Supporting information for: ATP hydrolysis mechanism in kinesin studied by combined quantum-mechanical / molecular-mechanical metadynamics simulations

Matthew J. McGrath,[*,†] I-F. Will Kuo,[‡] Shigehiko Hayashi,[¶] and Shoji Takada[*,†]

Dept of Biophysics, Graduate School of Science, Kyoto University, Sakyo, Kyoto 606-8502, Japan, Physical and Life Sciences Directorate, Lawrence Livermore National Livermore, California 94550, USA, and Dept of Chemistry, Graduate School of Science, Kyoto University, Sakyo, Kyoto 606-8502, Japan

E-mail: mcgrath@theory.biophys.kyoto-u.ac.jp; takada@biophys.kyoto-u.ac.jp

Abstract

Kinesin is a molecular motor that hydrolyzes adenosine triphosphate (ATP) and moves along microtubules against load. While motility and atomic structures have been well characterized for various members of the kinesin family, not much is known about ATP hydrolysis inside the active site. Here, we study ATP hydrolysis mechanisms in the kinesin-5 protein

[*]To whom correspondence should be addressed
[†]Kyoto University - Biophysics
[‡]Lawrence Livermore National Laboratory
[¶]Kyoto University - Chemistry

Eg5 by using combined quantum mechanics/molecular mechanics metadynamics simulations. Approximately 200 atoms at the catalytic site are treated by a dispersion corrected density functional and, in total, 13 metadynamics simulations are performed with their cumulative time reaching ~0.7 ns. Using the converged runs, we compute free energy surfaces and obtain a few hydrolysis pathways. The pathway with the lowest free energy barrier involves a two-water chain and is initiated by the $P_\gamma - O_\beta$ dissociation concerted with approach of the lytic water to $P_\gamma O_3^-$. This immediately induces a proton transfer from the lytic water to another water, which then gives a proton to the conserved Glu270. Later, the proton is transferred back from Glu270 to HPO_4^- via another hydrogen bonded chain. We find that the reaction is favorable when the salt bridge between Glu270 in switch II and Arg234 in switch I is transiently broken, which facilitates Glu270 ability to accept a proton. When ATP is placed in the ADP-bound conformation of Eg5, the ATP-Mg moiety is surrounded by many water molecules and Thr107 blocks the water chain, which together make the hydrolysis reaction less favorable. The observed two-water chain mechanisms are rather similar to those suggested in two other motors, myosin and F_1-ATPase, raising the possibility of a common mechanism.

Table of Contents Graphic

Keywords: Eg5; QM/MM; ATP hydrolysis; metadynamics

Supporting Information

Collective Variables

All of the collective variables used in this work are shown in Figures 1 to 5. Figures 1 to 3 are based on the crystal structure of Parke et al.[1] (ATP.N). Figures 4 and 5 are based on the crystal structure of Turner et al.[2] (ADP.N). All of the colvars in Figure 1 begin from the same initial configuration. The same is true for Figures 2, 3 and 5, while all three colvars in Figure 4 start from different initial structures. We have attempted to illustrate the significant differences between the colvars and the initial structures in the schematic drawings. In particular, the Arg234–Glu270 salt bridge is intact in ATP.1, ATP.2, ATP.3, ADP.2, ADP.3, ADP.4, and ADP.5, partially broken in ATP.7, ATP.8, and ADP.1, and completely broken in ATP.4, ATP.5, and ATP.6. One difference shown in the main paper but not illustrated here (due to space) is the openness of the ADP.N structures compared the ATP.N structures.

The differences in the colvars are summarized in Table 1. A few differences cannot be readily summarized in the tabular format, although they can be seen in the figures. These include the differences between ATP.1 and ATP.2, ATP.4 and ATP.5, as well as between ADP.2 and ADP.5. The difference between ATP.1 and ATP.2, as well as the difference between ADP.2 and ADP.5, is that ATP.2 and ADP.5 use the same distances as their counterparts, and then subtract the corresponding bonded distances in the initial structure. This was done to see if this style of colvar made for more efiňĄcient sampling. The difference between ATP.4 and ATP.5 is the choice of the third water molecule in the chain (there are two possibilities in that structure).

Functionals with exact exchange are considered to be the standard in enzymatic systems,[3] however exact exchange is cost-prohibitive for modern molecular simulation codes. Pure functionals, on the other hand, have been shown to be incorrect, while BLYP with empirical dispersion corrections has been shown to perform better than other pure functionals for biological molecules.[4] The second generation correction, while showing very good macroscopic properties for liquid water

under ambient conditions,[5–7] does not perform as well for other compounds.[8] The third generation of dispersion correction has recently been shown to reduce the error of using a pure functional for chemical reactions to that of uncorrected exact exchange functionals,[9] which makes it the best choice for sampling free energy surfaces with quantum chemistry.

Metadynamics

Generally in metadynamics runs, the primary goal is to calculate free energy proïňĄles/surfaces. For the free energy surface deïňĄned by the colvars, one can directly obtain it as the negative of the cumulative biasing potential. In order to estimate the error in the free energy surface of metadynamics simulation, one can apply various formulas.53 This error is the difference between the true free energy surface and the negative of the biasing potential, and depends on the diffusion coefiňĄcient of the collective variable as well as the parameters of the added Gaussian functions. Because the current system involves numerous chemical reactions, however, the free energy surfaces deïňĄned by the colvars are useful but not sufiňĄcient. We need to calculate free energy surfaces for various reaction coordinates that are different from the colvars, which requires an alternative method. As mentioned in the main text, the technique of Bonomi et al.[10] can be used to estimate the error introduced by limited sampling times in our free energy simulations. Using the metadynamics hills method, it can be shown that for a simulation of infinite length, the negative of sum of the applied hills is equal to the underlying free energy surface. The histogram reweighting technique of Bonomi et al.[10] also gives the underlying free energy surface. By comparing the results from the two methods (which are both equal to the underlying free energy surface in the limit of an infinite simulation), we can estimate the error introduced by having finite length simulations. This is shown in Figure 7, and the mean unsigned error between the two curves is the 3 kcal/mol mentioned in the main text.

