Sources of Cell-to-cell Variability in Canonical Nuclear Factor-κB (NF-κB) Signaling Pathway Inferred from Single Cell Dynamic Images

**Mridul K. Kalita**1, Khachik Sargsyan1,2, Bing Tian1, Adriana Paulucci-Holthauzen3,4, Habib N. Najm3,2, Bert J. Debusschere5,2, and Allan R. Brasier1,2

Department of Medicine, the Optical Imaging Laboratory, and the Sealy Center for Molecular Medicine, University of Texas Medical Branch, Galveston, Texas 77555-1060 and Sandia National Laboratories, Livermore, California 94551-0969

The canonical nuclear factor-κB (NF-κB) signaling pathway controls a gene network important in the cellular inflammatory response. Upon activation, NF-κB/RelA is released from cytoplasmic inhibitors, from where it translocates into the nucleus, subsequently activating negative feedback loops producing either monophasic or damped oscillatory nucleo-cytoplasmic dynamics. Although the population behavior of the NF-κB pathway has been extensively modeled, the sources of cell-to-cell variability are not well understood. We describe an integrated experimental-computational analysis of NF-κB/RelA translocation and NF-κB signaling pathway control is stimulus- and duration-specific (11, 13). NF-κB controls a core component of early response genes, important in cellular intercommunication controlling innate immunity and anti-apoptotic responses. Upon activation, NF-κB signaling pathway control is stimulus- and duration-specific (11, 13). NF-κB controls a core component of early response genes, important in cellular intercommunication controlling innate immunity and anti-apoptotic responses.
**Cellular Variability in NF-κB Signaling**

In addition, it was observed that NF-κB also potently activates expression of IκBα and TNFAIP3/A20, two proteins mediating discrete steps of the negative feedback regulation of the NF-κB signaling pathway. IκBα expression recaptures NF-κB/RelA back into the cytoplasm, regenerating the cytoplasmic IκBα pool (14, 15), whereas the deubiquitinase TNFAIP3/A20 binds and inactivates IKK (16), terminating IκBα proteolysis. Thus, NF-κB regulates multiple inhibitory feedback loops that ensure only transient activation of inflammation (17).

The understanding of dynamic regulation of the canonical signaling pathway has been enhanced by theoretical approaches. Deterministic ordinary differential equation (ODE) models have been developed using population-derived data to examine the specific functional roles of IκBα, -β, and -ε isoforms in regulating the temporal control of NF-κB (18) and the role of the NF-κB-TNFAIP3/A20 feedback loop (19) (reviewed in Ref. 20). New advances in dynamic microscopy have enabled single cell imaging experiments (21–23). These have supplemented population-based experiments to indicate that the dynamic behavior of NF-κB nuclear translocation can vary significantly between cells. In response to tonic TNFα stimulation, an initial synchronous pulse of NF-κB translocation is followed either by nuclear persistence (monophasic activation) or in a minority of cells by dampened oscillatory profiles (22, 24). Individual cells producing asynchronous oscillatory responses produce population robustness, a phenomenon that is predicted to reduce tissue sensitivity to paracrine stimuli (26).

Inspection of the dynamic imaging studies indicates that cell populations exhibit heterogeneity in the kinetics of the initial NF-κB response profile. The sources producing cell-to-cell variation are largely unknown. Sensitivity analysis in simulations has indicated that cellular geometry, represented by the ratio of cytoplasmic to nuclear volume, a parameter termed $k_v$, has a significant influence on the initial spike of NF-κB because these molecules translocate into the nucleus. Using a clustering approach on trajectory simulations sampled over the parameter space, six different dynamic types of NF-κB translocation behavior were identified (27). However, the experimental measurement of cellular geometry in single cell measurements has not yet been accomplished. Others have proposed stochastic responses at the level of gene expression, although these responses largely influence the oscillatory modes (28, 29).

Despite theoretical advances in NF-κB pathway modeling, individual reaction rates and initial parameters are underdetermined. To extend our understanding of the pathway, we examined dynamic regulation of the NF-κB pathway in an experimentally validated model of airway epithelial cells stably expressing an enhanced green fluorescent protein (EGFP)-tagged NF-κB/RelA molecule. Quantitative measurements of NF-κB translocation surprisingly indicate that only a small fraction (<20%) of cytoplasmic RelA is translocation-competent, being distributed in IκBα, -β, and -ε isoform-containing complexes. Single cell dynamic imaging revealed that RelA translocation dynamic was monophasic in these cells. Cellular geometry was experimentally measured using axial sectioning. Incorporation of the cellular geometry and fractional NF-κB translocation into the priors for the Bayesian inference (30) aided the determination of joint probability distribution for the relevant parameters in the translocation model. Our analysis suggests a subpopulation of cells with significantly faster NF-κB decay profiles due to differences in the IκBα translation and degradation dynamics.

**EXPERIMENTAL PROCEDURES**

**Materials**—Recombinant TNFα (Pepro Tech) was added directly to the culture medium at a final concentration of 30 ng/ml.

**Cell Lines and Extracts**—The pCX4-pur-EGFP-RelA-expressing plasmid was constructed by cloning EGFP-RelA cDNA into pCX4-pur vector (31). The EGFP-RelA cDNA was isolated from pcDNA-EGFP-RelA (32) by Agel and XbaI restriction endonuclease digestion and ligated into pCX4-pur vector. Bosc 23 cells were cotransfected with pCX4-pur-EGFP-RelA and the amphotropic packaging plasmid pCL-10A1 using Lipofectamine 2000 (Invitrogen). After 3 days, retrovirus in the conditioned medium was collected, filtered, and used to infect A549 cells at 30–50% confluence in the presence of 8 μg/ml Polybrene. After 2 days of infection, the culture medium was changed to contain 2 μg/ml puromycin for selection. Puromycin-resistant clones were isolated and transferred to 6-well cell culture plates. Only strongly green fluorescent clones were selected further. A549 EGFP-RelA cells were grown in a medium containing 90% Dulbecco’s modified Eagle’s medium, 10% heat-inactivated fetal bovine serum, 4 mM L-glutamine, 0.1 mM non-essential amino acids, 1 mM sodium pyruvate, 100 mg/ml G418, 100 units/ml penicillin G sodium, and 100 mg/ml streptomycin sulfate in a humidified atmosphere of 5% CO₂.

Cytoplasmic and sucrose cushion nuclear extracts were prepared as described (11, 33). Protein concentration was measured using Coomassie Blue staining using bovine serum albumin as a standard (Bio-Rad) or by amino acid analysis for absolute protein quantification.

For confocal fluorescence microscopy, 25-mm round microscope cover glasses (Fisher) were precoated with collagen solution (Roche Applied Science) in a culture plate. A549 EGFP-RelA cells were plated 1 day prior to the experiment. For imaging, the cover glass was mounted into a microscopic incubation chamber where cells were maintained at 37°C with humidified 5% CO₂.

**Immunostaining and Confocal Immunofluorescence Microscopy**—A549 cells were plated on rat tail collagen-treated cover glasses and stimulated for the indicated times. The cells were fixed with 4% paraformaldehyde in PBS and incubated with 0.1 m ammonium chloride for 10 min. Cells were permeabilized with 0.5% Triton-100, followed by incubation in blocking buffer (5% goat serum, 0.1% IGEPA CAL-630, 0.05% NaN₃, and 1% BSA) and incubated with anti-RelA antibody (c-20, sc-372, Santa Cruz Biotechnology, Inc. (Santa Cruz, CA)) in incubation buffer (0.1% IGEPA CAL-630, 0.05% NaN₃, and 2% BSA) overnight at 4°C. After washing, cells were stained with Alexa Fluor 555-conjugated goat anti-rabbit IgG (Invitrogen) in incubation buffer for 1 h. The cells were visualized by Zeiss fluorescence LSM-510 confocal microscope at ×63 magnification.
Quantitative Real-time PCR—Total RNA was extracted using acid guanidium phenol extraction (Tri Reagent; Sigma). For NF-κB-dependent gene analyses, 1 μg of RNA was reverse transcribed using SuperScript III in a 20-μl reaction mixture. One-half μl of cDNA product was amplified in a 20-μl reaction mixture containing 10 μl of SYBR Green Supermix (Bio-Rad) and 0.4 μM each of forward and reverse gene-specific primers, aliquoted into 96-well, 0.2-mm thin wall PCR plates, and covered with optical quality sealing tape. The plates were denatured for 90 s at 95 °C and then subjected to 40 cycles of 15 s at 94 °C, 60 s at 60 °C, and 1 min at 72 °C in an iCycler (Bio-Rad). For each sample, the cycles generated were recorded. The PCR assays were denatured for 10 min at 95 °C, followed by 40 cycles of 15 s at 95 °C and 60 s at 60 °C. After PCR was performed, PCR products were subjected to melting curve analysis to ensure that a single amplification product was produced. Quantification of changes in gene expression was performed using the ΔΔCt method using uninfected cells as a calibrator. The sequences of forward and reverse primers were as follows: for IkBa, 5′-TCGCAGTGG ACCTGCAAAAT-3′ and 5′-TGAGC-TGGT AGGGAGAATAGC-3′; for IkBe, 5′-GGAACAATCT CACTCTCTGACC-3′ and 5′-TCAGCTTCTCAT CCGAAGCAG-3′; for GroB, 5′-AGGGGTTC GCCGTTCTCGGA-3′ and 5′-CCGAGGGAG CCGGATTG-3′; for NF-κB2, 5′-GGAGG- TGAG GGGGTACTGTAT-3′ and 5′-ATTACCGGGGTC GCAA-GCA-3′; for IL-8, 5′-ACTGAGAAGTG ATGAGAGTG-3′ and 5′-AACCTCTGCA CCCAGTTTTC-3′; for A20, 5′-TGGCTAGCCGGC GAGTGAG-3′ and 5′-AGGGGTTC GCCGTTCTCGGA-3′ and 5′-AGGGGTTC GCCGTTCTCGGA-3′ and 5′-AAATTCGGTAC AT-3′, 1.4 numerical aperture oil immersion objective. For each of the TNFα-stimulated and non-stimulated fixed cells, images were collected using two lines of excitation, 488 and 633 nm (both at 5% of transmission), and two different channels of emission. The intensity of emission for EGFP-RelA was obtained by using a 505–530-nm band pass filter (Ch2-T1), and the intensity of fluorescence of To-Pro3 was obtained by using a 650-nm low pass filter (Ch3-T2), respectively. All images were collected using 4-frame Kalman averaging with a pixel time of 2.51 μs and a pixel size of 200 × 200 × 300 nm, hence z-steps of 0.3 μm at an optical slice of 0.7 μm. Scan zoom was 1.4, and frame size was 512 × 512. Images were imported to Metamorph software (Molecular Devices, Downingtown, PA) for processing. Initially, background was subtracted from all images based on the signal of non-cellular regions. Regions were traced around the images collected on Ch3-T2 (nuclei) and on Ch2-T1 (cell body). Areas of the desired regions were then calculated, and the volume for that particular slice was obtained by multiplying region areas by 0.3 μm as the z-dimension. The total volumes of the cell compartments (cells and nuclei) were then computed by summing the volumes for the individual slices.

