Supplemental Information for "Theory of long binding events in single molecule controlled rotation experiments on F_1-ATPase"

S. Volkán-Kacsó[a] **and R. A. Marcus**[a]

[a]Noyes Laboratory of Chemical Physics, California Institute of Technology, MC 127-72, Pasadena, CA 91125, USA

This manuscript was compiled on June 14, 2017

1. Further methods and derivations

Site assignment of binding and release events. A challenging aspect of data analysis in the controlled rotation experiments with fluorescent nucleotides is assigning binding and release events to the correct β subunit of the three. To confirm the site assignment by Adachi et al. a systematic study was performed by reassigning events to the other two subunits in all possible combinations, using the rate constant estimates provided in their work (e.g., forward versus backward rotation, short versus long events and various concentrations). Two criteria were used to assign the outcomes: 1. the smoothness of the function and 2. the degree of "hysteresis", i.e. the difference of the same rate estimated while rotating in the forward versus the backward direction. The site assignment by Adachi et al. "scored" best, although three other possible assignments came relatively close.

We recall that in the controlled rotation experiment the binding and release events were monitored both by rotating the γ shaft in the hydrolysis and synthesis direction, i.e., by increasing and decreasing θ's, respectively. Rate constants for both binding and release were extracted from both rotational directions in the work of Adachi et al..[1] These reported rate constants extracted from the positive rotation approximately overlap with those extracted from the negative rotation trajectories, and the difference between the two has been attributed by Adachi et al. to uncertainties in their measurements.

Derivation of Eq. 1 from rate equation with time-dependent rate constants. Let $k_1(t)$ and $k_2(t)$ be the time-dependent rate constants for consecutive steps A $\rightarrow$ B and B $\rightarrow$ C, respectively. For example, A could be an ATP bound state, B an ADP+Pi state and C a Pi bound state, and so k_1 would be the hydrolysis rate and k_2 the ADP release rate. The probability of being in A (or B or C) is p_1 (or p_2 or p_3, respectively) and the rate equations governing the system are

$$\begin{aligned} dp_1/dt &= -k_1(t)p_1, & \text{[S1]}\\ dp_2/dt &= k_1(t)p_1 - k_2(t)p_2, & \text{[S2]}\\ dp_3/dt &= k_2(t)p_2 & \text{[S3]} \end{aligned}$$

When the system is initially at time t_0 in state A, we have

$$p_2(t_0) = p_3(t_0) = 0, \quad p_1(t_0) = 1. \qquad \text{[S4]}$$

The quantity of interest is $\psi_{1\rightarrow 3}(t_2|t_0)$ the waiting time distribution of the composite process to be calculated next from the B $\rightarrow$ C lifetime density in Eq. S3,

$$\psi_{1\rightarrow 3}(t_2|t_0) = dp_3/dt = k_2(t_2)p_2(t_2). \qquad \text{[S5]}$$

The rate equations can be written as

$$\begin{aligned} p_1(t) &= e^{-\int_{t_0}^{t} k_1(t')dt'}, & \text{[S6]}\\ dp_2(t)/dt &= k_1(t)e^{-\int_{t_0}^{t} k_1(t')dt'} - k_2(t)p_2(t). & \text{[S7]} \end{aligned}$$

One can verify that the solution of Eq. S7 is

$$p_2(t) = \int_{t_0}^{t} k_1(t')e^{-\int_{t_0}^{t'} k_1(t'')dt''} e^{-\int_{t'}^{t} k_2(t'')dt''} dt'. \qquad \text{[S8]}$$

We note that from these integral limits it naturally follows that $t_0 < t' < t$ where t' is the time of an event A $\rightarrow$ B, and t of the event B $\rightarrow$ C. We write conditional survival probabilities for the simple processes, A $\rightarrow$ B as $p_1^c(t_1|t_0)$ and B $\rightarrow$ C as $p_2^c(t_2|t_1)$. We note that $p_1^c(t_1|t_0) = p_1(t_1)$. The waiting time distributions for these conditional probabilities are, as defined in the present article,

$$\psi_1(t'|t_0) = -dp_1^c(t'|t_0)/dt', \quad \psi_2(t|t') = -dp_2^c(t|t')/dt, \qquad \text{[S9]}$$

where for these simple processes we find from equations analogous to Eq. 1 from the main text,

$$\begin{aligned} \psi_1(t'|t_0) &= k_1(t')e^{-\int_{t_0}^{t'} k_1(t'')dt''}, & \text{[S10]}\\ \psi_2(t|t') &= k_2(t)e^{-\int_{t'}^{t} k_2(t'')dt''}. & \text{[S11]} \end{aligned}$$

Hence, substituting these results into Eq. S8 we find that

$$\begin{aligned} k_2(t)p_2(t) &= \psi_{1\rightarrow 3}(t_2|t_0)\\ &= \int_{t_0}^{t} k_1(t')e^{-\int_{t_0}^{t'} k_1(t'')dt''} k_2(t)e^{-\int_{t'}^{t} k_2(t'')dt''} dt'\\ &= \int_{t_0}^{t} \psi_2(t|t')\psi_1(t'|t_0)dt'. & \text{[S12]} \end{aligned}$$

