changes from the active site to the converter [7,9,11] but also in stabilizing the proper configuration of the active site for efficient ATP hydrolysis. This result can be tested experimentally by mutagenesis studies. We predict that mutation of these residues into non-polar amino acids such as Alanine will hamper the formation of a closed configuration of the active site that is hydrolysis-competent, akin to the closed post-rigor state generated *in silico* in this study. Therefore, in contrast to the decoupling mutants discussed in the literature,[35,36], which involve mutations further down the relay helix, mutations of residues near the N-terminus of the relay helix (e.g., Gln468, Asn472, Tyr573) should not only abolish motility but also significantly impair the ATPase activity. In this context, we note that in published studies,[75,76] mutation efforts that aim to identify residues essential to ATP hydrolysis have been limited to SwI/SwII motifs.

In a broader context, there is increasing interest in dissecting the role of slow ($\mu s - ms$) motions in enzyme catalysis.[39–41] For example, single molecule measurements[42] explicitly illustrated the presence of dynamic disorder in enzyme catalysis. However, the structural details behind the dynamic disorder are rarely clear and the magnitude of catalytic rate fluctuation is often small (~ 10). Similarly, a recent discussion[43] recommends the concept of catalytic network in enzymes with multiple conformations that occur serially and in parallel in the mechanism, while the practical challenge is to figure out the most relevant conformational degrees of freedom that have the most significant impact on the chemical event in the specific system under study. In this regard, molecular motors provide a remarkable opportunity for probing the role of slow structural transitions on catalysis because the effect is likely significant considering the functional importance of tightly regulating the timing of the chemical step in these systems. However, insight regarding the causality among events of different scale and physicochemical nature (e.g., hydrolysis vs. structural transitions in the active site and distant regions) is difficult to come by with either experimental studies alone or standard computational techniques that probe either only structural changes[9–11,21] or chemical activity in a single x-ray structure. In this study, the identity of the structural changes that have a direct impact on the ATP hydrolysis activity and the mechanism through which these structural changes are coupled to other conformational rearrangements remote from the active

site are analyzed for myosin as a representative molecular motor. The key missing link between structural properties of the motor and ATP hydrolysis activity is provided by extensive QM/MM calculations with carefully selected configurations of the system. The use of motor domains with very similar active site configurations but different overall structures makes it possible to directly probe the importance of distant structural changes on the ATPase activity, which was not possible with our previous study that utilized two original x-ray structures that differ also in the active site.[8] Therefore, this study highlights the unique value of computations that *simultaneously* probe structural changes and chemical activity,[41] which is useful for revealing detailed mechanism for other interesting issues in biomolecular motors such as the coordination between different ATPase sites in kinesin and myosin V.

Materials and Methods

Classical molecular dynamics simulations

Equilibrium simulations of the active site with a stochastic boundary condition and explicit solvent are carried out using the CHARMM program (c32a2 version) [44]. Three systems are simulated: post-rigor with ATP bound (denoted as the 1FMW simulation), pre-powerstroke with ATP bound (denoted as the 1VOM simulation) and post-rigor with ATP bound in a closed active site pocket (denoted as the 1FMW(closed) simulation; see below). As the starting coordinates for the 1FMW and 1VOM simulations, the X-ray structures solved by Rayment and co-workers for the *Dictyostelium discoideum* myosin II motor domain are used; the converter subdomain is in the down orientation and the active site is open in 1FMW [20] (resolution 2.15 Å), while the converter is up and the active site is closed in 1VOM [16] (resolution 1.90 Å). The missing residues (1, 203–208, 500–508, 622–626 in 1FMW, and 1, 205–208, 711, 716–719, 724–730 in 1VOM) are modelled by using the program ModLoop [45,46]. The protein atoms and Mg^{2+} -ATP in the active site are described with the all-atom CHARMM force field for proteins [47] and the water molecules are described with the TIP3P model [48]. Hydrogen atoms are added with HBUILD [49]. In all cases the system is partitioned into a 32 Å inner region centered at the geometric center of ATP, P-loop, Switch I, and Switch II (~ 12,800 atoms) with the remaining portion of the system in the outer region. Newtonian equations-of-motion are solved for the Molecular Dynamics (MD) region (within 28 Å), and Langevin equations-of-motion are solved for the buffer region (28–32 Å) with a temperature bath of 300 K [50]. All the water molecules in the inner region are subject to a weak GEO type of restraining potential to keep them inside the inner sphere with the MMFP module of CHARMM. The GEO restraining potential is in the form of a quartic polynomial on each oxygen atom in water: $k \times \Delta^2(\Delta^2 - V_P)$, with $\Delta = r - r_{\text{off}}$; k is the restraining quartic force constant (0.5 kcal/(mol·Å⁴)), r is the distance of the oxygen from the center of the simulation sphere, r_{off} is the cutoff distance (32.0 – 1.5 = 30.5 Å) below which the GEO restraint is set to zero, V_P is an offset value taken to be 2.25. These parameters lead to a restraining potential on water that smoothly turns on at 30.5 Å, reaches a well at 31.5 Å with a depth of –0.625 kcal/mol, and then quickly rises to be repulsive beyond 32.0 Å. All protein atoms in the buffer region are harmonically restrained with force constants determined directly from the B-factors in the PDB file [50]. Langevin atoms are updated heuristically during the simulation to consistently treat protein groups and water molecules that may switch regions during the simulation. All bonds involving hydrogen are constrained using the SHAKE algorithm [51] with a relative geometric tolerance of 10^{-10} , and the time step is set to 1 fs. Non-bonded interactions within the inner sphere are treated with an extended electrostatics model, in which groups beyond 12 Å interact as multipoles [52]. The entire system is heated gradually to 300 K and equilibrated for ~ 160 ps prior to the production simulations. To account for the electrostatics between the inner and outer region atoms and the effect of solvation, the generalized solvent boundary potential (GSBP) approach developed by Roux and co-workers is used [53,54]. In GSBP, the contribution from distant protein charges (screened by the bulk

