Manuscript submitted to **Biophysical** *Reports*
**Report**

# Bayesian Inference Reveals Permissive and Nonpermissive Channel Closings in CFTR

Alexander S. Moffett[1,†], Guiying Cui[2], Peter J. Thomas[3], William D. Hunt[4], Nael A. McCarty[2], Ryan S. Westafer[5], and Andrew W. Eckford[1,*]

[1]Dept. of Electrical Engineering and Computer Science, York University, 4700 Keele Street, Toronto, ON M3J 1P3, Canada
[2]Emory + Children's Center for Cystic Fibrosis and Airways Disease Research, Emory University School of Medicine and Children's Healthcare of Atlanta, Atlanta, GA 30322, USA
[3]Dept. of Mathematics, Applied Mathematics, and Statistics, Case Western Reserve University, 10900 Euclid Avenue, Cleveland, OH 44106, USA
[4]School of Electrical and Computer Engineering, Georgia Institute of Technology, 777 Atlantic Drive NW, Atlanta, GA 30332, USA
[5]Georgia Tech Research Institute, 400 10th St NW, Atlanta, GA 30318, USA
[†]Current position: Center for Theoretical Biological Physics and Department of Physics, Northeastern University, Boston, MA 02115, USA
[*]Correspondence: aeckford@yorku.ca

ABSTRACT The closing of the gated ion channel in the Cystic Fibrosis Transmembrane Conductance Regulator (CFTR) can be categorized as nonpermissive to reopening, which involves the unbinding of ADP or ATP, or permissive, which does not. Identifying the type of closing is of interest, as interactions with nucleotides can be affected in mutants or by introducing agonists. However, all closings are electrically silent and difficult to differentiate. For single-channel patch clamp traces, we show that the type of the closing can be accurately determined by a Bayesian inference algorithm, which we demonstrate using both simulated and lab-obtained patch clamp traces.

SIGNIFICANCE Membrane ion channels are embedded in the plasma membranes of many eukaryotic cells, and the current through these channels can be measured using a patch clamp apparatus. The opening and closing of an ion channel is dependent on a sequence of conformational changes between structural states, called kinetic microstates. These microstates are crucial to understanding the dynamics of the channel, and are a subject of intense theoretical and experimental interest. In CFTR, ATP binding occurs only in certain sequences of state transitions, while toxins, inhibitors, and point mutations are known to have a direct impact on these transitions. However, experiments to directly observe hidden (i.e., electrically equivalent) states are complex, and challenges remain in characterizing the states in their dynamic context.

In contrast to experimental methods, Bayesian inference is a simple and well-known technique which has long been used as a solution to hidden-variable inference problems, including patch clamp traces. Applying Bayesian inference to CFTR ion currents, our algorithm accurately distinguishes between permissive channel closings (without unbinding of nucleotide) and nonpermissive closings (with unbinding of nucleotide), providing insight into an otherwise hidden, but physiologically important, process. Our method is flexible, and can be used to complement or improve other contemporary analytical or experimental techniques.

## INTRODUCTION

Cystic fibrosis (CF) is a life-threatening genetic disease affecting the respiratory and digestive systems, caused by mutations to the cystic fibrosis transmembrane conductance regulator (CFTR) anion channel (1, 2). CFTR is a "broken" member of the ATP-binding cassette (ABC) transporter class, in that CFTR acts as an ATP-gated ion channel rather than an active transporter as is the function of other ABC transporters. CFTR consists of a single polypeptide chain, with two transmembrane domain (TMD)-nucleotide binding domain (NBD) pairs connected through a region called the R domain (3). Each of the two NBDs contribute to both of the two known binding sites for ATP, although only one of these sites facilitates the hydrolysis of ATP to ADP. The two TMDs form a gated channel which is controlled by the state of the two intracellular NBDs.