Proof of Eq. 10. We begin with the rate equation that describes the time evolution of the survival probability $p_{hyd}(\theta|\theta_0)$ in the pre-hydrolysis state: if ATP binding occurred at θ_0 then $p_{hyd}(\theta|\theta_0)$ is the probability that it survived until θ as ATP without undergoing hydrolysis.

$$\omega dp_{hyd}(\theta|\theta_0)/d\theta = -k_{hyd}(\theta)p_{hyd}(\theta|\theta_0). \quad [S13]$$

Here, $k_{hyd}(\theta)$ is the rate for conversion of ATP to ADP+Pi at θ. It depends only on the value of θ and not on a prior θ_0 at which ATP binding occurred. This assumption follows the arguments put forward by both Noji and Kinosita[1, 2]: If the events occur with only one site occupied then all events in the subunit hosting the site are completely determined by the present rotor angle (the rotor angle is either static or rotated at slow rates). In the controlled rotation experiments these conditions are met, since at the low ATP concentration 2 subunits are always empty and all the activity occurs in the 3$^{\rm rd}$ one.

The θ-based lifetime distribution ψ of events whereby binding of ATP occurs at θ_0 and hydrolysis subsequently at θ is analogous to the first half of Eq. 3 of the text,

$$\psi_{hyd}(\theta|\theta_0) = -dp_{hyd}(\theta|\theta_0)/d\theta. \quad [S14]$$

Integration of Eq. (S14) yields,

$$\int_{\theta_0}^{\theta} \psi_{hyd}(\theta'|\theta_0)d\theta' = -p_{hyd}(\theta|\theta_0) + 1, \quad [S15]$$

where we have used, by definition, $p_{hyd}(\theta_0|\theta_0) = 1$. We then have

$$p_{hyd}(\theta|\theta_0) = 1 - \int_{\theta_0}^{\theta} \psi_{hyd}(\theta'|\theta_0)d\theta'. \quad [S16]$$

From Eq. (S13) using Eq. (S14) we obtain

$$k_{hyd}(\theta) = \omega\psi_{hyd}(\theta|\theta_0)/p_{hyd}(\theta|\theta_0). \quad [S17]$$

Eqs. (S16) and (S17) yield

$$k_{hyd}(\theta) = \omega\psi_{hyd}(\theta|\theta_0)/[1 - \int_{\theta_0}^{\theta} \psi_{hyd}(\theta'|\theta_0)d\theta'], \quad [S18]$$

which is Eq. 10 in the text.

Proof that the differentiation with respect to θ_0 of LHS and RHS of Eq. 10 yield 0. Eq. 10 (without the hyd subscripts) can be written as a function of ψ and p, (the latter is the survival probability as defined in the MS),

$$k(\theta) = \frac{\omega\ \psi(\theta|\theta_0)}{\left[1-\int_{\theta_0}^{\theta}\psi(\theta'|\theta_0)d\theta'\right]} = \frac{\omega\ \psi(\theta|\theta_0)}{p(\theta|\theta_0)}. \quad [S19]$$

Taking the derivatives with respect to θ_0 for both RHS and LHS,

$$0 = \omega\ \frac{\frac{\partial\psi(\theta|\theta_0)}{\partial\theta_0}p(\theta|\theta_0) - \psi(\theta|\theta_0)\frac{\partial p(\theta|\theta_0)}{\partial\theta_0}}{p(\theta|\theta_0)^2}. \quad [S20]$$

Let us now calculate the derivatives in the numerator on the RHS. In the text we have deduced,

$$p(\theta|\theta_0) = e^{-1/\omega\int_{\theta_0}^{\theta}k(\theta')d\theta'} \quad [S21]$$

$$\psi(\theta|\theta_0) = -\frac{dp}{d\theta} = \frac{k(\theta)}{\omega}e^{-1/\omega\int_{\theta_0}^{\theta}k(\theta')d\theta'}. \quad [S22]$$

Taking the partial derivative with respect to θ_0 yields,

$$\frac{\partial p(\theta|\theta_0)}{\partial\theta_0} = \frac{k(\theta_0)}{\omega}e^{-1/\omega\int_{\theta_0}^{\theta}k(\theta')d\theta'} = \frac{k(\theta_0)}{\omega}p(\theta|\theta_0) \quad [S23]$$

$$\frac{\partial\psi(\theta|\theta_0)}{\partial\theta_0} = \frac{k(\theta)k(\theta_0)}{\omega^2}e^{-1/\omega\int_{\theta_0}^{\theta}k(\theta')d\theta'} = \frac{k(\theta_0)}{\omega}\psi(\theta|\theta_0). \quad [S24]$$

The numerator on the RHS of Eq. S20 then is

$$\frac{\partial\psi(\theta|\theta_0)}{\partial\theta_0}p(\theta|\theta_0) - \psi(\theta|\theta_0)\frac{\partial p(\theta|\theta_0)}{\partial\theta_0} = \frac{k(\theta_0)}{\omega}\psi(\theta|\theta_0)p(\theta|\theta_0) - \frac{k(\theta_0)}{\omega}\psi(\theta|\theta_0)p(\theta|\theta_0) = 0. \quad [S25]$$