Western Immunoblots—Cytoplasmic or nuclear extracts from a constant number of cells were boiled in Laemmli buffer, fractionated on 10% SDS-PAGE, and electrotransferred to polyvinylidene difluoride membranes (Millipore, Bedford, MA). Membranes were blocked in 5% milk, Tris-buffered saline, 0.1% Tween (TBS-T) for 1 h and immunoblotted with the indicated primary Ab for 1 h at 4 °C (Santa Cruz Biotechnology, Inc.). Membranes were washed four times in TBS-T. Immune complexes were detected by infrared dye (IRD)-labeled goat anti-rabbit Ab and imaged in an Odyssey infrared imager (Li-Cor, Lincoln, NE). Standard curves were constructed for recombinant RelA to estimate RelA abundance. Homogeneous recombinant GST-RelA(488–551) was quantified by amino acid analysis. Serial dilutions of the purified GST-RelA(488–551) protein were assayed to generate standard curves for each immunoblot.

Immunoprecipitations—Cytoplasmic extracts were subjected to depleting immunoprecipitation using 4 μg of rabbit anti-IkBα, -β, or -γ Abs, and immune complexes were captured on protein A-agarose beads (40-μl slurry, Sigma-Aldrich). This amount was determined in separate experiments to deplete >90% of the target protein. The abundance of RelA with each IkB isoform was then determined in Western immunoblots using anti-RelA antibody to quantify RelA associated with each IkB isoform. Depletion of the respective IkB isoform was demonstrated by Western immunoblot of the flow-through of the immunoprecipitate. For the sequential immunoprecipitation experiment, 4 μg of rabbit anti-IkBα Ab was incubated with 2 mg of cytoplasmic lysate. The immune complexes were then removed using 40 ml of Protein A magnetic beads (Dynal, Invitrogen). The supernatant was then sampled for Western immu-

Cellular Variability in NF-κB Signaling

Quantitative Real-time PCR—Total RNA was extracted using acid guanidium phenol extraction (Tri Reagent; Sigma). For NF-κB-dependent gene analyses, 1 μg of RNA was reverse transcribed using SuperScript III in a 20-μl reaction mixture. One-half μl of cDNA product was amplified in a 20-μl reaction mixture containing 10 μl of SYBR Green Supermix (Bio-Rad) and 0.4 μM each of forward and reverse gene-specific primers, aliquoted into 96-well, 0.2-mm thin wall PCR plates, and covered with optical quality sealing tape. The plates were denatured for 90 s at 95 °C and then subjected to 40 cycles of 15 s at 94 °C, 60 s at 60 °C, and 1 min at 72 °C in an iCycler (Bio-Rad). The PCR assays were denatured for 10 min at 95 °C, followed by 40 cycles of 15 s at 95 °C and 60 s at 60 °C. After PCR was performed, PCR products were subjected to melting curve analysis to ensure that a single amplification product was produced. Quantification of changes in gene expression was performed using the ΔΔCt method using uninfected cells as a calibrator. The sequences of forward and reverse primers were as follows: for IkBa, 5′-TCGCAGTGG ACCTGCAAAAT-3′ and 5′-TGAGC-TGGT AGGGAGAATAGC-3′; for IkBe, 5′-GGAACAATCT CACTCTCTGACC-3′ and 5′-TCAGCTTCTCAT CCGAAGCAG-3′; for GroB, 5′-AGGGGTTC GCCGTTCTCGGA-3′ and 5′-CCGAGGGAG CCGGATTG-3′; for NF-κB2, 5′-GGAGG- TGAG GGGGTACTGTAT-3′ and 5′-ATTACCGGGGTC GCAA-GCA-3′; for IL-8, 5′-ACTGAGAAGTG ATGAGAGTG-3′ and 5′-AACCTCTGCA CCCAGTTTTC-3′; for A20, 5′-TGGCTAGCCGGC GAGTGAG-3′ and 5′-AGGGGTTC GCCGTTCTCGGA-3′ and 5′-AGGGGTTC GCCGTTCTCGGA-3′ and 5′-AAATTCGGTAC AT-3′, 1.4 numerical aperture oil immersion objective. For each of the TNFα-stimulated and non-stimulated fixed cells, images were collected using two lines of excitation, 488 and 633 nm (both at 5% of transmission), and two different channels of emission. The intensity of emission for EGFP-RelA was obtained by using a 505–530-nm band pass filter (Ch2-T1), and the intensity of fluorescence of To-Pro3 was obtained by using a 650-nm low pass filter (Ch3-T2), respectively. All images were collected using 4-frame Kalman averaging with a pixel time of 2.51 μs and a pixel size of 200 × 200 × 300 nm, hence z-steps of 0.3 μm at an optical slice of 0.7 μm. Scan zoom was 1.4, and frame size was 512 × 512. Images were imported to Metamorph software (Molecular Devices, Downingtown, PA) for processing. Initially, background was subtracted from all images based on the signal of non-cellular regions. Regions were traced around the images collected on Ch3-T2 (nuclei) and on Ch2-T1 (cell body). Areas of the desired regions were then calculated, and the volume for that particular slice was obtained by multiplying region areas by 0.3 μm as the z-dimension. The total volumes of the cell compartments (cells and nuclei) were then computed by summing the volumes for the individual slices.

Western Immunoblots—Cytoplasmic or nuclear extracts from a constant number of cells were boiled in Laemmli buffer, fractionated on 10% SDS-PAGE, and electrotransferred to polyvinylidene difluoride membranes (Millipore, Bedford, MA). Membranes were blocked in 5% milk, Tris-buffered saline, 0.1% Tween (TBS-T) for 1 h and immunoblotted with the indicated primary Ab for 1 h at 4 °C (Santa Cruz Biotechnology, Inc.). Membranes were washed four times in TBS-T. Immune complexes were detected by infrared dye (IRD)-labeled goat anti-rabbit Ab and imaged in an Odyssey infrared imager (Li-Cor, Lincoln, NE). Standard curves were constructed for recombinant RelA to estimate RelA abundance. Homogeneous recombinant GST-RelA(488–551) was quantified by amino acid analysis. Serial dilutions of the purified GST-RelA(488–551) protein were assayed to generate standard curves for each immunoblot.

Immunoprecipitations—Cytoplasmic extracts were subjected to depleting immunoprecipitation using 4 μg of rabbit anti-IkBα, -β, or -γ Abs, and immune complexes were captured on protein A-agarose beads (40-μl slurry, Sigma-Aldrich). This amount was determined in separate experiments to deplete >90% of the target protein. The abundance of RelA with each IkB isoform was then determined in Western immunoblots using anti-RelA antibody to quantify RelA associated with each IkB isoform. Depletion of the respective IkB isoform was demonstrated by Western immunoblot of the flow-through of the immunoprecipitate. For the sequential immunoprecipitation experiment, 4 μg of rabbit anti-IkBα Ab was incubated with 2 mg of cytoplasmic lysate. The immune complexes were then removed using 40 ml of Protein A magnetic beads (Dynal, Invitrogen). The supernatant was then sampled for Western immu-
Cellular Variability in NF-κB Signaling

n blot for RelA, and the remainder was incubated with 4 mg of anti-IκBβ Ab. The process was repeated for removal of NF-κB1 and -2, as indicated.

