Published in final edited form as:

Comput Biol Chem. 2010 February ; 34(1): 19–33. doi:10.1016/j.compbiolchem.2009.11.002.

Resolution and uniqueness of estimated parameters of a model of thin filament regulation in solution

Srboljub M. Mijailovich[a,*], **Xiaochuan Li**[a], **Juan C. del Álamo**[a,b], **R. Hugh Griffiths**[c], **Vojislav Kecman**[d], and **Michael A. Geeves**[c]

[a]Dept. of Environmental Health, Harvard School of Public Health, Boston, MA 02115, United States

[b]Dept. of Mech. and Aerospace Eng., University of California San Diego, La Jolla, CA 92093, United States

[c]Dept. of Biosciences, University of Kent, Canterbury, Kent CT2 7NJ, UK

[d]Dept. of Computer Science, Virginia Commonwealth University, Richmond, VA 23284, United States

Abstract

The estimation of chemical kinetic rate constants for any non-trivial model is complex due to the nonlinear effects of second order chemical reactions. We developed an algorithm to accomplish this goal based on the Damped Least Squares (DLS) inversion method and then tested the effectiveness of this method on the McKillop-Geeves (MG) model of thin filament regulation. The kinetics of MG model is defined by a set of nonlinear ordinary differential equations (ODEs) that predict the evolution of troponin-tropomyosin-actin and actin-myosin states. The values of the rate constants are estimated by integrating these ODEs numerically and fitting them to a series of stopped-flow pyrene fluorescence transients of myosin-S1 fragment binding to regulated actin in solution. The accuracy and robustness of the estimated rate constants are evaluated for DLS and two other methods, namely quasi-Newton (QN) and simulated annealing (SA). The comparison of these methods revealed that SA provides the best estimates of the model parameters because of its global optimization scheme. However it converges slowly and does quantify the uniqueness of the estimated parameters. On the other hand the QN method converges rapidly but only if the initial guess of the parameters is close to the optimum values, otherwise it diverges. Overall, the DLS method proves to be the most convenient method. It converges fast and was able to provide excellent estimates of kinetic parameters. Furthermore, DLS provides the model resolution matrix, which quantifies the interdependence of model parameters thereby evaluating the uniqueness of their estimated values. This property is essential for estimating of the dependence of the model parameters on experimental conditions (e.g. Ca^{2+} concentration) when it is assessed from noisy experimental data such as pyrene fluorescence from stopped-flow transients. The advantages of the DLS method observed in this study should be further examined in other physicochemical

*Corresponding author at: Molecular and Integrative Physiological Sciences, Department of Environmental Health, Bldg. I, Room 1310D, Harvard School of Public Health, 665 Huntington Avenue, Boston, MA 02115, United States. Tel.: +1 617 432 4814; fax: +1 617 432 4710. smijailo@hsph.harvard.edu (S.M. Mijailovich).

systems to firmly establish the observed effectiveness of DSL vs. the other parameter estimation methods.

Keywords

Damped Least Squares; Ca^{2+} sensitivity; Resolution matrix; Tropomyosin-troponin-actin; Stopped-flow

1. Introduction

The kinetics of a chemically reacting system is usually modelled using ordinary differential equations that are parameterized by a reduced set of reaction rate constants. Precise knowledge of these constants is required to characterize the dynamical behavior of the system. The estimation of the reaction rate constants from experimental data constitutes an inverse problem and is much more complicated than solving the model for a given set of rate constants. Inverse problems arising in chemical kinetics can be addressed by discrete inverse theory (Menke, 1989). The most common methods for parameter estimation on chemical kinetics are least squares methods that minimize an objective function iteratively by working with its gradient (Farinha et al., 1997; Lisy and Simon, 1998; Tadi and Yetter, 1998). In these methods the objective function is defined as the error between the experimental observations and the predictions of the model, which depends on the model parameters. Several other approaches have been developed to solve this difficult minimization problem. One of them is the so-called “trust region”, which uses successive quadratic programming (SQP), as proposed by Arora and Biegler (2004). Also in use are various stochastic search techniques, including simulated annealing (Eftaxias et al., 2002; Goffe et al., 1994) and genetic algorithms (Terry and Messina, 1998).