Correcting for missed events and for replacing T_0 with T. In our previous treatment,[3] we provided a method to calculate the number of missed binding and release events in the analysis of single-molecule fluorescence trajectories. The missed events occur due to the finite time resolution of the apparatus, dictated by shot-noise from the limited photon counts from a single emitter, causing the blurring of bright ($\sigma = 1$) and dark ($\sigma = 0$) states.[4] The method uses, as an input, the binding and release rate constants using the functional forms from theory. The advantage of this method is that (1) if theory provides rate constant prediction the calculations need to be performed only once, (2) if no such predictions are available, the functional form of having 2 or less independent parameters per Region (Fig. 1.A) provides a tractable parameter fitting problem. In the calculations of the theoretical counterpart of the reported rates for binding in Figs. 1.A and S1 we also replaced T_0 with T thus taking into account for the error due to this replacement. Since the calculations use the occupancy in the angular bins, in the present calculation for the occupancy we used the solution of the time-dependent equation 7. We note that in our previous treatment in Ref. 3 the occupancy was approximated with the equilibrium probability of being in state $\sigma = 1$, $k_f/(k_f + k_b)$. The present calculations provide a method to calculate the occupancy for an arbitrary rotation rate ω when the binding and release rate constants are available for the entire 360° .

Within a bin of Δt the k_f and k_b are treated as constants, hence the survival probabilities S_0 in state 0 and S_1 in state 1, are exponential, $S_0(t) = \exp(-k_f t)$ and

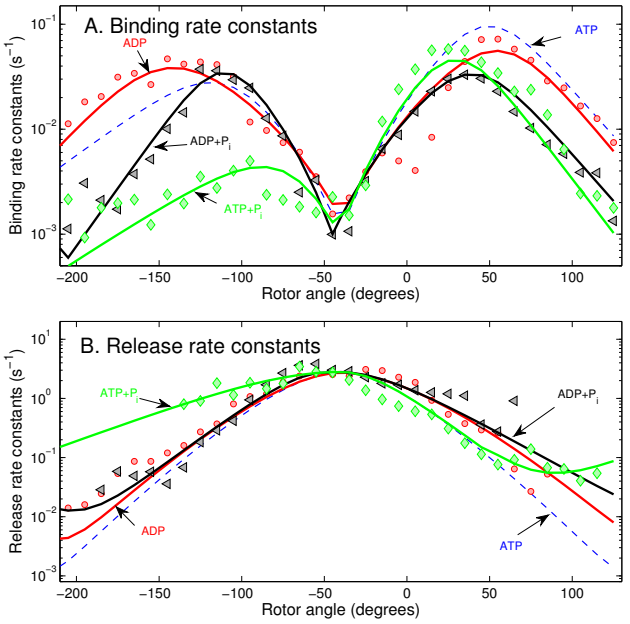

Fig. S1. Reported binding (A) and release (B) rate constants as a function of the θ angle for nucleotide binding in the presence of P_i in solution. For reference the theoretical values for ATP binding and release without solution P_i shown on Fig. 1 are plotted as blue dashed lines.

$S_1(t) = \exp(-k_b t)$. The probabilities that the states have lived a lifetime between 0 and t, i.e., that it decayed before t, are $[1-S_0(t)]$ and $[1-S_1(t)]$. Their time derivatives, the probability densities of lifetimes, the so-called "waiting time distributions" $\rho_0(t)$ and $\rho_1(t)$, are given by

$$\rho_0(t) = k_f \exp(-k_f t), \quad \rho_1(t) = k_b \exp(-k_b t). \qquad [S26]$$

The local time t is the time elapsed since the system entered interval j. To avoid very short binding intervals that might arise from shot-noise, in the analysis of Adachi et al.[1] bound states shorter than $3\tau = 0.1$ s were discarded. Then, the relevant events to be counted are those that occur within a time Δt of the interval j, and are separated by at least 3τ. To do so, we first define $p(\sigma_0, n)$ the probability of a succession of n events of occupancy change occurring at times $\{t_1, t_2, ...\}$ within an interval, if the system was in state σ_0 at $t = 0$. The probability that n events are missed due to shorter than 3τ on-times within an interval j is denoted by $p_{\rm miss}(\sigma_0, n)$. A contribution for missed events comes from missed 'cross-boundary events', $p'_{\rm miss}(\sigma_0, n)$, which are transitions occurring in a bin with the previous or subsequent transition occurring in another interval, and which are missed due to the time spent between the two events is shorter than 3τ. For example, one such event is when a $0 \to 1$ transition occurs during the $j^{\rm th}$ bin, and the subsequent transition occurs in a subsequent interval, but the time spent in state 1 is too short to be detected.

