

From Average Transient Transporter Currents to Microscopic Mechanism—A Bayesian Analysis

Published as part of *The Journal of Physical Chemistry B* virtual special issue “Gregory A. Voth Festschrift”.

August George and Daniel M. Zuckerman*



Cite This: *J. Phys. Chem. B* 2024, 128, 1830–1842



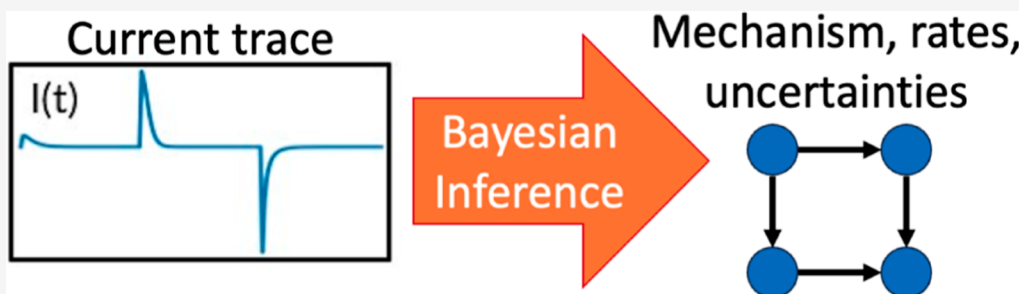
Read Online

ACCESS |

Metrics & More

Article Recommendations

Supporting Information



ABSTRACT: Electrophysiology studies of secondary active transporters have revealed quantitative mechanistic insights over many decades of research. However, the emergence of new experimental and analytical approaches calls for investigation of the capabilities and limitations of the newer methods. We examine the ability of solid-supported membrane electrophysiology (SSME) to characterize discrete-state kinetic models with >10 rate constants. We use a Bayesian framework applied to synthetic data for three tasks: to quantify and check (i) the precision of parameter estimates under different assumptions, (ii) the ability of computation to guide the selection of experimental conditions, and (iii) the ability of our approach to distinguish among mechanisms based on SSME data. When the general mechanism, i.e., event order, is known in advance, we show that a subset of kinetic parameters can be “practically identified” within ~ 1 order of magnitude, based on SSME current traces that visually appear to exhibit simple exponential behavior. This remains true even when accounting for systematic measurement bias and realistic uncertainties in experimental inputs (concentrations) are incorporated into the analysis. When experimental conditions are optimized or different experiments are combined, the number of practically identifiable parameters can be increased substantially. Some parameters remain intrinsically difficult to estimate through SSME data alone, suggesting that additional experiments are required to fully characterize parameters. We also demonstrate the ability to perform model selection and determine the order of events when that is not known in advance, comparing Bayesian and maximum-likelihood approaches. Finally, our studies elucidate good practices for the increasingly popular but subtly challenging Bayesian calculations for structural and systems biology.

INTRODUCTION

Transporters are a type of biological molecular machine that help regulate cellular homeostasis by pumping molecules across a membrane and maintaining ion gradients.^{1,2} As such, transporters play an essential role in cellular processes such as the uptake of nutrients and expelling of waste. These protein systems operate in a stochastic molecular environment, which suggests that they could exhibit some degree of stochasticity in their mechanism, a concept recently emerging as “pathway heterogeneity”,^{3–6} but which has been implicit in reports of noninteger stoichiometry of secondary active transport over many years.^{7–11} In a recent example, analysis of the small *Escherichia coli* multidrug transporter¹² points to utilization of different sequences of biochemical steps that enable a wide range of behaviors including 2:1 and 1:1 transport stoichiometry.^{4,13} Despite significant study, the precise mechanisms of transport,

i.e., the states visited, the allowed order(s) of transitions, and the associated rate constants, remain difficult to characterize.

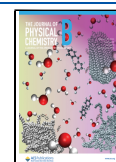
Traditional electrophysiology techniques^{14,15} have performed an essential role in many important discoveries related to membrane transport, such as determining the kinetics of ion channels^{16,17} and neuronal transmission,¹⁵ and have a long-standing history within the transporter research community.^{18,19} They provide dynamic information regarding the transport of ions across a membrane under various external conditions,

Received: October 23, 2023

Revised: January 31, 2024

Accepted: February 1, 2024

Published: February 19, 2024



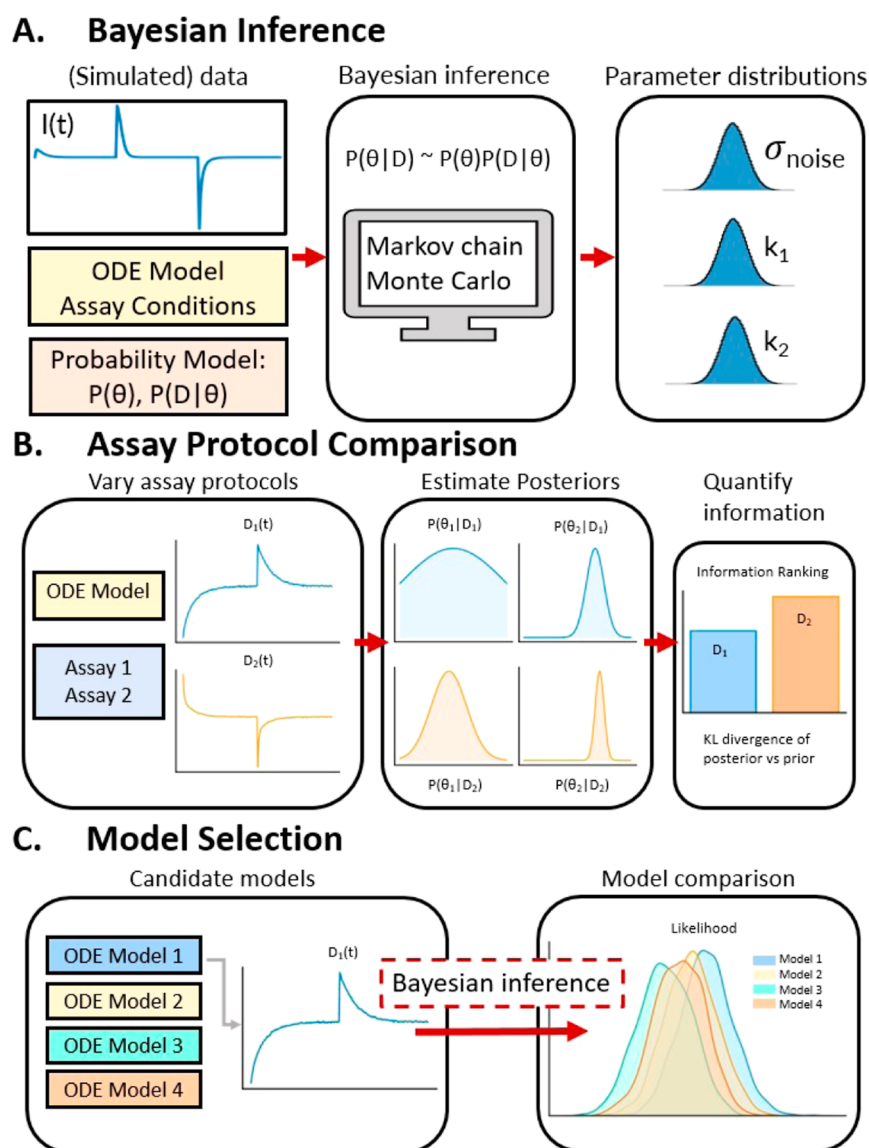


Figure 1. Overview of data analysis pipelines for transporter research. (A) BI is used with synthetic SSME assay data to generate parameter estimate distributions for reaction rate constants and other nuisance parameters (e.g., noise variance). (B) Different assay conditions are simulated from the same model, with the difference between the posterior and prior distributions quantified by the Kullback–Leibler (KL) divergence in order to approximate the information contained in each data set. (C) Likelihoods of different transporter reaction cycles are compared using the same data set for model selection.

which can be used to estimate specific kinetic parameters, such as net transport rates. Solid-supported membrane electrophysiology (SSME)^{20,21} extends traditional electrophysiology methods by introducing a fixed membrane embedded with reconstituted proteoliposomes that are perturbed under different external concentrations. This approach yields an averaged and aggregated transient current trace with a large signal relative to the noise and is stable across multiple perturbation concentrations in sequence, such as the three-stage reversal assay (see methods and work by Thomas et al.).²² However, it is unknown how well SSME can recapture microscopic information from the generated macroscopic data set. That is, how much information is contained in these data sets?

We address the question of inferring mechanistic details from SSME data by using Bayesian inference (BI). We are not aware of prior applications of BI to transporter parameter or mechanism inference. BI provides a well-established and robust

framework for estimating model parameters and their uncertainties from noisy data sets, as well as for distinguishing among models. Briefly, BI²³ is a powerful statistical method that generates a “posterior” probability distribution of model parameters, given the data set and prior beliefs about the model. The approach is computationally expensive compared to alternative frequentist methods such as maximum likelihood estimation (MLE)²³ but provides a comprehensive posterior distribution that contains most likely parameter estimates, uncertainties, and correlations, in addition to enabling comparison among models. In practice, the posterior distribution of parameters is estimated using Markov chain Monte Carlo (MCMC),^{24–26} which may prove challenging to converge with many systems of interest that are multidimensional and embody complex sampling landscapes, akin to rough energy landscapes. Here, we utilize a recently developed sampling approach that combines powerful ideas from physics (annealing) and from

machine-learning (normalizing flows)²⁷ to overcome convergence challenges.