Ideally, the parameter set that minimizes the objective function is a manifold of dimension zero in which the minimum of error is a single distinguished point in the error landscape. In this case, the error increases as the parameter values vector moves away from this point in any arbitrary direction. However, in systems where multiple parameters need to be estimated, the optimal parameter set may span a manifold of dimensions equal or greater than one, forming “valleys” or “hyper-valleys” in the error landscape. As a result, the model output may be independent of changes in one parameter or a certain combination of parameters, which prevents the univocal estimation of these parameters. A sensitivity analysis is therefore necessary to explore the variations in the model output with perturbations in the parameters, which are defined by the sensitivity coefficients (Frank, 1978; Tomovic and Vukobratovic, 1972; Varma et al., 1999). In fact, the parameters of the model can be determined only if the sensitivity coefficients are non-zero and linearly independent (Beck and Arnold, 1977). More rigorously, Tang et al. (2005) used the “physically bounded Gauss-Newton” (PGN) method by utilizing the sensitivity matrix as a global map to determine the unknown kinetic parameters of complex reaction networks that contain dozens of species and hundreds of reactions.

The primary focus of this study is on the iterative Damped Least Squares (DLS) method which estimates kinetic parameters and provides a model resolution matrix (**RM**) (Menke,

1989). The DLS method, which is also known as the Levenberg-Marquardt method (Kecman, 2001; Nocedal and Wright, 2006), utilizes the sensitivity matrix as a search map in the same way as the PGN method, and provides information about the parameter uniqueness via the model **RM**. We applied the DLS method to determine the parameters of a kinetic model by fitting experimental data from the isotherm of the myosin-S1 binding to regulated F-actin in solution. The fidelity of the DLS method was evaluated by comparing its results with those obtained from two other algorithms: simulated annealing, and quasi-Newton. The predictions of the kinetics model are calculated by a probabilistic algorithm based on the McKillop-Geeves (MG) three-state model of thin filament regulation in solution (McKillop and Geeves, 1993). According to this model, the actin-associated regulatory protein complexes consisting of tropomyosin (Tm) and troponin (Tn) switch between the three azimuthal positions on the actin filament surface thereby preventing or allowing myosin-S1 to (weakly or strongly) bind to actin. The kinetics of the interconversion of TmTn-actin states depends on calcium concentration and represents the primary mechanism of regulation of contraction in striated muscles. The estimated key parameters of the MG model and their dependence on calcium by the three-parameter methods are compared, and their uniqueness and robustness is discussed. The rate of convergence of the estimated parameters and the effectiveness of different parameter estimation methods is also evaluated.

2. Methods

All experimental data contain experimental error and natural fluctuations. In order to better understand how and to what degree these uncertainties affect estimated model parameters, we applied a probabilistic algorithm for the simulation of the MG three-state model of thin filament regulation in solution (McKillop and Geeves, 1993) developed by Chen et al. (2001). We tested the robustness of DLS parameter estimation by comparing the model predictions with two other widely used parameter estimation methods, namely, the quasi-Newton and simulated annealing. The model predictions required by the three estimation methods were calculated using the probabilistic formulation of the MG model (Chen et al., 2001). This method was preferred because of its speed since many model calculations may be needed to reach convergence.

2.1. The McKillop-Geeves three-state model of thin filament regulation

In vertebrate skeletal and cardiac muscles the interaction between myosin and actin is regulated by the actin-associated proteins, tropomyosin (Tm) and troponin (Tn), depending on the concentration of calcium (Ca^{2+}). The soluble fragment of the myosin molecule S1 is widely used for studying kinetics of myosin binding to regulated actin filaments in solution. This fragment, also known as the motor domain, contains all of the ATPase and actin binding properties of the parent myosin. In the absence of nucleotide, myosin-S1 forms a tight (rigor-like) bond to actin filaments. McKillop and Geeves (1993) proposed that the regulation of tropomyosin and troponin-containing thin filaments can be interpreted using three-states of the actin filament: (1) the “blocked” state, in which myosin-S1 binding to actin is prohibited; (2) the “closed” state, in which S1 can bind with actin, but cannot be isomerized further to next step; (3) and in the “open” state, where no limitation to S1