If p_0 and p_1 are the probabilities of being in state 0 and 1 at $t = 0$ then the reported rates can be obtained by subtracting the average number of missed events from the average total number of events,

$$k_f^{\rm rep}(\theta) \cong \Delta t^{-1} \left[p(0,1) + p(0,2) + p(1,2) - p_{\rm miss}(0,2) - p'_{\rm miss}(0,1) - p'_{\rm miss}(1,1)\right] \qquad [S27]$$

$$k_b^{\rm rep}(\theta) \cong (\Delta t p_0)^{-1} \left[p(1,1) + p(1,2) + p(0,2) - p_{\rm miss}(0,2) - p'_{\rm miss}(0,1) - p'_{\rm miss}(1,1)\right] \qquad [S28]$$

where the "rep" denotes "reported". In the calculations we neglect higher order contributions, *i.e.*, terms where $n \geqslant 3$, since they provide negligible contribution, but to treat experiments where the rates are significantly larger than $1/\Delta t$, they should also be included.

Next we calculate p_0, p_1, $p(\sigma_0, n)$, $p_{\rm miss}(\sigma_0, n)$ and $p'_{\rm miss}(\sigma_0, n)$ appearing in Eqs. S27 and S28, noting that the τ and $\Delta t = t$ are formally treated in the following derivation as variables, and are subsequently assigned the appropriate values. These probabilities are cast as expectation values with respect to $\tau_i = t_i - t_{i-1}$. In particular, $p(\sigma_0, n) = p_{\sigma_0} \langle \Theta(t_n < t < t_{n+1}) \rangle_{\{\tau_i\}}$ and $p_{\rm miss}(0,2) = p_0 \langle \Theta(t_2 < t < t_3)\Theta(\tau_2 < \tau) \rangle_{\{\tau_i\}}$, where Θ is 1 if the argument is true and 0 otherwise and $\langle \cdot \rangle$ is integration with respect to all τ_i's. $p'_{\rm miss}(\sigma_0, 1)$, formally written as $p_0 \langle \Theta(t_1 < t < t_2)\Theta(\tau_2 < \tau) \rangle_{\{\tau_i\}}$, also lead to a convolution integrals,[1] and a Laplace transform ($t \to s$, $\tau \to u$) is used to facilitates their evaluation.

For $p(\sigma_0, n)$ the Laplace transform $t \to s$ (we use a tilde to denote the transform) yields,

$$\tilde{p}(\sigma_0, n|s) = \int_0^\infty p(\sigma_0, n|t) e^{-st} dt = p(\sigma_0) \left\langle \exp(-s t_n)[1 - \exp(-s\tau_{n+1})]/s \right\rangle . \qquad [S29]$$

Using $t_n = \tau_1 + ... + \tau_n$, for $\sigma_0 = 0$, an analytic expression follows,

$$\tilde{p}(0, n|s) = p(0) \times \begin{cases} \left[\tilde{\rho}_0(s)\tilde{\rho}_1(s)\right]^{n/2}, & n \text{ even}, \\ \left[\tilde{\rho}_0(s)\right]^{(n+1)/2} \left[\tilde{\rho}_1(s)\right]^{(n-1)/2}, & n \text{ odd}, \end{cases} \qquad [S30]$$

where $\tilde{\rho}_0(s) = k_f/(s + k_f)$ and $\tilde{\rho}_1(s) = k_b/(s + k_b)$. An expression for $p(1, n|s)$ results by analogy from Eq. (S30), by interchanging states 0 and 1. For the missed events, for each short event condition an additional variable is used, so for $n = 2$ one additional variable is needed ($\tau \to u$), leading to a double Laplace transform,

$$\tilde{p}_{\rm miss}(0, 2|u, s) = p_0 \tilde{\rho}_0(s)\tilde{\rho}_1(s) u^{-1} (s + k_f)^{-1}. \qquad [S31]$$

and

$$\tilde{p}'_{\rm miss}(0, 1|u, s) = p_0 (s\,u)^{-1} \tilde{\rho}_0(s)[\tilde{\rho}_1(u) - \tilde{\rho}_1(s + u)], \qquad [S32]$$

[1] For example, for $\sigma_0 = 0$, the time-domain expression is $p(0,2|t) = p_0 \int_0^\infty \int_0^\infty \int_0^\infty \rho_0(\tau_1)\rho_1(\tau_2)\rho_0(\tau_3)\delta(t - \tau_1 - \tau_2 - \tau_3)\, d\tau_3 d\tau_2 d\tau_1$ and $p_{\rm miss}(0,2|t) = p_0 \int_0^\infty \int_0^\tau \int_0^\infty \rho_0(\tau_1)\rho_1(\tau_2)\rho_0(\tau_3)\delta(t - \tau_1 - \tau_2 - \tau_3)\, d\tau_3 d\tau_2 d\tau_1$.

where analogous expressions apply for $\tilde{p}_{\rm miss}(1,2|s)$ and $\tilde{p}'_{\rm miss}(1,1|u,s)$.