Building on related work for ion channels^{28,29} and methods developed for computational systems biology,^{30,31} and molecular biophysics,^{32–34} we implement a BI data analysis pipeline (Figure 1) to address the gap in knowledge about membrane transporter mechanisms and SSME data sets. To our knowledge, this is the first application of BI for the transporter mechanism. We use ordinary differential equation (ODE) models of transporter reaction cycles and SSME-like conditions to generate synthetic data. The synthetic data enable a cost-effective and controlled environment with known ground truth values for methods development, but at the cost of potential simplifications and biases not found in experimental data. Synthetic data also can be used hand-in-hand with experimentation to design more informative experiments, as detailed below.

In this study, we demonstrate the effectiveness of BI for three important tasks in analyzing SSME data: parameter and uncertainty estimation; optimization of experimental conditions; and inference of mechanism, i.e., model selection. Using a (known) default model and data set, after validating our pipeline, we examine the practical identifiability issue: the precision with which each of >10 model rate constants can be determined. We consider the important, but often overlooked, effects of uncertainties in “known” experimental conditions, namely, species concentrations. We next consider an array of experimental conditions to ascertain which experiments and experiment combinations are the most informative for parameter inference. Finally, we use a Bayesian model selection strategy for four 1:1 transport cycles to ascertain the data required to distinguish the underlying model, when that is not known in advance. We also examined MLE results for model selection.

We find that the BI pipeline is well suited for these key tasks in SSME-based transporter research, generating highly informative posterior distributions across the various models and synthetic assays studied. The results reveal that most model rate constants can be determined within ~1 order-of-magnitude precision with suitable data, starting from a six order-of-magnitude range, but that certain rate constants are intrinsically less identifiable and may require independent measurements. The experiment optimization ranking reveals which SSME conditions contain the most information, typically with those with large pH/concentration changes, as well as combinations of different assays. Also, we found that high-information data sets enable distinguishing among four possible mechanistic models, whereas low-information data sets could not. These results illustrate the likely utility of BI in studying transporter mechanisms, giving insight into optimal experimental design and suggesting experiments that may be able to fully determine the reaction pathways of membrane transporters. Comparison of MLE calculations for model selection with BI results shows that certain MLE algorithms, properly tuned, can be successful, but BI may be a more robust approach overall. Finally, we also describe our experiences, and lessons learned, in MCMC sampling for BI calculations.

METHODS

In the following section, we describe the computational methods used to model membrane transporters, simulate SSME assays, and perform data analysis. A more comprehensive description of

the methods with implementation details can be found in the Supporting Information.

Modeling Membrane Transporters. Building on the foundational work by Mitchell³⁵ and Jardetzky,³⁶ we model membrane transport using the alternating access model. We represent these models using biochemical networks, with the dynamics determined by ODEs governed by mass action kinetics.³⁷ There are four idealized 1:1 antiporter models³⁸ that transport a single ion (H) and substrate (S) in opposite directions across the membrane. With two conformations, outward-facing (OF) and inward-facing (IF), there are six reaction states, with a unique set of reactions and conformational states distinguishing the models, as illustrated in Figure 7. These differences lead to a unique order of reaction events. For example, in cycles 1 and 3, k_1^i corresponds to an ion binding rate constant, but in cycles 2 and 4, k_1^i corresponds to a substrate unbinding rate constant. The different physical processes arise from the different states used in the model, such as with an unbound OF state used for cycles 1 and 3, and a doubly bound OF conformation in cycles 2 and 4. We primarily use the “Cycle 1” model in this study, with the exception of the model selection results, which utilize all four models.

The rate constants use an Arrhenius-like formulation,^{39,40} which naturally accommodates a dynamic membrane potential using an idealized capacitor model (see Supporting Information eqs 1 and 2). However, we found empirically that the membrane voltage had a negligible effect on the rate constants and observed current for the SSME conditions studied (see Supporting Information Figures 1 and 2), which in turn were motivated by experimental conditions.²² (Under more physiological conditions, i.e., larger membrane potentials, we confirmed that our formulation produced a significant adjustment in the rate constants as expected.) Therefore, we use a fixed membrane voltage of zero in order to reduce the computational complexity of our model. We also assume that the rate constants do not vary under different pH values. Furthermore, to ensure detailed balance, one of the 12 rate constants is necessarily constrained by thermodynamic consistency.⁴¹ Their governing ODEs along with further simulation and parameter details are described in the Supporting Information.

Generating Synthetic SSME-Like Data. As described previously, SSME experiments²⁰ consist of many proteoliposomes with reconstituted proteins deposited on a solid membrane. This system is placed in a bath of chemical species and, when perturbed, creates a gradient that drives ion transport across the membrane that is measured. After an initial equilibration stage, the concentrations are perturbed (i.e., activated) and the system relaxes to a new steady-state condition. In a three-stage reversal assay,²² the external concentrations are adjusted back to the initial values in the final stage, switching the gradient and driving transport in the opposite direction (i.e., reversal). Importantly, due to the stability of SSME, multiple assays can be performed in sequence under different perturbation amounts. Figure 3 shows an idealized diagram of an SSME experiment, with synthetic data traces shown in Figures 4, 5, and Supporting Information Figure 3.

We numerically simulate these assays by integrating the transporter cycle of the ODEs, with conditions changed for equilibration, activation, and reversal stages at $t = 0, 1$, and 2 s (and additional stages as needed). During each assay stage, the external concentrations of the ion and substrate in the bath solution are held fixed. We use a stiff ODE solver (CVODES)⁴²

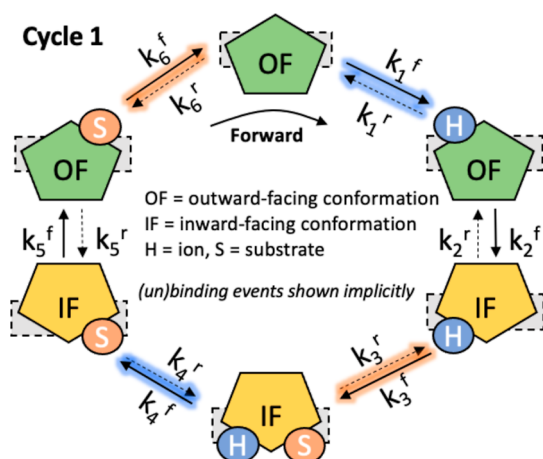


Figure 2. 1:1 transport cycle (antiport/exchange). Starting from the topmost fully unbound OF state and proceeding clockwise, an outside proton (H) binds, followed by conformational eversion to the IF state, then binding of an inside substrate (S), and so on. Note that binding and unbinding events are implicit in the state transitions. The time evolution of each state is described by a system of ODEs given in [Methods](#), with rate constants shown at transitions. This “Cycle 1” model is used to generate the synthetic data used in the study.

with a low tolerance to improve numerical stability. The net current is calculated from the change in internal ion concentrations of a single liposome, converting from the change in molar concentration to current and multiplying by the number of liposomes in the experiment.

$$I_{\text{net}}(t) = I_{\text{liposome}}(t)N_{\text{liposomes}} = \left(\frac{d[H_{\text{in}}^+]}{dt} \text{Vol}_{\text{in}} N_{\text{Av}} z \right) N_{\text{liposomes}} \quad (1)$$

where Vol_{in} is the internal volume of a single liposome embedded with transporters, N_{Av} is Avogadro’s constant, z is the elementary charge of an H^+ ion, and $N_{\text{liposomes}}$ is the total number of liposomes in the SSME assay. Here, we assume a known constant liposome volume and number liposomes as well as uniformity across the aggregate of liposomes. We note that while our data generating function in [eq 1](#) is a relatively simple macroscopic observable with few parameters, the change in the ion concentration $\frac{d[H_{\text{in}}^+]}{dt}$ is coupled to the other differential equations that have latent parameters (i.e., microscopic rate constants).

BI and Parameter Estimation. The predicted SSME-like data are used in the Bayesian pipeline to estimate the probability of the model parameters given the data, as given by Bayes’ theorem

$$P(\theta|D) = \frac{P(D|\theta) \cdot P(\theta)}{P(D)} \quad (2)$$

where

$P(\theta|D)$ is the posterior probability of the model parameters θ given the observed data D .

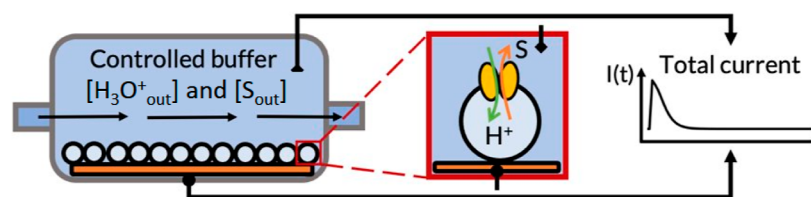
$P(D|\theta)$ is the likelihood of observing the data D given the model parameters θ .

$P(\theta)$ is the prior probability of the model parameters θ .

$P(D)$ is the evidence, the probability of the data D .