**Mathematical Modeling Reactions and Parameters**—The model involves a two-compartment kinetics of the NF-κB signaling pathway: IKK and NF-κB as activator proteins and A20 and IκBα as inhibitor proteins (Fig. 3). Because all proteins in the system are present in large numbers, as reported under “Results,” the effects of stochasticity in individual molecular reactions on the time evolution of each cell are expected to be minor. Therefore, an ODE model is chosen to describe the NF-κB translocation kinetics based on Ref. 19. The mathematical representation consists of a system of 14 ODEs and 27 reactions that account for complex formation of NF-κB/A20 and TNFAIP3/A20-IKK as well as IκBα degradation and RelA and IκBα nuclear and cytoplasmic transport (see the supplemental material for a listing of the equations and the associated nominal rate constants). The IKK complex and the dimer NF-κB are considered as single proteins; hence, the dynamics leading to formation of these complexes will be disregarded. It is assumed that the cytoplasmic complex IKK may exist in one of three forms: neutral (IKKn), active (IKKα), or inactive (IKKi). We also assume that each form of IKK undergoes degradation with the same degradation rate (kdeg); therefore, the total amount of IKK (i.e. IKKn + IKKα + IKKi) remains constant at equilibrium. The IKKα can form transient complexes with IκBα proteins or IκBα-NF-κB complexes. Based on Western immunoblot analysis, we have introduced another form of NF-κB, termed “reservoir NF-κB,” which does not translocate into the nucleus. An indirect measure of total NF-κB that is translocation-competent is termed “active NF-κB.” The reaction between the two reservoirs of NF-κB and active NF-κB were assumed to be very slow compared with the translocation dynamics, and their rates of translocation were therefore set to zero in the current analysis. The mRNA synthesis rate is the sum of a steady term and of an inducible term and is the same for both the A20 and IκBα proteins. The ODE model is presented in the supplemental material.

**Sensitivity Analysis Methodology**—Sensitivity analysis was used to rank model variables that were most important in controlling the pathway dynamics. We investigated the sensitivity of 30 parameters using a polynomial chaos representation. The 30 parameters include all 27 reaction rates, initial NF-κB concentration (NF-κB bound to IκBα), total NF-κB concentration, and kio, ratio (supplemental material) (35).

**Bayesian Inference**—Bayesian framework leads to probabilistic estimates of the model parameters in sparse and noisy data sets. The approach relies on Bayes’ formula,

\[
P(M|D) = \frac{P(D|M)P(M)}{P(D)} \quad (\text{Eq. 1})
\]

where the prior probability \( P(M) \) and the posterior probability \( P(M|D) \) represent degrees of belief about model \( M \), before and after having the particular data \( D \), respectively (30). The likelihood function, viewed as a function of the model, \( \mathcal{L}(M) = P(D|M) \), relates the prior distribution to the posterior, up to normalization by the evidence \( P(D) \). Below we describe how each component is defined in the context of our problem.

For estimating the parameters \( \kappa \) and \( \delta \) in the model:

\[
m(t) = \frac{I_n(t)}{I_n(t) + I_c(t)} = \frac{N_n(t)}{N_n(t) + N_c(t)} = \frac{\text{NF-κB}_n(t)}{\text{NF-κB}_n(t) + k_i\text{NF-κB}_i(t)} \quad (\text{Eq. 2})
\]

**Data D**—The measured absolute intensities in nucleus \( (I_n) \) and cytoplasm \( (I_c) \) are proportional to the number of NF-κB molecules in the nucleus \( (N_n) \) and cytoplasm \( (N_c) \), respectively (i.e. \( N_n = \alpha I_n \) and \( N_c = \alpha I_c \)). However, because \( \alpha \) is unknown, we consider only the ratio of nuclear intensities to the total ones and relate them to the model variables (i.e. concentrations) in the following way (including time dependence), where NF-κBn(t) and NF-κBi(t) are the total nuclear and total cytoplasmic concentrations for NF-κB in any form. To be precise, \( m(t) \) is the model output for the nuclear intensity fraction, whereas the corresponding data set \( D = \{d(t)\} \) of measurements taken at different time points \( t_i, i = 1, 2, \ldots, N \) forms the data that are used in Bayes’ formula.

**Model M**—The model in this context is a vector of some of the ODE model parameters (i.e. rate constants, initial conditions, the volume ratio \( k_{io} \), and the total concentration,

\[
T_c = \frac{\text{NF-κB}_n(t) + k_i\text{NF-κB}_i(t)}{1 + k_i} = \text{const.} \quad (\text{Eq. 3})
\]

**Prior P(M)**—In the absence of any specific shape information, we assume a uniformly distributed prior for the logarithms of the rate constants across a reasonable physical range for each of them. The prior information for the total concentration \( T_c \) was obtained experimentally and assumed to be a log normal distribution. We assume a log normal prior distribution for the volume ratios \( k_{io} \), as well, with parameters that are found by fitting the “area” data obtained from z-stack imaging data.

**Evidence P(D)**—The evidence merely plays a role of normalizing constant and does not affect the posterior estimation, as will become clear below.

**Likelihood P(D|M)**—The likelihood construction is viewed as a function of the model \( M \) and is equal to the probability of having a particular data set given the model. This conveniently incorporates errors produced by measurement as well as errors associated with the fact that the model is a simplification of the real network. This discrepancy is assumed to be gaussian-distributed and independent for each time snapshot. The underlying gaussian error model is assumed to have zero mean and an S.D. value that is linearly related to the measured data,

\[
é(t) \sim N(0, \sigma^2)
\]

where \( \sigma^2 = s_p + s_qd(t) \).

The linear dependence coefficients, \( s_p \) and \( s_q \), are hyperparameters. In the single cell inference case, the hyperparameters are inferred along with the model parameters, whereas in the multicell inference, we have fixed them (to keep the dimensionality of the Markov chain Monte Carlo (MCMC) algorithm from being too high) to values that are within the range of the posterior estimates found from single cell inference. Namely, these parameters are set to \( s_p = 0.01 \) and \( s_q = 0.03 \). The
assumed discrepancy model (Equation 4) then leads to the likelihood formula,

$$ L(M) = \prod_{j=1}^{N} \left( \frac{1}{\sigma_j \sqrt{2\pi}} e^{-\frac{(m_j - d_j)^2}{2\sigma_j^2}} \right) $$  \hspace{1cm} (Eq. 5)

where $N$ is the number of data points. The product is a direct consequence of the independent measurement assumption. Furthermore, in order to sample from the posterior distribution via Bayes’ formula, an MCMC technique has been used, with the original Metropolis algorithm (36) enhanced by adaptive proposal distributions (37) (see the supplemental material). Also, to ensure positive values for all of the parameters, MCMC was performed on the logarithms of the values of the parameters.

Single Cell Inference—Bayesian inference was performed for each cell based on respective cell data. In the single cell analysis, the dimensionality of the MCMC chain is eight parameters (i.e., three reaction kinetic rate constants ($c_{4a}$, $c_{6a}$, and $i_1$) as well as the volume ratio $k_{\nu}$, initial NF-κB-1xkB concentration, and total NF-κB concentration $T_{\nu}$, along with hyperparameters $s_p$ and $s_q$ were inferred). We have obtained a reasonable chain mixing and acceptance rate of ~50% for $N_{\text{MCMC}} = 80,000$ steps by using an initial gaussian proposal distribution of width 0.05 (i.e., $C_0 = 0.05S_p$; see supplemental material) for each parameter during the first $n_s = 5,000$ steps and adaptive MCMC covariance scaling factor $\gamma = 0.05$ for the remaining 75,000 steps.

Simultaneous Inference of Multiple Cells—For all three multicell analyses (i.e., 20-, 16-, and 4-cell analyses), we allowed the reaction kinetic rate constants ($c_{4a}$, $c_{6a}$, and $i_1$) to be common among cells during the inference, whereas other parameters ($k_{\nu}$, initial NF-κB, and $T_{\nu}$) were made specific to each cell and were inferred. The rationale behind this strategy is that the probability of a cell property to vary from other cells is higher for its cell volume, its compartmental volume ratios, and its NF-κB concentration rather than its physiological reaction rates. In order to avoid a very high dimensional MCMC, the noise model parameters were set to $s_p = 0.01$ and $s_q = 0.03$ for all cells. In 20-cell inference, for instance, this resulted in a 63-dimensional space for the inference problem. The rest of the MCMC parameters were kept the same as in the single cell analysis.

RESULTS

Validation of the Cellular Model—To measure dynamic changes in RelA translocation at the single cell level, we constructed human A549 alveolar epithelial cells stably expressing an EGFP-tagged NF-κB/RelA. Because ectopic EGFP-RelA expression may influence the dynamics of the system response, we characterized the cells to determine the relative levels of ectopic to endogenous RelA and to assess translocation dynamics. Western immunoblots were performed of control and stimulated cytoplasmic and nuclear extracts (these fractionation methods have been extensively validated, with the nuclear fraction free of cytoplasmic contamination, measured by β-tubulin staining, and the cytoplasmic fraction free of nuclear contamination, measured by β-laminin staining (33, 38)). Here, in unstimulated cell cytoplasm, the 96-kDa EGFP-RelA band could be identified at a ~1:1 ratio with that of 65-kDa endogenous RelA in A549 wild type (WT) cells (Fig. 1A). Upon TNFα stimulation, both EGFP-RelA and endogenous RelA translocated into the nucleus. Interestingly, TNFα stimulation did not result in significant cytoplasmic depletion of RelA in either wild type or EGFP-RelA-expressing cells (Fig. 1A). This behavior is similar to other cell types where the fraction of cytoplasmic protein translocating into the nucleus was analyzed (15).