In the method the Laplace-space expressions are then inverted back to time-domain by elementary inversion, with the aid of symbolic computation.[5] The one-event contributions in Eqs. S27 and S28 are given as

$$p(0,1) = p_0 k_f/(k_f - k_b)[\exp(-k_b t) - \exp(-k_f t)], \quad [S33]$$
$$p(1,1) = p_1 k_b/(k_b - k_f)[\exp(-k_f t) - \exp(-k_b t)], \quad [S34]$$

and the two-event contributions,

$$p(0,2) = p_0 k_f k_b/(k_f - k_b)^2 \times \left(-e^{-k_f t} + e^{-k_b t} - k_f t e^{-k_f t} + k_b t e^{-k_f t}\right), \quad [S35]$$
$$p(1,2) = p_1 k_f k_b/(k_f - k_b)^2 \times \left(e^{-k_f t} - e^{-k_b t} + k_f t e^{-k_b t} - k_b t e^{-k_b t}\right). \quad [S36]$$

The leading terms in the missed events read

$$p_{\rm miss}(0,2) = k_f k_b^2 e^{-k_f t - k_b 3\tau} (k_f + k_b)^{-1} (k_f - k_b)^{-2} (e^{k_f 3\tau} - e^{k_b 3\tau} + k_f t e^{k_f 3\tau} - k_f t e^{k_b 3\tau} - k_b t e^{k_f 3\tau} + k_b t e^{k_b 3\tau} - k_f 3\tau\, e^{k_f 3\tau} + k_b 3\tau\, e^{k_f 3\tau}). \quad [S37]$$

For the cross-boundary missed events, the Laplace inversion yields for the leading terms,

$$p'_{\rm miss}(0,1) + p'_{\rm miss}(1,1) = e^{-k_b 3\tau - k_f t} - \frac{k_f e^{-k_f t}}{k_f - k_b} + \frac{k_b e^{-k_b 3\tau}}{k_f - k_b}. \quad [S38]$$

The probability terms p_0 and p_1 are evaluated by solving the time-dependent rate equation if the binding and release rates are known for the whole 360° range in a subsequent section.[2]

Method for the calculation of site occupancy for rotation with a constant rate. Controlled rotation experiments can provide binding and release rate constants for the full 360° profile. Both $k_{on}(\theta)$ and $k_{off}(\theta)$ are needed to solve the time-dependent rate equation for the occupancy for both short and long binding events. Adachi et al. limited their analysis to events where the occupancy σ switches between 0 and 1, and so the occupancy can be defined as the probability $p_1(t)$ at any given time t of being in state $\sigma = 1$. The probability of being in state $\sigma = 0$ is $p_0(t)$ and $p_0 + p_1 = 1$. If $\theta(0) = 0$, the evolution of the occupancy is given by

$$dp_1/dt = k_{on}(\omega t)(1 - p_1) - k_{off}(\omega t) p_1, \quad [S39]$$

where the rotation rate is positive $\omega > 0$ for the forward (hydrolysis) rotation and negative $\omega < 0$ for the reverse (synthesis) rotation. Using $k_{on}(\theta)$ and $k_{off}(\theta)$ from the text, Eq. (S39) can be solved numerically with an example shown in Fig. S2.

[2] In our previous treatment, that was limited to an angular range of $(-50^{\rm o}, 50^{\rm o})$ steady state conditions were assumed, which yielded $p_0 = k_b/(k_f + k_b)$ and $p_1 = k_f/(k_f + k_b)$.

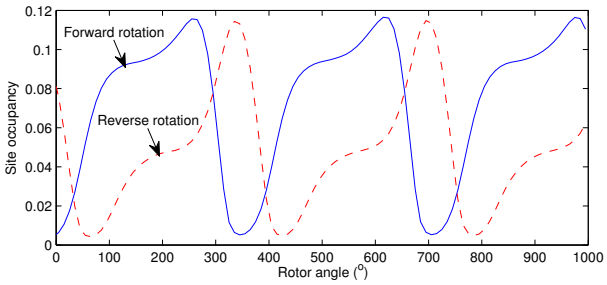

Fig. S2. Site occupancy as a function of rotor angle for forward (counter-clockwise) and reverse (clockwise) rotation calculated for ATP in the absence of P_i in solution.

We note that the rate constants in Eq. (S39) can correspond to different species for long events. In particular if ATP is in solution and the rotation is in the hydrolysis direction then for ATP binding rates ADP release rates are used, or if ADP+P_i is in solution and the rotation is in the synthesis direction, than ATP release rates are used for the rate constants in Eq. S39. The probabilities $p_0(t)$ and $p_1(t)$ can be then used as input in theoretical calculations[3] to correct for missed events and to correct for the effect of replacing T_0 with T in the estimation of the reported (uncorrected) binding and release rate constants. In our previous treatment a quasi-equilibrium approximation was used for the occupancy in Region 1.[3] While that approximation applies for Regions 1 and 2 of Fig. 2, it may not be generally applicable for Regions 3 and 4 due to the finite rate of rotation. So when extending the analysis to the latter regions we propose using the results for $p_0(t)$ and $p_1(t)$ obtained from Eq. S39 rather than using the quasi-equilibrium values. For such calculations a future theory will be developed to treat the $k_{on}(\theta)$ and $k_{off}(\theta)$ vs. θ behavior in these ranges when more data become available.