binding to actin is imposed. In this study, the unit size of Tm-Tn complex is assumed to cover 7 actin monomers,[1] denoted as actin$_7$·TmTn (Maytum et al., 1999) (see Fig. 1). The repeat of TmTn every 7 units uniquely defines a TmTn unit that can rigidly move between the three states. Because Ca^{2+} binding to Tn significantly decreases the affinity of Tn to actin, the distribution between these three states is therefore affected by Ca^{2+} concentration. There are three myosin states in the model: one unbound state where myosin is in solution and two bound states where myosin is bound to actin. The bound states are denoted as weakly bound A-states and strongly bound, i.e. rigor-like states, as R-states (Fig. 1).

The MG model (Fig. 1) can be fully described by defining each state as combination of either, blocked, closed or open TmTn state, and the number actin sites with myosin bound in A- and R-stated within an actin$_7$·TmTn unit. The complete chemical kinetics of the MG model can be described by 45 states (see Fig. 2 for details) and 45 corresponding chemical kinetics equations, as previously described by Chen et al. (2001). These equations are

$$\begin{aligned}
\frac{dp_1(t)}{dt} &= -k_{+B}p_1 + k_{-B}p_2,\\
\frac{dp_2(t)}{dt} &= k_{+B}p_1 - (k_{-B} + 7k_{+1}c + k_{+T})p_2 + k_{-1}p_3 + k_{-T}p_{10},\\
&\vdots\\
\frac{dp_{45}(t)}{dt} &= k_{+2}p_{44} - 7k_{-2}p_{45},
\end{aligned} \tag{1}$$

where $p_i(t)$ is the fraction TmTn units in state i, which is defined by position of Tm (i.e. in block, closed or open state) and a particular combination of actin unoccupied sites and S1 bound in A- or R-states within the actin$_7$·TmTn unit. The equilibrium constants of the model are defined as $K_B = k_{+B}/k_{-B}$, $K_T = k_{+T}/k_{-T}$, $K_1 = k_{+1}/k_{-1}$, and $K_2 = k_{+2}/k_{-2}$, where k_{+B}, k_{+T}, k_{+1}, and k_{+2} are forward rate constants, and k_{-B}, k_{-T}, k_{-1}, and k_{-2} are backward rate constants. The system of Eq. (1) is nonlinear because the forward transition rate between the free actin state and myosin bound to actin in the Astate depends on the concentration of unbound S1. In Eq. (1) we denoted S1 concentration as c = [S1], thus the effective myosin binding rate constant, $k_{+1}c$, is in units s^{-1}. The concentration of S1 in solution, i.e. the concentration of unbound myosin, decreases as myosin binds to actin and, therefore, decreases the effective rate of myosin binding. The concentration of free myosin-S1 in solution is equal to the initial concentration of free myosin-S1, [$S1_O$], minus concentrations of myosin to bound to actin in either A- or in R-state. Thus the binding rate depends on probabilities of all myosin bound states, p_{bound}. The equations of the MG model are solved numerically by Gear's backward differentiation formulas (up to order five) (Hindmarsh, 1972). All 45 equations can be easily solved for a wide range of parameter combinations. However, in some extreme cases in which some of the parameters take much larger values than others, the set of Eq. (1) becomes numerically "stiff" and it is difficult to solve using

[1]Although there is evidence of longer range cooperativity in the A_7TmTn filament, the simpler approach will be used here focusing this study rather on parameter estimation methods than on full functional representation of the thin filament regulation by MG model.

standard methods. In those cases, Monte Carlo simulations should be used in order to solve the resulting stiff set of equations efficiently.