Table 2 gives some information on the convergence of the various metadynamics runs. The

first column ("Product") indicates the simulation time at which the product state is first seen. This product state is defined by the colvar itself, and often includes the transfer of a proton to a protein residue. The next column indicates when the simulation crosses back into the reactant state, again dictated by the colvars (all bonds are in exactly the same state as they were when the simulation began). The final column gives the total simulation time. If the run never re-crosses into the reactant state (or if it never reaches the product state), a "–" is given. From this table it can be seen that all initial runs except for ADP.4 at least crossed into the product state. Run ATP.5.SCALED does not, but the initial run ATP.5 does. Many of the runs do not recross, however. From examining the trajectories, proton transfers not explicitly included in the colvar appear to trap the system in a given configuration. For example, once the ATP bond breaks, in some runs a proton will spontaneously transfer to the β-phosphate group of the ADP. Since this transfer is not included in the colvar definition, the hills method has difficulty driving the system back across the barrier. It is also interesting to note that two out of five ADP.N runs only crossed into the product state very late in the run, while most of the eight ATP.N runs crossed relatively early. While not conclusive evidence, this suggests that those reaction mechanisms in the ADP.N states are more unfavorable.

Additional Results

The times listed in Table 2 were determined by visual inspection of the trajectories. This effort was guided by examination of the collective variables into which the hills were being placed. Figure 6 plots the value of the main collective variable of several runs as a function of the simulation time. The value of the collective variable in the reactant state is always higher than that of the product state, based on the definitions above. It can be seen that in two of the plotted runs (ATP.2 and ADP.3), the colvar moves to the product value and returns to the reactant value. Examination of the trajectories in these areas revealed the correct structures, and so the runs were rescaled and restarted with smaller hills. In the SCALED runs (the lighter, dotted lines), ADP.3 explores both

the product and the reactant values, while ATP.2 spends most of its time in the reactant region. Near the end of the ATP.2.SCALED run (around 28 ps), it finally crosses over into the product well. ATP.3 is able to cross into the product state fairly early, but it then gets trapped. The same happens to ATP.2 after recrossing the barrier the third time. None of these configurations are included in the analysis, since in those regions it appears the collective variables are insufficient to properly describe the system.

Representative structures for ATP.2 and ATP.4 are shown in Figures 8 and 9. One noticeable feature of ATP.4 is the number of different structures found in the reactant and product wells. It appears that ATP.4 is able to support the formation of hydroxide and hydronium ions more easily than other structures, leading to more structural diversity. The reason for this is possibly due to the proximity of the third water molecule in the colvar to the positively charged Arg234 residue. In ATP.4, this water is situated between 3.2-5.1 Å from the branch carbon atom of Arg234, while in ATP.5 the third water lies between 3.8–6.4 Å from this same carbon. The closer approach of the water in ATP.4 may allow easier formation of the hydroxide ion. As with those structures shown in the main article for ATP.5 and ADP.3, the structures in the vicinity of the transition state show elongated or broken ADP-P_i bonds, consistent with a concerted or dissociative mechanism. ATP.2 also shows formation of the $H_2PO_4^-$ ion, which is more stable than the "product" defined by the collective variable (HPO_4^- with a protonated Glu270). Interestingly, ATP.4 never reaches this state. This may be due to the differences in the ATP.4 and ATP.5 structure after the reaction occurs, which is explored in more detail below. The third observation made in the main article (water as the attacking nucleophile instead of hydroxide) is seen here as well. This becomes more interesting when one examines various reactant structures in ATP.4. In particular, the hydroxide ion does form in the active site, sometimes very close to the ATP. However, these states are not stable and when nucleophilic attack finally happens, it is by a full water molecule. As mentioned in the main paper, this is probably due to the negative charge on both the hydroxide and ATP.

Two kinds of geometric data are plotted in Figures 10 to 15: distances defining the chemical

reaction, and the coordination number of several heavy atoms in the active site. The latter is given by:[11]

$$CN(r) = \sum_{\mathrm{i,j}} \frac{1-(r_{\mathrm{ij}}/r_0)^n}{1-(r_{\mathrm{ij}}/r_0)^m} \tag{1}$$

where r_{ij} is the distance between atom i and atom j, r_0 is a "cutoff" distance, and m and n are parameters which control the sharpness of the cutoff. We found that $m = 24$ and $n = 12$ dampen out the fluctuations to the extent that the major transitions can more easily been seen. For P–O and O–H (looking at the number of oxygens and hydrogens around phosphorous and oxygen atoms, respectively), we found values of r_0 of 2.2 and 1.3 Å, respectively, gave the expected coordination numbers for stable molecules (e.g. $CN = 2$ for the oxygen atom in water).

Mechanistic details of the ATP hydrolysis reaction can be seen in Figures 10 and 11 for runs ATP.2 and ATP.4, including the evolution of coordinate numbers of select atoms and cartoon schematics for all four successful unscaled runs in Figures 12 to 15. The important atoms in the reaction are given in Figures 12 and 13. For ATP.2, there is no difference between the first reaction mechanism shown in Figure 12 and the collective variable in Figure 1. The reaction proceeds very much as one would expect, according to the coordination numbers of various atoms given in the lower panel of Figure 12 and key reaction distances given in Figure 10. The water clearly attacks the terminal phosphate before losing a proton (the attack happens in the top panel of Figure 10 at 10.9 ps, while the proton transfer happens around 11.8 ps). The proton of W2 is then transferred to Glu270 at 11.9 ps. The most interesting event after this is the formation of the $H_2PO_4^-$ ion at 19.6 ps, which can be seen in the CN graph in Figure 12 as well as the lower panel of Figure 10.