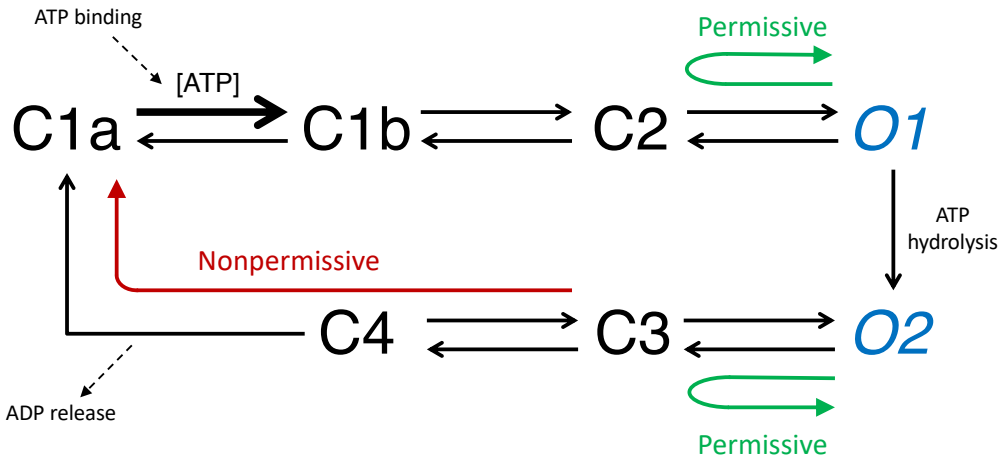


Figure 1: CFTR cycle model. Possible state transitions are indicated with black arrows. The C1a → C1b transition is sensitive to ATP concentration, and is indicated with a bold arrow and labelled [ATP]. States with open and closed ion channels are indicated in blue italics and black, respectively. Closings for which the reopening occurs in the same state are called *permissive*, depicted with green arrows; closings from O2 where the reopening is O1 are called *nonpermissive*, depicted with a red arrow. See also reference (4) and the SI.

Of key interest are *permissive* and *nonpermissive* closings of the CFTR ion channel, in the sense of permissive to rapid re-opening the channel. Considering the kinetic model in Figure 1 (see (4)), nonpermissive closings involve the release of ADP in order to enable the binding of ATP (5, 6), and include the irreversible C4 → C1a transition. On the other hand, permissive closings do not include this transition, meaning that they do not involve the release of ADP or ATP, leading to faster re-opening. However, as these transitions occur on states in the same conductance level, the two types of closing cannot be directly distinguished by the patch clamp. An analogous problem is to distinguish the pre- and post-hydrolytic states, O1 and O2. While we assume there are only two conductance states, O and C, some mutants have different conductance between O1 and O2 (7, 8), which allows the hydrolytic transition to be observed directly.

The problem of distinguishing hidden features of a signal is a natural application for Bayesian inference, in which the objective is to obtain the *a posteriori* probability $p(s \mid y)$ of a hidden random variable $s$ given observations $y$. In this case, $s$ represents the underlying kinetic state, while $y$ represents the conductance state, O or C. Many algorithms exist for calculating Bayesian inference, either exactly or approximately (9), such as the sum-product algorithm (10). Applications of inference algorithms are found in diverse areas such as bioinformatics (11, 12), biophysics (13, 14), telecommunications, and more recently in machine learning (15).

Previous work has used Bayesian inference to estimate hidden kinetic states from patch clamp measurements (16), The main contribution of this report is to apply and extend a Bayesian inference algorithm, showing that it can accurately distinguish between permissive and non-permissive channel closings in CFTR patch clamp traces, without any training or prior knowledge about parameter values. We demonstrate that this is an accurate and robust technique which reveals the hidden details of CFTR kinetics, particularly the precise timing of nucleotide unbinding from the channel.

## MATERIALS AND METHODS

### Receptor model

We use a physical model of CFTR described in Figure 1 (see (4)). In this 7-state model, states C1a, C1b, C2, C3, and C4 are fully closed, so we assume that ions are completely unable to pass through when CFTR is in these states, and thus, they have the same conductance level. States O1 and O2 are open states in which chloride can flow through CFTR. (The states are labelled so that the first letter indicates whether the channel is closed (C) or open (O).) Beginning with C1a, with a single ATP bound at the first ATP binding site, which is incapable of hydrolysis (17), the reversible transition to C1b occurs when a second ATP binds so that both binding sites are occupied. Because this step involves ATP binding, the rate depends on the concentration of ATP. The reversible transitions from C1b to C2 and from C2 to O1 are conformational changes resulting in an open CFTR , much like any ligand-gated channel such as the acetylcholine receptor (18). The transition from O1 to O2 is the first of two irreversible steps in the cycle, with one NBD-bound ATP undergoing hydrolysis to ADP. CFTR can then

undergo reversible conformational changes from O2 to C3 and from C3 to C4, resulting in a closed pore. Finally, the second irreversible step occurs in the transition from C4 to C1a, where the NBD-bound ADP unbinds from CFTR, leaving one apo and one filled ATP binding site. This 7-state model is in agreement with the 4-state simplified cyclic gating model of (19) and models distinguishing multiple closed and/or open states (20), (21).