We use a standard log-likelihood function that assumes Gaussian experimental noise

A. Simplified diagram of experiment



B. Three stage assay: equilibration, activation, reversal

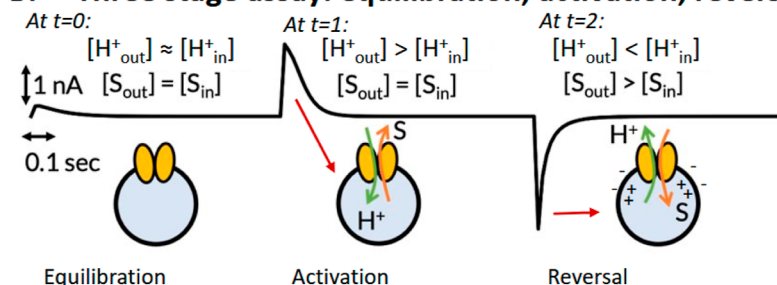


Figure 3. Schematic of SSME experiments. In SSME experiments generally (A), transporters embedded on liposomes transfer ions under a driving gradient, inducing a current. The bottom panel (B) shows a three-stage reversal assay: after an initial equilibration phase with matching concentrations inside and outside liposomes, the concentrations of the external species are perturbed to new fixed values in an activation phase to induce a transient current of ions across the membrane, and the external perturbation is subsequently reversed. During these perturbations, the electrochemical potential difference of the ion drives transport that relaxes in an exponential-like curve.

$$L(\theta, \sigma^2|D) = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_{i=1}^n (D_i - D_{\text{pred}}(\theta)_i)^2 \quad (3)$$

Here, our synthetic observation data over multiple time points, $D = (D_1, D_2, \dots, D_n)$, are generated from eq 1 by solving the governing ODEs with a set of reference (“ground truth”) reaction rate constants, compartment volumes, number of liposomes, and initial concentrations and then adding Gaussian white noise. During inference calculations, one set of trial parameters at a time, i.e., the vector θ , is used to generate the “predicted” data, D_{pred} , and evaluated in (eq 3) using a candidate Gaussian noise variance σ^2 . Additional nuisance parameters, for concentration uncertainties and experimental bias, are subsumed into θ .

The reaction rate constants are represented on a log 10 scale with wide uniform prior distributions covering 6 orders of magnitude to reduce bias in our analysis. Furthermore, we incorporate several nuisance parameters for the noise variance, initial species concentrations, and a multiplicative biasing factor for the overall current, each with uniform priors.

We use the “pocoMC” package for Monte Carlo sampling, which combines a physically motivated adaptive annealing protocol with a machine-learning accelerated preconditioned Monte Carlo sampler developed by Karamanis et al.^{27,43} The annealed importance sampling framework,^{44,45} in the context of BI, employs a pseudotemperature parameter to transition from effectively infinite temperature (uniform sampling) to the posterior distribution of interest. At each intermediate “temperature”, a neural network is trained using normalizing flows⁴⁶ via specialized autoencoder neural networks^{47,48} to simplify the geometry of the sampling space. We found this method to have improved performance over alternative Bayesian and MLE methods (see Supporting Information Figure 18–21).

Information Quantification and Experiment Optimization. For experiment optimization and recommendation, we are primarily interested in screening for protocols that yield high-information data and reduce the variance of our parameter estimates. In a Bayesian context, we quantify the information gained from a given data set by evaluating the Kullback-Leibler (KL) divergence⁴⁹ between the posterior and prior distributions.⁵⁰ This approach quantifies the distance between two probability distributions. We use this to evaluate the difference between the updated beliefs once the data have been observed (posterior distribution) and the prior beliefs before the data were observed (prior distribution). For a large number of samples, the discrete form of the KL divergence is

$$D_{\text{KL}}(P||Q) \approx \frac{1}{N} \sum_{i=1}^N \log\left(\frac{P(x_i)}{Q(x_i)}\right) \quad (4)$$

where

- $D_{\text{KL}}(P||Q)$ is the KL divergence from Q to P ,
- $P(x_i)$ is the posterior probability of the sample i ,
- $Q(x_i)$ is the prior probability of the sample i ,
- N is the total number of samples.

To overcome the issue of low sample number and the ‘noise’ in the posterior distribution, we utilize a Gaussian mixture model⁵¹ to estimate a smooth approximation of the estimated

posterior, which is tuned using the Bayesian information criterion.⁵² This workflow is illustrated in Figure SI 4.

Additionally, we quantify the improved precision of parameter estimates between two posterior distributions using the sum of the standard deviations. Specifically, we compare the sum of standard deviations across the marginalized posterior for different data sets

$$\text{sum of standard deviations for data set } i = \sum_{j=1}^M \sigma_{i,j} \quad (5)$$

where $\sigma_{i,j}$ is the standard deviation of the samples for the j th parameter (M total) based on the posterior for the i th data set. Note that these standard deviations of the posterior marginals differ from σ of eq 3 which models noise in the data values.

Model Selection. In order to compare between possible transporter reaction cycles, we examine the log-likelihoods (3), which in our formulation describe the scaled mean-squared error of the residuals. With BI, we generate a distribution of log-likelihoods, which correspond to the log probability of the data given the model parameters. Note that the likelihood function is simply proportional to the posterior in our implementation because of the uniform priors used. The ability to discern the most likely model depends on the separation of these log-likelihood distributions: if all models are equally likely, then their log-likelihood distributions will be overlapping, and if one model is more likely, then its log-likelihood distribution will contain the maximum and be separated from the others. Additionally, the model evidence, $P(D)$ can be estimated using sequential Monte Carlo methods and provides an alternative metric for model comparison⁵³ (shown in Supporting Information).

Implementation. We use the Tellurium package⁵⁴ in python⁵⁵ to build human-readable systems biological markup language⁵⁶ files using Antimony⁵⁷ and simulate the ODEs using libroadrunner.⁵⁸ For improved reproducibility, we use a.yaml⁵⁹ configuration file to specify the relevant model and data files, as well as the simulated assay conditions and model calibration settings. BI with preconditioned Monte Carlo is done using the pocoMC package,⁶⁰ affine invariant ensemble sampling using the emcee package⁶¹ and MLE is done using the Scipy Optimize package.⁶² Plots for figures are generated using Matplotlib,⁶³ and numerics are done using numpy.⁶⁴ For BI and MLE, multiple replicas are used to check for convergence. The code is available on GitHub: github.com/ZuckermanLab/Bayesian_Transporter, with data available upon request.

RESULTS

We present results for transporter parameter inference from SSME data, for comparison of the informativeness of different experimental protocols for parameter inference and for determination of the transport mechanism (event order). We focus on 1:1 exchangers (antiporters) based on the classical alternating access paradigm,^{35,36} described via a system of ODEs along with Gaussian noise. Throughout, we use synthetic data for 1:1 ion/substrate exchange generated using the ODE model for “Cycle 1” (Figure 2) described in Methods. The model has 12 rate constants, of which 11 are independent due to the thermodynamic cycle constraint.⁴¹ Synthetic data enable validation of the calculations and by itself represent a significant challenge for MCMC sampling. Alternative mechanisms for 1:1 transport, i.e., different event orders, are considered below.

The synthetic SSME current data models a three-stage protocol,²² shown in Figure 3, with current values from the