ATP.4 is a bit more unexpected. The reaction path depicted in Figure 13 is not the same as the colvar in Figure 2. In particular, Ser232 acts as a proton relay and the third water molecule is not used. It is interesting that this free energy difference is not appreciably different from that of the two-water mechanism (ATP.5), within the uncertainty of the simulations. The plot of the coordination number during the period of the reaction (Figure 13) reveals that W2 and Glu270

spend some time sharing a proton before W1 transfers its proton to form HPO_4^{-2}, after which the coordination number of all the residues is much more stable. Further details in Figure 11 show that W1 attacks ATP at $t = 21.0$ ps and loses its proton at $t = 21.1$ ps (top panel), while Ser232 and W2 finish transferring their protons simultaneously at $t = 21.1$ ps (bottom panel). As was also seen in the plot of the CN, W2 and Glu270 spend about 0.2 ps sharing a proton immediately before the reaction. It should be noted here that the proton transfers very quickly back and forth between the two oxygen atoms, and does not hover at an intermediate distance; that appearance is simply due to the averaging procedure employed to smooth out the curves. The reason that the $H_2PO_4^-$ is not seen in ATP.4 but is seen in ATP.5 seems to be the location of the transferring water (W3 in Figure 2 in the main text) with respect to Thr107 and Glu270. In ATP.5, Glu270 is able to transfer a proton directly to W3, which is in the proper position to transfer a proton to Thr107, which can then transfer a proton to the phosphate. In ATP.4, when W3 is in the proper position to accept the proton from Glu270, it is much too far away to transfer it to Thr107. W3 only approaches Thr107 after W2 injects itself into the hydrogen bonding network between Glu270 and W3, and it seems the Glu270-W2-W3-Thr107-P_i transfer is not as favorable as Glu270-W3-Thr107-P_i.

In addition to these plots, we have included two movie files in this Supporting Information: one of run ATP.5 (ATP.5.mpg) and one of run ADP.3 (ADP.3.mpg). These two runs were chosen because ATP.5 has the lowest energy barrier and most stable product state, while ADP.3 is the only run from the crystal structure of Turner et al.[2] which successfully sampled the barrier; this is probably due to only two distances being included in the collective variable description, instead of three or more as is included in the others. The movies show the trajectories beginning from the metadynamics run, progressing to the product state, and terminating after returning to the reactant state. Several important observations can be made here that are also mentioned in the main paper. The first is that the mechanisms appear to be either concerted or dissociative, with the metaphosphate ion (PO_3^-) forming before attack by the nucleophile. The second is that water, not hydroxide, always attacks the metaphosphate ion. As mentioned in the main text, this could be due

to electrostatic repulsion between the two negatively charged ions. Finally, the product $H_2PO_4^-$ is seen in both runs, despite that formation of this product is not included explicitly in the collective variable description. This suggests that the singly charged inorganic phosphate is the true product state of the reaction. All of these conclusions are explored more methodically in the main text of this article.

Two movies are available for download for the Supporting Information. The first (ATP.5.mpg) is an animated trajectory of run ATP.5 around the time of the proton transfers. Atoms which are explicitly included in the collective variables are indicated by spheres. The triphosphate tail of ATP is shown on the left, while residues Arg234 and Glu270 are shown on the top-right and bottom-right, respectively. Cleavage of the ADP-P bond, attack of the water, and proton transfer to the Glu270 all happen simultaneously around 14 s. This is followed by proton transfer back to the inorganic phosphate at around 18 s, the reverse reaction at 22 s, and reprotonation of HPO_4^{-2} at 25 s.

The second (ADP.3.mpg) is an animated trajectory of run ADP.3 around the time of the proton transfers. Atoms which are explicitly included in the collective variables are indicated by spheres. The triphosphate tail of ATP is shown on the left, while residues Arg234 and Glu270 are shown on the top-right and bottom-right, respectively. Cleavage of the ATP bond can be seen starting at around seven seconds, while attack by water occurs at around 15 seconds and proton transfer to the backbone carbonyl at 16 s. The animation finishes with an inorganic phosphate.

Table 1: A summary of the differences between the collective variables in all thirteen metadynamics runs. “Intact”, “Partially intact”, and “Broken” for the second column indicate that two, one, and zero hydrogen bonds, respectively, are made between Arg234 and Glu270 in the initial structures. The third column indicates the number of water molecules involved in the proton transfer chain. If a protein residue acts as a proton shuttle, this is indicated by “+ Res”.

Run	State of Arg234–Glu270 salt bridge	Final location of proton transfer	No. water molecules in water chain	Initial structure
ATP.1	I	Glu270	2	ATP.1
ATP.2	I	Glu270	2	ATP.1
ATP.2.SCALED	I	Glu270	2	ATP.1
ATP.3	I	HPO_4^{-2}	2	ATP.1
ATP.4	B	Glu270	3	ATP.4
ATP.4.SCALED	B	Glu270	3	ATP.4
ATP.5	B	Glu270	3	ATP.4
ATP.5.SCALED	B	Glu270	3	ATP.4
ATP.6	B	HPO_4^{-2}	2	ATP.4
ATP.7	PI	Glu270	3	ATP.7
ATP.8	PI	HPO_4^{-2}	2	ATP.7
ADP.1	PI	HPO_4^{-2}	1+Ser233	ADP.1
ADP.2	I	Glu270	1+Thr107	ADP.2
ADP.3	I	Gln106	1	ADP.3
ADP.3.SCALED	I	Gln106	1	ADP.3
ADP.4	I	HPO_4^{-2}	2	ADP.4
ADP.5	I	Glu270	1+Thr107	ADP.2

