First Round of Reviewer Comments

Reviewer: 1

Comments to the Author

The manuscript by Shekhar et al. describes a joint experimental-computational study on the V-ATPase rotary mechanism. A newly discovered crystallographic state is described and studied by molecular simulation to connect with other known states. My review will focus on the computational aspects of the paper.

There is considerable room for improvement in the analysis and description of MD results. Most notably, some of the key findings appear to be described as more confident or definite than they probably are, at least based on the analysis shown. I realize the journal calls for a condensed writeup but the authors must present high-quality analysis.

Most importantly, the authors should describe the uncertainty associated with all their findings. If the sampling does not permit a reliable estimate of uncertainty, that should be stated.

Here are some more specific comments:

Is the analysis for Fig. 4 based solely on contacts (by some distance threshold)? That should be stated. The text ‘speaks’ as if allosteric were perfectly captured by the analysis, whereas I am not sure that allosteric principles are fundamentally understood in the present day.

There is no uncertainty analysis associated with the fairly small differences shown in Fig. S2. Are they significant?

The paper uses two advanced simulation methods, REST and metadynamics. Since the report is meant for a more general audience, the authors should briefly describe the methods and note their limitations.

One of the least persuasive calculations reported is the calculation and comparison of kinetics for inward vs outward directions. In my view, kinetics for complex systems estimated from one-dimensional free energy profiles are highly questionable. The authors should address this point. The timescales reported are 0.34 (two significant figures!) and 0.5 ms. The authors should estimate uncertainties for these values. If they are correct as given, then it’s clear both pathways will be used nearly equally. But as I understand it, no unbinding through the inner pathway is observed ... and yet the authors conclude on p. 5 that the inward kinetics will govern the process. Why? What is the evidence the inward pathway even occurs?
Another set of calculations that appears problematic to me is for “transfer entropy” (confusingly called a different name, mutual information, in methods). Presumably the data presented result from the unreadable Fig. S8. Is there signal there or just noise? Uncertainty must be quantified. Also mutual information, as I understand, does not convey directionality/causation but only correlation. The text should be amended to clarify this. Any statements should be clearly justified by the data.

On p. 8, when the authors state that the ‘cumulative time of catalytic dwell … takes no more than 0.5 ms’, they should indicate the basis for this. If it’s from an uncertain calculation, that should be indicated.

The notation and definitions around eq (2) are unclear or sloppy. Please update carefully.

Eq (3) has an extraneous parenthesis in numerator.

Fig S4 is missing a reference and should describe how the plot was obtained.

It wasn’t completely clear from the text which aspects of the model in Fig. S1 were new. That should be stated more clearly and indicated graphically.

I suggest to visually highlight the new state in Fig. 3.

I didn’t understand the use of the word ‘relegated’ on p. 4.

Reviewer: 2

Comments to the Author

The authors determined the structure of V1-ATPase in presence of ADP.AlF4- by X-ray crystallography and, hence an important missing intermediate which will help elucidating critical aspects of the function of this prototypic molecular machine. The structure they obtained has two ADP.AlF4--bound catalytic sites and one empty catalytic site. Compared to the structure of 2(AMP.PNP)V1 obtained previously, the authors suggest that for the present structure, one of the ADP.AlF4--bound catalytic sites in 2(ADP.AlF4--)V1 corresponds to a state that enables phosphate release after ATP hydrolysis (Pi-release dwell).

Substantiating this claim, the authors also performed all-atom MD simulations, starting from the structure of 2(AMP.PNP)V1 with AMP.PNP in the tight site replaced by ADP.Pi. In the simulations, statistically significant deviations from both the catalytic dwell model and the ATP-binding dwell model were seen, together with some similarity with the structure of 2(ADP.AlF4-)V1. This result is suggested as corroboration that the structure of 2(ADP.AlF4-)V1 showed a distinct dwell state that should occur after the catalytic dwell and before the ATP-binding dwell.

Additional replica exchange simulations with solute tempering (REST2) and well-tempered Funnel metadynamics identified two potential transition pathways, connecting the newly identified Pi-release dwell with the ATP-binding dwell, namely, inward (internally towards the stalk) and outward (away from the stalk). Analysis of the energetics and kinetics of these pathways suggests a rate-limiting step as well as an allosteric regulation mechanism that can control the transition from the Pi-release dwell to the ATP-binding dwell.
In summary, this study tentatively identified a new catalytic intermediate of V1-ATPase and provided new insights about a part of rotatory catalysis, namely, the transition from the catalytic dwell to the ATP-binding dwell. These findings represent important advances in our understanding of mechano-chemical energy conversion and molecular motors in general, and will certainly be relevant for a larger part of the readership.

Note that, due to lack of expertise, I cannot comment on the quality of the crystallography, and will focus at the theoretical parts.

Major issues:

(1) Maybe I missed it, but I failed to find the evidence for one of the conclusion of kinetics-controlling pathway (“Therefore, we suggest that Pi is released before ADP, and the duration of the Pi-release dwell will be controlled by the kinetics of the inward release pathway”, page 5, bottom left corner). Can the authors provide support for this claim?

(2) It seems confusing to me that the authors compared their 2(ADP.AlF4-)V1 structure to the transition state structure of F1-ATPase observed by Menz et al. (R. Menz et al., Cell, 2001). The authors interpreted one of the ADP.AlF4--bound catalytic site in 2(ADP.AlF4-)V1 as the Pi-release dwell (post-hydrolysis, pre-product release), whereas in the transition state structure of F1-ATPase, the ADP.AlF4--bound catalytic sites are interpreted as catalytic transition states of ATP hydrolysis, and it is the ADP+sulfate-bound catalytic site that is interpreted as the post-hydrolysis, pre-product release state. Hence, it appears that similar terms were used to refer to potentially different states, which might be misleading. If so, could the authors further clarify their reason for designating the newly observed structure as Pi-release dwell?