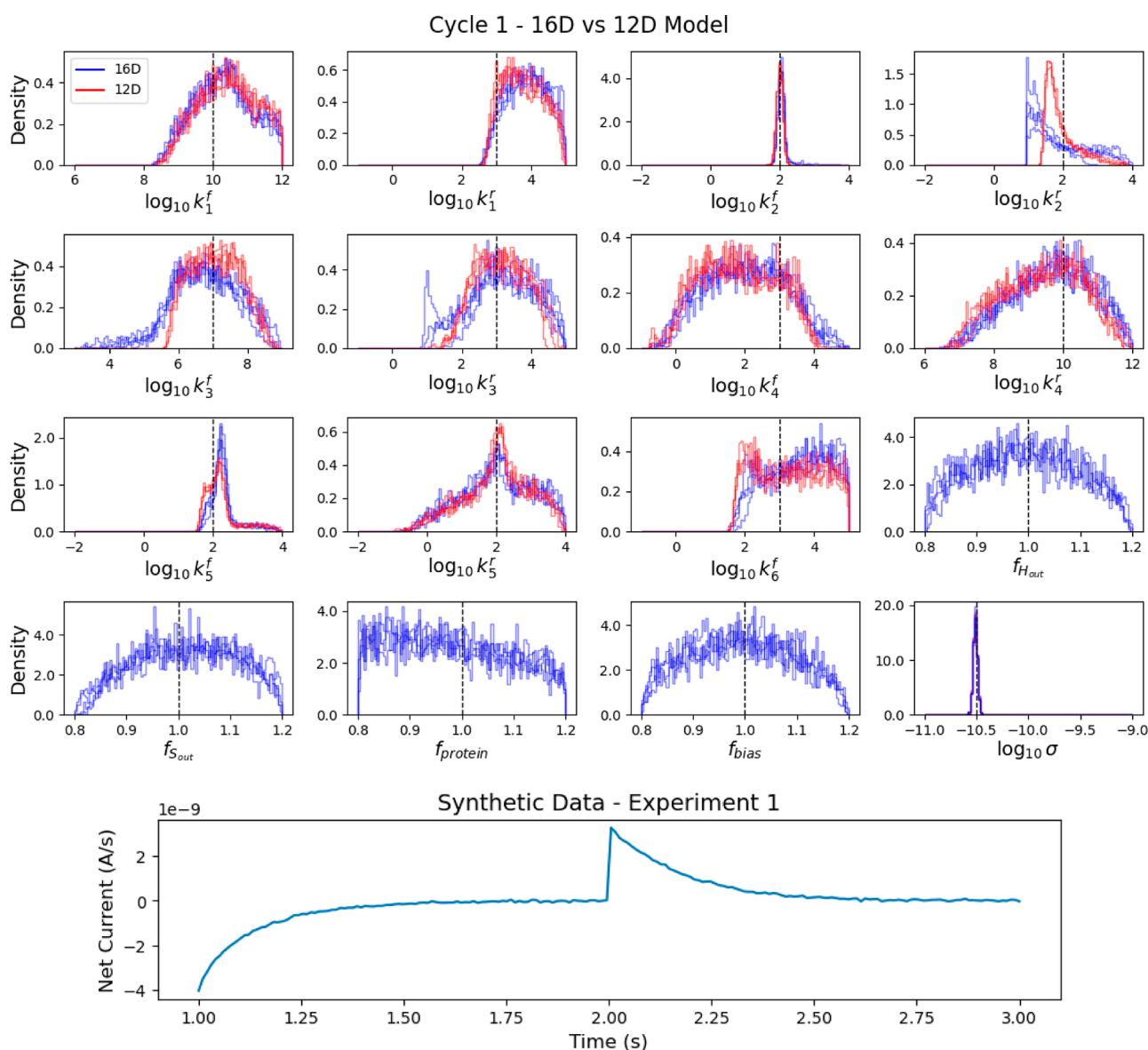


Figure 4. Parameter inference, comparing the effect of simple and more realistic parametrizations (Experiment 1). The estimated parameter distributions for two different model assumptions are shown based on multiple MCMC sampling replicas, with ground truth values (vertical lines) for reference. Here, the 12D model (red) parameters consists of 11 rate constants and Gaussian noise standard deviation. The 16D model (blue) additionally includes uncertainty in the initial assay concentrations and a measurement bias term. The horizontal range for each distribution matches the width of the uniform prior, notably 6 orders of magnitude for rate constants. Only data from “Experiment 1” were analyzed here, as shown in the bottom panel for reference. The 12D model used 3 MCMC sampling replicas, while the 16D model used 4 sampling replicas.

activation and reversal phases used in analysis. After the system is equilibrated in the initial stage, the external solution is adjusted to a new fixed composition in the second “activation” stage, creating an electrochemical potential difference that drives transport and relaxes in an exponential-like manner to a new steady state. In the third “reversal” stage, the external solution reverts to the initial concentrations, driving the system toward the equilibrium point of the first stage and generating a corresponding current trace. The results shown are based on computations neglecting voltage effects for simplicity because even when rate constants are formulated to account for voltage (Supporting Information Section 1.2), the modeled conditions—which are based on SSME experiments²²—lead to negligible $\sim\mu\text{V}$ -scale voltages (Supporting Information Section 1.3). Different experimental protocols considered here corre-

spond to activation by different external substrate concentrations and pH values.

Importantly, our BI calculations start from extremely noninformative prior ranges: rate constants are assumed to be unknown, and equally likely, within a range of 6 orders of magnitude. This breadth of parameter space makes the calculations challenging but is important in avoiding implicit foreknowledge of the true parameter values. MCMC sampling is performed with the pocoMC package, which uses an unbiased annealing process in combination with the “normalizing flows” approach.²⁷

Parameter Estimation and Comparison of Limited and More Realistic Experimental Uncertainties. A central issue in the analysis of experimental data is accounting for measurement noise and potential bias, not only for the output of the experiment, e.g., the current in electrophysiology, but also

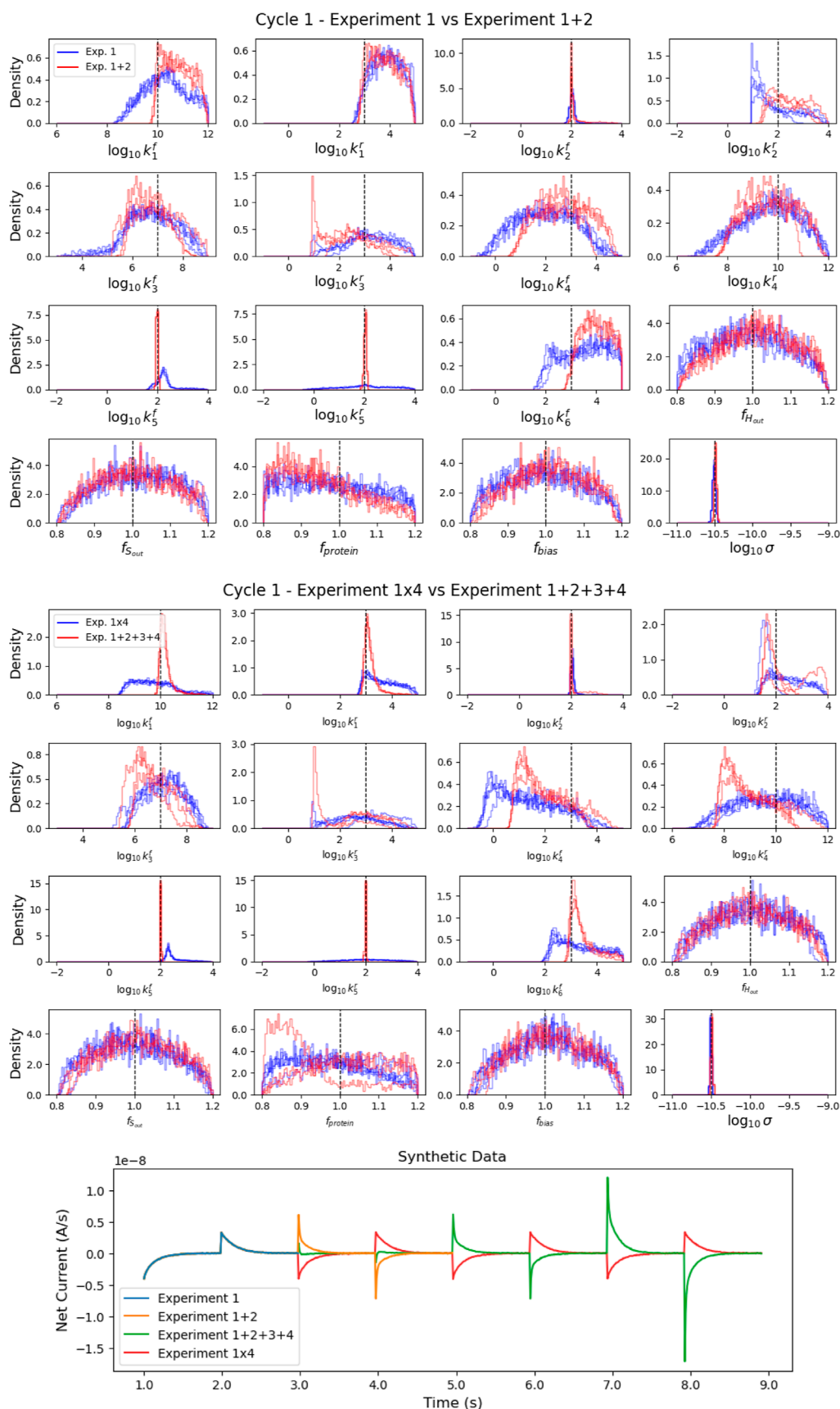


Figure 5. Posterior marginal distribution comparisons. The estimated parameter distributions for various synthetic assay protocols are shown based on multiple MCMC sampling replicas, with ground truth values (vertical lines) for reference. (Top) A single three-stage perturbation assay (blue, Exp. 1) is compared against a sequence of two three-stage perturbations assays (red, Exp. 1 + 2). The introduction of a second unique three-stage perturbation assay into the data set yields significant reductions in estimate variances. The experiment 1 protocol was analyzed via 4 MCMC sampling replicas, while

Figure 5. continued

the experiment 1 + 2 protocol used 3 sampling replicas. (Middle) A single three-stage perturbation assay (blue, Exp. 1 $\times$ 4) is replicated four times in sequence, and compared against a sequence of four unique three-stage perturbation assays (red, Exp. 1 + 2 + 3 + 4). The introduction of additional three-stage perturbation assays into the data set yields a significant reduction in estimate variance as compared to technical replicas. The experiment 1 $\times$ 4 protocol used 5 sampling replicas, while the experiment 1 + 2 + 3 + 4 protocol used 3 sampling replicas. (Bottom) Traces of the synthetic SSME-like data used.