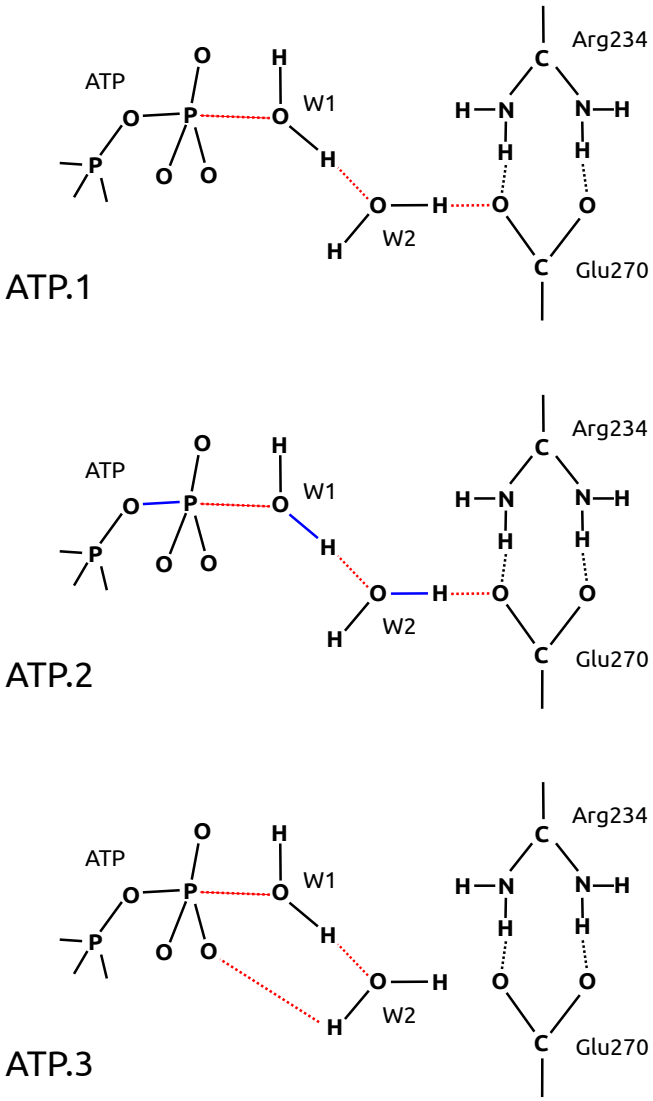

Figure 1: The collective variables for runs 1, 2, and 3, starting from the crystal structure of Parke et al.[1]. Red and blue lines indicate that the distances are added or subtracted from the colvar, respectively. Solid and dotted lines indicate bonds formed in the initial (reactant) structure and the final (product) structure, respectively.

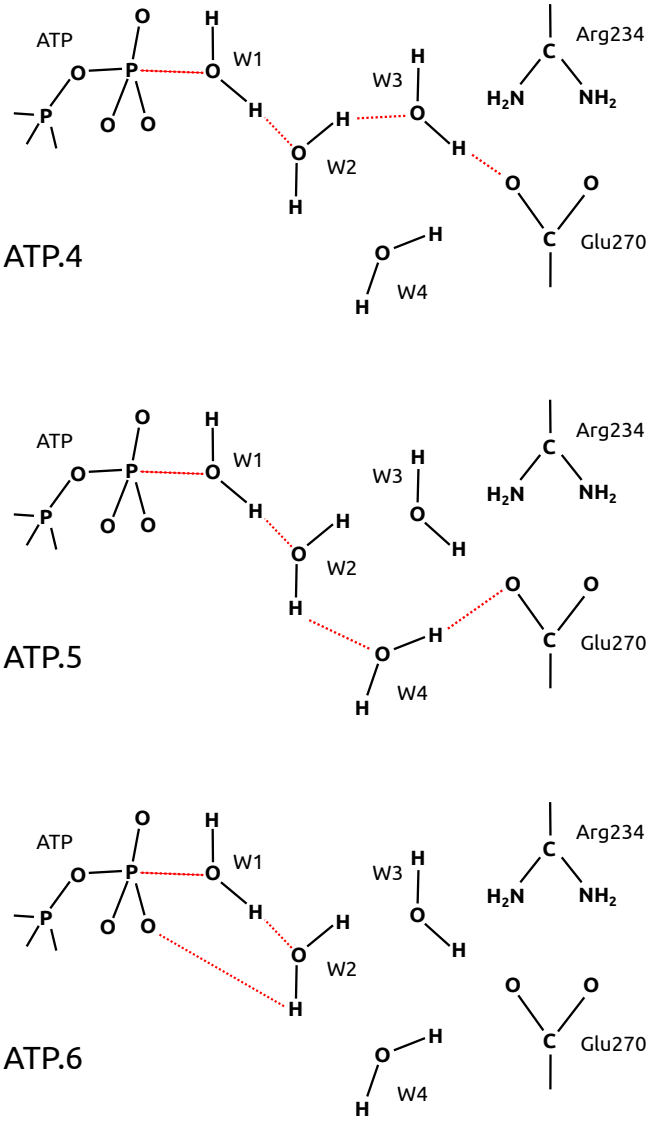

Figure 2: The collective variables for runs 4, 5, and 6, starting from the crystal structure of Parke et al.[1]. Legend is the same is in Figure 1.

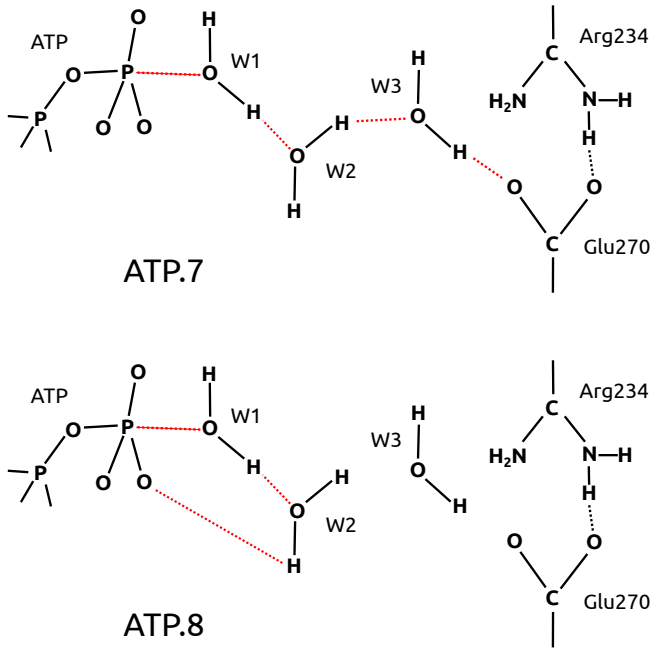

Figure 3: The collective variables for runs 7 and 8, starting from the crystal structure of Parke et al.[1]. Legend is the same is in Figure 1.