In the experimental trajectories the nucleotide occupancy is always less than 0.3 for all rotor angles and conditions.[1] The calculated maximal occupancy shown in Fig. S3 as a function of the turnover angle $\theta_{\rm TO}$ indicates that eventually, for large enough $\theta_{\rm TO}$ ($\sim 75 - 85^{\rm o}$ depending on the species) the occupancy will exceed 0.3, if one ignored the turnover in Fig. 2 and assumed instead an exponential dependence of rate constants vs. θ in the turnover region. Thus, (1) the < 0.3 occupancy condition yields an effective constraint that should be always met when extracting model parameters and (2) the turnover is a true feature which cannot be attributed to experimental artifacts, such as missed events.

Numerical implementation of the integrations in Eqs. 1 and 4. We use a simple quadrature process to evaluate the integrals, implemented in MatLab[5]. To accelerate the calculations, vectorization of the repeated summations is achieved by using the "cumsum" function and vector operations native to MatLab. Parallelization is implemented using the "parfor" cycle. Typical run times

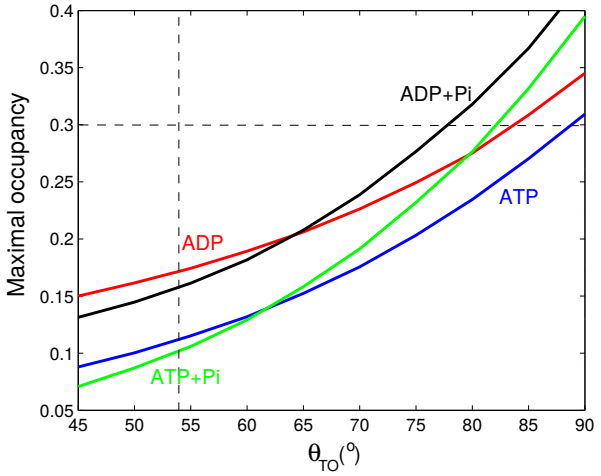

Fig. S3. Maximal site occupancy as a function of the turnover angle $\theta_{\rm TO}$ for the right side branch calculated for all four nucleotide conditions used in experiment.

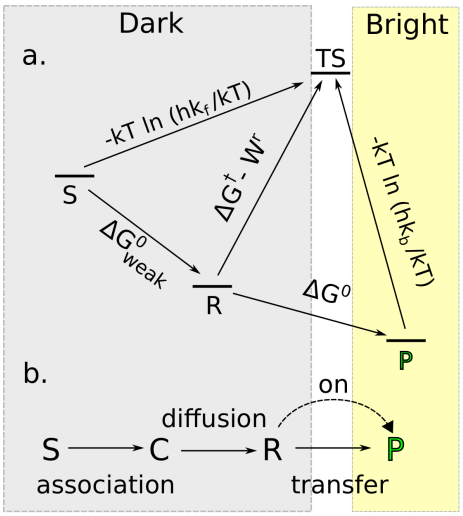

Fig. S4. Free energy diagram for group transfer process (a) and the sequence of processes in nucleotide binding (b). The fluorescence emission from Cy3-nucleotide is assumed to turn on upon the transfer from the weak binding site (R) to the strong binding site in the pocket (P).

are in the order of 1-2 seconds on a single Xeon E3-1220 processor.

2. Remarks on the diffusion and transfer mechanism

In the present paper we account for the initial process that leads up to the reactant state (R) of the group transfer. Reminiscent of established models[6, 7] in the binding step, starting with a nucleotide at an outside position in the ATPase, then entering the channel leads to the final state of the nucleotide in the catalytic pocket. The latter is modelled as a diffusion-activation process (Fig. S4). A diffusive process occurs inside the entrance of the channel, seen in the subunit structure, after which the group transfer (reactive process) leads to the final bound state (P) of the nucleotide in the β subunit pocket. In the theory we postulated[8] that the R state is a "weak" binding state that was identified in the experimental measurements of the affinities in the β subunits.[9] Here we treat the diffusive process as having a θ-dependent rate constant $k_{\rm D}(\theta)$. At larger angles of $\theta > 50^{\rm o}$ the entrance of the host β subunit is partially closed, hindering diffusion, and a transition to reach the R state instead of being activation-controlled becomes diffusion-controlled. The rate of this process in the theory will be therefore monotonically decreasing as a function of increasing θ in this region of θ's. In the present application exponential dependence on θ for the diffusion is assumed, as described in a later section.