(3) Some conclusion rest on the observation of a significant drift (larger RMSD) seen in the MD simulations. I am not sure if I saw the necessary control, i.e., the RMSD of a simulation for which no drift is expected. Also, in order to see whether the RMSD histograms in Fig S3 result from a drift or from mere fluctuations, RMSD time traces would be helpful.

Minor:

(4) Please provide error estimates / error bars for the free energy calculations, Figs 5C,D and S4.

(5) The statement “such conformational differences were much lesser (between 0.1-0.6 Å) with the empty and tight sites” is unclear, but it seems to be an important argument for the conclusion of designating a Pi-release dwell.
Author’s Response to Peer Review Comments:

Tempe, March 25th, 2022

Dear Professor Editor,

on behalf of my coauthors, Mrinal Shekhar, Chitrik Gupta, Kano Suzuki and co-corresponding author Takeshi Murata, I would like to submit revisions manuscript, entitled Revealing a hidden intermediate of rotatory catalysis with X-ray crystallography and Molecular simulations, to be considered for publication in the ACS Central Sciences as a research article.

The reviews have offered us an opportunity to repeat a few more replicas of metadynamics simulations, and put together additional statistics for error analysis of the reported thermodynamic and kinetic estimates of molecular motor action. In addition, we have revisited the literature to find more experimental data, corroborating with our simulated mechanism of hexameric ATPase’s product release pathways. We believe these revisions indeed add to the confidence of our computational results. Detailed rebuttals are provided separately.

Trusting that the present manuscript meets the standards of ACS Central Sciences. Sincerely,

Abhishek Signharoy, Ph.D.

Assistant Professor, School of Molecular Sciences Biodesign Center for Applied Structural Discovery
Arizona State University

Web: https://web.asu.edu/abhi.

Phone: 812-369-3268
The manuscript by Shekhar et al. describes a joint experimental-computational study on the V-ATPase rotary mechanism. A newly discovered crystallographic state is described and studied by molecular simulation to connect with other known states. My review will focus on the computational aspects of the paper.

There is considerable room for improvement in the analysis and description of MD results. Most notably, some of the key findings appear to be described as more confident or definite than they probably are, at least based on the analysis shown. I realize the journal calls for a condensed writeup but the authors must present high-quality analysis.

Most importantly, the authors should describe the uncertainty associated with all their findings. If the sampling does not permit a reliable estimate of uncertainty, that should be stated.

We thank the Reviewer for seeking further detail on the error and uncertainty analyses of the reported computations, and for an extremely thorough review. To this end, we have now repeated the metadynamics simulations of the so-called inward P$_i$-release pathway two more times that has taken the cumulative metadynamics sampling to 2000 ns (for the outward release path) + 3 x 400 ns (for the inward release path) = 3.2 µs. Adding 3 µs of equilibrium MD and ~500 ns of REST2 simulations, we have now simulated this 0.5 M atom system for close to 7 µs (Table S1). Subsequent analyses of these simulations have allowed us to provide confidence metrics to the energy and kinetics estimates, as well as on the mutual information calculations.

Here are some more specific comments:

*Is the analysis for Fig. 4 based solely on contacts (by some distance threshold)? That should be stated. The text ‘speaks’ as if allostery were perfectly captured by the analysis, whereas I am not sure that allosteric principles are fundamentally understood in the present day.*

Following the Reviewer’s suggestion, we have now refrained from explicitly inferring allosteric principles from the cut-off based network models of Fig. 4. Instead, we have interpreted the network to determine correlated dynamics between the A and B subunits of the hexameric ATPase. On Page 4, we note:

“As illustrated in Fig. 4A, prior to hydrolysis, the ATP remains bound to the G235, G237, K238, R262 and F425 residues of the A subunit of the tight pocket and the Arg-finger R350 residue of the B subunit (denoted A$^\prime$ and B$^\prime$, and a similar site-wise nomenclature is followed for all other AB pairs). Consequently, the correlated dynamics of the A$^\prime$ and B$^\prime$ subunits is regulated by the interface-bound ATP (Fig. 4C). On breaking the covalent bond to the terminal phosphate, the ADP stays connected to the A subunit, while the P$_i$ interacts primarily with the Arg-finger of the B-subunit. Thus, the communication across the bound ATP is lost (Fig. 4D), resulting in a looser ADP·P$_i$-bound AB interface. This looseness
of the ADP·P_i bound interface is reflected in elevated fluctuations of the A' subunit post hydrolysis (Fig. S2).”

Also, in the Methods section (Page 10), we have stated the distance threshold used in the computations:
“Edges are drawn between nodes whose residues are within a cutoff distance of 4.5 Å for at least 75% of an MD trajectory.”

*There is no uncertainty analysis associated with the fairly small differences shown in Fig. S2. Are they significant?*

Error bars are now added to the RMSF analysis using a 10-fold resampling procedure, wherein we have removed 50% of the data randomly and repeated the computation ten times. The differences in flexibility between the ATP- and ADP-bound proteins are found to be higher than the standard deviation within each resampling result. Thanks to the Reviewer’s suggestion, this analysis adds confidence to our inferences on the looseness of the protein-protein interfaces following ATP hydrolysis, particularly in regions close to the ATP binding site.