solvent) in the outer region $\Delta W_{elec}^{(io)}$ is represented in terms of the corresponding electrostatic potential in the inner region, $\varphi_s^{(io)}$,

$$\Delta W_{elec}^{(io)} = \sum_{\alpha \in inner} q_{\alpha} \varphi_s^{(io)} \quad (1)$$

where q_{α} is the partial charge on atom α . The dielectric effect on the interactions among inner region atoms is represented through a reaction field term,

$$\Delta W_{elec}^{(ii)} = \frac{1}{2} \sum_{mn} Q_m M_{mn} Q_n \quad (2)$$

where $Q_{m/n}$ is the generalized multipole moment and M_{mn} is the element of the reaction field matrix, $\mathbf{M}$. The static field due to the outer region atoms $\varphi_s^{(io)}$, is evaluated with the linear Poisson-Boltzmann equation using a focusing scheme that places a 70 Å cube of fine grid (0.4 Å) into a larger 120 Å cube of coarse grid (1.2 Å). The reaction field matrix, $\mathbf{M}$, is evaluated using 400 spherical harmonics. In the Poisson-Boltzmann calculations, the protein dielectric constant of $\epsilon_p = 1$, the water dielectric constant of $\epsilon_w = 80$, and 0.0 M salt concentration are used. The optimized radii of Roux and Nina [55,56] based on experimental solvation energies of small molecules as well as the calculated interaction energy with explicit waters are adopted to define the solvent-solute dielectric boundary.

The 1FMW(closed) simulation is for a conformation in which SwII is displaced to form a closed active site in the post-rigor structure, and it is prepared with the following procedure. The starting structure for equilibration is taken from the our previous umbrella sampling simulation [10] for the active site of 1FMW; specifically, the last snapshot in the window with a $\Delta RMSD = 2.0$ Å is used. The $\Delta RMSD$ is measured for SwI and SwII with reference to the x-ray structures of 1FMW and 1VOM (i.e., $\Delta RMSD = RMSD_{1FMW}^{SwI/II} - RMSD_{1VOM}^{SwI/II}$), and a value of 2.0 Å corresponds to an active site (in terms of SwI and SwII coordinates) that is very close to that in the 1VOM X-ray structure (see discussions below). To equilibrate this structure, MD simulations are performed for about 300 ps; to prevent any residual effects from the $\Delta RMSD$ constraint in the previous umbrella sampling simulations, two restraining potentials are applied. The first is a NOE type restraint on the Arg238-Glu459 salt bridge with a force constant of 10 kcal/mol/Å² and an upper bound of 2.0 Å between HH21/HH12 of Arg238 and OE2/OE1 of Glu459. The second one is a harmonic restraint between O^{3γ} of ATP and HN of Gly457 with a force constant 10 kcal/mol/Å² and the reference minimum at 1.8 Å. In production runs, no constraint is applied to the active site residues.

All three (1FMW, 1VOM and 1FMW(closed)) simulations have been carried out for at least 2 ns. For the 1VOM state, substantially longer (12 ns) simulations have also been carried out with a smaller (20 Å) inner region for comparison; the results for key properties of interest are essentially the same and therefore not included.

Reaction path and potential of mean force with combined QM/MM simulations

The QM/MM partitioning and the QM level (SCC-DFTBPR)—Reaction energetics are studied based on the combined quantum mechanics/molecular mechanics (QM/MM) potential energy surface [13,15,57]. The QM region includes 48 atoms at the self-consistent charge density functional tight binding (SCC-DFTB) level [58,59] with augmented 3rd order

expansion[60] and newly developed parameters for phosphate hydrolysis reactions (see below). The QM region includes the tri-phosphate and part of the ribose group of ATP, the catalytic water, the Mg^{2+} ion and all its ligands (side chains of Thr186, Ser237 and two water molecules) as well as the side chain of the conserved Ser236. Four link atoms are introduced to saturate the valence of the QM boundary atoms at the QM/MM interface, which interact with all MM atoms through electrostatic interactions except with the link host (four $C\alpha$ atoms here); no van der Waals interaction is considered for the link atoms. This simple link atom treatment has been shown to be appropriate when the charges around the link atom are small, [61,62] which is the case here. Indeed, the partial charges on the link-atoms are always small; the values for link atoms on the amino acid sidechains are typically below 0.02, while the partial charge for the link atom associated with ATP is slightly larger, in the range of 0.07–0.08. Therefore, there is no indication that the QM region is artificially polarized by the link atom scheme adopted here. Although a larger QM region including key residues such as Arg238 and Glu459 can also be used, we made the current choice so that the contributions from active site residues to the hydrolysis energetics can be analyzed using perturbation analysis[63] in a consistent manner. All MM atoms are described with the CHARMM 22 force field [47].