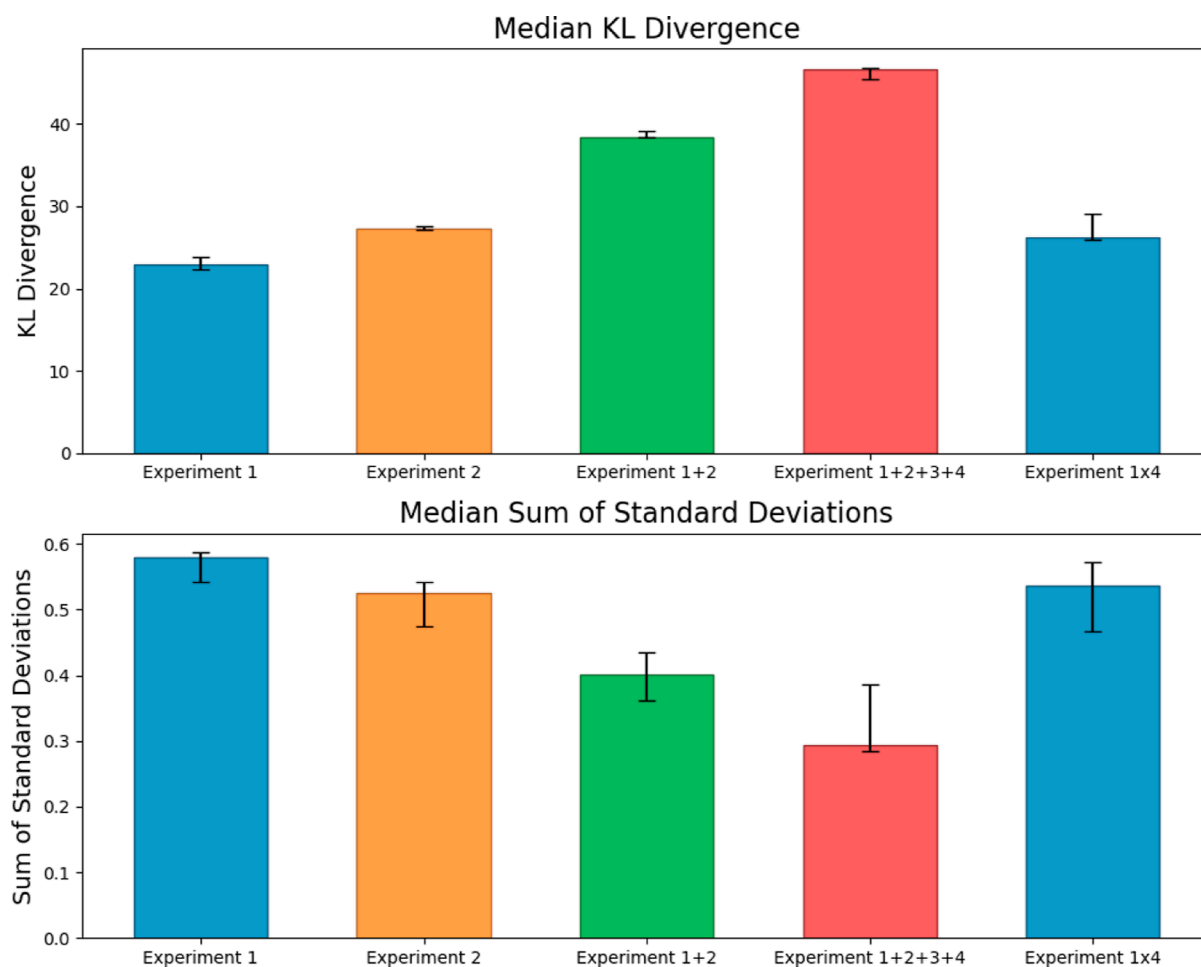


Figure 6. Quantifying information across experimental protocols. (Top) Median KL divergence between the posterior and uninformative uniform priors across multiple replicas, with error bars spanning the minimum to maximum values. (Bottom) The median sum of standard deviations of the parameter distributions, across multiple replicas is shown with error bars denoting the range from minimum to maximum. The introduction of additional experiments significantly increases the information gained as compared to a single experiment or technical replicas. The sum of standard deviations is correlated with the KL divergence and indicates an improvement in parameter precision as additional experiments are introduced. We note that KL divergence is dimensionless, but can be expressed in terms of nats. Similarly, the sum of standard deviations does not correspond to a physical quantity, but rather is a composite measure used to compare the relative estimated precision across varying conditions.

in “known” control parameters such as the concentrations of molecular and ionic species. Previous work analyzing calorimetry data has demonstrated the importance of accounting for possible measurement error in concentrations and, in some cases, shown that BI evidently can correct concentrations when a physical model for the system is known.^{33,65}

We therefore characterized the parameters for the Cycle 1 antiporter (Figure 2) using both a simplified and more realistic set of parameters. We first considered a 12-parameter set (12D model) consisting of the 11 independent rate constants of Cycle 1 plus a single “nuisance” parameter σ for the noise in the current measurement. We also studied a 16D model, which included four additional nuisance parameters: f_{bias} representing a systematic linear bias in the current to account for uncertainty

intrinsic to SSME measurements⁶⁶ and f_{protein} , f_{Hout} , and f_{Sout} representing uncertainties in concentrations of transporters, ion (or proton, H), and substrate (S) following our prior work.³³ The four additional parameters are represented as dimensionless scaling factors; see [Methods](#).

Our initial analysis is based on the (synthetic) “Experiment 1” data set, which represents an SSME three-stage protocol run a single time. This experiment consists of an initial equilibration phase (external pH 7) followed by an activation phase (external pH 7.3) and a final relaxation phase (external pH = 7). The external substrate concentration is held fixed at 1 mM for the duration of the experiment. Additional synthetic experiments, analyzed below, vary the external substrate concentrations. See the [Methods](#) and [Supporting Information](#) for further details. As

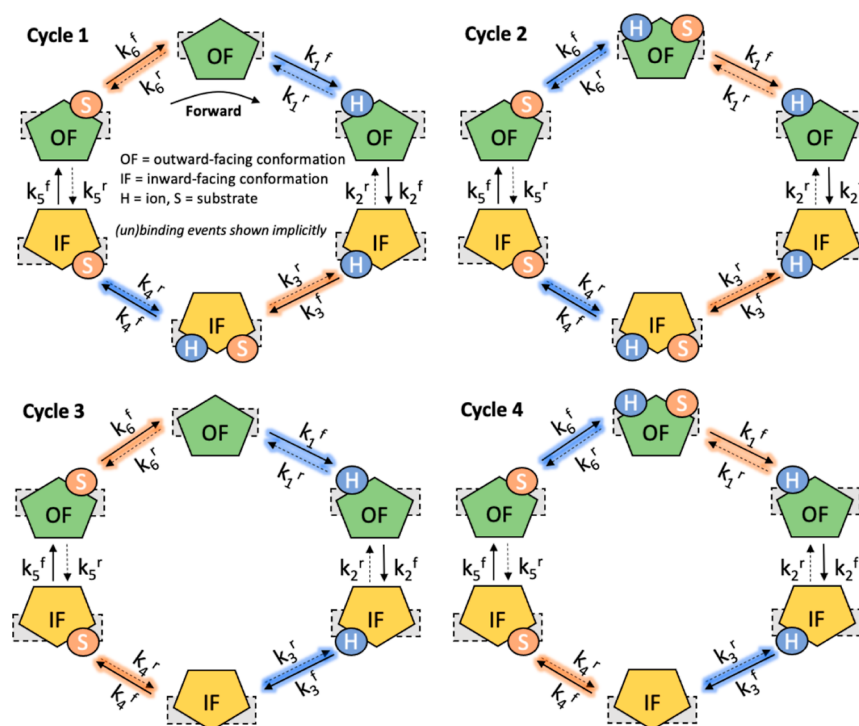


Figure 7. Four tightly coupled 1:1 antiporter reaction cycles. Each cycle transports an ion (H) and a substrate (S) in opposite directions via alternating access (OF and IF) conformations. Each network has a unique set of six reaction states, which represent different reaction event orders. The ion (un)binding reactions are highlighted in blue, with the substrate (un)binding reactions highlighted in orange, with each cycle having a unique combination of ion and substrate reactions. The binding and unbinding (i.e., dissociation) events are shown implicitly for improved visual clarity.

will be seen, the Experiment 1 data set is relatively uninformative by comparison to other sets.

The estimated posterior distribution is shown in Figure 4, projected as one-dimensional (1D) marginals of the multidimensional joint distribution. These marginal distributions determine the “credibility regions” (uncertainty ranges) of each parameter individually, providing low dimensional insight into a complex multidimensional distribution. We note that the modes of each individual marginal distribution do not necessarily correspond to the global optimal parameter set(s) due to the complex correlations among variables in the full joint posterior distribution. That is, the most globally likely parameter set could be offset from the peaks of the 1D marginals. Overall, only a few parameters can be determined precisely from the simple Experiment 1 data set analyzed; however, below, we will see significant improvement in parameter identifiability with improved data sets. When BI is able to find distributions that are narrow compared to the prior (full ranges shown), those include the known true parameter values. Furthermore, general consistency among independent MCMC replicas suggests that the posteriors are well sampled.

Most important here is the comparison between 12D and 16D parametrizations, which is indicative of the “cost” of including realistic nuisance parameters for experimental uncertainties, as well as the feasibility of MCMC sampling in the more complex case. Although the distributions for the rate constants are slightly broader in the 16D case, we see that inclusion of the additional nuisance parameters does not dramatically degrade parameter identifiability. Furthermore, the sampling is slightly worse in the 16D case, but the replicas are consistent enough overall to distinguish which parameters are practically identifiable, i.e., determined within ~ 1 order of magnitude. These results are further supported by examining the sum of standard deviations,

which indicates a modest decrease in the estimated precision of the model parameters when using the 16D model (shown in Figure SI 13).

Going forward, we employ only the more realistic 16D parametrization.

Information Quantification Across Simulated SSME Assay Conditions and Data Sets. Next, we examine optimizing the design of experiments by quantifying the gain in information and parameter precision when different data sets are employed, reflecting different experimental conditions and/or replicates. This is done by using the 16D model described above to generate data sets from several different assay protocols (detailed in Supporting Information) based on four different sets of experimental conditions. As before, multiple replicas of a sequential Monte Carlo²⁷ BI algorithm with broad uniform priors and a standard Gaussian-noise log-likelihood are run for each data set.