**Fig. S2. P_i release results in a looser AB interface.** Root Mean Square Fluctuation (RMSF) of subunits A and B, forming the ADP/ATP-binding site, shown for the ADP.P_i-bound A' pocket (black) and ATP-bound A' pocket (purple). The binding site is more rigid (lower RMSF) in the ATP-pocket before the hydrolysis reaction. Residues in the vicinity of the ATP binding pocket (G235, G237, K38, R262, F245, R350) are highlighted. Data is collected from the final 100 ns of each of the 3 x 500 ns MD simulations (Table S1). Error bars are added to the RMSF analysis using a 10-fold resampling procedure, wherein we have removed 50% of the data randomly and repeated the computation ten times.

*In my view, kinetics for complex systems estimated from one-dimensional free energy profiles are highly questionable. The authors should address this point. The timescales reported are 0.34 (two significant figures!) and 0.5 ms. The authors should estimate uncertainties for these values. If they are correct as given, then it’s clear both pathways will be used nearly equally. But as I understand it, no unbinding through the inner pathway is observed … and yet the authors conclude on p. 5 that the inward kinetics will govern the process. Why? What is the evidence the inward pathway even occurs?*
We have now performed a convergence analysis of the 1-dimensional PMF derived from the 2D free energy profile and presented it in Fig. S7. Error analysis of the inward pathway benefits from running two new metadynamics simulations.

**Fig. S7:** one-dimensional projection of the free energy of phosphate release along the funnel axis for (A) outward pathway (B) inward pathway. Time series of the phosphate position along the funnel axis captured via funnel metadynamics for (C) outward release (D) inward release. Green band represents the crystal structure position of the phosphate. (F and H) To monitor the convergence of the free energy profile, energy difference between the highest barrier and the lowest minima (labeled using * in panels E and G) was monitored during the course of funnel metadynamics. Akin to work by Okazaki and Hummer (Ref. 18 in article), this barrier height is considered converged when it changes by ≤ 1.5 Kcal/mol over a course of time ≥ 50 ns e.g. after 900
ns in the outward trajectory and 300 ns in the inward trajectory. Thereafter, the error in the A and B panels is computed as the standard deviation of the free energy estimate during this converged part of the trajectory.

Using the error bars on the PMF we have determined the range of kinetics, reported on page 5:

“Free-energy profiles along both pathways have comparable local energy barriers of height 6-7 kcal/mol for initial product release (Fig. S7). However, the inward release pathway has a number of unbinding intermediates (Fig. 5 and Movie 1), which are missing from the outward release mechanism. The transient trapping of P_i in these intermediates suggests that in the solvated protein environment the inward release path out of the ATPase motor can be slower than the outward release mechanism, which we examine using the 1D free energy profiles and Eq. 2 (see Methods). The outward path takes an estimated 0.3±0.1 ms for releasing P_i into the bulk solvent following an 18 Å long pathway. In comparison, inward release requires 1.3±0.4 ms following a 15 Å long pathway, which also detaches P_i from the catalytic site.”

Following the concerns of the Reviewer on deriving 1D profiles, we have added the following comments, stating clearly the limitations of this approach and how the computed rate estimates make biophysical sense in the context of V-type ATPase motors. In agreement with the Reviewer, we have also stated on page 6 that both the inward and outward pathways are relevant to the mechanism of product release.

“We note that this analysis is qualitative, given the kinetic treatment of an 1D energy profile misses the contributions of alternate pathways and hidden barriers on the rate, especially when the profile does not necessarily represent a minimum free energy pathway. Yet, the barriers seen in the metadynamics simulations, despite reflecting our bias on the choice of displacement-based reaction coordinates, remain well within 13 kcal/mol – an upper bound of energy barriers for molecular dynamics within ATP motors set by the net energy released from ATP hydrolysis 9. These values are also comparable to the ones reported for F-type ATPase in 18. So, inferring from the presence of multiple barriers with heights of 6-7 kcal/mol, we attribute the lifetime of the product release dwell to be at least 1.3 ms, enabling the diffusion of P_i out of the ATP pocket. Driving the P_i further out of the protein via the inward path can take even longer, which will be computationally more expensive to model. Our qualitative comparison of the rates still suggests that P_i is released before ADP (Fig. S5), and the duration of the P_i-release dwell will be controlled by the kinetics of the inward release pathway. Approximately 3-fold slower than the outer pathway, this inner route of P_i release remains thermodynamically accessible and will offer a bottleneck for ATP activity by the V-type motor.”

We have now added some peripheral support on the presence of an inward pathway by finding in the literature, the feasibility of noncatalytic phosphate binding sites in F-type ATPases. Though the exact location of these noncatalytic sites is not known, their existence at the interface of subunit A and the central stalk is in line with the intermediate P_i binding site we observe during the inward release pathway. So, we note on page 6:

“The inward pathway requires lesser rearrangement of the product-containing AB interface than the outer one, thus allowing for a small aperture for rapid P_i release from the pocket. However, the P_i still remains non-specifically bound to parts of the stalk-A interface. This result of incomplete P_i unbinding finds support from the experimental observation of ‘noncatalytic’ phosphate binding sites in F-type ATP synthase, also somewhere in A-subunit 46. Nucleotide binding to these sites is expected to play a regulatory role 47. Though the exact locations of the noncatalytic P_i sites are yet unknown, our simulation
now shows the possibility of electrostatically driven secondary P\textsubscript{i} binding in V-ATPase, beyond the primary ATP site and traditional P\textsubscript{i}-binding loop (Movie S1 and Fig. S8). Two more repeats of the metadynamics along the inward path reproduced a similar observation.”

Fig. S8: Noncatalytic Phosphate binding site at the interface of subunit A\textsuperscript{b} and the central stalk seen during the inward P\textsubscript{i} release pathway from subunit A\textsuperscript{t}, including close up of the residues involved in the pocket.