The SCC-DFTB approach is an approximate density functional method,[58] and has been successfully applied to various biological problems [64,65]. Its reasonable balance between the computational speed and accuracy make it possible to carry out the large number of reaction path (minimum energy path) and potential of mean force calculations that distinguish the current work from previous QM/MM studies of myosin, which were all based on a few minimum energy paths.[8,31,66] In the current study, a parameterization of SCC-DFTB including the 3rd-order extension[60] and fitted using a set of phosphate hydrolysis reactions in the gas phase is adopted; the parameterization is referred to as SCC-DFTBPR to distinguish it from the standard parameterization[58] and the details for the development and benchmark will be reported separately. In Table 1, however, we summarize the data set used for fitting and key results for the evaluation of the model. Although there are still notable errors compared to high-level *ab initio* or DFT calculations, the SCC-DFTBPR is comparable to the best semi-empirical method available in the literature for phosphate chemistry.[67,68] Compared to the AM1(d) parameterization,[67] which is very specific for transphosphorylation reactions, there are much fewer reaction-specific parameters for SCC-DFTBPR (one Hubbard derivative parameter for each element type); in fact, as seen in Table 1, even for reactions in the QCRNA database, many of which are rather different in nature from the reactions used to parameterize SCC-DFTBPR, the performance of SCC-DFTBPR is very respectful. Considering that the most important issue in the current work is the variation of the ATP hydrolysis energetics with different conformations of the active site, which is a relative quantity, the impact of the systematic errors in SCC-DFTBPR on the result is likely small.

In most QM/MM simulations, especially the PMF calculations, an inner region of 20 Å in the GSBP framework[54,69] is used. For MEP calculations, however, calculations with a larger 32 Å inner region have also been carried out for comparison; as shown in Tables 1, 2 of Supporting Information, neither the barriers nor the key geometrical parameters are sensitive to the size of the inner region.

Potential of Mean Force (PMF) calculations—To make a direct comparison with available experimental results,[30,6] PMF simulations have been carried out for ATP hydrolysis in the 1VOM state, which is well accepted to be the proper conformational state for efficient ATP hydrolysis. [16,17]

For the first step of the reaction (Fig. 2a), which involves Ser236 as the proton relay, two variables are chosen as the reaction coordinate and the corresponding 2-dimensional PMF is generated using umbrella sampling and the weighted histogram analysis[70–72]. The first

coordinate is simply the $P^{\gamma}\text{-O}^{\text{lytic}}$ distance and the second coordinate involves a linear combination of two antisymmetric stretches that describe the proton transfers involving the lytic water and Ser236 (i.e., $r_{O^{\text{Ser}}\dots H^{\text{lytic}}} - r_{O^{\text{lytic}}\dots H^{\text{lytic}}} + r_{O^{2\gamma}\dots H^{\text{Ser}}} - r_{O^{\text{Ser}}\dots H^{\text{Ser}}}$). In total 32 windows are used and 250 ps simulations are performed for each window. Convergence of the PMF is monitored by examining the overlap of reaction-coordinates distributions sampled in different windows and by evaluating the effect of leaving out segments of trajectories. For example, Fig. S1a in Supporting Information shows the PMF calculated using only the first 150 ps of each window; the result is very comparable to that shown in Fig. 2b, which indicates that the sampling carried out is sufficient for the current purpose.

For the second step of the reaction, which involves largely a rotation of the $O^{2\gamma}H$ group of the γ phosphate, umbrella sampling calculations are carried out using a single reaction coordinate, which is the anti-symmetric stretch involving $O^{2\gamma}H$ and $O^{3\beta}$ of ATP. The PMF, however, is constructed as a two-dimension map where the $O^{3\gamma}\text{-P}^{\gamma}$ distance is also used as another reaction coordinate; no umbrella potential for the $O^{3\beta}\text{-P}^{\gamma}$ distance is necessary because the corresponding free energy is rather flat. The amount of data used in the WHAM analysis include 21 windows, with each window sampling 250 ps; similar to the PMF for the first step, the convergence is evaluated by leaving out segments of trajectories. As shown in Fig. S1b in Supporting Information, using only the first 150 ps of each window gives results very comparable to that shown in Fig. 2d.

Average structures from windows that correspond to key regions on the PMFs (reactant, TS1, INT1 and product) are summarized in Fig. S2 in Supporting Information. For both the reactant (ATP + Wat) and TS1, also superimposed are averaged structures from MEP calculations. As discussed in **Results**, the positions of key residues and water molecules are very similar in the PMF and MEP calculations, except for Ser181 and the water molecule hydrogen-bonded to the Ser181 side chain. Since the contribution of Ser181 is understood based on perturbative analysis and its orientation is *consistent* in all MEP calculations regardless of the configuration of the active site (i.e., 1VOM vs. 1FMW(closed)), it is justified to use MEP to study the configurational dependence of ATP hydrolysis activity instead of the much more demanding (and not straightforward to analyze) PMF calculations.

Minimum Energy Path (MEP) calculations—To analyze the dependence of the ATP hydrolysis activity on the configuration of the active site, a large set of MEP calculations are carried out using snapshots from equilibrium simulations of the 1VOM and 1FMW(closed) states. Since the PMF calculations indicate that the rate-limiting step is the first step, MEP calculations are only carried out for that step. Similar to the PMF calculations, the mechanism involving Ser236 as the proton relay is studied for both the 1VOM and 1FMW(closed) states.