As noted in the introduction, we are interested in determining whether the channel closings are *permissive* or *nonpermissive*. Nonpermissive closings include the irreversible C4 to C1a transition, in which ADP unbinds and a binding site is available for ATP, while permissive closings do not. Thus, considering Figure 1:

- A permissive closing has the same initial and final open states, i.e. O1 → C2 . . . C2 → O1 or O2 → C3 . . . C3 → O2; and
- A nonpermissive closing has different initial and final open states, i.e., O2 → C3 . . . C2 → O1. From Figure 1, the only way to do this is to proceed through the C4 → C1a transition.

The states under the ellipsis (. . .) can be any valid sequence of *closed* states from Figure 1, not necessarily the same state.

The kinetic microstates of CFTR can be modelled using a master equation of the form

$$\frac{dP}{dt} = PR, \tag{1}$$

where $P$ is a row vector with length equal to the number of kinetic microstates, and $R$ is a square matrix of kinetic rates for each possible state transition. In this formulation, $P_i$ is the probability that a receptor is in microstate $i$, while $R_{ij}$ is the transition rate from state $i$ to state $j$. The rate matrix $R$ and the full master equation are given in the SI.

## Patch clamp signal model

Formally, our system contains a set $\mathcal{V}$ of observable conductance states, a set $\mathcal{S}$ of hidden kinetic microstates, and a mapping $m : \mathcal{S} \to \mathcal{V}$ from microstates to conductance states. For CFTR, we have

$$\mathcal{V} = \{0, 1\} \tag{2}$$

$$\mathcal{S} = \{\mathrm{C1a}, \mathrm{C1b}, \mathrm{C2}, \mathrm{C3}, \mathrm{C4}, \mathrm{O1}, \mathrm{O2}\} \tag{3}$$

$$m(s) = \begin{cases} 0, & s \in \{\mathrm{C1a}, \mathrm{C1b}, \mathrm{C2}, \mathrm{C3}, \mathrm{C4}\} \\ 1, & s \in \{\mathrm{O1}, \mathrm{O2}\} \end{cases} \tag{4}$$

The conductance states $\{0, 1\}$ correspond to the ion channel's current when closed and open, respectively.

The patch clamp observes the channel current through additive noise, and samples these observations to form discrete-time signals. Let $y = [y_1, y_2, \ldots, y_n] \in \mathbb{R}^n$ represent the sequence of observations for a single channel, and let $s = [s_1, s_2, \ldots, s_n] \in \mathcal{S}^n$ represent the corresponding microstates. Then at the $k$th sample, the patch clamp measures

$$y_k = m(s_k) + n_k, \tag{5}$$

where $n_k$ forms a sequence of independent, identically distributed Gaussian random variables with zero mean and variance $\sigma^2$.

The transitions of microstates $s_{k-1} \to s_k$ are modelled as a discrete-time Markov chain (22): let $Q = [Q_{ij}]$ represent a $|\mathcal{S}| \times |\mathcal{S}|$ transition probability matrix, with $Q_{ij} = \Pr(s_k = j \mid s_{k-1} = i)$. Given the rate matrix $R$, and a discrete time step $\Delta t$, $Q$ is given by

$$Q = I + R\Delta t + o(\Delta t), \tag{6}$$

where $I$ is the identity matrix of the same size as $R$, and $\lim_{h\to 0}(o(h)/h) = 0$. The value of $\Delta t$ should be small enough that we can assume $Q = I + R\Delta t$; a rule of thumb is that $Q$ should be diagonally dominant. Note that $1/\Delta t$ is the sampling rate.

## Inference, parameter estimation, and simulation

We use the sum-product algorithm (which implements Bayesian inference, see (10)) to obtain the *a posteriori* distribution $p(s_k \mid y)$. Meanwhile, the transition probability matrix $P$ and noise variance $\sigma^2$ are unknown *a priori* and must be estimated from the data. The expectation-maximization (EM) algorithm (23) is a standard tool for this kind of simultaneous inference-estimation task. We employ a variant of the EM algorithm, known as the factor graph EM algorithm (24, 25), which is intended for

use alongside sum-product inference algorithms. The complete details of our algorithm are described in the supplemental information (SI).