Another set of calculations that appears problematic to me is for “transfer entropy” (confusingly called a different name, mutual information, in methods). Presumably the data presented result from the unreadable Fig. S8. Is there signal there or just noise? Uncertainty must be quantified. Also mutual information, as I understand, does not convey directionality/causation but only correlation. The text should be amended to clarify this. Any statements should be clearly justified by the data.

Following two new metadynamics simulations of the inward release pathway, we have now redone the mutual information analysis (Fig. S9), and found a strong signal in the vicinity of the Arginine finger R350 residue (page 6).

Fig. S9: Bar Plot depicting the average mutual information (MI) between phosphate position along the funnel axis and the backbone RMSD of residues in subunits A and B in e, b and t states. The metadynamics simulation of the inward release pathways was repeated thrice, and the MI information was extracted from Eq. 3 to determine the mean and the standard deviation. Region close to R350 is highlighted to show a strong information exchange occurs between the local conformation and global
hinge-bending movement of the B subunit, particularly in the ADP.P\(_i\)-bound \(t\) state. This correlation decreases systematically from \(B_t\) to \(B_b\) to \(B_e\).

“A transfer entropy or so-called mutual information analysis (see Methods) along the inward pathway further reveals that information is exchanged between the backbone RMSD of the AB binding interface residues and the reaction coordinate vector of the product release (Fig. S9)\(^{10}\). In particular, coupling of the Arg finger R350 conformations and the residues in its vicinity with the hinge movements of the ADP·P\(_i\)-bound \(B_t\) subunit is roughly 2-fold more pronounced than the coupling between any other B-subunit residues with the global interface changes. This correlation between local and global changes mediated specifically by R350 decreases from the \(B_t\) to \(B_b\) to \(B_e\) conformations. Based on this systematic difference in sidechain information between the \(B_t\) and \(B_e\) sites, it is inferred that disengagement of the \(P_i\) from R350 will break the correlation between local and global conformations of the B subunit. Since the breaking of such information channels is energetically expensive \(^{31}\), the unbinding of \(P_i\) from the Arg finger is expected to offer a rate-determining barrier for the transformation of the \(P_i\)-release dwell into the ATP-binding dwell.”

We have also added text to ensure a proper interpretation of the mutual information values:

“Given two random variables \(X\) and \(Y\), mutual information is an information theory metric that quantifies the interdependence between \(X\) and \(Y\) in terms of correlation (and not directionality). Mutual information (MI) is commonly expressed as…”

**On p. 8, when the authors state that the ‘cumulative time of catalytic dwell … takes no more than 0.5 ms’, they should indicate the basis for this. If it’s from an uncertain calculation, that should be indicated.**

Following the Reviewer’s cogent suggestion we have now stated on page 8:

“Finally, comparing all the available kinetic data, we note that a 120° rotation back to the catalytic dwell takes \(\sim 1.4\) ms\(^{25}\) is commensurate to the cumulative time of catalytic dwell \(\rightarrow P_i\) release dwell \(\rightarrow P_i\) release \(\rightarrow ATP\)-binding dwell transition that takes at least between 0.3±0.1 to 1.3±0.4 ms.”

**I didn’t understand the use of the word ‘relegated’ on p. 4.**

It was a typo that is now replaced by ‘regulated’.

**Eq (3) has an extraneous parenthesis in numerator.**

**Fig S4 is missing a reference and should describe how the plot was obtained.**

**It wasn’t completely clear from the text which aspects of the model in Fig. S1 were new. That should be stated more clearly and indicated graphically. I suggest to visually highlight the new state in Fig. 3.**

All these points have now been clarified. We greatly appreciate the Reviewer for this attention to details.
Reviewer: 2

In summary, this study tentatively identified a new catalytic intermediate of V1-ATPase and provided new insights about a part of rotatory catalysis, namely, the transition from the catalytic dwell to the ATP-binding dwell. These findings represent important advances in our understanding of mechano-chemical energy conversion and molecular motors in general, and will certainly be relevant for a larger part of the readership.

We thank the Reviewer for the positive view on the impact and reach of our contribution. Please find our responses to the concerns herewith.

Maybe I missed it, but I failed to find the evidence for one of the conclusion of kinetics-controlling pathway (“Therefore, we suggest that Pi is released before ADP, and the duration of the Pi-release dwell will be controlled by the kinetics of the inward release pathway”, page 5, bottom left corner). Can the authors provide support for this claim?

Between pages 5 and 6 we note:

“Free-energy profiles along both pathways have comparable local energy barriers of height 6-7 kcal/mol for initial product release (Fig. S7). However, the inward release pathway has a number of unbinding intermediates (Fig. 5 and Movie 1), which are missing from the outward release mechanism. The transient trapping of Pi in these intermediates suggests that in the solvated protein environment the inward release path out of the ATPase motor can be slower than the outward release mechanism, which we examine using the 1D free energy profiles and Eq. 2 (see Methods). The outward path takes an estimated 0.3±0.1 ms for releasing Pi into the bulk solvent following an 18 Å long pathway. In comparison, inward release requires 1.3±0.4 ms following a 15 Å long pathway, which also detaches Pi from the catalytic site.”