Starting from each snapshot randomly collected from the MD simulations, adiabatic mapping calculations are first carried out to generate the approximate path; multiple forward and backward adiabatic mapping calculations are carried out to ensure complete structural relaxation. In these calculations, the adapted basis Newton-Raphson approach is used for energy minimization with a gradient threshold of 0.05 kcal/mol/Å for the last several (~ 4) adiabatic mapping cycles. The approximate reaction path from the adiabatic mapping is then used as the initial guess for the conjugated peak refinement (CPR) calculations,[73] which locate the true saddle point(s) along the reaction path. The energy barrier is defined as the total energy difference between the saddle point and the reactant (ATP·H₂O) state.

To gain insights into the variation of MEP barriers from different snapshots, perturbative analysis of the barrier height[8] is carried out in which the partial charges of specific residues are turned off and the change in the barrier height is calculated based on single point QM/MM energy calculations. Inspection of the results (see Supporting Information) suggest that the

protein/solvent contributions (especially the relative trends between results for different active site configurations) are dominated by the active site residues and water molecules. Therefore, only results for these groups are included (see Fig. S3 in Supporting Information).

Approximations and validity of models—The current study shares with typical molecular simulation studies the limitations associated with non-polarizable force fields and approximate QM methods. Since we are mostly interested in relative quantities, i.e., the dependence of hydrolysis barrier/energetics on the conformational properties of the motor domain, we expect that the results are less sensitive to the inaccuracies in the force field and QM method.

There are two approximations specific to this study. First, the closed post-rigor state is generated *in silico*; second, only the associative pathway involving Ser236 as the relay group is analyzed. Regarding the first approximation, we emphasize that the generated structure has been characterized by potential of mean force simulation in our recent study,[10] which showed that the closed configuration is nearly iso-energetic to the experimentally (x-ray crystallography) characterized open configuration of the post-rigor state. Therefore, the closed post-rigor state employed in this study is not a high-energy conformation. Since Switch-II closure has been proposed to be tightly coupled to structural changes in nearby regions,[9–11,21] it is likely that the closed post-rigor conformation modeled here does *not* represent a kinetically significant intermediate as characterized in fluorescence experiments.[22,6] This does not, however, invalidate the use of the modeled structure as a unique way to probe how structural properties of the motor domain influence the ATP hydrolysis activity.

Regarding the issue of hydrolysis mechanism, not all possibilities[23,24] have been investigated here. In particular, we assume that Ser236 plays an active role in the rate-limiting step. The importance of involving a proton-relay group has been emphasized in the phosphoryl transfer literature based on stereochemical considerations,[28,29] which motivated our model; the fact that Ser236 is a highly conserved residue in myosin has also been discussed using such a catalytic mechanism.[74] Mutation studies[75,76] found, however, that the ATPase activity is not much reduced in the S236A mutant compared to the wild type myosin; recent structural studies (Rayment, private communication) found that the S236A mutant is structurally also very similar to the wild type. Therefore, it is possible that Ser236 does not, in fact, play an important role in hydrolysis in the wild type myosin, although it is also possible that an alternative mechanism is operative in the mutant. Further studies are on the way to pin point the catalytic mechanism and nature of the rate-limiting transition state in the wild type and the S236A mutant myosin. Since the charge development during the hydrolysis of ATP is expected to be similar regardless of the precise role of Ser236, we expect that our conclusion regarding the importance of active site stability is not limited to the specific mechanism analyzed in the current QM/MM calculations. Ultimately, the validity of our analysis and model needs to be tested by the experimental mutation studies proposed in **Concluding Discussions**.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work is supported by a grant from the National Institutes of Health (R01-GM071428). Q.C. acknowledges an Alfred P. Sloan Research Fellowship. Computational resources from the National Center for Supercomputing Applications at the University of Illinois are greatly appreciated.