Given $p(s_k \mid y)$, we define a confidence threshold $C$, $0 \leq C < 1$, and use the following decision rule to obtain the estimate $\hat{s}_k$ for each $s_k$:

$$\hat{s}_k = \begin{cases} \arg\max_{s_k \in \mathcal{S}} p(s_k \mid y), & p(s_k \mid y) > C, \\ \emptyset, & \text{otherwise.} \end{cases} \tag{7}$$

That is, $\hat{s}$ is the maximum *a posteriori* (MAP) estimate if the estimate exceeds $C$; otherwise, $\hat{s}$ is null ($\emptyset$). Setting $C = 0$ obtains the MAP estimate for all states $s_k$, while setting $C > 0$ reduces the probability of false alarm.

We test our inference algorithm via Monte Carlo simulation, by generating instances of discrete-time Markov chains with transition probability matrix $P$ (equation (6)) and adding noise (equation (5)). To evaluate the algorithm, we consider probability of false alarm $P_{\mathrm{FA}}$ and probability of missed detection $P_{\mathrm{MD}}$, respectively also known as type I and type II errors. We first make a list $L_{\mathrm{GT}}$ of closings (i.e., any sequence $\mathrm{OX} \rightarrow \mathrm{CX} \ldots \mathrm{CX} \rightarrow \mathrm{OX}$, with OX and CX representing any open or closed state, respectively, and where all states under the ellipsis (...) are closed states), and a list $L_{\mathrm{GT}}^{(\mathrm{np})}$ of nonpermissive closings, in the ground truth state sequence $s_k$. We make similar lists $L_{\mathrm{E}}$ and $L_{\mathrm{E}}^{(\mathrm{np})}$ for the sequence of estimated states $\hat{s}_k$. Represent the length of a list as, e.g., $|L_{\mathrm{GT}}|$. Now let $n_{\mathrm{FA}}$ be the number of closings that appear in $L_{\mathrm{E}}^{(\mathrm{np})}$ but not in $L_{\mathrm{GT}}^{(\mathrm{np})}$, and let $n_{\mathrm{MD}}$ be the number of closings that appear in $L_{\mathrm{GT}}^{(\mathrm{np})}$ but not in $L_{\mathrm{E}}^{(\mathrm{np})}$. Then

$$P_{\mathrm{FA}} = \frac{n_{\mathrm{FA}}}{|L_{\mathrm{E}}^{(\mathrm{np})}|}, \quad P_{\mathrm{MD}} = \frac{n_{\mathrm{MD}}}{|L_{\mathrm{GT}}^{(\mathrm{np})}|}. \tag{8}$$

### Patch clamp measurements

Single CFTR channels were studied in inside-out patches pulled from *Xenopus* oocytes injected with cRNA encoding the wildtype channel, as previously described (26). Briefly, to enable removal of the vitelline membrane, oocytes were placed in a bath solution containing (in mM) 200 monopotassium aspartate, 20 KCl, 1 $MgCl_2$, 10 EGTA, and 10 HEPES, pH 7.2 adjusted with KOH. Gigaohm seals were formed with patch pipettes pulled from borosilicate glass and filled with solution containing (in mM) 150 N-methyl-D-glutamine (NMDG) chloride, 5 $MgCl_2$, and 10 TES buffer, pH 7.5. After excision of the patch, CFTR channels were activated by bath solution containing 150 mM NMDG chloride, 1.1 mM $MgCl_2$, 2 mM Tris-EGTA, 10 mM TES buffer, 1 mM MgATP, and 127 U/ml PKA, pH 7.5. Currents were recorded at $V_M$ = -100 mV using an Axopatch 200B amplifier, with filtering at 0.1-1 kHz. The sampling rate of the patch clamp data was 2 kHz.

### Code

Code is available on GitHub, and was used to generate all results in this paper. We do not include simulated patch clamp traces in this paper, but simulations can be generated using the code we provide. The GitHub repository also includes a Jupyter notebook and raw data for generating all results in this paper. See: http://github.com/andreweckford/PatchClampFactorGraphEM/

## RESULTS

### Analysis using simulated patch clamp measurements

Here we give results obtained from Monte Carlo simulations of the CFTR ion channel. To generate simulated patch clamp results, we use example model parameters for the seven-state CFTR model given in Table 1. These example values are *not* provided to the inference algorithm, so performance does not in general depend on their accuracy.