“We note that this analysis is qualitative, given the kinetic treatment of an 1D energy profile misses the contributions of alternate pathways and hidden barriers on the rate, especially when the profile does not necessarily represent a minimum free energy pathway. Yet, the barriers seen in the metadynamics simulations, despite reflecting our bias on the choice of displacement-based reaction coordinates, remain well within 13 kcal/mol – an upper bound of energy barriers for molecular dynamics within ATP motors set by the net energy released from ATP hydrolysis”. These values are also comparable to the ones reported for F-type ATPase in 18. So, inferring from the presence of multiple barriers with heights of 6-7 kcal/mol, we attribute the lifetime of the product release dwell to be at least 1.3 ms, enabling the diffusion of Pi out of the ATP pocket. Driving the Pi further out of the protein via the inward path can take even longer, which will be computationally more expensive to model. Our qualitative comparison of the rates still suggests that Pi is released before ADP (Fig. S5), and the duration of the Pi-release dwell will be controlled by the kinetics of the inward release pathway. Approximately 3-fold slower than the outer pathway, this inner route of Pi release remains thermodynamically accessible and will offer a bottleneck for ATP activity by the V-type motor.”

We have now added some peripheral support on the presence of an inward pathway by finding in the literature, the feasibility of noncatalytic phosphate binding sites in F-type ATPases. Though the exact location of these noncatalytic sites is not known, their
existence at the interface of subunit A and the central stalk is in line with the intermediate Pi binding site we observe during the inward release pathway. So, we note on page 6:

“The inward pathway requires lesser rearrangement of the product-containing AB interface than the outer one, thus allowing for a small aperture for rapid Pi release from the pocket. However, the Pi still remains non-specifically bound to parts of the stalk-A interface. This result of incomplete Pi unbinding finds support from the experimental observation of ‘noncatalytic’ phosphate binding sites in F-type ATP synthase, also somewhere in A-subunit.46 Nucleotide binding to these sites is expected to play a regulatory role.47 Though the exact locations of the noncatalytic Pi sites are yet unknown, our simulation now shows the possibility of electrostatically driven secondary Pi binding in V-ATPase, beyond the primary ATP site and traditional Pi-binding loop (Movie S1 and Fig. S8). Two more repeats of the metadynamics along the inward path reproduced a similar observation.”

Fig. S8: Noncatalytic Phosphate binding site at the interface of subunit A and the central stalk seen during the inward Pi release pathway from subunit A, including close up of the residues involved in the pocket.

It seems confusing to me that the authors compared their 2(ADP.AIF4-)V1 structure to the transition state structure of F1-ATPase observed by Menz et al. (R. Menz et al., Cell, 2001). The authors interpreted one of the ADP.AIF4--bound catalytic site in 2(ADP.AIF4-)V1 as the Pi-release dwell (post-hydrolysis, pre-product release), whereas in the transition state structure of F1-ATPase, the ADP.AIF4--bound catalytic sites are interpreted as catalytic transition states of ATP hydrolysis, and it is the ADP+sulfate-bound catalytic site that is interpreted as the post-hydrolysis, pre-product release state. Hence, it appears that similar terms were used to refer to potentially different states, which might be misleading. If so, could the authors further clarify their reason for designating the newly observed structure as Pi-release dwell?

We have now removed the confusing sentence. Instead, we have stated how knowledge on the inhibition of F-ATPase by AIF4-- allowed us to choose the same moiety for inhibiting V-ATPase. We further note that the F- and V-ATPases are known to follow different rotatory mechanisms. Thus, it is difficult to infer the functional significance of the ADP•AIF4--bound V-type model by just drawing analogy with the AIF4--inhibited models of the F-ATPase. Rather, after crystallographically determining the inhibited AIF4--inhibited V
ATPase structure, we place it in the context of other known V-type models to infer the functional significance as a P\textsubscript{i}-release intermediate. So we state on page 3:

“It has been reported that the ATPase activity of bovine mitochondrial F\textsubscript{1}-ATPase is inhibited in the presence of ADP and aluminum fluoride\textsuperscript{12}. In this study, we purified EhV\textsubscript{1} in the presence of ADP and AlF\textsubscript{4}\textsuperscript{−}, analogous to the case of the bovine F\textsubscript{1}-ATPase (see Methods). The ATP hydrolysis activity of the purified EhV\textsubscript{1} was not observed, suggesting that the EhV\textsubscript{1} was indeed inhibited by binding of ADP\cdot AlF\textsubscript{4}\textsuperscript{−} in the nucleotide-binding site(s). We crystallized the inhibited EhV\textsubscript{1} and obtained a crystal structure at 3.8 Å resolution on an R factor of 22.7% and a free R factor of 26.6% (Table 1). However, we note that although the number of dwell states in EhV\textsubscript{1} and mammalian F\textsubscript{1} appears to be the same, these V\textsubscript{1} and F\textsubscript{1} motors show clear differences in the release order of cleavage products, rotational arrest points, dynamics and conformational changes\textsuperscript{3}. Thus, the ADP\cdot AlF\textsubscript{4}\textsuperscript{−}-inhibited V\textsubscript{1} and F\textsubscript{1} motors need not yield similar functional states. A physical interpretation of the stationary AlF\textsubscript{4}\textsuperscript{−}-inhibited V\textsubscript{1} structure is hence derived by comparison with only the other known crystallographic V-type models (and not the F-type ones), followed by computer simulations.”

Some conclusions rest on the observation of a significant drift (larger RMSD) seen in the MD simulations. I am not sure if I saw the necessary control, i.e., the RMSD of a simulation for which no drift is expected. Also, in order to see whether the RMSD histograms in Fig S3 result from a drift or from mere fluctuations, RMSD time traces would be helpful.