Table 2: The time when each run first visits the product state (defined by the collective variable) and recrosses to the reactant state. The total simulation time for each run is given in the final column.

Run	Product [ps]	Reactant [ps]	Total time [ps]
ATP.1	8.4	–	57.0
ATP.2	12.0	29.5	56.9
ATP.2.SCALED	27.1	–	27.6
ATP.3	37.5	–	55.6
ATP.4	20.9	28.2	58.0
ATP.4.SCALED	18.0	–	26.4
ATP.5	15.5	32.3	57.7
ATP.5.SCALED	–	–	22.9
ATP.6	18.6	–	55.7
ATP.7	7.3	–	56.6
ATP.8	36.8	–	43.7
ADP.1	40.6	–	45.0
ADP.2	45.1	–	45.4
ADP.3	7.7	19.0	44.8
ADP.3.SCALED	4.2	12.3	27.1
ADP.4	–	–	44.9
ADP.5	9.6	–	37.0

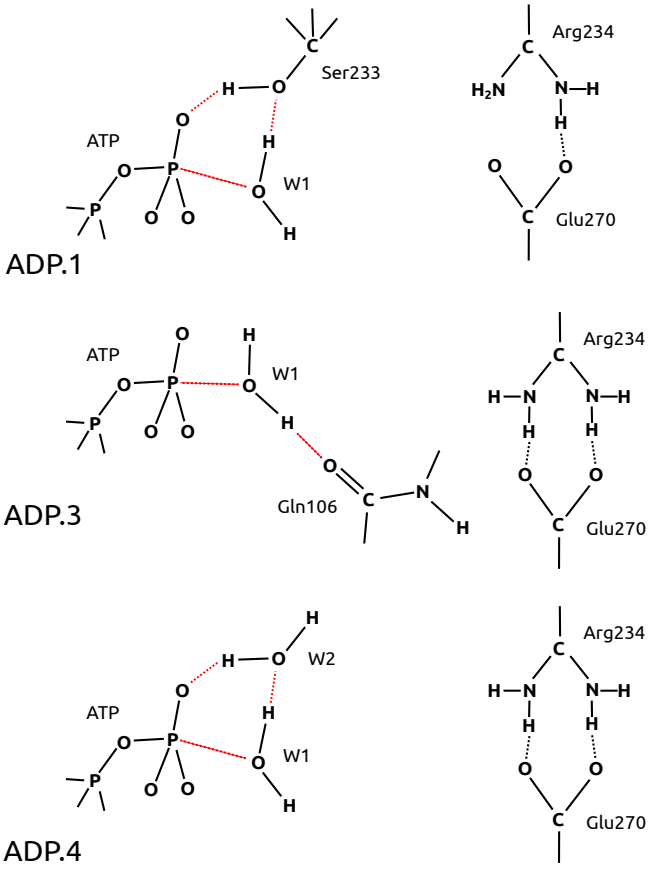

Figure 4: The collective variables for runs 1, 3, and 4, starting from the crystal structure of Turner et al.[2]. Legend is the same is in Figure 1.

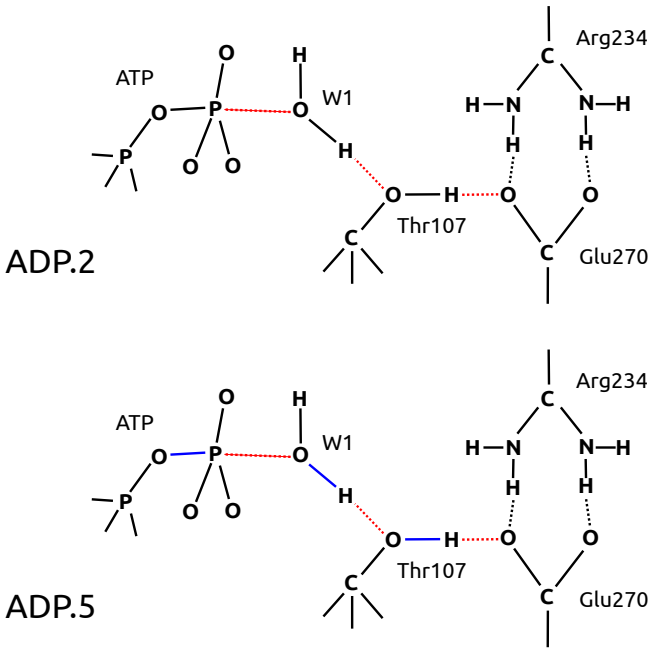

Figure 5: The collective variables for runs 2 and 5, starting from the crystal structure of Turner et al.[2]. Legend is the same is in Figure 1.

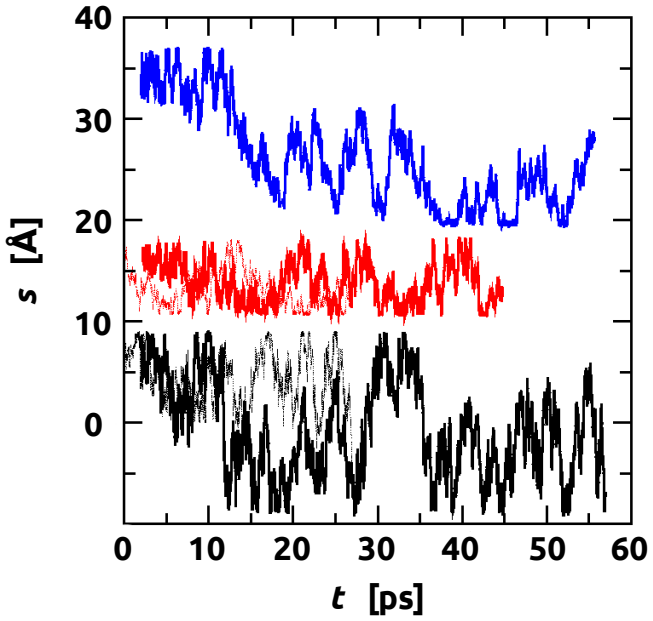

Figure 6: A plot of some collective variables of various runs as a function of the simulation time. Black, red, and blue represent ATP.2, ADP.3, and ATP.3, respectively, with dotted lines depicting runs which had been rescaled. The lines are shifted by an arbitrary amount for clarity.