*Properties of nonpermissive closings:* In Figure 2, we used our simulator to generate ground truth state sequences $s_k$; we then count the number of nonpermissive closings $|L_{\mathrm{GT}}^{(\mathrm{np})}|$ and the total number of closings $|L_{\mathrm{GT}}|$, and divide by time to get rates. In the figure, we see that rate of total closings varies from roughly 0.5-0.7 closings/s, while nonpermissive closings occur at a much lower rate. While the rate of closings depends on ATP concentration due to the cyclical nature of CFTR gating, the ratio of nonpermissive to total closings remains nearly independent of concentration. These observations are consistent with the dynamics of our model: a low ATP concentration will lead to a longer dwell time in state C1a, thus increasing the interval between openings without affecting the rates of transitions at the boundary between open and closed. Moreover, nonpermissive closings are relatively rare, with a ratio of roughly 1 nonpermissive closing per 12 total closings.

*Missed detection and false alarm probability:* In Figure 3, we give missed detection and false alarm probabilities, $P_{\mathrm{MD}}$ and $P_{\mathrm{FA}}$, using our algorithm (see equation (8)). Setting the confidence threshold $C = 0$ gives the MAP estimate of each state (see

| Origin state \ Destination state | C1a | C1b | C2 | O1 | O2 | C3 | C4 |
|---|---|---|---|---|---|---|---|
| C1a | | $9.0 \cdot 10^3$ (M s)$^{-1}$ [ATP] | | | | | |
| C1b | 5.0 s$^{-1}$ | | 7.7 s$^{-1}$ | | | | |
| C2 | | 5.8 s$^{-1}$ | | 4.9 s$^{-1}$ | | | |
| O1 | | | 10.0 s$^{-1}$ | | 7.1 s$^{-1}$ | | |
| O2 | | | | | | 3.0 s$^{-1}$ | |
| C3 | | | | | 7.0 s$^{-1}$ | | 6.0 s$^{-1}$ |
| C4 | 1.7 s$^{-1}$ | | | | | 12.8 s$^{-1}$ | |

Table 1: Parameters for the CFTR channel, corresponding to the model in Figure 1. A blank entry indicates that the transition is impossible. [ATP] indicates molar concentration of ATP.

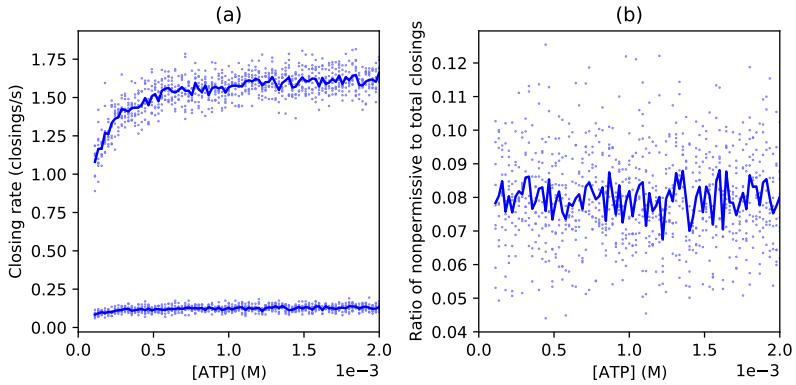


Figure 2: Prevalence of nonpermissive closings. *Subfigure (a)*: Rate of all channel closings (top line) and nonpermissive closings (bottom line) versus ATP concentration. *Subfigure (b):* Ratio of nonpermissive to total closings versus ATP concentration. Dots represent the outcomes of each simulation run, lines represent mean value at each concentration. Sampling rate = 100 Hz, rate parameters from Table 1.

equation (7)); we see a low $P_{\text{MD}}$, meaning few nonpermissive closings are missed, but $P_{\text{FA}}$ is relatively high, i.e. around 50%. (Since nonpermissive closings are rare, cf. Figure 2b, this is still far better than random guessing.) With a higher confidence threshold of $C = 0.8$, $P_{\text{FA}}$ is reduced, at the expense of increased $P_{\text{MD}}$. This is explained by noting that lower-confidence state estimates are discarded, so those closings will be missed by the algorithm. This demonstrates that $C$ can be adjusted to trade off $P_{\text{FA}}$ against $P_{\text{MD}}$. The performance of the algorithm is dependent on ATP concentration, with error rates increasing as concentration increases.

## Application to experimentally-obtained patch clamp measurements

In Figure 4, we show the application of our algorithm to lab-obtained CFTR patch clamp measurements. We show two examples, corresponding to two different experiments. For techniques used to obtain these measurements, see the Methods section.