We have now plotted necessary controls in Fig. S3 to strengthen this point on the difference between RMSD histograms of A subunit conformations.
Fig. S3. RMSD distribution (A) of the simulated A' pocket before and after P_i release measured with respect to the X-ray structure of the A' pocket from newly discovered AlF_3-bound “P_i-release” dwell (PDB: 7VW7), and that with respect to the A' pocket of previously resolved “ATP-binding” dwell (PDB: 5KNB in Fig. 1). Data for the pre-P_i released state is determined from MD and subsequent REST2 simulations; data for the post-P_i released state is determined by relaxing the final model from the 2000 ns-long outward pathway of the Metadynamics simulation further using 100 ns long regular MD. We observe that before the P_i release, the simulated pocket remains similar to the P_i-release conformation, but after the P_i release, the same pocket drifts towards 5KNB. RMSD histogram of ATP- (red dotted) and ADP.P_i-bound (red solid) A' pocket is determined with respect to the respective starting models of the MD simulations (ATP-bound 3VR6 and ADP.P_i-bound 3VR6) to draw a baseline on what are the lowest possible RMSD values that infer similarity in structure. These baseline data show that the global changes of the A' pocket after the P_i release, despite being ~2-2.5 Å on the RMSD scale, are still larger than those prior to the P_i release of ~1.2 Å when the pocket contains either ATP or ADP.P_i. So, the conformational difference between the ATP and ADP.P_i-bound A' pockets (i.e. between the catalytic and P_i-release dwells) is lesser than the difference between the ADP.P_i- and ADP-bound A' pockets (i.e. between the P_i-release and ATP-binding dwells). RMSD distribution of the ADP.P_i-bound pocket with respect to the ATP-bound A' pocket of the “catalytic dwell” (PDB: 3VR6) of Fig. 1 is plotted a black solid line. The difference between the ADP.P_i-bound A' pocket w.r.t 3VR6+ATP (black line) > ADP.P_i-bound A' pocket w.r.t 7VW7+ADP.AlF_3 (solid-blue barplot) ≥ the baseline RMSD of ADP.P_i-bound A' pocket w.r.t 3VR6+ADP.P_i (red solid). Consistent with Fig. 2, this trend suggests that A' with ADP.P_i deviates from A' with ATP, justifying dwell states for these two conformations. (B) Distribution of inter-residue distances (from Fig. 2) at the ADP.P_i-bound A'B' and ATP-bound A'B' interfaces during the REST2 simulation are also presented, capturing the similarity with distances reported from the 3VR6 and 7VW7 X-ray structure. While the ATP-bound A'B' interface distance is commensurate with both crystal structures, the ADP.P_i-bound A'B' interface is closer to 7VW7, showing a bias for the P_i-dwell-like conformations.

Minor:

(4) Please provide error estimates / error bars for the free energy calculations. The error bars are now provided in the SI.
Fig. S7: one-dimensional projection of the free energy of phosphate release along the funnel axis for (A) outward pathway (B) inward pathway. Time series of the phosphate position along the funnel axis captured via funnel metadynamics for (C) outward release (D) inward release. Green band represents the crystal structure position of the phosphate. (F and H) To monitor the convergence of the free energy profile, energy difference between the highest barrier and the lowest minima (labeled using * in panels E and G) was monitored during the course of funnel metadynamics. Akin to work by Okazaki and Hummer (Ref. 18 in article), this barrier height is considered converged when it changes by \( \leq 1.5 \) Kcal/mol over a course of time \( \geq 50 \) ns e.g. after 900 ns in the outward trajectory and 300 ns in the inward trajectory. Thereafter, the error in the A and B panels is computed as the standard deviation of the free energy estimate during this converged part of the trajectory.

(5) The statement “such conformational differences were much lesser (between 0.1-0.6 Å) with the empty and tight sites” is unclear, but it seems to be an important argument for the conclusion of designating a Pi-release dwell.

We have now added more clarifications to this statement on page 4:

“Illustrated in Fig. 2E, we find that the conformations of the conserved residues of E261 and R262 of A subunit and the Arg-finger (R350 in B subunit) are deviated by 1.1–1.8 Å, which is probably induced by binding of AlF\(_4\) molecule instead of the gamma-phosphate of AMP·PNP; such conformational differences were much lesser (between 0.1-0.6 Å) with the empty and bound sites of the 2\((\text{ADP·AlF}_4)V_1\) structure relative to 2\((\text{AMP·PNP})V_i\), as seen in Fig. 2C-D. Furthermore, the C-terminal domain of the AB pair shows slightly open conformation that may allow Pi release but not as wide to allow the release of the ADP (Fig. 2E). From these findings, we designated the AB pair, which was more open than the tight, as \(\text{ADP·Pi-bound}\) form, and interpreted that the structure corresponds to the state of waiting for Pi release (denoted as “Pi-release dwell”) following ATP hydrolysis.”
Name: Peer Review Information for "Revealing a hidden intermediate of rotatory catalysis with X-ray crystallography and Molecular simulations"

Second Round of Reviewer Comments

Reviewer: 1

Comments to the Author

The manuscript by Shekhar et al. has been significantly improved but still contains important (though easily fixable) flaws. The basic issue is the text continue to attribute *much* more quantitative certainty to their calculations than is warranted by the data. Let me be clear that this does not mean the calculations cannot contribute importantly to the structural biology analysis - I believe they have contributed valuable insights. But we must not overstate what calculations currently can do, as this serves neither the computational nor experimental communities, i.e., the future readers of this interesting study.