References

1. Alberts, B.; Bray, D.; Lewis, J.; Raff, M.; Roberts, K.; Watson, JD. Molecular biology of the cell. Garland Publishing, Inc.; 1994.
2. Vale RD, Milligan RA. The way things move: looking under the hood of molecular motor proteins. *Science* 2000;288:88–95. [PubMed: 10753125]
3. Eisenberg E, Hill TL. Muscle contraction and free energy transduction in biological systems. *Science* 1985;227:999–1006. [PubMed: 3156404]
4. Gao Y, Karplus M. Biomolecular motors: the F₁-ATPase paradigm. *Curr Opin Struct Biol* 2004;14:250–259. [PubMed: 15093841]
5. Shih WM, Gryczynski Z, Lakowicz JR, Spudich JA. A FRET-based sensor reveals large ATP hydrolysis-induced conformational changes and three distinct states of the molecular motor myosin. *Cell* 2000;102:683–694. [PubMed: 11007486]
6. Málnási-Csizmadia A, Pearson DS, Kovács M, Woolley RJ, Geeves MA, Bagshaw CR. Kinetic resolution of a conformational transition and the ATP hydrolysis step using relaxation methods with a *Dictyostelium* myosin II mutant containing a single tryptophan residue. *Biochemistry* 2001;40:12727–12737. [PubMed: 11601998]
7. Geeves MA, Holmes KC. The molecular mechanism of muscle contraction. *Adv Prot Chem* 2005;71:161–193.
8. Li GH, Cui Q. Mechanochemical coupling in myosin: A theoretical analysis with molecular dynamics and combined QM/MM reaction path calculations. *J Phys Chem B* 2004;108:3342–3357.
9. Fischer S, Windshugel B, Horak D, Holmes KC, Smith JC. Structural mechanism of the recovery stroke in the myosin molecular motor. *Proc Natl Acad Sci USA* 2005;102:6873–6878. [PubMed: 15863618]
10. Yu H, Ma L, Yang Y, Cui Q. Mechanochemical coupling in myosin motor domain, I. Equilibrium active site simulations. *PLoS Comput Biol* 2007;3:0199.
11. Yu H, Ma L, Yang Y, Cui Q. Mechanochemical coupling in myosin motor domain, II. Analysis of critical residues. *PLoS Comput Biol* 2007;3:0214.
12. Koppole S, Smith JC, Fischer S. The structural coupling between ATPase activation and recovery stroke in the myosin II motor. *Structure* 2007;15:825–837. [PubMed: 17637343]
13. Warshel A, Levitt M. Theoretical studies of enzymic reactions - dielectric, electrostatic and steric stabilization of carbonium-ion in reaction of lysozyme. *Journal of Molecular Biology* 1976;103(2): 227–249. [PubMed: 985660]
14. Lipkowitz, KB.; Boyd, DB.; Gao, J., editors. Reviews in Computational Chemistry VII. VCH; New York: 1995.
15. Field MJ, Bash PA, Karplus M. A combined quantum-mechanical and molecular mechanical potential for molecular-dynamics simulations. *Journal of Computational Chemistry* 1990;11(6):700–733.
16. Smith CA, Rayment I. X-ray structure of the Magnesium(II) · ADP · Vanadate complex of the *Dictyostelium discoideum* myosin motor domain to 1.9 Å resolution. *Biochemistry* 1996;35:5404–5417. [PubMed: 8611530]
17. Geeves MA, Holmes KC. Structural mechanism of muscle contraction. *Annu Rev Biochem* 1999;68:687–728. [PubMed: 10872464]
18. Onishi H, Ohki T, Mochizuki N, Morales MF. Early stages of energy transduction by myosin: Roles of Arg in switch I, of Glu in switch II, and of the salt-bridge between them. *Proc Natl Acad Sci USA* 2002;99:15339–15344. [PubMed: 12429851]
19. Rayment I, Rypniewski WR, Schmidtbase K, Smith R, Tomchick DR, Benning MM, Winkelmann DA, Wesenberg G, Holden HM. 3-dimensional structure of myosin subfragment-1: A molecular motor. *Science* 1993;261:50–58. [PubMed: 8316857]
20. Bauer CB, Holden HM, Thoden JB, Smith R, Rayment I. X-ray structures of the Apo and MgATP-bound states of *Dictyostelium discoideum* myosin motor domain. *J Biol Chem* 2000;275:38494–38499. [PubMed: 10954715]
21. Koppole S, Smith JC, Fischer S. Simulations of the myosin II motor reveal a nucleotide-state sensing element that controls the recovery stroke. *J Mol Biol* 2006;361:604–616. [PubMed: 16859703]
22. Málnási-Csizmadia A, Woolley RJ, Bagshaw CR. Resolution of conformational states of *Dictyostelium* myosin II motor domain using tryptophan (W501) mutants: Implications for the open-

- closed transition identified by crystallography. *Biochem* 2000;39:16135–16146. [PubMed: 11123942]
23. Admiraal S, Herschlag D. Mapping the transition state for atp hydrolysis: implications for enzymatic catalysis. *Chem Biol* 1995;2:729–739. [PubMed: 9383480]
 24. Glennon TM, Villa J, Warshel A. How does gap catalyze the gtpase reaction of ras?: A computer simulation study. *Biochem* 2000;39:9641–9651. [PubMed: 10933780]
 25. Guthrie RD, Jencks WP. IUPAC recommendations for the representation of reaction mechanisms. *Acc Chem Res* 1989;22:343–349.
 26. Hassett A, Blattler W, Knowles JR. Pyruvate kinase: Is the mechanism of phospho transfer associative or dissociative? *Biochem* 1982;21:6335–6340. [PubMed: 7150563]
 27. Aqvist J, Kolmodin K, Florian J, Warshel A. Mechanistic alternatives phosphate monoester hydrolysis: what conclusions can be drawn from available experimental data? *Chem & Biol* 1999;6:R71–R80. [PubMed: 10074472]
 28. Cleland WW, Hengge AC. Enzymatic mechanisms of phosphate and sulfate transfer. *Chem Rev* 2006;106:3252–3278. [PubMed: 16895327]
 29. Gerratana B, Sowa GA, Cleland WW. Characterization of the transition-state structures and mechanisms for the isomerization and cleavage reactions of uridine 3'-m-nitrobenzyl phosphate. *J Am Chem Soc* 2000;122:12615–12621.
 30. Trentham DR, Eccleston JF, Bagshaw CR. Kinetic-analysis of ATPase mechanisms. *Q Rev Biophys* 1976;9:217–281. [PubMed: 183232]
 31. Schwarzl SM, Smith JC, Fischer S. Insights into the chemomechanical coupling of the myosin motor from simulation of its ATP hydrolysis mechanism. *Biochem* 2006;45:5830–5847. [PubMed: 16669626]
 32. Riccardi D, Koenig P, Guo H, Cui Q. proton transfer in carbonic anhydrase is controlled by electrostatics rather than the orientation of the acceptor. *Biochem* 2008;47:2369–2378. [PubMed: 18247480]
 33. Wang HY, Oster G. Energy transduction in the F₁ motor of ATP synthase. *Nature* 1998;396:279–282. [PubMed: 9834036]
 34. von Hippel PH, Delagoutte E. A general model for nucleic acid helicases and their “coupling” within macromolecular machines. *Cell* 2001;104:177–190. [PubMed: 11207360]
 35. Sasaki N, Ohkura R, Sutoh K. *Dictyostelium myosin II* mutations that uncouples the converter swing and ATP hydrolysis cycle. *Biochem* 2003;42:90–95. [PubMed: 12515542]
 36. Ito K, Uyeda QP, Suzuki Y, Sutoh K, Yamamoto K. Requirement of domain-domain interaction for conformational change and functional ATP hydrolysis in myosin. *J Biol Chem* 2003;278:31049–31057. [PubMed: 12756255]
 37. Eaton WA, Henry ER, Hofrichter J, Mozzarelli A. Is cooperative oxygen binding by hemoglobin really understood? *Nature Structural Biology* 1999;6:351–358.
 38. Lee AW, Karplus M. Structure-specific model of hemoglobin cooperativity. *Proc Acad Natl Sci USA* 1983;80:7055–7059.
 39. Benkovic SJ, Hammes-Schiffer S. A perspective on enzyme catalysis. *Science* 2003;301:1196–1202. [PubMed: 12947189]
 40. Boehr DD, Dyson HJ, Wright PE. An nmr perspective on enzyme dynamics. *Chem Rev* 2006;106:3055–3079. [PubMed: 16895318]
 41. Rod TH, Radkiewicz JL, Brooks CL. Correlated motion and the effect of distal mutations in dihydrofolate reductase. *Proc Natl Acad Sci USA* 2003;100:6980–6985. [PubMed: 12756296]
 42. Lu HP, Xun LY, Xie XS. Single-molecule enzymatic dynamics. *Science* 1998;282:1877–1882. [PubMed: 9836635]
 43. Benkovic SJ, Hammes GG, Hammes-Schiffer S. Free energy landscape of enzyme catalysis. *Biochem* 2008;47:3317–3321. [PubMed: 18298083]
 44. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. CHARMM - A program for macromolecular energy, minimization, and dynamics calculations. *J Comput Chem* 1983;4:187–217.
 45. Fiser A, Do RKG, Sjöli A. Modeling of loops in protein structures. *Prot Sci* 2000;9:1753–1773.