To confirm these data, comparative confocal imaging experiments were conducted. In unstimulated cells, EGFP-RelA was primarily distributed in the cytoplasm. In response to TNFα stimulation, EGFP-RelA translocated into the nucleus (exclusive of nucleioli) with kinetics similar to that observed for the endogenous RelA protein in A549 WT cells (Fig. 1B). Additionally, we noted that although the nuclear staining was increased in response to stimulation, detectable amounts of RelA remained cytoplasmic for both EGFP-RelA and endogenous RelA in A549 WT cells (Fig. 1B).

We and others have shown that NF-κB/RelA controls the expression of a network of genes whose expression occurs in temporally distinct waves, with some genes activated rapidly (IkBa, GroB, TNFAIP3/4A20, IL-8, and IL-6) and others being activated slowly (IkBe, NF-κB2, and Naf-1) in response to TNFα stimulation (11, 13). To further characterize our cellular model, the profiles of representative members of the NF-κB-dependent gene network were analyzed in EGFP-RelA and A549 WT cell responses to tonic TNFα stimulation. Members of the “early” NF-κB-dependent subnetwork were strongly induced by TNFα, peaking within 1 h of stimulation in a manner that was identical for both the EGFP-RelA and the A549 WT cells (Fig. 1C). Similarly, the “late” subnetwork was activated after 6–9 h, with kinetics and magnitude that were distinguishable between EGFP-RelA and A549 WT cells (Fig. 1C). These data indicate that the EGFP-RelA is expressed at nearly endogenous levels and does not detectably affect the kinetics of RelA translocation or expression of downstream NF-κB-dependent genes.

Identification of a Cytoplasmic NF-κB Reservoir—The qualitative Western immunoblots and the confocal microscopy suggested that RelA remained cytoplasmic, even during peak NF-κB translocation in response to a saturating TNFα dose (30 ng/ml). To confirm this observation, we fractionated resting and TNFα-stimulated cells into cytoplasmic and nuclear extracts and measured changes in RelA abundance using quantitative Western immunoblot, where serial dilutions of purified recombinant RelA were co-fractionated for use as a standard curve (Fig. 2A). From this analysis, we estimated that A549 WT cells express ~288,000 molecules of endogenous RelA per cell, and nearly 23,000 translocate, a number constituting ~10% of the total cytoplasmic molecules available. These data indicate that a significant fraction of RelA remains cytoplasmic, a finding confirmed using an independent aptamer pull-down technique to quantitatively analyze RelA in these same cells (39).

Cytoplasmic RelA is complexed in various amounts with 1xkBα, β-γ-ε, and -t and Bcl-3 proteins in a cell type-specific fashion (5, 6). To determine the relative distribution of RelA associated with these 1xkB isoforms in A549 cells, cytoplasmic
extracts from unstimulated A549 WT cells were immunoprecipitated with IκBα, -β, -γ, -ε, and -ι and Bcl-3 antibodies, and associated RelA was determined by quantitative Western immunoblot (Fig. 2B). Here, the majority of RelA associates with IκBα in a 4:1.5:1 ratio with IκBβ and IκBε, with no detectable association with IκBγ, IκBε, or Bcl-3 (Fig. 2B).

To determine whether this IκBα/β/ε-associated RelA represented the majority of the cytoplasmic RelA pool, we conducted a depleting immunoprecipitation experiment. Cytoplasmic extracts from unstimulated A549 cells were incubated individually with IκBα, -β, -γ, -ε, or -ι or Bcl-3 Abs, and RelA-associated complexes were removed by binding protein A-agarose beads. The remaining, non-IκB isofrom-complexed RelA was measured in the supernatant by quantitative Western blot (Fig. 2C). Although IκBα was completely depleted from the cytoplasmic extract (Fig. 2C, middle), the majority of RelA remained in the supernatant, again confirming that a minority of the total cytoplasmic RelA is complexed with IκBα, -β, and -ε.

Finally, a sequential immunodepletion experiment was conducted to study the association of RelA with NF-κB1/p50, NF-κB2/p52, and IκB(α and β) proteins (a schematic representation is shown in Fig. 2D). Quantitation of the images indicates that 48% of RelA is associated with NF-κB1/p50 and 15% is associated with NF-κB2/p52 (with the remainder of RelA associating with IκB isofroms; Fig. 2E). Because NF-κB1 and -2 are not proteolyzed by the TNF-induced IκKB pathway (12, 40), these complexes largely account for the inert cytoplasmic sequestration of RelA. Together, we interpret these data to indicate that a significant fraction of RelA in A549 cells is not associated with IκB isofroms and therefore is not translocation-competent in response to TNFα.

As an independent estimation of the fraction of NF-κB translocation, we performed serial axial images of control and TNFα-stimulated EGFP-RelA-expressing cells. Integrated fluorescence intensity was summed for the entire cell, and the fraction that was in the nuclear compartment was expressed as a fraction of the total fluorescence intensity. These data also indicated that 8–14% of the total EGFP-RelA signal translocated into the nucleus after stimulation (Fig. 2F), a finding that confirmed the cellular fractionation experiments (Fig. 2A).

Measurement of Cytoplasmic/Nuclear Volume Ratio—In mathematical representations of the NF-κB pathway, the cellular geometry parameter γ, the ratio of cytoplasmic to nuclear...
volume, plays an important role in pathway dynamics, determining the nuclear concentration of translocated RelA (19, 27). Because this parameter has not been experimentally determined, we performed volume measurements using confocal microscopy. In these measurements, control and TNF-stimulated EGFP-RelA cells were fixed and imaged. Multiple 0.7-μm-thick sequential images were taken along the vertical (z) dimension with an optical slice separation of 0.3 μm. These slices were assembled, and $k_v$ values were computed. These data show that the $k_v$ values vary from cell to cell, ranging from 2 to 4.5 with a mean of 3. The fitted lognormal prior distribution with this experimental data with 90% confidence interval is shown in Fig. S1. Moreover, $k_v$ is not significantly affected by TNFα treatment (see Table 1).
**Cellular Variability in NF-κB Signaling**

**Model Formulation**—Both confocal and Western blot experiments indicated the existence of a reservoir of non-IκB-associated RelA that does not translocate to the nucleus. Based on these convergent experimental results, we modified a two-level negative feedback deterministic model to include “active” and “inactive” NF-κB forms (Fig. 3). The inactive form is assumed to be sequestered in a reservoir, with NF-κB1 or -2, and is assumed to be in slow exchange with the active translocation-competent RelA, bound by IκB complexes. The rate of exchange between these two NF-κB reservoirs are assumed to be very small on the time scales of the TNFα stimulation experiments and were therefore set to zero. The inactive reservoir and active NF-κB together account for the total cellular concentration of NF-κB measured by the dynamic imaging experiments.

**TABLE 1**

| Stimulation   | k
|----------------|-----------------|
| C-1           | 4.278           |
| C-2           | 4.192           |
| C-3           | 2.265           |
| C-4           | 3.609           |
| C-5           | 2.433           |
| TNF-1         | 1.094           |
| TNF-2         | 1.309           |
| TNF-3         | 3.082           |
| TNF-4         | 3.394           |
| TNF-5         | 2.213           |
| TNF-6         | 2.515           |
| TNF-7         | 2.815           |

**FIGURE 3. Schematic model of NF-κB activation network.** While in an inactivated state, NF-κB is located in the cytosol complexed with the inhibitory protein IκB proteins, represented here as the single IκBα species. Upon TNFα stimulation, IKK is activated, which catalyzes degradation of IκBα, releasing NF-κB. The activated free NF-κB is translocated into the nucleus, where it induces expression of downstream genes, such as A20 and IκBα (boldface lines, c1 and c4). These two proteins act as negative feedback loops to turn off the NF-κB signaling (dotted lines and translation reactions, c2 and c3). Experimentally shown in this study, a large amount of NF-κB does not translocate to the nucleus upon stimulation and acts as a reservoir of NF-κB. The exchange between the reservoir and translocable NF-κB is shown with dashed lines; however, their exact mechanism and rate kinetics (P1, P2) are yet to be determined and are assumed to be relatively slow relative to the overall NF-κB translocation kinetics. The total concentration of NF-κB is determined by quantitative Western blots in this study to be 0.12 μM. The initial translation-competent fraction of NF-κB is represented by NF-κB/IκBα complex. See **supplemental material** for specific parameter values.

EGFP-RelA Cells Show Monophasic Translocation Profiles—EGFP-RelA transfectants were imaged using dynamic confocal microscopy. Sequential images were first obtained to establish basal conditions. The cells were then stimulated with a saturating dose of TNFα (30 ng/ml), a dose that produces maximal effects on RelA translocation and target gene expression. Serial confocal images were taken every 6 min, and EGFP fluorescence was quantified and expressed as a fractional translocation. Inspection of the time series for images up to 6 h shows that the RelA translocation profile is monophasic (not shown), a finding consistent with the nuclear translocation of endogenous RelA in our previous studies (11). The profile for the first 72 min of stimulation is shown in Fig. 4.