The main issue that should be clarified further is the kinetics estimates for the two pathways and whether they are quantitatively meaningful. I am still skeptical they are meaningful. Right now, the exit times are quoted as 0.3 +/- 0.1 ms and 1.3 +/- 0.4 ms. My opinion based on extensive experience is that there is no unbiased computational method currently capable of providing this level of precision in kinetics estimates for a moderate size protein, let alone a system of the complexity of the V ATPase. Perhaps an order-of-magnitude estimate can be achieved. Estimates derived via a 1D PMF with constant diffusion must be considered less reliable. I appreciate that the authors have now described the results as “qualitative,” though I think “semi-quantitative” is fair. The authors have not stated how the error bars were calculated. These should be generated from completely independent calculations and described as an underestimate of uncertainty given the methodological basis. The problem with the current description is that the numerical values are presented as intrinsically reliable. The manuscript would be improved by more cautious wording reflecting the multiple layers of uncertainties and assumptions. These concerns would also impact the discussion of quantitative results in the Discussion. I think it’s better to say less than to imply that calculations are better than they are.

On a related though methodologically distinct point, the error bars derived for RMSF and shown in Fig. S2 appear to be significantly optimistic. The authors performed a kind of bootstrap subsampling based on 10 draws of 50 percent of the data. Presumably (although I don’t see that described) they approximated some confidence interval or computed a standard error treating the 10 draws as if they were independent. But the draws are absolutely not independent! They all come from the same highly correlated time series. A procedure like this will artificially narrow uncertainties. The authors could convince themselves by starting with two small data sets of, say, Gaussian numbers drawn from *identical* distributions. The two sets will always have slightly different means, and a procedure like theirs can then produce artificially narrow error bars suggesting they are different. For their RMSF data,
the authors need to use an established procedure (e.g., block averaging). It may be that there are no significant differences, and that just has to be accepted. If the authors want to use a new statistical procedure that should be justified on a clear statistical/mathematical/physical basis.

Minor. The axes of Fig. 5 panels C and D are labeled identically but the meaning of the axes evidently is not identical. Can there be some renaming so readers don’t have to get confused?

Reviewer: 2

Comments to the Author

The authors have addressed all my concerns satisfactorily.

Author's Response to Peer Review Comments:

Dear Professor Editor,

Tempe, May 8th, 2022

on behalf of my coauthors, Mrinal Shekhar, Chitrak Gupta, Kano Suzuki and co-corresponding author Takeshi Murata, I would like to submit revisions manuscript, entitled Revealing a hidden intermediate of rotatory catalysis with X-ray crystallography and Molecular simulations, to be considered for publication in the ACS Central Sciences as a research article.

We have now worked on the two remaining minor criticisms of Reviewer 1, as detailed in the attached rebuttal:

First, we have now performed Brownian Dynamics (BD) simulations to address the uncertainties on the kinetic estimates. Despite numerical errors, which will almost always remain due to the finite timescale of molecular simulations, and more so for systems of size as the V-type ATPase, the BD approach offers statistically significant estimates of the sib-millisecond-scale unbinding rates. The newly estimated kinetic trends remain in agreement with the past results that were determined using more approximate mean first passage time computations. So, our conclusions on the plausible rate determining step of ATPase’s catalysis cycle remain unperturbed. To perform these simulations in time, we have now added one more author to the manuscript, namely Dr. Chun Kit Chan from ASU.

Second, we have also performed block averaging of the data to ensure that our reported fluctuations, and the scientific inferences made thereof are not impacted by the correlation in the time-series data.

Taken together, this round of revisions gave us the opportunity to perform new error analyses for adding more credence to our scientific conclusions. Trusting that the present manuscript meets the standards of ACS Central Sciences we submit the revisions.

Sincerely,

Abhishek Singharoy, Ph.D.

Assistant Professor, School of Molecular Sciences
The manuscript by Shekhar et al. has been significantly improved but still contains important (though easily fixable) flaws. The basic issue is the text continue to attribute *much* more quantitative certainty to their calculations than is warranted by the data. Let me be clear that this does not mean the calculations cannot contribute importantly to the structural biology analysis - I believe they have contributed valuable insights. But we must not overstate what calculations currently can do, as this serves neither the computational nor experimental communities, i.e., the future readers of this interesting study.

We thank the Reviewer for finding our revisions a significant improvement, and the calculations potentially valuable. We have worked towards addressing the remaining two specific concerns as follows.

My opinion based on extensive experience is that there is no unbiased computational method currently capable of providing this level of precision in kinetics estimates for a moderate size protein, let alone a system of the complexity of the V ATPase. Perhaps an order-of-magnitude estimate can be achieved. Estimates derived via a 1D PMF with constant diffusion must be considered less reliable. I appreciate that the authors have now described the results as “qualitative,” though I think “semi-quantitative” is fair. The authors have not stated how the error bars were calculated. These should be generated from completely independent calculations and described as an underestimate of uncertainty given the methodological basis. The problem with the current description is that the numerical values are presented as intrinsically reliable. The manuscript would be improved by more cautious wording reflecting the multiple layers of uncertainties and assumptions. These concerns would also impact the discussion of quantitative results in the Discussion. I think it’s better to say less than to imply that calculations are better than they are.

We agree with the Reviewer, and as noted, had already toned down our message in the first round of revisions. Now, we further address the issues on the reliability and error analysis associated with the kinetic estimates by introducing Brownian Dynamics (or BD) computations with the 2D PMFs derived from the funnel metadynamics simulations. Described in the Methods, these computations were performed on the GPU-accelerated Atomic-Resolution Brownian Dynamics (ARBD) software. First, a probe particle is defined with the diffusion coefficient of 1.5 Å²/ ns and 5.10 Å²/ ns for dissociation towards and away from the stalk on the basis of previous diffusion calculation of P_i by Okazaki et.al (Ref 18). Second, the product-release landscape was projected on a 2D grid with spacing same as that of the original PMFs. Third, using gradients of this profile and the stated diffusion coefficient the equation of motion is integrated of over a timestep of 1 ps (for the more rugged inward release) and 3 ps (for the less rugged outward release). Each simulation was repeated 100 times, starting from random initial positions on the free energy landscape. We monitor the average time taken by the diffusive particle to exit the binding pocket and reach either in the bulk solvent (20 Å along the funnel axis on the outward pathway) or the stalk-A subunit interface (15-16 Å along the funnel axis on the inward pathway). We also compare these timing with that of a random walker on a flat landscape with diffusion coefficient identical to P_i.