Prior to applying our algorithm to the patch clamp signal, we perform a preprocessing step, block averaging (taking the sample average over non-overlapping blocks) for blocks of 50 samples, and scaling (multiplying by a constant, here 1.75, so that most signal features are in the range from 0 to 1). The block averaging step is performed to reduce the noise at high frequencies, which contains very little useful information about the signal, while preserving the features of interest at lower frequencies, improving the performance of the algorithm. We show the preprocessed signal in the middle plots of Figure 4, overlaid with vertical lines indicating the closings found by our algorithm, both permissive and nonpermissive. In the bottom plots of Figure 4, we give the state estimates $\hat{s}_k$ found by the algorithm, with different colors indicating whether or not the initial and final estimated states both exceed a confidence threshold of $C = 0.8$. From the raw data and preprocessed traces, abrupt transitions from high to low current correspond well to the detected closings.

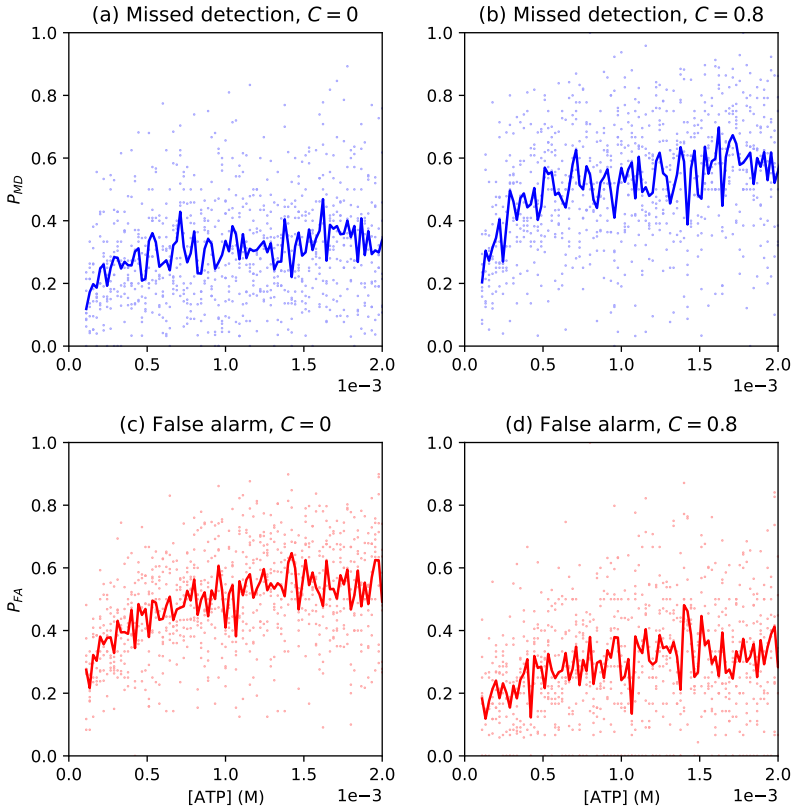


Figure 3: Missed detection and false alarm probabilities. *Subfigures (a), (b):* Missed detection probability versus ATP concentration for confidence threshold $C = 0$ and $C = 0.8$, respectively. *Subfigures (c), (d):* False alarm probability versus ATP concentration for $C = 0$ and $C = 0.8$, respectively. Dots represent each simulation run, while lines represent the average at each concentration. Sampling rate = 100 Hz, 20000 samples, 400 EM iterations, $\sigma^2 = 0.02$, rate parameters from Table 1.

## Discussion

We have presented an algorithmic tool for revealing the precise microstate kinetics of CFTR. Our method is simple, accurate, and robust, and can be applied to any patch clamp measurement of a single channel; moreover, it requires no training phase. Our inference-based approach can also be extended for use in combination with other techniques as long as a stochastic model exists for them.

By revealing permissive and nonpermissive closings, we can precisely estimate the timing of each nucleotide unbinding event, a key step in CFTR's kinetic model. Furthermore, our method may be used to study the effect of reagents that are known to affect hidden-state kinetics of CFTR, such as scorpion venom (4); future experiments in this direction might also be used to analyze pharmaceuticals that target CFTR. More generally, beyond permissive and nonpermissive closings, the features of our algorithm give the designer of an experiment a novel and fine-grained algorithmic tool to discover changes to the behaviour of receptor proteins. For example, this method could be used to determine the fine-grained, microstate-by-microstate effects of particular agonists or mutations.