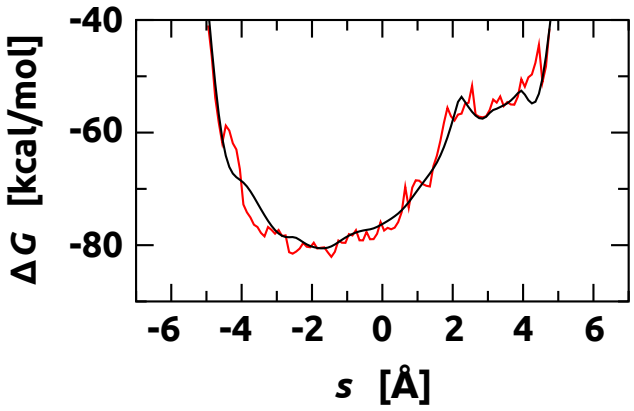

Figure 7: A plot of the free energy surface at $t = 56.8$ ps for run ATP.2 as a function of the original two dimensional colvars compressed into a single dimension. The black and red lines were computed from the deposited hills and the reweighting method of Bonomi et al.[10], respectively.

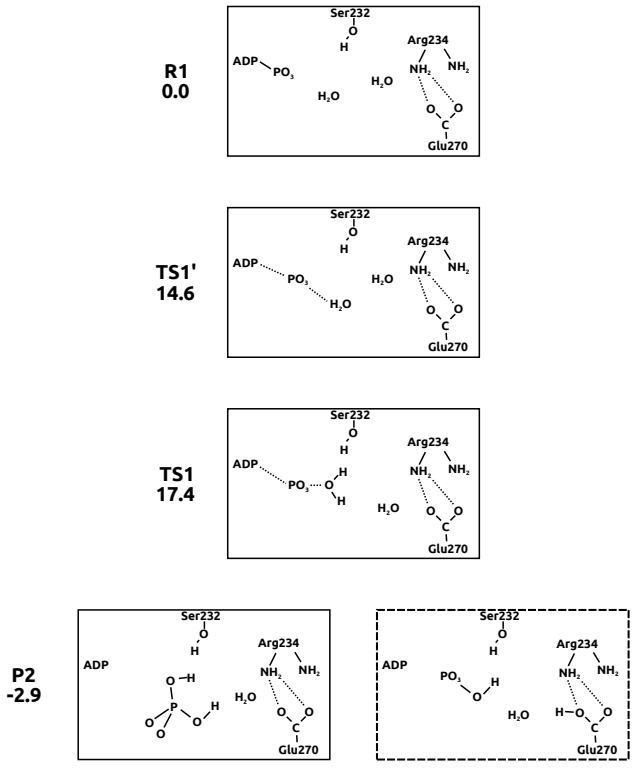

Figure 8: Schematic representations of the reactant, product, and transition state structures marked on Figure 5 in the main text for run ATP.2. A solid box indicates the most stable structure at each point. A prime indicates a less important state, i.e. a more stable transition state or a less stable reactant/product.

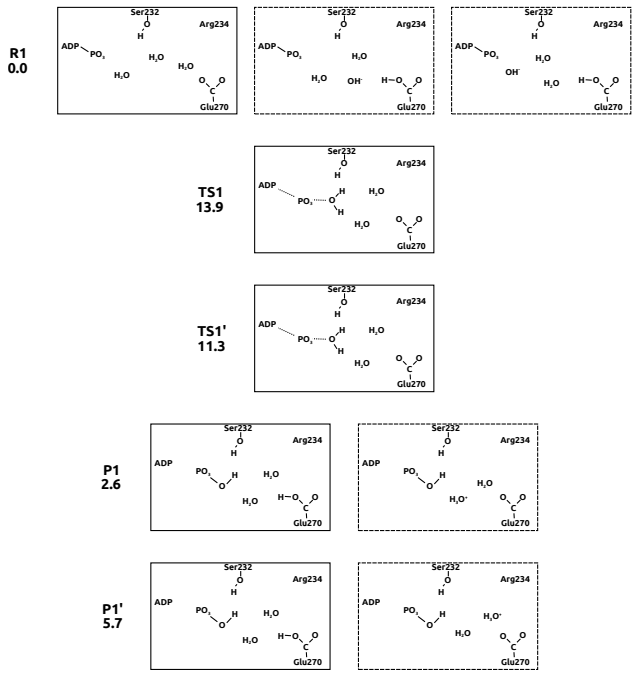

Figure 9: Schematic representations of the reactant, product, and transition state structures marked on Figure 5 in the main text for run ATP.4. A solid box indicates the most stable structure at each point. A prime indicates a less important state, i.e. a more stable transition state or a less stable reactant/product.