Further analysis of the RelA translocation response clearly indicated that the time until peak nuclear translocation was highly variable, with 25% cells showing a peak translocation between 41 and 50 min after stimulation. The basal and peak NF-κB values differed between cells. In addition, we noted distinct profiles of decay of the monophasic peak, examined below. Together these data indicate the presence of cellular variability in the canonical RelA translocation profile in its non-oscillatory activation mode.

**Refinement of Initial Concentrations of NF-κB as the IκBαNF-κB Complex**—Initial ODE simulations using previous model formulations assuming that the total NF-κB was translation-competent did not agree with the basal levels of nuclear NF-κB for the cells under study. These higher initial
concentrations resulted in oscillatory behavior more frequently than did low initial concentrations for the same cell geometry (data not shown). This observation suggested an overestimation of the initial NF-κB concentration in the translocation-competent form, represented by the IκBα-NF-κB complex. Therefore, this initial NF-κB concentration was inferred in our modified model along with the total NF-κB concentration (including the reservoir NF-κB). Prior information on these concentrations was obtained from Fig. 2A (also see lognormal prior distribution, Fig. S2), where the average total NF-κB concentration in the cell population was calculated as 0.12 μM and the active NF-κB in the form of IκBα-NF-κB complex was estimated to be 0.035 μM. We note that this concentration is nearly half the value of the concentration used in the Lipniacki formulation (0.06 μM) (19) or nearly one-third of that used in the Hoffmann model (0.1 μM) (18).

**Parameter Sensitivity Analysis**—Because the dynamic imaging experiments generate only 15 data points per cell, inferring all 30 parameters in our ODE model from the experimental data for each cell would lead to an ill conditioned inverse problem. Instead, only a subset of the model parameters needs to be inferred. To determine which parameters to infer, a sensitivity analysis was performed to assess the relative influence of parameters affecting the monophosphoryl nuclear NF-κB response. To characterize this response quantitatively, we chose four characteristic output quantities: the stationary value for the characteristic output quantities, the timing of the maximum value, the rate of IκBα degradation, and the rate of IκBα-inducible mRNA synthesis (see supplemental material for a discussion of the sensitivity analysis method used). Besides the three parameters describing the initial state and geometry (κ_a, T_c, and initial concentration of NF-κB), the rate of degradation of the bound form of the IκBα protein in the IκBα-NF-κB complex (c_{d_1}), the rate of IκBα protein production (c_{p_1}), the IκBα mRNA degradation rate (c_{g_1}), and the rate of IκBα-inducible mRNA synthesis (c_{s_1}) showed the strongest influence on the dynamic behavior of the system. Given the strongly correlated effect of

![Graph of NF-κB translocation profiles of individual cells.](image)

**FIGURE 4.** NF-κB translocation profiles of individual cells. Shown is the nuclear/total ratio of fluorescence intensity for EGFP-RelA cells sequentially imaged for indicated times after TNFα stimulation. Each line represents a single cell tracing. Time-point 0 represents when TNFα stimulation was done.

![Graph of output quantities of interest used for sensitivity analysis.](image)

**FIGURE 5.** The output quantities of interest used for sensitivity analysis. A, all quantities are related to the nuclear/total NF-κB ratio values; Z_i is the stationary value of the unstimulated period, Z_2 is the time of the peak, Z_3 is the peak value itself, and Z_4 is the final value. B, an overall sensitivity measure S_k = \|\alpha_i\| + \|\alpha_j\| + \|\alpha_k\| + \|\alpha_l\| is plotted, where \alpha_i is the sensitivity of the kth output quantity of interest with respect to the ith input parameter, where k = 1, 2, 3, and 4, and i = 1, 2, ..., 30.

---

**Single Cell Bayesian Inference of Model Parameters for 20 Individual Cells**—Parameter inference was performed with the experimental noise characteristics s_p and s_q for all 20 measured NF-κB translocation profiles individually. In the forward model simulations, the initial NF-κB-IκBα concentration was set to its corresponding parameter value, and the amount of NF-κB in the inactive reservoir was computed to satisfy the parameter T_c. Initial concentrations for other NF-κB-containing molecular species were set to 0, and concentrations for other molecular species were set to an appropriate initial guess. The simulation was then run for 20 h in an unstimulated regime in order to obtain a steady state that mimics the experimental configuration prior to stimulation. After this initial equilibration, stimulation with TNFα was switched on in the model, and the simulation was continued for the duration of the experiments. Fig. 6A shows the experimental data and the ODE model prediction...
with the maximum a posteriori (MAP) inferred parameter values for one sample cell, cell 8. The figure also shows the model prediction based on the mean values of the parameters, together with the associated posterior predictive uncertainty that incorporates the experimental error with the uncertainty associated with the finite number of data (36). The corresponding MCMC chains for six inferred model parameters are shown (Fig. 6B).
To quantify the uncertainty in and the correlations between the inferred parameters, we analyzed the joint posterior distributions of the six inferred parameters for the given cells (Fig. 6C). All of the conclusions below hold in general for all cells, although they are only illustrated on the single, eighth cell. The posterior distribution summarizes the current state of knowledge about all of the inferred parameters based on the available data. These posteriors were estimated from the last 40,000 samples of the corresponding MCMC chain (Fig. 6A) to ensure that the samples were drawn from a chain that had reached a stationary state.

The posterior distributions incorporate the information from the prior distributions and the likelihood function. In the absence of any quantitative information on rate constants, the prior distributions typically provide only reasonable physical bounds on the values of these parameters. The likelihood function, on the other hand, is based on fitting the model to the experimental data; hence, generally the posterior values are driven by the likelihood function. This explains, for example, the strong positive posterior correlations between the total NF-κB concentration $T_{c}$ and initial IkBa-NF-κB concentration (Init) (Fig. 6C). Indeed, as the total concentration increases, in order for the model to fit the data comparably well, the model should use higher initial concentration because the data are based on ratios of concentrations only. Similarly, we notice high positive posterior correlations between $c_{4a}$ and $c_{6a}$, reflecting the fact that IkBa mRNA translation and IkBa degradation from the IkBa-NF-κB complex have opposite effects on the nuclear/total NF-κB ratio (Fig. 6C). Increasing $c_{4a}$ tends to increase the amount of translocated NF-κB by accelerating the translocation, whereas increasing $c_{6a}$ reduces the amount of translocation of NF-κB by accelerating the negative feedback through IkBa, discussed below.

The single cell analysis produced good fits with the experimental data based on 80,000 MCMC steps for each of the 20 cells (Fig. 6D). The shaded region corresponding to the posterior predictive uncertainty on the inferred model predictions for each of the 20 cells suggests that the model has been able to capture most of the inherent variation of the experimental data. The MAP estimates for the inferred parameters are shown in Table 2 and Fig. 6E together with the corresponding mean estimates and the associated error bars. The greatest variation in the MAP values for the six inferred parameters across these 20 cells is seen for the $c_{4a}$ and $c_{6a}$ rate constants. Note that the inferred MAP values for $k_{v}$ and $T_{c}$ are very similar for all cells and are close to the maxima of their corresponding prior distributions. This indicates that the inference is driven mainly by the initial information (i.e. the priors for these parameters). Namely, the MCMC chain tends to move the MAP estimate toward the maximum of the prior distribution if the corresponding parameter does not have much effect on the likelihood function. Obviously, although the ODE model is able to

\[ \text{TABLE 2} \]

MAP values of the eight inferred parameters in the single cell Bayesian inference based on translocation data for each of the 20 cells (see also Fig. 6E).