Posterior distributions comparing different data sets are shown in Figure 5 and provide strong evidence for improved parameter identifiability compared with the data set considered previously in Figure 4. As data from more diverse experimental conditions are included, the posterior marginals become narrower. In the best case examined, the combination of Experiments 1 + 2 + 3 + 4 leads to the effective identification (within ~ 1 order of magnitude) of six out of 11 rate constants. Figure 5 also shows the effect of including “technical replicates” of Experiment 1, which leads to better precision for many rate constants, but only to a slight degree. Generally, rate constants for conformational transitions were better determined than (un)binding rates, and there was not a major difference between ion and substrate parameter identifiability. It is notable that two-dimensional posterior marginals (Figure SI 12) show strong correlations among the IF binding and unbinding steps,

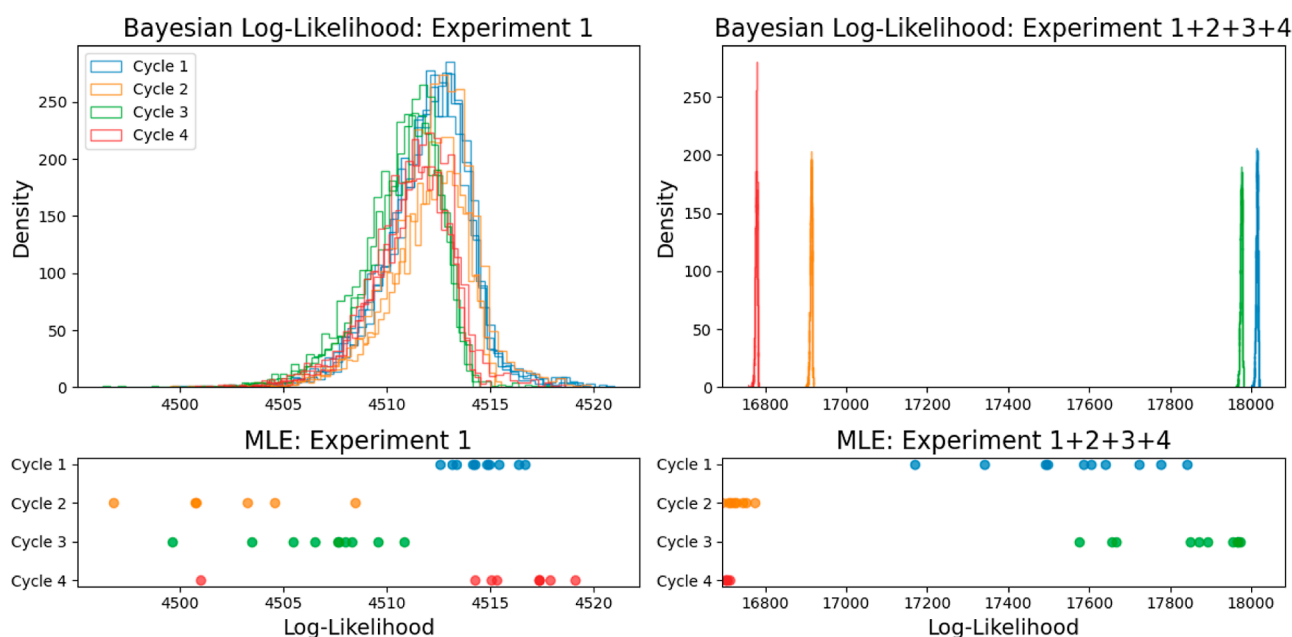


Figure 8. Transporter mechanism selection. (Top row) The Bayesian log-likelihood distributions of four transporter reaction cycles using less (left) and more (right) informative data sets, based on multiple MCMC sampling replicas. The less informative data set yields a large overlap between the model likelihood distributions, while the more informative data set yields a large divergence between the likelihoods, suggesting a single most likely model given the data. (Bottom row) The maximum log-likelihood values determined from repeated optimization trials for four transporter reaction cycles using less (left) and more (right) informative data sets, across multiple replicas. The MLE optimization approach generates a range of likelihood values across models and data sets, but fails to reliably converge on the maximum likelihood. We note that the “Cycle 1” (blue) transporter mechanism is used as a ground truth and is expected to have the largest likelihood.

effectively identifying the dissociation constants for steps 3 and 4. In the present data, by comparing MCMC replicates for the same experimental protocol, we see that sampling is not fully converged, but it does appear to be sufficient to support conclusions about parameter identifiability.

We can quantify information content of the various protocols via the KL divergence⁴⁹ of the posterior distribution relative to the prior distribution⁵⁰ for each data set. This effectively estimates the separation between two distributions and is used as a proxy for the information gained from our data set as compared to an uninformative prior. Furthermore, we calculate the sum of standard deviations of each parameter in the marginalized posterior distributions for each data set to quantify the precision of the parameter estimates. These results are shown in Figure 6. As expected from the posterior distributions, the combination of unique experimental conditions contains the more information and higher precision than a single experiment alone or technical replicas. Also, these results show a correlation between the “information gained” and practical identifiability of parameters.

Transporter Mechanism Identification (Model Selection). We now examine whether SSME current traces are sufficient to select among different event orders, i.e., mechanisms. We consider the four 1:1 antiport pathways shown in Figure 7, which includes the previously considered cycle (Figure 2) along with three additional mechanisms; see Methods for details. For comparison to BI calculations, we use MLE²³ with the differential evolution algorithm,⁶⁷ which was found empirically to perform well on a test (simpler) data set after hyper-parameter tuning (see Supporting Information). We examine the log-likelihoods for each model and data set, using both Bayesian and MLE methods, as shown in Figure 8. Finally, these results are compared with the estimated marginal

likelihood (i.e., model evidence) values generated during sampling.

Inferring which mechanism generated a given data set, the task of “model selection” requires the most informative data set explored for parameter inference. The BI results show a general overlap between each reaction cycle log-likelihood distribution when the less informative data set (Expt 1) is used, but a clear separation between distributions when the informative data set (Expt 1 + 2 + 3 + 4) is used. The results correctly identify “Cycle 1” as the most likely model by many orders of magnitude when using the informative data set, but with the uninformative data set, all the models have a similarly high likelihood. The maximum likelihood results fail to reliably converge to the expected maximum likelihood across the conditions studied, despite tuning of hyperparameters as described in the Supporting Information. Additional Bayesian posterior distributions for the cycles and data sets are shown in Figure SI 5–11.

Our results comparing the log-likelihoods are further supported when examining the marginal likelihood (i.e., model evidence) for each model under an informative and uninformative data set, as shown in the Supporting Information Figure 14. We find that the marginal likelihoods for each model have small relative differences when using an uninformative data set. In contrast, when an informative data set is used, there is a significant relative difference between the marginal likelihoods for each model-with the ground truth model having the highest evidence.

DISCUSSION

Understanding detailed mechanisms of molecular processes is a central goal of modern structural biology, and here, we have applied Bayesian inference (BI) to study the mechanism of driven biological transport, apparently for the first time. We have

examined the recently developed SSME approach, which is designed for the (relatively) high throughput study of transporters with high signal-to-noise.^{22,66} Synthetic SSME data allowed us to study one of the simplest realistic systems, a 1:1 antiporter based on ground-truth reference values, and also enabled facile assessment of the potential value of additional experiments; as detailed below, BI computations even for synthetic data represent a significant challenge. Our approach accounted for important, often overlooked facets of experimental uncertainty, including both concentration measurement uncertainty³³ and potential systematic measurement bias for electrophysiological currents.⁶⁶

Our study revealed both striking findings and challenges for quantifying the mechanism in transporters. First, it is surprising how much information about individual model parameters resides in SSME data, despite the simple exponential-like visual appearance of the data. The BI results (Figure 5) show that 6 out of 11 rate constants of a 1:1 transport model are “identifiable”—estimated within an order of magnitude (or less)—starting from a highly permissive six order-of-magnitude prior range for every rate constant. Furthermore, the correlation structure of the BI posterior (Figure SI 12) for the remaining unidentified rate constants implicitly determines two dissociation constants, i.e., rate-constant ratios, and provides effective guidance on which parameters should be determined from independent experiments. With the experimental data used, the conformational transition rate constants generally are better determined than on- and off-rates, but neither substrate nor ion (un)binding rates have a clear advantage for identifiability.

In comparing mechanistic pathways in a “model selection” task, the results (Figure 8) suggest that a well-chosen set of experiments can enable successful model selection, more than a single experiment or technical replicates, revealing the sequence of mechanistic steps. These results demonstrate the potential of BI to determine an unknown transporter mechanism from SSME data, or to eliminate unlikely candidate mechanisms. While we modeled each cycle individually, the Bayesian framework is compatible with and could be extended for mixed and hierarchical models (i.e., combinations of multiple pathways). As such, this study provides a foundation to systematically determine the mechanisms of more complex transporter cycles such as the proposed model for EmrE.¹³

The technical challenges involved with both BI and maximum likelihood (MLE) computations for synthetic data sets were significant. The less-than-ideal agreement among replica MCMC runs (Figure 5) highlights the sampling challenge despite our use of a sophisticated and highly parallelized MC sampling approach (pocoMC),^{27,60} which required approximately 24 h of computing per replicate. During the course of this study, we examined quite a few MCMC methods, none of which could provide the performance of pocoMC; for reference, we show a comparison to the affine-invariant ensemble sampler⁶¹ in Figures SI 20 and 21. Likewise, we examined a series of maximum likelihood methods, and most of the methods failed to optimize our systems even after tuning of hyperparameters (Figures SI 15–18); the MLE data shown in Figure 8 are from the differential evolution algorithm,⁶⁷ which were the best performer. We found that the BI methods were the most computationally expensive but had a comparable efficiency (number of likelihood calculations per second) as MLE methods such as differential evolution (see Figure SI 19), while generating full posterior distributions rather than point estimates. On the whole, our data on what might be considered

a simple model with synthetic data clearly demonstrate the technical challenges of BI^{30,68} and highlight the need for careful evaluation of MCMC sampling and MLE optimization.