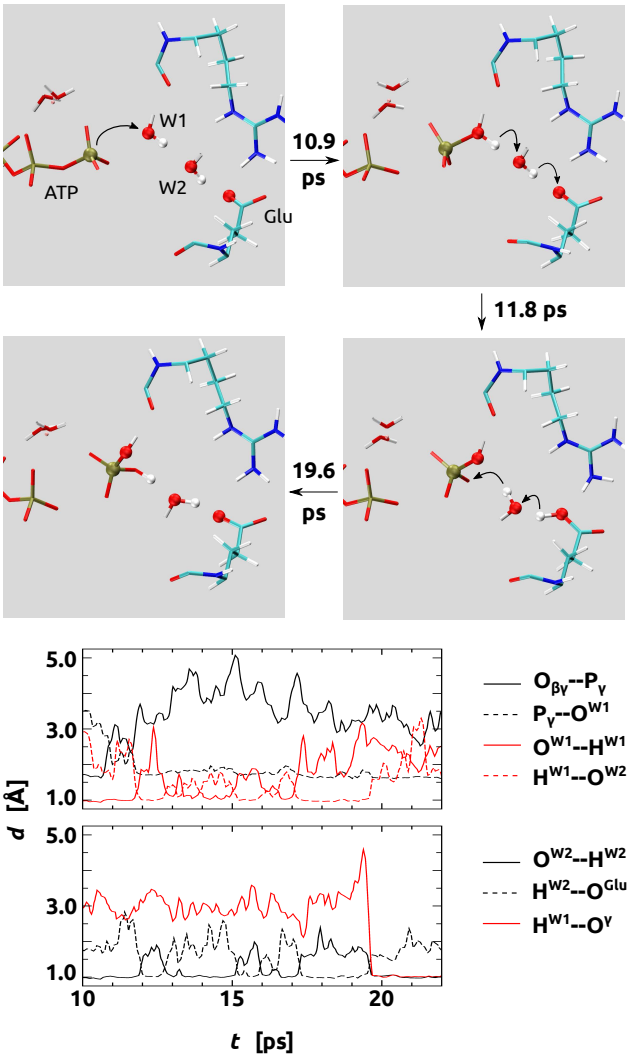

Figure 10: The value of various distances during run ATP.2 as a function of simulation time, focusing on the ATP dissociation reaction. The atom labels are shown in Figure 12.

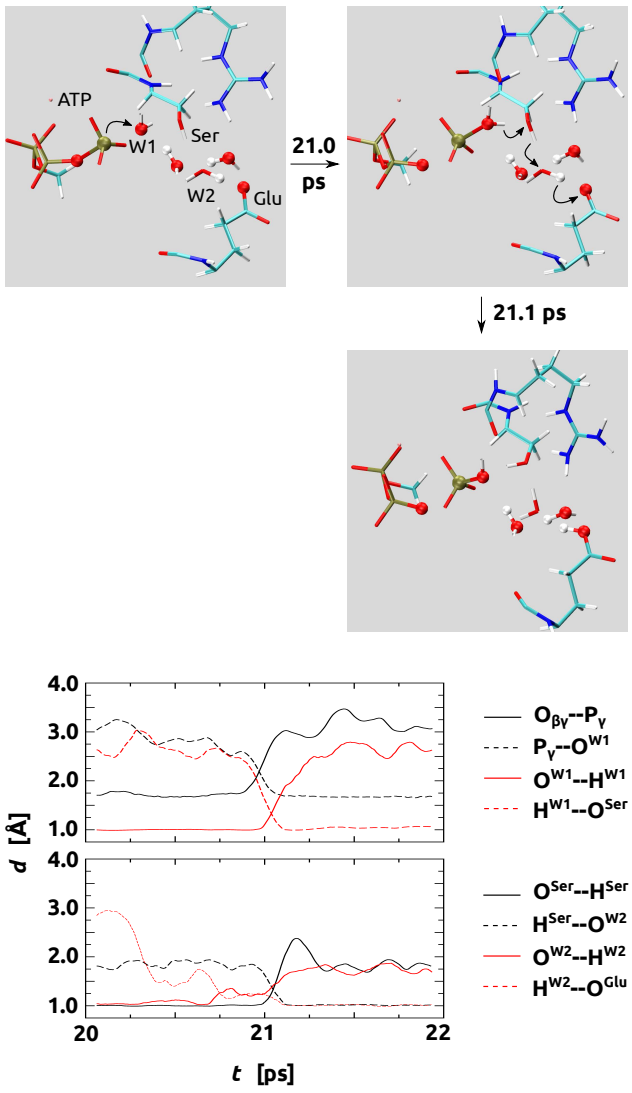

Figure 11: The value of various distances during run ATP.4 as a function of simulation time, focusing on the ATP dissociation reaction. The atom labels are shown in Figure 13.

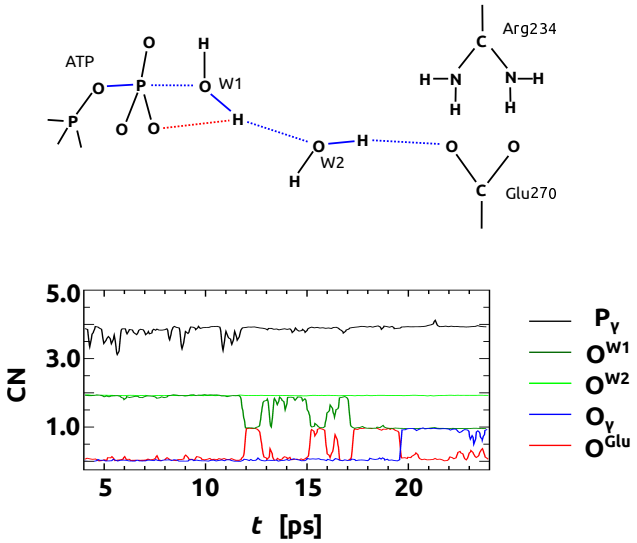

Figure 12: The approximate reaction scheme of ATP.2. Blue distances indicate the reaction at around 12 ps, while red indicates the reaction around 20 ps (in addition to the H^{W2}–O^{W2} bond reforming). The value of various coordination numbers of key atoms during run ATP.2 as a function of simulation time is shown in the lower panel.