46. Fiser A, Sýali A. ModLoop: Automated modeling of loops in protein structures. *Bioinformatics* 2003;19:2500–2501. [PubMed: 14668246]
47. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiórkiewicz-Kuczera J, Yin D, Karplus M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 1998;102:3586–3616.
48. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys* 1983;79:926–935.
49. Brunger AT, Karplus M. Polar hydrogen positions in proteins: Empirical energy placement and neutron-diffraction comparison. *Proteins* 1988;4:148–156. [PubMed: 3227015]
50. Brooks CL III, Karplus M. Solvent effects on protein motion and protein effects on solvent motion: Dynamics of the active-site region of lysozyme. *J Mol Biol* 1989;208:159–181. [PubMed: 2769750]
51. Ryckaert JP, Ciccotti G, Berendsen HJC. Numerical integration of cartesian equations of motion of a system with constraints - Molecular dynamics of n-alkanes. *J Comput Phys* 1977;23:327–341.
52. Steinbach PJ, Brooks BR. New spherical-cutoff methods for long-range forces in macromolecular simulation. *J Comput Chem* 1994;15:667–683.
53. Beglov D, Roux B. Finite representation of an infinite bulk system: Solvent boundary potential for computer-simulations. *J Chem Phys* 1994;100:9050–9063.
54. Im W, Berneche S, Roux B. Generalized solvent boundary potential for computer simulations. *J Chem Phys* 2001;114:2924–2937.
55. Nina M, Beglov D, Roux B. Atomic radii for continuum electrostatics calculations based on molecular dynamics free energy simulations. *J Phys Chem B* 1997;101:5239–5248.
56. Nina M, Im W, Roux B. Optimized atomic radii for protein continuum electrostatics solvation forces. *Biophys Chem* 1999;78:89–96. [PubMed: 17030305]
57. Gao JL, Truhlar DG. Quantum mechanical methods for enzyme kinetics. *Annual Review of Physical Chemistry* 2002;53:467–505.
58. Elstner M, Porezag D, Jungnickel G, Elsner J, Haugk M, Frauenheim T, Suhai S, Seifert G. Self-consistent-charge density-functional tight-binding method for simulations of complex materials properties. *Physical Review B* 1998;58(11):7260–7268.
59. Cui Q, Elstner M, Kaxiras E, Frauenheim T, Karplus M. A qm/mm implementation of the self-consistent charge density functional tight binding (scc-dftb) method. *Journal of Physical Chemistry B* 2001;105(2):569–585.
60. Yang Y, Yu H, York D, Cui Q, Elstner M. Extension of the self-consistent-charge tight-binding-density-functional (scc-dftb) method: third order expansion of the dft total energy and introduction of a modified effective coulomb interaction. *J Phys Chem A* 2007;111:10861–10873. [PubMed: 17914769]
61. Reuter N, Dejaegere A, Maigret B, Karplus M. Frontier bonds in qm/mm methods: A comparison of different approaches. *Journal of Physical Chemistry A* 2000;104(8):1720–1735.
62. König PH, Hoffmann M, Frauenheim T, Cui Q. A critical evaluation of different QM/MM frontier treatments with SCC-DFTB as the QM method. *J Phys Chem B* 2005;109:9082–9095. [PubMed: 16852081]
63. Cui Q, Karplus M. Catalysis and specificity in enzymes: A study of triosephosphate isomerase (*tim*) and comparison with methylglyoxal synthase (*mgs*). *Adv Prot Chem* 2003;66:315–372.
64. Riccardi D, Schaefer P, Yang Y, Yu HB, Ghosh N, Prat-Resina X, König P, Li GH, Xu DG, Guo H, Elstner M, Cui Q. Development of effective quantum mechanical/molecular mechanical (qm/mm) methods for complex biological processes. *Journal of Physical Chemistry B* 2006;110(13):6458–6469.
65. Elstner M, Frauenheim T, Suhai S. An approximate dft method for qm/mm simulations of biological structures and processes. *THEOCHEM* 2003;632:29–41.
66. Grigorenko BL, Rogov AV, Topol IA, Burt SK, Martinez HM, Nemukhin AV. Mechanism of the myosin catalyzed hydrolysis of atp as rationalized by molecular modeling. *Proc Acad Natl Sci USA* 2007;104:7057–7061.