For rebuttal purposes we have prepared the following summary figure of the results, presented in the following. This entire analysis is now presented in Fig. S7, and reported on page 5 and 6 of the manuscript. In the last paragraph of Discussions, we have also added precautionary text as the Reviewer suggests.
Caption. 2D PMFs of a flat landscape (A) and that of the outward (B) and inward (C) $P_i$ release paths. Unbinding transition time distributions derived from BD trajectories on these landscapes (D-F). Each distribution representing time estimates collected over 100 different BD simulations. The BD simulations of the outward pathway was repeated using 3 different PMF estimates collected over the last 50 ns, 100 ns and 150 ns of the metadynamics trajectory to see the impact of the ~1 kcal/mol numerical error on the corresponding time estimates. Traversing the inner-release pathway is on an average an order of magnitude slower than the outer-release path. In comparison, diffusion on a flat landscape is much faster taking 75 - 375 ns, and is in the ballpark of the 120 ns analytical estimate determined by dividing the area under panels A or C with $P_i$’s diffusion coefficient of 1.5 Å²/ps. Finally, we present a Table of all the key BD parameters, and the new kinetic estimates derived for the unbinding time distributions. The P-values imply that the unbinding time distribution for the outer and inner pathways are distinct from those on the flat landscape.

This approach brings forth some key conceptual and numerical advantages over our last attempt of computing mean first passage time from 1D PMFs, which is outlined.

First, by performing BD simulations, we bypass the need to project the 2D PMFs onto 1D profiles, and the associated risks of assuming the projections to be ‘implicitly reliable’. Rather, we simulate the explicit dynamics of a particle diffusing on the entire 2D landscape, and estimate the transition times required by the particle to diffuse from bound to free states. To perform this simulation, we take advantage of existing data on diffusion coefficients of the phosphate ($P_i$) group inside a commensurate protein environment of F-type ATPases, wherein similar collective coordinates were employed to monitor the dynamics of $P_i$ (Ref. 18 – Okazaki and Hummer, 10.1073/pnas.1305497110). The kinetic estimates now become explicit in that, they are determined directly from the number of steps taken by a BD trajectory. Also, no additional assumptions are made on the dimensionality of the reduced space beyond the ones already used for performing the metadynamics simulations.
Second, given the PMFs, any number of independent BD simulations can now be performed to estimate the mean and error on the passage times. For example, we have used 100 replica of BD simulations per PMF. Within estimated errors, the ratio of the rates of inward vs. outward \( P_i \)-release becomes statistically significant - we find that on an average across the 100 BD trajectories of the inward-release path is 10-fold slower than the outer one.

Finally, using a flat free energy landscape, we can analytically estimate the time taken to diffuse between two points, and then calibrate the BD timestep to match this analytical result. According the 1-3 ps timesteps were chosen; larger timesteps made the integrator unstable. It is still possible, as the Reviewer notes that our estimates are an order of magnitude off the experimental measurements. But the BD treatment ensures that the trends in computed rate-estimates are internally consistent.

On a related though methodologically distinct point, the error bars derived for RMSF and shown in Fig. S2 appear to be significantly optimistic. The authors performed a kind of bootstrap subsampling based on 10 draws of 50 percent of the data. Presumably (although I don’t see that described) they approximated some confidence interval or computed a standard error treating the 10 draws as if they were independent. But the draws are absolutely not independent! They all come from the same highly correlated time series. A procedure like this will artificially narrow uncertainties. The authors could convince themselves by starting with two small data sets of, say, Gaussian numbers drawn from *identical* distributions. The two sets will always have slightly different means, and a procedure like theirs can then produce artificially narrow error bars suggesting they are different. For their RMSF data, the authors need to use an established procedure (e.g., block averaging). It may be that there are no significant differences, and that just has to be accepted. If the authors want to use a new statistical procedure that should be justified on a clear statistical/mathematical/physical basis.

Following the Reviewer’s thoughtful suggestion, we have performed averaging of the MD data with block sizes of 2, 4, 8 and 16, and re-did the RMSF analysis and comparison. Despite changes in the per-residue fluctuations, the difference between ADP and ATP pockets have remained, particularly in the highlighted binding pocket regions (in the residue number 230-270 range).
Fig. S2. P<sub>i</sub> release results in a looser AB interface. Root Mean Square Fluctuation (RMSF) of subunits A and B, forming the ADP/ATP-binding site, shown for the ADP.P<sub>i</sub>-bound A<sub>i</sub> pocket (black) and ATP-bound A<sub>i</sub> pocket (purple). The binding site is more rigid (lower RMSF) in the ATP-pocket before the hydrolysis reaction. Residues in the vicinity of the ATP binding pocket (G235, G237, K38, R262, F245, R350) are highlighted. Data is collected from the final 100 ns of each of the 3 x 500 ns MD simulations (Table S1). Error bars are added to the RMSF analysis using a 10-fold resampling procedure, wherein we have removed 50% of the data randomly and repeated the computation ten times. The RMSF and resampling were repeated for five different block-sizes (wherein the trajectory was strided by 2, 4, 8 and 16-folds) to seek whether correlation in the time-series has affected the fluctuations, and more importantly, whether the difference in fluctuations between the ADP.P<sub>i</sub>-bound and ATP-bound A<sub>i</sub> pockets are independent of these correlations.