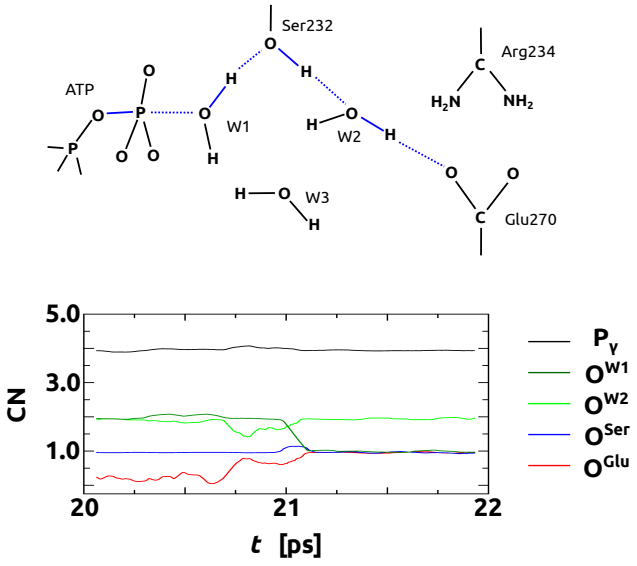

Figure 13: The approximate reaction scheme of ATP.4. Blue distances indicate the primary ATP hydrolysis reaction mechanism. The value of various coordination numbers of key atoms during run ATP.4 as a function of simulation time is shown in the lower panel.

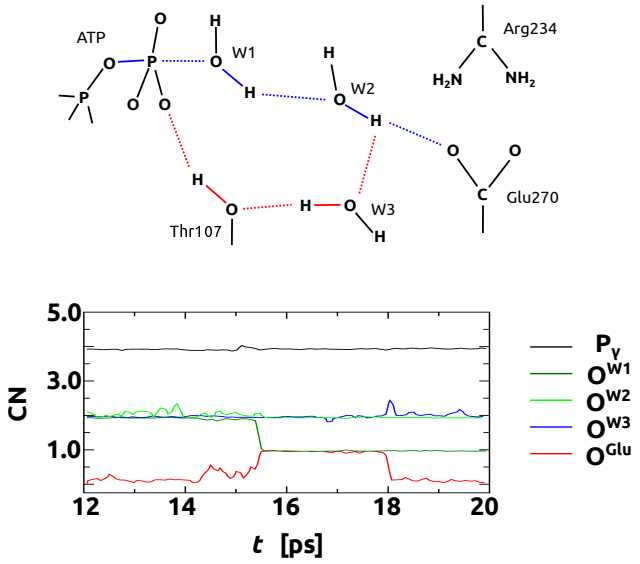

Figure 14: The value of various coordination numbers of key atoms during run ATP.5 as a function of simulation time, focusing on when the important reactions occur. The approximate reaction scheme is shown in the lower panel. Blue distances indicate the first reaction (before approximately 16 ps), while the red indicates the second reaction (at approximately 18 ps).

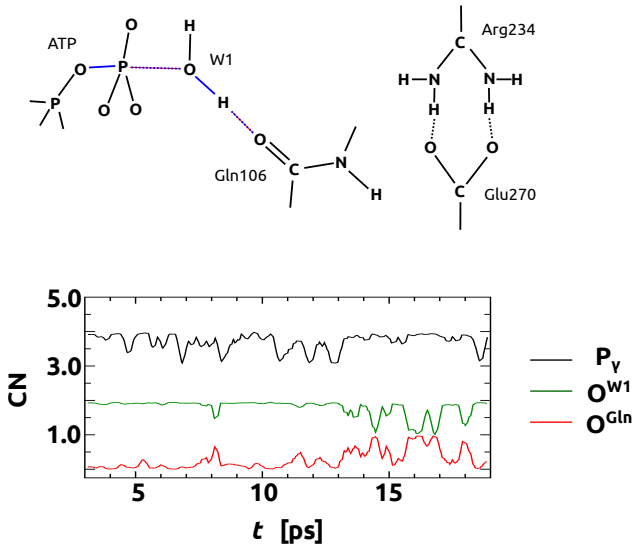

Figure 15: The value of various coordination numbers of key atoms during run ADP.3 as a function of simulation time, focusing on when the important reactions occur. The approximate reaction scheme is shown in the lower panel. Blue distances indicate the first reaction (before approximately 16 ps), while the red indicates the second reaction (at approximately 18 ps).

References

(1) Parke, C. L.; Wojcik, E. J.; Kim, S.; Worthylake, D. K. *J. Biol. Chem.* **2010**, *285*, 5859–5867.

(2) Turner, J.; Anderson, R.; Guo, J.; Beraud, C.; Fletterick, R.; Sakowicz, R. *J. Biol. Chem.* **2001**, *276*, 25496–25502.

(3) Friesner, R. A.; Guallar, V. *Annu. Rev. Phys. Chem.* **2005**, *56*, 389–427.

(4) Antony, J.; Grimme, S. *Phys. Chem. Chem. Phys.* **2006**, *8*, 5287–5293.

(5) Schmidt, J.; VandeVondele, J.; Kuo, I.-F. W.; Sebastiani, D.; Siepmann, J. I.; Hutter, J.; Mundy, C. J. *J. Phys. Chem. B* **2009**, *113*, 11959–11964.

(6) Kühne, T. D.; Pascal, T. A.; Kaxiras, E.; Jung, Y. *J. Phys. Chem. Lett.* **2011**, *2*, 105–113.

(7) Baer, M. D.; Mundy, C. J.; McGrath, M. J.; Kuo, I.-F. W.; Siepmann, J. I.; Tobias, D. *J. Chem. Phys.* **2011**, *135*, 124712.

(8) McGrath, M. J.; Kuo, I.-F. W.; Ghogomu, J. N.; Mundy, C. J.; Siepmann, J. I. *J. Phys. Chem. B* **2011**, *115*, 11688–11692.

(9) Grimme, S. *Wiley Inter. Rev.: Comput. Mol. Sci.* **2011**, *1*, 211–228.

(10) Bonomi, M.; Barducci, A.; Parrinello, M. *J. Comput. Chem.* **2009**, *30*, 1615–1621.

(11) Laio, A.; Gervasio, F. L. *Rep. Prog. Phys.* **2008**, *71*, 126601.