If $k_f(\theta)$ is the forward binding rate constant introduced in our previous treatment for the group transfer process, then in the diffusion-activation mechanism of binding k_f will be the rate of activation process and we use the well-known steady-state relation[6, 7]

$$\frac{1}{k_{\rm on}(\theta)} = \frac{1}{k_{\rm D}(\theta)} + \frac{1}{k_f(\theta)}. \qquad [S40]$$

In the θ-range of $-50^{\rm o} < \theta < 50^{\rm o}$, where stalling and controlled rotation experiments overlap $k_f \ll k_d$ and so in Eq. (S40) $k_{\rm on}(\theta) \cong k_f(\theta)$, i.e., k_f is rate-determining. However, at larger angles, the narrowing of the opening of the binding channel seen in X-ray crystallography structures for a $\sim 80^{\rm o}$ structure) renders the R state energetically more difficult to reach. As a result if $\theta > 50^{\rm o}$, $k_{\rm D}$ becomes slower than k_f, so $k_{\rm D}$ will be the rate-determining step, so Eq. (S40) reduces to $k_{\rm on}(\theta) \cong k_{\rm D}(\theta)$.

If for the diffusion-activation mechanism $k_{\rm Db}$ and k_b are the a "back" diffusion and the back transfer (*i.e.*, activation) rate constants respectively, then for the "off" rate constant $k_{\rm off}$ the equation

$$\frac{1}{k_{\rm off}(\theta)} = \frac{1}{k_{\rm Db}(\theta)} + \frac{1}{k_b(\theta)} \qquad [S41]$$

applies. For the group transfer rate constants k_f and k_b θ-dependent expressions were derived. Presently, we do not predict the values of the $k_{\rm D}$'s and $k_{\rm Db}$'s, which also determine the turnover angles, rather we extract them by fitting to controlled rotation data, as described in more detail below.

We note that at any θ the forward diffusion rate constant $k_{\rm D}(\theta)$ divided by the reverse diffusion rate constant $k_{\rm Db}(\theta)$ is equal to the equilibrium constant, and that the forward activation reaction rate constant $k_f(\theta)$ divided by the reverse activation reaction rate constant $k_b(\theta)$ is equal to the same equilibrium constant. So, using equation S41 the net forward rate constant $k_{\rm on}(\theta)$ divided by the net reverse rate constant $k_{\rm off}(\theta)$ is equal to the same equilibrium constant at any θ, but the present data are too scattered to be interpreted at this time (cf. Fig. S5). There is also the question that we plan to investigate as to whether in these systems of long time events $k_{\rm on}(\theta)$ and $k_{\rm off}(\theta)$ refer to inverse processes.

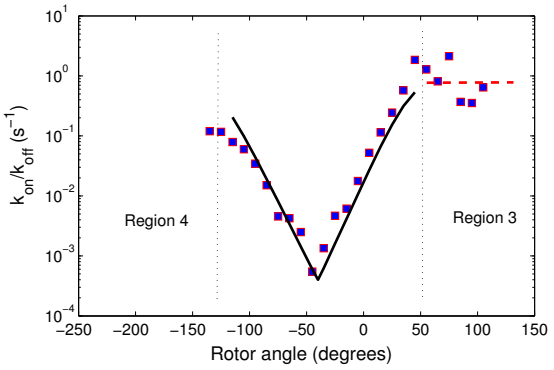

Fig. S5. kon/koff vs. rotor angle plot for ATP binding. Squares are from experiment, continuous lines from theory and the dashed line is a guide. Excessive scatter is seen beyond the turnover (Region 3).

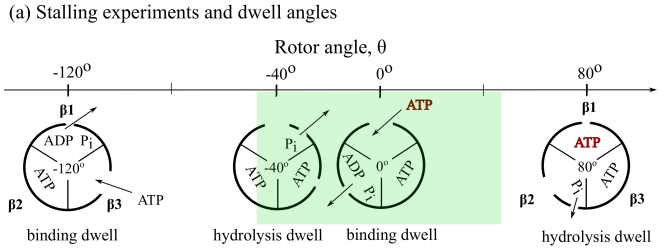

Fig. S6. Shaft rotation, nucleotide binding/release activity and conformational changes of a ring subunit in single molecule experiments of stalling (a), controlled rotation (b) and free stepping rotation (c). In (a-c) the system is represented at four different values of the rotor angle θ, corresponding to dwell angles at -120^o, -40^o, 0^o and 80^o. θ is defined relative to subunit β_1, i.e., $\theta = 0$ for the binding dwell with an empty β_1.

The near-symmetry in the θ-dependence and its relation to the structure and the stepping rotation. The γ shaft is eccentric in the plane of the $\alpha_3\beta_3$ ring, due to the presence of additional helices in the region that project out of the ring, and to some extent due to the natural bending of its coiled-coil structure. The overall geometry shows in the structure (Fig 2.B) a quasi mirror symmetry with a plane that cuts through the eccentric part of the protruding subunits of the γ-structural motifs. The ring subunits beneath this protrusion are found, in their relaxed states, in two conformations, "closed" and "tight", while those facing away from the bulge are found in the open state.[10] An open β structure is associated with a state of the binding dwell which in F_1-ATPase rotation experiments is identified as $\theta = 0^o$. Hence, in the neighbourhood of 0^o value, i.e., in Regions 1 and 2 the relaxed β is in an open state. Meanwhile, on the opposite side, in a range around $\theta \cong 180^o$, in Regions 3 and 4 the subunit is forced closed by the γ.