| Cell | $c_{4a}$ | $c_{6a}$ | $t_s$ | $k_v$ | Init | $T_{c}$ | $s_p$ | $s_q$ |
|------|---------|---------|-------|-------|------|--------|------|------|
| 1    | 0.533   | 3.844 $\times 10^{-4}$ | 6.968 $\times 10^{-4}$ | 2.781 | 0.050 | 0.111 | 0.011 | 0.021 |
| 2    | 0.221   | 1.680 $\times 10^{-4}$ | 1.354 $\times 10^{-4}$ | 2.924 | 0.044 | 0.109 | 0.016 | 0.016 |
| 3    | 0.941   | 1.235 $\times 10^{-3}$ | 2.970 $\times 10^{-3}$ | 2.663 | 0.059 | 0.114 | 0.005 | 0.075 |
| 4    | 0.230   | 2.456 $\times 10^{-4}$ | 9.997 $\times 10^{-4}$ | 3.099 | 0.038 | 0.112 | 0.010 | 0.041 |
| 5    | 0.153   | 1.063 $\times 10^{-4}$ | 8.128 $\times 10^{-4}$ | 2.339 | 0.037 | 0.116 | 0.005 | 0.029 |
| 6    | 0.086   | 1.635 $\times 10^{-3}$ | 1.652 $\times 10^{-3}$ | 2.638 | 0.052 | 0.113 | 0.015 | 0.056 |
| 7    | 0.771   | 4.370 $\times 10^{-4}$ | 7.308 $\times 10^{-4}$ | 2.228 | 0.034 | 0.108 | 0.008 | 0.023 |
| 8    | 0.967   | 9.301 $\times 10^{-4}$ | 1.227 $\times 10^{-3}$ | 1.987 | 0.043 | 0.111 | 0.008 | 0.025 |
| 9    | 0.133   | 3.575 $\times 10^{-4}$ | 4.578 $\times 10^{-4}$ | 3.039 | 0.072 | 0.107 | 0.013 | 0.044 |
| 10   | 0.252   | 4.077 $\times 10^{-4}$ | 9.137 $\times 10^{-4}$ | 2.523 | 0.040 | 0.105 | 0.005 | 0.022 |
| 11   | 0.257   | 2.122 $\times 10^{-3}$ | 1.389 $\times 10^{-3}$ | 2.754 | 0.042 | 0.113 | 0.014 | 0.039 |
| 12   | 0.010   | 3.088 $\times 10^{-4}$ | 1.725 $\times 10^{-4}$ | 2.936 | 0.021 | 0.111 | 0.005 | 0.021 |
| 13   | 0.826   | 1.065 $\times 10^{-4}$ | 6.741 $\times 10^{-4}$ | 2.567 | 0.046 | 0.120 | 0.008 | 0.030 |
| 14   | 0.493   | 3.637 $\times 10^{-4}$ | 1.347 $\times 10^{-3}$ | 2.573 | 0.030 | 0.105 | 0.011 | 0.021 |
| 15   | 0.004   | 2.952 $\times 10^{-4}$ | 1.224 $\times 10^{-3}$ | 2.988 | 0.032 | 0.105 | 0.004 | 0.032 |
| 16   | 0.255   | 1.491 $\times 10^{-4}$ | 8.544 $\times 10^{-4}$ | 2.800 | 0.041 | 0.117 | 0.013 | 0.017 |
| 17   | 1.026   | 1.285 $\times 10^{-4}$ | 2.962 $\times 10^{-4}$ | 2.902 | 0.060 | 0.110 | 0.004 | 0.020 |
| 18   | 0.697   | 5.414 $\times 10^{-4}$ | 3.990 $\times 10^{-4}$ | 2.875 | 0.060 | 0.106 | 0.010 | 0.013 |
| 19   | 0.133   | 1.175 $\times 10^{-3}$ | 8.733 $\times 10^{-4}$ | 2.705 | 0.089 | 0.126 | 0.013 | 0.008 |
| 20   | 0.228   | 7.646 $\times 10^{-4}$ | 1.538 $\times 10^{-4}$ | 3.263 | 0.036 | 0.116 | 0.007 | 0.022 |

FIGURE 6. Fit of inferred model with respect to experimental data and MCMC chain. A, MAP fit (red line) and the posterior mean fit (blue line) of the inferred model with the experimental data (black dots) for the first 72 min of TNFα stimulation for cell 8. The shaded region corresponds to the posterior predictive uncertainty that incorporates the measurement noise and the lack of knowledge (i.e. the uncertainty associated with the fact that only limited data are available). The data points until time $t = 0$ min represent the steady state of the system, where there is no TNFα stimulation. The arrow at $t = 0$ marks the TNFα stimulation time point. B, the corresponding MCMC chain for the logarithms of six model parameters used in the inference run. Only every 100th chain state is shown for clarity. C, marginal and joint posterior PDFs of the six inferred parameters for cell 8. The marginal posteriors for each individual parameter are shown on the diagonal, whereas the contour plots below the diagonal show the marginalized joint posterior distributions between the corresponding parameters. The posteriors for the individual parameters are unimodal with well defined maxima. The joint posteriors show a positive correlation between Init and $T_{c}$ and between $c_{4a}$ and $c_{6a}$. In order to properly sample the stationary state of the MCMC chain, a total of $n = 40,000$ samples from the end of the chain were retained to estimate the posterior distributions. D, model fit for 20 individual cell translocation profiles. Bayesian inference was used against translocation data of each of the 20 cells individually. The red lines represent the model fit corresponding to the MAP values of the parameters shown in Table 2. The black dots correspond to the experimental time course data. The shaded regions correspond to the posterior predictive uncertainties. Cell numbers are indicated. E, all eight inferred parameters are shown. The black bars show the mean values of the parameters together with the S.D. (error bars) based on the last 40,000 steps of the MCMC. The white bars correspond to the MAP estimates.
reproduce the translocation dynamics observed in each of the 20 cells, some of the inferred parameters may not be biologically realistic. One possible reason for obtaining unrealistic parameters in the inference is the limited amount of experimental data available to constrain the inferred parameters. The ratio of experimental data versus the number of parameters to be inferred can be improved by recognizing that some parameters, such as reaction rate constants, can be expected to be the same in all cells. We therefore performed an inference of the ODE model parameters on all 20 cells simultaneously.

**Multicell Bayesian Inference of Model Parameters Using Data from 20 Cells Simultaneously**—For the inference of parameters in a single ODE model based on multiple-cell NF-κB translocation profiles simultaneously, the reaction rate constants \( c_{a3}, c_{6a3}, \) and \( i_v \) were assumed to be the same among all 20 cells, whereas each cell was allowed its own, cell-specific values for the volume ratio \( k_a \), the initial NF-κB-IκBα concentration, and the total NF-κB concentration \( T_c \). We believe these assumptions are justified because the biochemical processes governing the IκBα translation \( (c_{a3}) \), the IκBα proteolysis \( (c_{6a3}) \), and NF-κB nuclear import rate \( (i_v) \), are unlikely to change from cell to cell, whereas our experimental data indicate that the NF-κB concentrations and cell volumes do vary across the cell population. Thus, with a total of 20 cells, there were 63 (3 \( \times \) 20 + 3) parameters to infer. Note that although this presents a higher dimensional inference problem than in the previous case of single cell model inferences, it is better constrained because in the single cell inferences, a total of 120 (6 \( \times \) 20) parameters had to be inferred against the same amount of experimental data. As shown in Fig. 7A, predictions with the inferred parameters for this multiple cell model show good agreement with the respective experimental data for 16 of the 20 cells. Also, it is for these 16 cells only that the multicell inference leads to fits that are similar to the inferred single cell model predictions for the corresponding cells (Fig. 6D). Predictions with the globally inferred model parameters for cells 8, 13, 14, and 20 deviated from the experimental data considerably. For these cells, experimental data suggest a much faster rate of nuclear-to-total NF-κB decay than indicated by the inferred model. Faster or slower decaying signals compared with the regular kinetics are reflected in the inferred values of \( c_{a3} \) and \( c_{6a3} \). The reaction rate \( c_{a3} \) implies faster appearance of IκBα in the system due to translation, thereby inactivating the active NF-κB signal, which can be seen as quick decay from the translocation peak in four cells (cell 8, 13, 14, and 20) compared with others.

The MAP estimates for the parameters inferred with this multicell analysis are listed in Table 3 and Fig. 7B, together with the mean estimates and the corresponding error bars. Comparing Fig. 6D and Fig. 7A, it is clear that the inferred parameters show a lot more cell-to-cell variability in the multicell inference than in the case where each cell’s dynamics were fit individually. Also, smaller error bars for the multicell case are explained by the fact that the multicell inference is better constrained than the single cell inferences (i.e. it has more data available per inferred parameter or, in other words, fewer degrees of freedom to fit the data).

The greater cell-to-cell variation is expected for the inferred parameters because forcing the reaction rates to be shared between all cells now only leaves three parameters instead of six to account for cell-to-cell variability in the observed translocation dynamics. In this context, it is illustrative to study the role of the prior distribution in the Bayesian inference. As shown in Fig. 6E (panel corresponding to \( k_a \)), the inferred values for \( k_a \) in the single cell inference are between 2 and 3 for all cells (i.e. very close to the maximum of the prior distribution for \( k_a \) (supplemental material). This means that in this case, the prior performs a regularizing role in the inference. Because there are not many data to constrain the parameters, the inference problem is borderline underdetermined, meaning that there is not enough information in the data to properly infer all parameters. The Bayesian inference in this case tends to select values for \( k_v \) that follow the prior distribution (i.e. prior information is used to constrain the parameter values to make up for the lack of information in the observational data). For the multicell inference (Fig. 7B), most of the inferred values for \( k_v \) appear near the edges of the prior distribution for \( k_v \). In this case, there is a higher ratio of data points versus parameters to be inferred, and the posterior distribution is determined more by the likelihood function than by the prior. In order to fit cell-to-cell variations in the observational data, the inferred parameter values also vary from cell to cell. The prior distribution in this case performs more of a bounding role, keeping the inferred parameters from moving outside of the range of feasible values observed in the z-stacking experiments that the prior for \( k_v \) was derived from. The fact that almost all inferred \( k_v \) values are near the edges of the \( k_v \) prior distribution suggests that biologically unrealistic parameter \( k_v \) values would be needed in order to generate a better fit with the observed data, which is an indication that the inferred ODE model may not be the most suitable for the full set of 20 cells or that other parameters, besides the ones inferred in this analysis, also need to be retuned.

**Inference of Model Parameters Considering Two Subpopulations**—The above multicell inference analysis suggests that there are two subpopulations in the cells studied here. Notably, four cells (8, 13, 14, and 20) showed fast decay in the NF-κB translocation signal, with a clear inflection point in the decay signal, whereas the rest of the cells exhibit relatively slower decay dynamics. To explore this possibility, the multicell Bayesian inference was repeated, but this time for the fast decay subpopulation (cells 8, 13, 14, and 20) and the other, slow decay, subpopulation separately. The results, summarized in Fig. 8A, show much better fits with the data for all cells in the subpopulation-based analysis than in the full 20-cell multicell inference. Their mean estimates are presented in Fig. 8B.