## AUTHOR CONTRIBUTIONS

ASM, NAM, and AWE wrote the article. AWE designed the research, wrote the simulation code, and carried out the simulations. NAM and GC performed the patch clamp experiments. PJT, WDH, and RSW edited the article and contributed research ideas.

## DECLARATION OF INTEREST

The authors declare no competing interests.

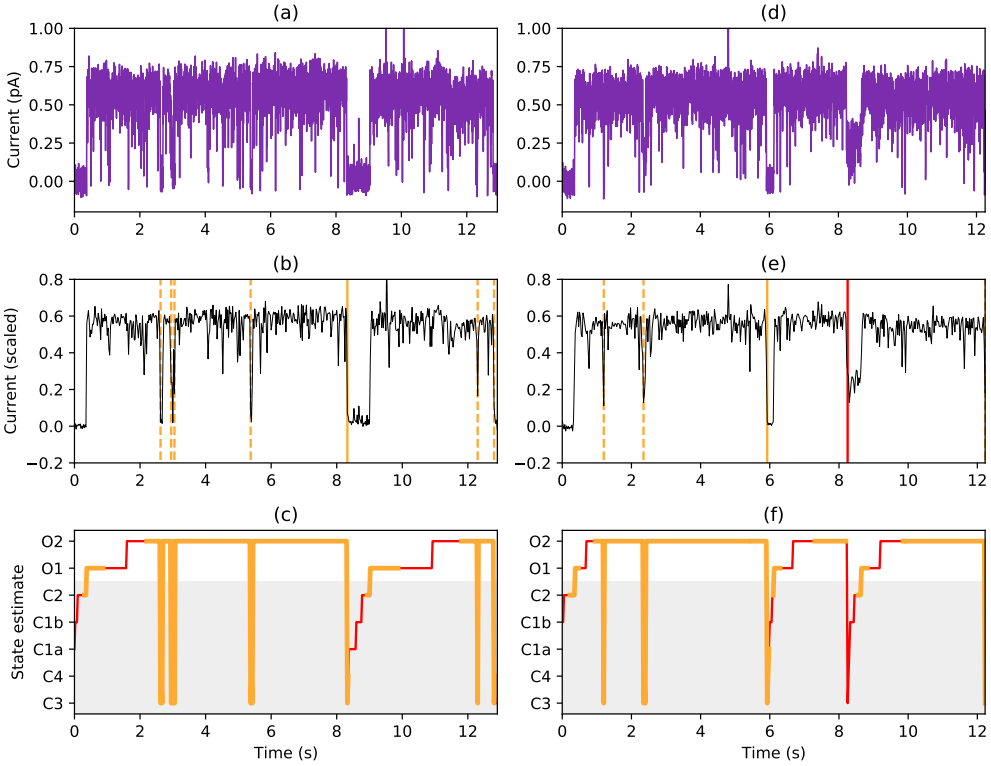


Figure 4: CFTR patch clamp measurements (raw and preprocessed) along with the corresponding inferred states for two different experiments, one in each column. *Top row (subfigures (a), (d)):* Measured patch clamp current. *Middle row (subfigures (b), (e)):* Patch clamp current signal after filtering and preprocessing; this signal is provided to the factor graph EM algorithm. Detected closings are depicted on these figures, with dashed / solid vertical lines respectively indicating permissive / nonpermissive closings; orange / red lines represent detected closings that, respectively, exceed / do not exceed $C = 0.8$. *Bottom row (subfigures (c), (f)):* Inferred state after 400 EM iterations; orange / red lines represent state estimates that, respectively, exceed / do not exceed $C = 0.8$.

## ACKNOWLEDGMENTS

ASM and AWE were funded by DARPA via the RadioBio program under grant number HR001117C0125. PJT was supported by NSF via grant number DMS-2052109 and by Oberlin College. WDH, NAM, and RSW were funded in part by DARPA via the RadioBio program under grant number HR001117C0124.

## REFERENCES

1. Csanády, L., P. Vergani, and D. C. Gadsby, 2019. Structure, gating, and regulation of the CFTR anion channel. *Physiological Reviews* 99:707–738.

2. Rey, M. M., M. P. Bonk, and D. Hadjiliadis, 2019. Cystic fibrosis: Emerging understanding and therapies. *Annual Review of Medicine* 70:197–210.

3. Zhang, Z., and J. Chen, 2016. Atomic structure of the cystic fibrosis transmembrane conductance regulator. *Cell* 167:1586–1597.