67. Nam K, Cui Q, Gao J, York DM. A specific reaction parameterization for the am1/d hamiltonian for transphosphorylation reactions. *J Theo Comp Chem* 2007;3:486–504.
68. Arantes GM, Loos M. Specific parametrisation of a hybrid potential to simulate reactions in phosphatases. *Phys Chem Chem Phys* 2006;8:347– 353. [PubMed: 16482277]
69. Schaefer P, Riccardi D, Cui Q. Reliable treatment of electrostatics in combined QM/MM simulation of macromolecules. *J Chem Phys* 2005;123Art. No. 014905
70. Roux B. The calculation of the potential of mean force using computer-simulations. *Computer Physics Communications* 1995;91(1–3):275–282.
71. Kumar S, Bouzida D, Swendsen RH, Kollman PA, Rosenberg JM. The weighted histogram analysis method for free-energy calculations on biomolecules .1. the method. *Journal of Computational Chemistry* 1992;13(8):1011–1021.
72. Kumar S, Rosenberg JM, Bouzida D, Swendsen RH, Kollman PA. Multidimensional free-energy calculations using the weighted histogram analysis method. *Journal of Computational Chemistry* 1995;16(11):1339–1350.
73. Fischer S, Karplus M. Conjugate peak refinement - an algorithm for finding reaction paths and accurate transition-states in systems with many degrees of freedom. *Chemical Physics Letters* 1992;194(3):252–261.
74. Rayment I. The structural basis of the myosin ATPase activity. *J Biol Chem* 1996;271:15850–15853. [PubMed: 8663496]
75. Shimada T, Sasaki N, Ohkura R, Sutoh K. Alanine scanning mutagenesis of the switch I region in the ATPase site of *Dictyostelium discoideum* myosin II. *Biochem* 1997;36:14037–14043. [PubMed: 9369475]
76. Li XD, Rhodes TE, Ikebe R, Kambara T, White HD, Ikebe M. Effects of mutations in the γ -phosphate binding site of myosin on its motor function. *J Biol Chem* 1998;273:27404–27411. [PubMed: 9765269]
77. Elstner M, Cui Q, Munih P, Kaxiras E, Frauenheim T, Karplus M. Modeling zinc in biomolecules with the self consistent charge-density functional tight binding (scc-dftb) method: Applications to structural and energetic analysis. *J Comp Chem* 2003;24:565. [PubMed: 12632471]
78. Giese TJ, York DM, et al. QCRNA 1.0: a database of quantum calculations for RNA catalysis. 2006;25:423–433.

Fig. 1.

Illustration of main structural features of the (a–b) post-rigor and (c–d) pre-powerstroke x-ray structures (PDB code 1FMW,[20] and 1VOM[16], respectively) of the myosin motor domain. In (b,d), the key structural motifs of the active site (Switch I, II and P-loop) and those exhibiting major differences between the two states (relay helix, relay loop, wedge loop and the SH1/SH2 helices) are shown. In (a,c), key hydrogen-bonding interactions in the active site region are shown explicitly; the arrows in (a) indicate essential interactions to form upon the post-rigor→pre-powerstroke transition. Active site water molecules are not included for clarity. Unless explicitly stated, the active site is “open” for the post-rigor state and “closed” for the pre-powerstroke state.

Fig. 2.

ATP hydrolysis in the post-rigor and pre-powerstroke states of the myosin motor domain. (a) The schematics of the reaction pathway followed in the QM/MM simulations, which involves Ser 236 as the proton relay. (b) The two-dimensional potential of mean force (PMF, in kcal/mol) for the first step of ATP hydrolysis in the pre-powerstroke state; the x-axis is the distance between the lytic water oxygen and P^γ of ATP, and the y-axis is a collective coordinate that describes the relayed proton transfer involving the lytic water, Ser 236 and the γ-phosphate of ATP. (c) Overlay of the active site for the transition state region in the PMF simulations and ensemble MEP calculations for the pre-powerstroke state. (d) The two-dimensional PMF for the second step of ATP hydrolysis in the pre-powerstroke state; the x-axis is the antisymmetric stretch involving the O^{2γ}H group and O^{3β} of ATP, and the y-axis is the P^γ-O^{3β} distance.

Fig. 3.