Our experimental data indicated the existence of a reservoir of translocation-competent NF-κB that represented only a fraction of the total NF-κB. We therefore calculated the fraction of the translocation-competent NF-κB given as the abundance of the NF-κB-IκBα complex for each of the 20 individual cells based on the parameter estimates. These distributions are plotted in Fig. 9 and agree well with our experimental observations (cf. Figs. 1A and 2A and F).

**Simulations of c4a Variations**—Our analysis of the multicellular parameters suggests that dynamics in the regulator feedback loop, such as cellular variations in IκBα translation rate \( (c_{a3}) \), can affect the decay rate of the monophasic NF-κB profile.
To illustrate this prediction, simulations using variations in $c_{4a}$ were performed (Fig. 10). With increasing IκBα translation rate, more IκBα protein is synthesized per unit time than nominally. The newly synthesized IκBα saturates free NF-κB and traps it into the IκBα-NF-κB complex, reducing free NF-κB available for nuclear translocation. As a result, the initial peak of nuclear NF-κB is reduced, and despite the presence of a tonic TNF signal, so much IκBα is produced that the secondary NF-κB response can no longer be seen. Conversely, as $c_{4a}$ is reduced relative to nominal values, more free NF-κB is available to translocate, and less IκBα is present, permitting a secondary translocation response (gray to orange to blue curves). These

FIGURE 7. Multicell Bayesian inference (20 cells). A, model fit for 20 individual cell translocation profiles. Bayesian inference was used against translocation data of each of the 20 cells simultaneously. The red lines represent the model fit corresponding to the MAP values of the parameters shown in Table 3. The black dots correspond to the experimental time course data. Cell numbers are indicated. Single cell-independent inference model fits (red lines) and multicell simultaneous inference (blue lines) model fits are compared. The multicell inference results in cells 8, 13, 14 and 20 (marked with blue frames) deviate from those of the single cell because they fail to capture the fast decay dynamics. This motivates the division of the cells into two subpopulations and separate multicell inference in each of the subpopulation by itself. The black dots represent experimental data. B, all six inferred parameters are shown. The black bars show the mean values of the parameters together with the S.D. (error bars) based on the last 40,000 steps of the MCMC. The white bars correspond to the MAP estimates. Note that the rate constants $c_{4a}, c_{6a}$, and $i_1$ are shared for all cells.
**Cellular Variability in NF-κB Signaling**

**TABLE 3**

| Inferred values of six parameters used in multicell Bayesian inference of 20 cells together |
|-----------------------------------------|
| Rate constants are shared for all 20 cells with MAP values: $c_{m}=0.119 \text{s}^{-1}$, $c_{m}=1.781 \times 10^{-3} \text{s}^{-1}$, $\delta_{i}=1.149 \times 10^{-3} \text{s}^{-1}$. For more illustrative bar plots, see Fig. 7B. |
| Cell | $k_{v}$ | $T_{c}$ | Init |
|------|--------|--------|------|
|      | µM    | µM    |      |
| 1    | 4.319 | 0.165 | 0.054 |
| 2    | 4.317 | 0.085 | 0.032 |
| 3    | 2.846 | 0.157 | 0.042 |
| 4    | 4.020 | 0.092 | 0.027 |
| 5    | 0.863 | 0.099 | 0.039 |
| 6    | 2.251 | 0.075 | 0.032 |
| 7    | 2.511 | 0.172 | 0.040 |
| 8    | 3.881 | 0.191 | 0.051 |
| 9    | 1.575 | 0.098 | 0.063 |
| 10   | 1.739 | 0.167 | 0.061 |
| 11   | 3.609 | 0.131 | 0.045 |
| 12   | 2.788 | 0.081 | 0.018 |
| 13   | 1.849 | 0.134 | 0.035 |
| 14   | 2.073 | 0.131 | 0.032 |
| 15   | 1.259 | 0.076 | 0.034 |
| 16   | 5.284 | 0.089 | 0.026 |
| 17   | 0.986 | 0.160 | 0.056 |
| 18   | 3.162 | 0.087 | 0.029 |
| 19   | 3.847 | 0.058 | 0.035 |
| 20   | 2.687 | 0.075 | 0.023 |

data suggest that differences between these subpopulations can be attributed to the dynamics of the regulatory feedback loops and are a good objective for further studies.

**DISCUSSION**

Despite the theoretical treatment of the NF-κB signaling pathway, there are few data available for a rigorous inference of the model parameters or to account for major sources of cellular variability in NF-κB pathway dynamics. In this study, we applied a modified model of the canonical NF-κB pathway that accounts for a distinct subpopulation of translocation-competent NF-κB. We applied Bayesian inference to estimate important parameters using time-resolved NF-κB translocation measurements at a single cell level. Bayesian inference is well suited for this approach, being able to tolerate large noise levels and to provide a joint probability distribution for correlated parameters. Experimental measurements of cytoplasmic and nuclear volumes and fractional NF-κB translocation were used to inform the prior distribution of the cellular geometry parameter $k_{v}$ and the translocation-competent NF-κB-IκBα complex. This Bayesian inference allowed a better understanding of the single cell dynamics and insight into the major sources of cellular variability in NF-κB signaling.

There are several new insights generated from our study. First, we note that the total abundance of RelA varies considerably from cell to cell, a finding that may contribute to significant cell-to-cell variation in the activity profile of the pathway. This is observed both in our stable EGFP-RelA transfectants and for endogenous RelA in non-transfected cells, indicating external random effects on the NF-κB pathway in each cell. This point is important because models produced based on population-averaged experimental data will not accurately capture the variation of individual cellular responses.

Another significant finding from this study is that a large proportion of cytoplasmic NF-κB is not induced by a saturating concentration of TNFα to translocate into the nucleus. This result is quite surprising because TNFα is one of the most potent NF-κB activators currently known (25). We provide data that indicate the inactive complex is due to association with the NF-κB1 and -2 precursors, cytoplasmic ankyrin repeat-containing proteins that sequester a greater molar amount of RelA than that associated with individual IκB isoforms. As a result of these experimental measurements, we modified the deterministic model to include two forms of NF-κB: one as an inactive reservoir in the cytoplasm and the second as a translocation-competent reservoir as an IκBα:NF-κB complex. Our inclusion of NF-κB reservoir in the model resulted in dramatically better fitting of the model with experimental single cell data. More work will be required to understand how the reservoir exchanges with the translocation-competent IκBα:NF-κB pool and how stimuli that induce NF-κB1/2 processing affect the dynamics of the pathway.

Although it has been generally assumed that the cellular geometry affects the dynamic NF-κB response, this critical parameter has been assumed in previous modeling studies. $k_{v}$ determines the concentration of nuclear NF-κB as it undergoes nuclear import, with the initial NF-κB concentration determining the kinetics of activation of the negative feedback loops. Our experimental measurements indicate that $k_{v}$ in A549 cells is distributed between 2 and 3, values that are lower than previous modeling assumptions. Also, $k_{v}$ does not dramatically change as a result of pathway activation.

Bayesian inference was then applied to this refined NF-κB signaling pathway to estimate parameters using the single cell imaging profiles. The choice of the six inferred parameters in the Bayesian methodology was dictated by a global sensitivity analysis of all 30 input parameters around the nominal values. It is important to note that the inferred concentrations of active NF-κB (i.e. in the form of the IκBα:NF-κB complex) are generally about half the concentration used in previous models. Incorporating prior measurement data on $k_{v}$ and the total NF-κB concentrations in the inference helped to constrain the inferred parameters to biologically realistic values and made the inference process better conditioned overall. The probabilistic nature of the Bayesian inference allowed us to develop the model on a single cell basis and assess the uncertainty in the inferred parameters.

Using inferred values of $k_{v}$ and the total NF-κB concentration ($T_{c}$) for all 20 studied cells, we estimated the fraction of reservoir inactive NF-κB concentration in these single cells. Fig. 9 clearly shows a high fraction of reservoir NF-κB in all cells as well as considerable cell-to-cell variability in this quantity. This may be one of the major sources of cell-to-cell variability because different cells have different amounts of total NF-κB and different fractions of initial active NF-κB, affecting the dynamics of NF-κB translocation and the downstream signaling initiation process.

Although the analysis above suggests several crucial parameters influencing NF-κB translocation dynamics that exhibit cell-to-cell variability, those parameters were not sufficient to capture the cell-to-cell variability observed in the present data set. A Bayesian inference of a single ODE model against data from all 20 cells simultaneously, assuming the rate constants are shared between all of the cells, whereas initial and geometric
properties of the cells are cell-specific, was unable to fit all cells. This multicell analysis clearly suggested the existence of two subpopulations of cells; four cells showed much faster decay in the nuclear/total NF-κB ratio dynamics than the rest of the cells. Separate multicell inference runs for these two subpopulations resulted in much improved model fits with the experimental data for all cells. With all other parameters statistically similar for the cells in both subpopulations, both $c_{4a}$ and $c_{6a}$ were significantly larger for the cells in the fast decay subpopulation than in the slow decay subpopulation.