While the synthetic SSME data studied here were motivated by experimental assay conditions and parameters^{22,69} and we accounted for concentration uncertainty and bias, experimental data will present new challenges. Higher ion/substate stoichiometry will introduce additional parameters and pathways, which may require model simplification.³⁹ As noted elsewhere, we used a simplified zero-voltage assumption due to the specific sets of conditions studied, justified by negligible peak voltages (Supporting Information Section 1.3). This is not a fundamental limitation, as the full, voltage-dependent formulation (Supporting Information Section 1.2) uses a dynamic membrane voltage adjustment term for the rate constants based on a capacitor model, which accounts for time-dependent membrane electrical properties^{70,71} found in ion channels. This formulation was consistent with physiological expectations under large membrane voltages (Supporting Information Section 1.3). A correction term (and parameter) for changes in rate constants due to significant differences in pH values may also be required. Furthermore, issues of transporter polarity and uniformity across liposomes, time delays from the mechanics of fluid mixing, and the effects of membrane capacitive coupling⁶⁶ may require additional “nuisance” parameters in models.

Our Bayesian approach included minimal assumptions about parameter values but did require a set of prior parameter ranges as well as a fixed reaction network structure with known states and transitions. While our choice of uniform priors introduces very little bias into the model, it may scale poorly for large dimensions or lead to overfitting if the model structure is not known, in which case, more informative data or priors may be required. For more complex systems, such as with increased ion/substate stoichiometry, there may be a mixture of pathways.^{4–6} The BI workflow can be adapted to model this in several ways, based on: one large network model containing all possible elementary reactions,³⁸ a mixture model of separate pathways, or a simplified representation of the model.³⁹ Unlike single-cycle models, more complex models may exhibit multimodal posterior distributions. The MCMC sampling algorithm employed, preconditioned Monte Carlo,²⁷ is well suited for sampling challenges as it is highly parallelized and uses an adaptive simulated annealing procedure capable of sampling multimodal distributions. Finally, as we have demonstrated, BI is able to compare mechanisms, which enables the identification of more and less probable models, given the available data.

Future work will aim to address the abovementioned challenges through close collaboration with experimentalists and any necessary extensions of our current modeling framework. We emphasize nevertheless that a Bayesian framework “self reports” on parameter and model uncertainty and hence can indicate when additional data are needed for BI and/or whether independent measurements of certain parameters are required. Although this study was limited to SSME-like data, the generic approach should be applicable to many other measurements.

CONCLUSIONS

This work has attempted to shed new light on the study of transporter proteins through the use of Bayesian inference (BI) and information theory to compare mechanisms, determine rate constants for elementary steps, and guide the experiment design. We performed a systematic BI study for 1:1 ion/substrate

secondary active transporters based on the emerging SSME experimental approach. Because the task of transporter mechanistic inference involves a combination of complex mechanisms, experiments under a potentially wide range of conditions, and numerically demanding analysis, we employed synthetic SSME data to understand the experimental information content and data requirements. Encouragingly, we found that a majority of the rate constants for individual mechanistic steps could be determined from well-designed SSME experiments and that the order of events could also be inferred. Although we found that BI was more reliable for model selection than maximum-likelihood approaches, our results also underscored the challenges of sampling for BI and the need for careful assessment. Taken as a whole, the study is a necessary step toward a more complete understanding of complex transporter reaction mechanisms, enabling effective model comparison, precise parameter estimation, and informed experimental design.

■ ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jpcb.3c07025>.

Transporter models, simulated SSME assays, probabilistic models used, quantification method, additional marginal posterior distribution figures, comparison of the sum of standard deviations for the 16D and 12D models, model selection methods, and comparison of MLE and BI methods (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Daniel M. Zuckerman – Department of Biomedical Engineering, Oregon Health and Science University, Portland, Oregon 97239, United States; orcid.org/0000-0001-7662-2031; Email: zuckermd@ohsu.edu

Author

August George – Department of Biomedical Engineering, Oregon Health and Science University, Portland, Oregon 97239, United States

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.jpcb.3c07025>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Katherine Henzler-Wildman, Nathan E. Thomas, and Merissa Brousseau for helpful discussions of SSME experiments; James Faeder, Jeremy Goecks, Peter Jacobs, Herbert Sauro, and Charles Springer for advice regarding the research and manuscript; Michael Grabe and John Rosenberg for valuable discussions of biological transport; and Jeremy Copperman, Aidan Estelle, Sagar Kania, and John Russo for help and advice with the project. This work was supported by NSF grant MCB-2119837.

■ REFERENCES

- (1) Alberts, B.; Johnson, A.; Lewis, J.; Morgan, D.; Raff, M.; Roberts, K.; Walter, P. *Molecular Biology of the Cell*, 6th ed.; Garland Science, Taylor and Francis Group: New York, NY, 2015.
- (2) Schultz, S. G. *Basic Principles of Membrane Transport*; Cambridge University Press: Cambridge, MA, 1980.
- (3) Mayes, H. B.; Lee, S.; White, A. D.; Voth, G. A.; Swanson, J. M. Multiscale kinetic modeling reveals an ensemble of Cl[−]/H⁺ exchange pathways in ClC-ec1 antiporter. *J. Am. Chem. Soc.* **2018**, *140*, 1793–1804.
- (4) Grabe, M.; Zuckerman, D. M.; Rosenberg, J. M. EmrE reminds us to expect the unexpected in membrane transport. *J. Gen. Physiol.* **2020**, *152*, No. e201912467.
- (5) Yue, Z.; Bernardi, A.; Li, C.; Mironenko, A. V.; Swanson, J. M. Toward a multipathway perspective: pH-dependent kinetic selection of competing pathways and the role of the internal glutamate in Cl[−]/H⁺ antiporters. *J. Phys. Chem. B* **2021**, *125*, 7975–7984.
- (6) Swanson, J. M. J. Multiscale kinetic analysis of proteins. *Curr. Opin. Struct. Biol.* **2022**, *72*, 169–175.
- (7) Terry, D. S.; Kolster, R. A.; Quick, M.; LeVine, M. V.; Khelashvili, G.; Zhou, Z.; Weinstein, H.; Javitch, J. A.; Blanchard, S. C. A partially-open inward-facing intermediate conformation of LeuT is associated with Na⁺ release and substrate transport. *Nat. Commun.* **2018**, *9*, 230.
- (8) Henderson, R. K.; Fendler, K.; Poolman, B. Coupling efficiency of secondary active transporters. *Curr. Opin. Biotechnol.* **2019**, *58*, 62–71.
- (9) Kettner, C.; Bertl, A.; Obermeyer, G.; Slayman, C.; Bihler, H. Electrophysiological Analysis of the Yeast V-Type Proton Pump: Variable Coupling Ratio and Proton Shunt. *Biophys. J.* **2003**, *85*, 3730–3738.
- (10) Huang, Y.; Chen, W.; Dotson, D. L.; Beckstein, O.; Shen, J. Mechanism of pH-dependent activation of the sodium-proton antiporter NhaA. *Nat. Commun.* **2016**, *7*, 12940.
- (11) Stelzl, L. S.; Fowler, P. W.; Sansom, M. S. P.; Beckstein, O. Flexible Gates Generate Occluded Intermediates in the Transport Cycle of LacY. *J. Mol. Biol.* **2014**, *426*, 735–751.
- (12) Schuldiner, S.; Granot, D.; Mordoch, S. S.; Ninio, S.; Rotem, D.; Soskin, M.; Tate, C. G.; Yerushalmi, H. Small is mighty: EmrE, a multidrug transporter as an experimental paradigm. *Physiology* **2001**, *16*, 130–134.
- (13) Robinson, A. E.; Thomas, N. E.; Morrison, E. A.; Balthazor, B. M.; Henzler-Wildman, K. A. New free-exchange model of EmrE transport. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, E10083–E10091.
- (14) Sakmann, B.; Neher, E. Patch clamp techniques for studying ionic channels in excitable membranes. *Annu. Rev. Physiol.* **1984**, *46*, 455–472.
- (15) Hodgkin, A. L.; Huxley, A. F. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.* **1952**, *117*, 500–544.
- (16) Colquhoun, D.; Hawkes, A. G. On the stochastic properties of single ion channels. *Proc. R. Soc. London, Ser. B* **1981**, *211*, 205–235.
- (17) Hille, B. *Ion Channels of Excitable Membranes*; Sinauer: Sunderland, MA, 2001; Vol. 507.
- (18) Loo, D.; Hazama, A.; Supplisson, S.; TuRK, E.; Wright, E. M. Relaxation kinetics of the Na⁺/glucose cotransporter. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 5767–5771.
- (19) Verkhratsky, A.; Parpura, V. History of electrophysiology and the patch clamp. In *Patch-Clamp Methods and Protocols*; Humana Press: New York, NY, 2014; Vol. 1–19.
- (20) Bazzzone, A.; Costa, W. S.; Braner, M.; Călinescu, O.; Hatahet, L.; Fendler, K. Introduction to solid supported membrane based electrophysiology. *J. Visualized Exp.* **2013**, *75*, No. e50230.
- (21) Schulz, P.; Garcia-Celma, J. J.; Fendler, K. SSM-based electrophysiology. *Methods* **2008**, *46*, 97–103.
- (22) Thomas, N. E.; Feng, W.; Henzler-Wildman, K. A. A solid-supported membrane electrophysiology assay for efficient characterization of ion-coupled transport. *J. Biol. Chem.* **2021**, *297*, 101220.
- (23) Gelman, A.; Carlin, J. B.; Stern, H. S.; Dunson, D. B.; Vehtari, A.; Rubin, D. B. *Bayesian Data Analysis*; CRC Press, 2013.
- (24) Hastings, W. K. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **1970**, *57*, 97–109.
- (25) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys.* **1953**, *21*, 1087–1092.