Comparison of hydrolysis barrier and properties of the lytic water in the pre-powerstroke state and closed post-rigor state simulations. (a) Minimum energy path (MEP) barriers for the first step of ATP hydrolysis calculated starting from snapshots collected from equilibrium simulations of the two states. The barriers are plotted against the Arg238-Glu459 salt bridge planarity (specified by the $NH_2^{Arg}-NH_1^{Arg}-OE1^{Glu}-OE2^{Glu}$ dihedral angle) and the differential distance between the lytic water and Wat2 in the reactant and transition state (TS1 in Fig. 2). The black dots indicate data for the pre-powerstroke state; the blue, green and red set indicate data from the closed post-rigor simulations with different behaviors of Wat2 (see text).

Fig. 4.

Structural features for ATP hydrolysis in the closed post-rigor state where SwII is displaced to close the active site. (a–b): representative active site structures for the hydrolysis transition state with a planar and twisted Arg238-Glu459 salt-bridge configuration, respectively. (c) The integrated number of water molecules in the active site during different segments of the MD simulations of the closed post-rigor state; results for the pre-powerstroke and open post-rigor states are also shown for comparison. (d) Key hydrogen-bonding interactions in the active site region of the closed post-rigor state; compare to Fig. 1a,c. The arrows indicate interactions that are broken when Switch II is displaced to close the active site in the post-rigor state; i.e., rearrangements in the N-terminus of the relay helix and wedge loop are required to form the stable active site as in the pre-powerstroke state (Fig. 1c).

Table 1
Summary for the performance of SCC-DFTBPR for the energetics and geometry of phosphate reactions.^a

Structures	Property	MAXE	Error RMSE	Analysis ^b MUE	MSE
Training set ^c					
All structures	Energetics ^d	-9.2 (-12.1) [6.1]	3.3 (4.0) [1.4]	2.4 (2.8) [1.0]	-0.4 (-0.5) [-0.2]
Stable states	P-O (Å)	0.41/-0.12	0.05/0.02	0.02/0.02	0.00/0.00
	O-P-O (°)	-11.8/7.8	2.3/1.9	1.4/1.3	-0.1/0.0
Saddle points ^e	P-O ^{lytic} (Å)	-0.40/-0.04	0.14/0.02	0.07/0.02	-0.03/0.01
	P-O ^{leaving} (Å)	-0.33/0.18	0.15/0.08	0.10/0.05	0.03/0.04
	P-O ^{others} (Å)	0.07	0.02	0.02	0.00
O-P-O (°)		11.5/6.5	2.9/2.5	2.0/1.9	0.0/-0.2
QCRNA set ^f					
Stable states	Energetics ^g	14.4 (14.8)	5.3 (5.3)	4.3 (4.5)	1.0 (-0.8)
	P-O (Å) ^h	0.45/0.19	0.07/0.04	0.03/0.02	0.00/0.00
	O-P-O (°) ^h	-9.3	2.2	1.5	0.0
Transition states	Energetics ^g	12.8	6.0	5.4	-1.6

^aIn SCC-DFTBPR, the atomic properties of P and pair-wise P-X repulsive potentials are developed following the standard procedures for SCC-DFTB parameterization.[58,77] The parameters specific to the phosphate hydrolysis reactions are the four Hubbard derivatives, one for each element type (P:-0.07; O:-0.22; C:-0.24; H:-0.08), and three element-independent Gaussian parameters for damping the Hubbard derivative for large charges (D_0 : -0.09; $3Q$: 16.1; $Q0$: 0.75). The terms including the Hubbard derivatives are important because they significantly impact proton affinities[60] and phosphate hydrolysis involves many proton transfer steps.

^bMAXE: Maximal Error; RMSE: Root-Mean-Square Error; MUE: Mean Unsigned Error; MSE: Mean Signed Error.

^cThe reactions used to parameterize SCC-DFTBPR include the hydrolysis of dimethyl monophosphate ester (DMP) and monomethyl monophosphate ester (MMP) with different protonation states; both dissociative and associative mechanisms are considered, together with the possibility of involving water as the proton relay group. All together, 37 gas-phase reaction energies (18 of which are energy barriers) involving 47 structures are included. The geometries are optimized at the B3LYP/6-31+G(d,p) level and energies at MP2 level with the “G3Large” basis set are used as reference.

^dThe numbers without parentheses are errors for single point SCC-DFTBPR energies at B3LYP/6-31+G(d,p) structures; those with parentheses are errors for SCC-DFTBPR energies following optimizations at the same level; those with brackets are errors for MP2/G3Large single point energies at the SCC-DFTBPR structures.

^eO^{lytic}, O^{leaving} and O^{others} indicate the nucleophilic oxygen, leaving oxygen and spectator oxygen atoms, respectively. When there are two entries, the number before the slash is the overall performance and the number after slash excludes one to two extreme deviations.

^f As additional benchmark, all (19) RNA model reactions with the overall charge of -1 are selected from the QCRNA database from the York group.[78] these include 56 reaction energies and 47 barriers. These systems were consistently calculated at the B3LYP/6-311++G(3df,2p) level for energy and B3LYP/6-31++G(d,p) for structure by the York group. Both single point energy calculations and geometry optimizations are carried out at the SCC-DFTBPR level; for transition states, only single point energies are considered.

^g Numbers without parentheses are the errors for SCC-DFTBPR single point energies at the reference (B3LYP/6-31++G(d,p)) structures, those with parentheses are the errors after geometry optimization at the SCC-DFTBPR level.

^h Optimized at the SCC-DFTBPR level for the stable states only; when there are two entries, the number before the slash is the overall performance and the number after slash excludes several extreme deviations.