The observed subpopulations therefore point at another mechanism of cell-to-cell variability. Besides the unlikely pos-

---

**FIGURE 8. Multicell Bayesian inference for subpopulations of 16 cells and four cells.**

A, model fit for 16 (slow decay subpopulation) and four (fast decay subpopulation, cells 8, 13, 14, and 20, marked by blue frames) individual cell translocation profiles. Within each subpopulation, Bayesian inference was used against translocation data of all of the cells of the subpopulation simultaneously. The red lines represent the model fit corresponding to the MAP values of the parameters. The black dots correspond to the experimental time course data. The shaded regions correspond to the posterior predictive uncertainties. Cell numbers are indicated. B, all six inferred parameters are shown. The black bars show the mean values of the parameters together with the S.D. value (error bars) based on the last 40,000 steps of the MCMC. The red bars correspond to the MAP estimates. Note that the rate constants $c_{4a}$, $c_{6a}$, and $i_1$ are shared for all cells of the corresponding subpopulation (fast decay cells or slow decay cells).
More focused experimentation, outside of the scope of the present study, will be required to identify and understand the source of the pathway deviation in the observed subpopulations.

In summary, this study reported and validated novel experimental observations indicating both active and inactive NF-κB in cells. The study also performed a novel Bayesian inference of ODE model parameters against single live cell translocation data, using data from experimentally measured cellular compartment volumes and NF-κB concentrations as prior information. We observed cell-to-cell differences in inferred values of compartment volumes, initial concentrations of NF-κB, and the translation rate of IκBα along with the rate of degradation of IκBα from the NF-κB/IκBα complex. All of these factors influence the overall NF-κB dynamics of cells and are sources of cell-to-cell variability in the observed NF-κB translocation dynamics.

REFERENCES
1. Knight, D. A., and Holgate, S. T. (2003) *Respirology* 8, 432–446
2. Tian, B., and Brasier, A. R. (2003) *Recent Prog. Horm. Res.* 58, 95–130
3. Brasier, A. R. (2006) *Cardiovasc. Toxicol.* 6, 111–130
4. Nowak, D. E., Tian, B., Jamaluddin, M., Boldogh, I., Vergara, L. A., Choudhary, S., and Brasier, A. R. (2008) *Mol. Cell. Biol.* 28, 3623–3638
5. Mercurio, F., DiDonato, J. A., Rosset, C., and Karin, M. (1993) *Genes Dev.* 7, 705–718
6. Beg, A. A., and Baldwin, A. S., Jr. (1993) *Genes Dev.* 7, 2064–2070
7. Poyet, J. L., Srinivasula, S. M., Lin, J. H., Fernandes-Alemni, T., Yamaoka, S., Tischis, P. N., and Alnemri, E. S. (2000) *J. Biol. Chem.* 275, 37966–37977
8. Mercurio, F., Zhu, H., Murray, B. W., Shevchenko, A., Bennett, B. L., Li, J., Young, D. B., Barbosa, M., Mann, M., Manning, A., and Rao, A. (1997) *Science* 278, 860–866
9. Spencer, E., Jiang, J., and Chen, Z. J. (1999) *Proc. Natl. Acad. Sci. U.S.A.* 96, 195–215
10. Tian, B., Nowak, D. E., Jamaluddin, M., Wang, S., and Brasier, A. R. (2005) *J. Biol. Chem.* 280, 17435–17448
11. Hoffmann, A., Levchenko, A., Garcia-Sastre, A., and Baltimore, D. (2002) *Science* 298, 1241–1245
12. Lipniacki, T., Paszek, P., Brasier, A. R., Luxon, B., and Kimmel, M. (2004) *J. Theor. Biol.* 228, 195–215
13. Wang, Y., Paszek, P., Horton, C. A., Kell, D. B., White, M. R., Broomhead, D. S., and Muldoon, M. R. (2011) *Mol. Syst. Biol.* 7, 258
14. Sun, S. C., Ganchi, P. A., Ballard, D. W., and Greene, W. C. (1993) *Science* 259, 1912–1915
15. Han, Y., and Brasier, A. R. (1997) *J. Biol. Chem.* 272, 9825–9832
16. Wertz, I. E., O'Rourke, K. M., Zhou, H., Eby, M., Aravind, L., Seshagiri, S., Wu, P., Wiesmann, C., Baker, R., Boone, D. L., Ma, A., Koonin, E. V., and Dixit, V. M. (2004) *Nature* 430, 694–699
17. Brasier, A. R. (2008) in *Cellular Signaling and Innate Immune Responses to RNA Virus Infections* (Brasier, A. R., Garcia-Sastre, A., and Lemon, S. M., eds) American Society for Microbiology, Washington, D. C.
18. Hoffmann, A., Levenson, A., Scott, M. L., and Baltimore, D. (2002) *Science* 298, 1241–1245
19. Carotti, F., Dower, S. K., and Qwarnstrom, E. (2000) *J. Biol. Chem.* 275, 41028–41034
20. Tian, B., and Brasier, A. R. (2003) *Recent Prog. Horm. Res.* 58, 95–130
21. Wang, Y., Paszek, P., Horton, C. A., Kell, D. B., White, M. R., Broomhead, D. S., and Muldoon, M. R. (2011) *Mol. Syst. Biol.* 7, 258
22. Mercurio, F., Zhu, H., Murray, B. W., Shevchenko, A., Bennett, B. L., Li, J., Young, D. B., Barbosa, M., Mann, M., Manning, A., and Rao, A. (1997) *Science* 278, 860–866
23. Krishna, S., Jensen, M. H., and Sneppen, K. (2006) *Proc. Natl. Acad. Sci. U.S.A.* 103, 10840–10845
24. Carlotti, F., Dower, S. K., and Qwarnstrom, E. (2000) *J. Biol. Chem.* 275, 41028–41034
25. Mercurio, F., DiDonato, J. A., Rosset, C., and Karin, M. (1993) *Genes Dev.* 7, 705–718
26. Beg, A. A., and Baldwin, A. S., Jr. (1993) *Genes Dev.* 7, 2064–2070
27. Poyet, J. L., Srinivasula, S. M., Lin, J. H., Fernandes-Alemni, T., Yamaoka, S., Tischis, P. N., and Alnemri, E. S. (2000) *J. Biol. Chem.* 275, 37966–37977
28. Mercurio, F., Zhu, H., Murray, B. W., Shevchenko, A., Bennett, B. L., Li, J., Young, D. B., Barbosa, M., Mann, M., Manning, A., and Rao, A. (1997) *Science* 278, 860–866
29. Spencer, E., Jiang, J., and Chen, Z. J. (1999) *Proc. Natl. Acad. Sci. U.S.A.* 96, 195–215
30. Tian, B., Nowak, D. E., Jamaluddin, M., Wang, S., and Brasier, A. R. (2005) *J. Biol. Chem.* 280, 17435–17448
31. Hoffmann, A., Levenson, A., Scott, M. L., and Baltimore, D. (2002) *Science* 298, 1241–1245
32. Lipniacki, T., Paszek, P., Brasier, A. R., Luxon, B., and Kimmel, M. (2004) *J. Theor. Biol.* 228, 195–215
33. Wang, Y., Paszek, P., Horton, C. A., Kell, D. B., White, M. R., Broomhead, D. S., and Muldoon, M. R. (2011) *Mol. Syst. Biol.* 7, 258
34. Sereny, I., O'Rourke, K. M., Zhou, H., Eby, M., Aravind, L., Seshagiri, S., Wu, P., Wiesmann, C., Baker, R., Boone, D. L., Ma, A., Koonin, E. V., and Dixit, V. M. (2004) *Nature* 430, 694–699
35. Brasier, A. R. (2008) in *Cellular Signaling and Innate Immune Responses to RNA Virus Infections* (Brasier, A. R., Garcia-Sastre, A., and Lemon, S. M., eds) American Society for Microbiology, Washington, D. C.
36. Hoffmann, A., Levenson, A., Scott, M. L., and Baltimore, D. (2002) *Science* 298, 1241–1245
37. Lipniacki, T., Paszek, P., Brasier, A. R., Luxon, B., and Kimmel, M. (2004) *J. Theor. Biol.* 228, 195–215
38. Cheong, R., Hoffmann, A., and Levenson, A. (2008) *Mol. Syst. Biol.* 4, 192
39. Wang, Y., Paszek, P., Horton, C. A., Kell, D. B., White, M. R., Broomhead, D. S., and Muldoon, M. R. (2011) *Mol. Syst. Biol.* 7, 23
40. Nelson, D. E., Ihekwaba, A. E., Elliott, M., Johnson, J., Rigney, C., Foreman, B. E., Nelson, G., See, V., Horton, C. A., Spiller, D. G., Edwards, S. W., McDowell, H. P., Unitt, J. F., Sullivan, E., Grimley, R., Benson, N., Broomhead, D., Kell, D. B., and White, M. R. H. (2004) *Science* 306, 704–708
41. Krishna, S., Jensen, M. H., and Sneppen, K. (2006) *Proc. Natl. Acad. Sci. U.S.A.* 103, 10840–10845
42. Carlotti, F., Dower, S. K., and Qwarnstrom, E. (2000) *J. Biol. Chem.* 275, 41028–41034