The closing of the hinge in β can produce different behaviors. There is seen be a difference in the coupling of the β subunit to the rotor, since the -40° and 0° states are both open, but not identical, a difference most evident in the binding of ATP when P_i is present in solution. For the left branch (Regions 1 and 2 on Fig. 2), there is a significant difference seen in Fig. S1 between the rate of ATP binding in the presence of P_i compared to all other cases (i.e., ADP with and without P_i and ATP with no P_i in solution): the symmetry with respect to θ_O is broken, the binding of ATP is not favourable if P_i is present in the solution.

The θ dependence of the diffusive process and the extension of the model to the left branch of rotor angles. For the C→R process in the scheme from Fig. S4, it is expected that during this diffusive process, the time $1/k_D$ needed to reach the initial weak binding site, depends on the degree of openness of the host subunit. Along the positive branch (Regions 1 and 3 in Fig. 1.A) the resulting narrowing of the channel the rate of diffusion becomes slower as the angle is being increased. This behavior naturally leads to a trend in k_D versus θ which is opposite to the trend of the forward reactive step, i.e., a monotonically decreasing function. The reverse rate constant of the diffusive k_{-D} is assumed to show a similar trend in its θ-dependence, is related to the above k_D by $k_D(\theta)/k_{-D}(\theta)$. For the functional form of $k_D(\theta)$ and $k_{Db}(\theta)$, in our analysis we assume approximately exponential θ-dependence, with negative trends, i.e., $a = d\ln k_d/d\theta < 0$ and $a_D = d\ln k_{Dd}/d\theta < 0$, respectively, formally written as

$$\ln k_D(\theta) = \ln k_D(\theta_{TO}) + a\,(\theta - \theta_{TO}), \quad [S42]$$

$$\ln k_{Db}(\theta) = \ln k_{Db}(\theta_{TO}) + a_b\,(\theta - \theta_{TO}), \quad [S43]$$

In our analysis we identify the otherwise arbitrary reference angle with θ_{TO}, the turnover angle of k_D for the positive branch in Fig. 1.A, defined as $k_f(\theta_{TO}) = k_D(\theta_{TO})$.

To treat the left branch of angles comprising of Regions 2 and 4 on Fig. 1.A we use the diffusion-reaction model described above for the right branch, with modifications noted as follows. Binding of nucleotides is invariably associated with the closing of the subunit but is now coupled to the rotation in such a way to produce a clockwise, i.e., negative rotation of γ. Hence all rate constant will have the opposite trend in their θ-dependence compared to the quantities describing the right branch. To differentiate between the right and the left branch, a 'prime' symbol is formally used to denote the quantities describing the latter, i.e., in Eq. (S42-S43) k'_f, k'_b, k'_D and k'_{Db} will denote the forward, backward, diffusion and backward diffusion rate constants. The values of these rate constants

plotted with their experimental counter parts are in part extracted by fitting to the experimental data and in part borrowed from the corresponding quantities in the right hand site. In particular, the slopes of the group transfer log rates (the bottleneck processes in Region 2) are borrowed from the right hand side (Region 1). However, $\ln k'_f(\theta)$ and $\ln k'_b(\theta)$ are extracted by fitting.

1. Adachi K, Oiwa K, Yoshida M, Nishizaka T, Kinosita, Jr. K (2012) Controlled rotation of the F1-ATPase reveals differential and continuous binding changes for ATP synthesis. *Nat. Comm.* 3:1022.
2. Watanabe R et al. (2012) Mechanical modulation of catalytic power on F-1-ATPase. *Nat. Chem. Biol.* 8(1):86–92.
3. Volkán-Kacsó S, Marcus RA (2016) Theory of controlled rotation experiments, predictions, tests and comparison with stalling experiments in F_1-atpase. *Proc. Natl. Acad. Sci USA* 113(43):12029–12034?
4. Volkán-Kacsó S (2014) Two-state theory of binned photon statistics for a large class of waiting time distributions and its application to quantum dot blinking. *J. Chem. Phys.* 140:224110.
5. (Version R2016b) *MATLAB. The language of technical computing.* (MathWork, Inc., 1984-2018).
6. Marcus RA (1960) Discussion comment on mixed reaction-diffusion controlled rates. *Discuss. Farad. Soc.* 29:129.
7. Noyes RM (1961) Effects of diffusion rates on chemical kinetics. *Progr. React. Kinet.* 1:129–60.
8. Volkán-Kacsó S, Marcus RA (2015) Theory for rates, equilibrium constants, and brønsted slopes in F_1-atpase single molecule imaging experiments. *Proc. Natl. Acad. Sci USA* 112(46):14230–14235.
9. Weber J, Senior AE (1997) Catalytic mechanism of F1-ATPase. *Biochim. Biophys. Acta* 1319:19–58.
10. Braig K, Menz RI, Montgomery MG, Leslie AG, Walker JE (2000) Structure of bovine mitochondrial F1-ATPase inhibited by mg2+adp and aluminium fluoride. *Structure* 8(6):567–573.