4. Fuller, M. D., Z.-R. Zhang, G. Cui, and N. A. McCarty, 2005. The block of CFTR by scorpion venom is state-dependent. *Biophysical Journal* 89:3960–3975.

5. Gout, T., 2012. Role of ATP binding and hydrolysis in the gating of the cystic fibrosis transmembrane conductance regulator. *Annals of Thoracic Medicine* 7:115.

6. Zhang, Z., F. Liu, and J. Chen, 2018. Molecular structure of the ATP-bound, phosphorylated human CFTR. *Proceedings of the National Academy of Sciences* 115:12757–12762.

7. Zhang, Z.-R., B. Song, and N. A. McCarty, 2005. State-dependent chemical reactivity of an engineered cysteine reveals conformational changes in the outer vestibule of the cystic fibrosis transmembrane conductance regulator. *Journal of Biological Chemistry* 280:41997–42003.

8. Zhang, J., and T.-C. Hwang, 2017. Electrostatic tuning of the pre-and post-hydrolytic open states in CFTR. *Journal of General Physiology* 149:355–372.

9. MacKay, D., 2003. Information Theory, Inference and Learning Algorithms. Cambridge University Press.

10. Kschischang, F. R., B. J. Frey, and H.-A. Loeliger, 2001. Factor graphs and the sum-product algorithm. *IEEE Transactions on Information Theory* 47:498–519.

11. Xiong, H. Y., Y. Barash, and B. J. Frey, 2011. Bayesian prediction of tissue-regulated splicing using RNA sequence and cellular context. *Bioinformatics* 27:2554–2562.

12. Meyer, X., L. Dib, D. Silvestro, and N. Salamin, 2019. Simultaneous Bayesian inference of phylogeny and molecular coevolution. *Proceedings of the National Academy of Sciences* 116:5027–5036.

13. Metzner, P., F. Noé, and C. Schütte, 2009. Estimating the sampling error: Distribution of transition matrices and functions of transition matrices for given trajectory data. *Physical Review E* 80:021106.

14. Potrzebowski, W., J. Trewhella, and I. Andre, 2018. Bayesian inference of protein conformational ensembles from limited structural data. *PLoS computational biology* 14:e1006641.

15. Barber, D., 2012. Bayesian reasoning and machine learning. Cambridge University Press.

16. Rosales, R., J. A. Stark, W. J. Fitzgerald, and S. B. Hladky, 2001. Bayesian restoration of ion channel records using hidden Markov models. *Biophysical journal* 80:1088–1103.

17. Infield, D. T., K. M. Strickland, A. Gaggar, and N. A. McCarty, 2021. The molecular evolution of function in the CFTR chloride channel. *Journal of General Physiology* 153:e202012625.

18. Albuquerque, E. X., E. F. Pereira, M. Alkondon, and S. W. Rogers, 2009. Mammalian nicotinic acetylcholine receptors: From structure to function. *Physiological Reviews* 89:73–120.

19. Csanády, L., and B. Töröcsik, 2014. Catalyst-like modulation of transition states for CFTR channel opening and closing: New stimulation strategy exploits nonequilibrium gating. *Journal of General Physiology* 143:269–287.

20. Vergani, P., S. W. Lockless, A. C. Nairn, and D. C. Gadsby, 2005. CFTR channel opening by ATP-driven tight dimerization of its nucleotide-binding domains. *Nature* 433:876–880.

21. Csanády, L., 2017. CFTR gating: Invisible transitions made visible. *Journal of General Physiology* 149:413–416.

22. Smith, G. D., 2002. Modeling the stochastic gating of ion channels. *In* Computational cell biology, Springer, 285–319.

23. Dempster, A. P., N. M. Laird, and D. B. Rubin, 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)* 39:1–22.

24. Dauwels, J., A. Eckford, S. Korl, and H.-A. Loeliger, 2009. Expectation maximization as message passing-Part I: Principles and Gaussian Messages. *arXiv preprint arXiv:0910.2832* .

25. Loeliger, H.-A., J. Dauwels, J. Hu, S. Korl, L. Ping, and F. R. Kschischang, 2007. The factor graph approach to model-based signal processing. *Proceedings of the IEEE* 95:1295–1322.

26. Infield, D. T., G. Cui, C. Kuang, and N. A. McCarty, 2016. Ion Channels and Transporters in Lung Function and Disease: Positioning of extracellular loop 1 affects pore gating of the cystic fibrosis transmembrane conductance regulator. *American Journal of Physiology-Lung Cellular and Molecular Physiology* 310:L403.