- (26) Jones, G. L.; Qin, Q. Markov chain Monte Carlo in practice. *Annu. Rev. Stat.* **2022**, *9*, 557–578.
- (27) Karamanis, M.; Beutler, F.; Peacock, J. A.; Nabergoj, D.; Seljak, U. Accelerating astronomical and cosmological inference with preconditioned Monte Carlo. *Mon. Not. R. Astron. Soc.* **2022**, *516*, 1644–1653.
- (28) Münch, J. L.; Paul, F.; Schmauder, R.; Benndorf, K. Bayesian inference of kinetic schemes for ion channels by Kalman filtering. *Elife* **2022**, *11*, No. e62714.
- (29) Calderhead, B.; Epstein, M.; Sivilotti, L.; Girolami, M. Bayesian approaches for mechanistic ion channel modeling. *Methods Mol. Biol.* **2013**, *1021*, 247–272.
- (30) Linden, N. J.; Kramer, B.; Rangamani, P. Bayesian parameter estimation for dynamical models in systems biology. *PLoS Comput. Biol.* **2022**, *18*, No. e1010651.
- (31) Shin, J.; Porubsky, V.; Carothers, J.; Sauro, H. M. Standards, dissemination, and best practices in systems biology. *Curr. Opin. Biotechnol.* **2023**, *81*, 102922.
- (32) Tavakoli, M.; Tsekouras, K.; Day, R.; Dunn, K. W.; Pressé, S. Quantitative Kinetic Models from Intravital Microscopy: A Case Study Using Hepatic Transport. *J. Phys. Chem. B* **2019**, *123*, 7302–7312.
- (33) Estelle, A. B.; George, A.; Barbar, E. J.; Zuckerman, D. M. Quantifying cooperative multisite binding in the hub protein LC8 through Bayesian inference. *PLoS Comput. Biol.* **2023**, *19*, No. e1011059.
- (34) Nguyen, T. H.; La, V. N. T.; Burke, K.; Minh, D. D. L. Bayesian regression and model selection for isothermal titration calorimetry with enantiomeric mixtures. *PLoS One* **2022**, *17*, No. e0273656.
- (35) Mitchell, P. A General Theory of Membrane Transport From Studies of Bacteria. *Nature* **1957**, *180*, 134–136.
- (36) Jardetzky, O. Simple Allosteric Model for Membrane Pumps. *Nature* **1966**, *211*, 969–970.
- (37) Voit, E. O.; Martens, H. A.; Omholt, S. W. 150 years of the mass action law. *PLoS Comput. Biol.* **2015**, *11*, No. e1004012.
- (38) George, A.; Bisignano, P.; Rosenberg, J. M.; Grabe, M.; Zuckerman, D. M. A systems-biology approach to molecular machines: Exploration of alternative transporter mechanisms. *PLoS Comput. Biol.* **2020**, *16*, No. e1007884.
- (39) Bisignano, P.; Lee, M. A.; George, A.; Zuckerman, D. M.; Grabe, M.; Rosenberg, J. M. A kinetic mechanism for enhanced selectivity of membrane transport. *PLoS Comput. Biol.* **2020**, *16*, No. e1007789.
- (40) Paz, A.; Claxton, D. P.; Kumar, J. P.; Kazmier, K.; Bisignano, P.; Sharma, S.; Nolte, S. A.; Liwig, T. M.; Nayak, V.; Wright, E. M.; et al. Conformational transitions of the sodium-dependent sugar transporter, vSGLT. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, E2742–E2751.
- (41) Hill, T. L. *Free Energy Transduction and Biochemical Cycle Kinetics*; Springer: New York, NY, 1989.
- (42) Serban, R.; Hindmarsh, A. C. CVODES: the sensitivity-enabled ODE solver in SUNDIALS. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*; ASME, 2005, pp 257–269.
- (43) Karamanis, M.; Seljak, U. Preconditioned Monte Carlo for Gradient-Free Bayesian Inference in the Physical Sciences. *Phys. Sci. Forum* **2023**, *9*, 23.
- (44) Huber, G. A.; McCammon, J. A. Weighted-ensemble simulated annealing: Faster optimization on hierarchical energy surfaces. *Phys. Rev. E* **1997**, *55*, 4822–4825.
- (45) Neal, R. M. Annealed importance sampling. *Stat. Comput.* **2001**, *11*, 125–139.
- (46) Papamakarios, G.; Nalisnick, E.; Rezende, D. J.; Mohamed, S.; Lakshminarayanan, B. Normalizing flows for probabilistic modeling and inference. *J. Mach. Learn. Res.* **2021**, *22*, 2617–2680.
- (47) Papamakarios, G.; Pavlakou, T.; Murray, I. Masked autoregressive flow for density estimation. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc., 2017; Vol. 30.
- (48) Germain, M.; Gregor, K.; Murray, I.; Larochelle, H. Made: Masked autoencoder for distribution estimation. In *Proceedings of the 32nd International Conference on Machine Learning*; PMLR, 2015, pp 881–889.
- (49) Kullback, S.; Leibler, R. A. On information and sufficiency. *Ann. Math. Stat.* **1951**, *22*, 79–86.
- (50) Huber, H. A.; Georgia, S. K.; Finley, S. D. Systematic Bayesian posterior analysis guided by Kullback-Leibler divergence facilitates hypothesis formation. *J. Theor. Biol.* **2023**, *558*, 111341.
- (51) Reynolds, D. Gaussian mixture models. In *Encyclopedia of Biometrics*; Springer: Boston, MA, 2009.
- (52) Neath, A. A.; Cavanaugh, J. E. The Bayesian information criterion: background, derivation, and applications. *Wiley Interdiscip. Rev.: Comput. Stat.* **2012**, *4*, 199–203.
- (53) Knuth, K. H.; Habeck, M.; Malakar, N. K.; Mubeen, A. M.; Placek, B. Bayesian evidence and model selection. *Digit. Signal Process.* **2015**, *47*, 50–67.
- (54) Choi, K.; Medley, J. K.; König, M.; Stocking, K.; Smith, L.; Gu, S.; Sauro, H. M. Tellurium: an extensible python-based modeling environment for systems and synthetic biology. *Biosystems* **2018**, *171*, 74–79.
- (55) van Rossum, G.; Drake Jr, F. L. *Python Tutorial*; Centrum voor Wiskunde en Informatica: Amsterdam, The Netherlands, 1995; Vol. 620.
- (56) Hucka, M.; Finney, A.; Sauro, H. M.; Bolouri, H.; Doyle, J. C.; Kitano, H.; Arkin, A. P.; Bornstein, B. J.; Bray, D.; Cornish-Bowden, A.; et al. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics* **2003**, *19*, 524–531.
- (57) Smith, L. P.; Bergmann, F. T.; Chandran, D.; Sauro, H. M. Antimony: a modular model definition language. *Bioinformatics* **2009**, *25*, 2452–2454.
- (58) Somogyi, E. T.; Bouteiller, J.-M.; Glazier, J. A.; König, M.; Medley, J. K.; Swat, M. H.; Sauro, H. M. libRoadRunner: a high performance SBML simulation and analysis library. *Bioinformatics* **2015**, *31*, 3315–3321.
- (59) Ben-Kiki, O.; Evans, C.; Ingerson, B. *YAML ain't markup language (YAML) version 1.1. Working Draft 2008-05*, 2008; Vol. 11.
- (60) Karamanis, M.; Nabergoj, D.; Beutler, F.; Peacock, J. A.; Seljak, U. pocoMC: A Python package for accelerated Bayesian inference in astronomy and cosmology. **2022**, arXiv preprint arXiv:2207.05660.
- (61) Goodman, J.; Weare, J. Ensemble samplers with affine invariance. *Commun. Appl. Math. Comput. Sci.* **2010**, *5*, 65–80.
- (62) Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **2020**, *17*, 261–272.
- (63) Hunter, J. D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95.
- (64) Van Der Walt, S.; Colbert, S. C.; Varoquaux, G. The NumPy array: a structure for efficient numerical computation. *Comput. Sci. Eng.* **2011**, *13*, 22–30.
- (65) Nguyen, T. H.; Rustenburg, A. S.; Krimmer, S. G.; Zhang, H.; Clark, J. D.; Novick, P. A.; Branson, K.; Pande, V. S.; Chodera, J. D.; Minh, D. D. Bayesian analysis of isothermal titration calorimetry for binding thermodynamics. *PLoS One* **2018**, *13*, No. e0203224.
- (66) Bazzzone, A.; Barthmes, M.; Fendler, K. *Methods in Enzymology*; Elsevier, 2017; Vol. 594, pp 31–83.
- (67) Storn, R.; Price, K. Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. *J. Global Optim.* **1997**, *11*, 341–359.
- (68) Kilic, Z.; Schweiger, M.; Moyer, C.; Pressé, S. Monte Carlo samplers for efficient network inference. *PLoS Comput. Biol.* **2023**, *19*, No. e1011256.
- (69) Hussey, G. A.; Thomas, N. E.; Henzler-Wildman, K. A. Highly coupled transport can be achieved in free-exchange transport models. *J. Gen. Physiol.* **2019**, *152*, No. e201912437.
- (70) Chen, D.; Xu, L.; Tripathy, A.; Meissner, G.; Eisenberg, B. Permeation through the calcium release channel of cardiac muscle. *Biophys. J.* **1997**, *73*, 1337–1354.
- (71) Grabe, M.; Lecar, H.; Jan, Y. N.; Jan, L. Y. A quantitative assessment of models for voltage-dependent gating of ion channels. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17640–17645.