Once we determine the vector of 45 states, $\mathbf{p}(\mathbf{k}, t)$, where vector $\mathbf{k} = (k_1,\ldots k_n)$ represents an array of rate transition constants and other relevant model parameters, we calculate the fraction of actin sites in each one of the three actomyosin states by summing the occupancy of the actin sites over all TmTn units, and normalizing by the total number of actin sites occupied by S1 in the R-state when the system is in equilibrium. The model predictions are tested against measurements of the pyrene fluorescence intensity during stopped-flow experiments (Boussouf et al., 2007a,b; Boussouf and Geeves, 2007). A drop in pyrene fluorescence is proportional to myosin binding to actin in the R-state. Thus, the calculated instantaneous fractions of actin sites that are not occupied or which are in the A-state (i.e. which are not in the R-state), denoted $\mathbf{g}(\boldsymbol{\kappa}, t)$, can be compared with corresponding experimental data, $\mathbf{d}^{obs}(t)$, at the same instant. Here, the vector $\boldsymbol{\kappa} = (\kappa_1, \ldots \kappa_m)$ represents the set of m free model parameters that need to be estimated. Note that $\boldsymbol{\kappa}$ is usually a subset of the complete parameter set $\mathbf{k}$, since some of the parameters are prescribed by the experimental protocol, in which case $m \leq n$. For example, the prescribed concentration of actin, myosin and calcium are the same in both the model simulations and those used in the experiments. Also, m is reduced if some parameters, such as several rate or equilibrium state transition constants, vary a little over the course of multiple experiments and can be measured independently.

2.2. Parameter estimation methods

2.2.1. Damped Least Squares (DLS)—Damped Least Square (i.e. Levenberg–Marquardt) inversion is a widely used method for the estimation of model parameters that has two important features: quantitative evaluation of the uniqueness of the estimated parameters and good parameter resolution (Menke, 1989). The DLS method is based on the iterative minimization of the mean-square error of the model predictions with respect to experimental observations. In this study, we estimate the rate transition constants of the McKillop-Geeves model (Fig. 1) by minimizing the variance between predicted history of the fraction of actin sites unoccupied by bound myosin in R-state, $\mathbf{g}(\boldsymbol{\kappa}, t)$, and the same fraction deduced from a time course of (normalized) pyrene fluorescence intensity, $\mathbf{d}^{obs}(t)$. We exclusively fitted data recorded during stopped-flow experiments for various concentrations of Ca^{2+} in which actin concentration is in excess of myosin-S1 concentration (Boussouf et al., 2007a,b; Boussouf and Geeves, 2007). The measurements performed in these experiments covered a range of independent variables wide enough to uniquely resolve the set of model parameters for each Ca^{2+} concentration. These observations $\mathbf{d}^{obs}(t)$ are represented by a set on N values of the pyrene fluorescence sampled at discrete time points.

The model predictions deviate from the experimental observations by $\mathbf{e}(\boldsymbol{\kappa}, t) = \mathbf{d}^{obs}(t) - \mathbf{g}(\boldsymbol{\kappa}, t)$, yielding the mean-square-error

$$E(\boldsymbol{\kappa}) = \frac{\left\|\mathbf{e}\right\|_2^2}{N} = \frac{1}{N}\sum_{i=1}^{N}\left(g_i(\boldsymbol{\kappa}) - d_i^{obs}\right)^2, \quad (2)$$

which represents an integral measure of the "goodness" of the model prediction fit to the observations. Minimization of Eq. (2) provides the set of free parameters $\boldsymbol{\kappa}$ that best fit the experimental data. Finding the best set of parameters which minimize Eq. (2) by the DLS method leads to the following iterative scheme

$$\begin{aligned} \Delta\boldsymbol{\kappa}^{r+1} &= \left[\mathbf{G}_r^T\mathbf{G}_r + \varepsilon^2\mathbf{I}\right]^{-1}\mathbf{G}_r^T\mathbf{e}^r = \mathbf{J}\mathbf{e}^r \\ \boldsymbol{\kappa}^{r+1} &= \boldsymbol{\kappa}^r - \Delta\boldsymbol{\kappa}^{r+1} \end{aligned} \quad (3)$$

where $\mathbf{e}^r$ is the model error for the set of parameters obtained in the rth iteration, $\boldsymbol{\kappa}^r$. $\mathbf{G}_r$ is the Jacobian matrix of the model predictions with respect to the parameters at iteration r,$(\mathrm{G}_r)_{i,j} = \partial\mathbf{g}(\boldsymbol{\kappa}^r, t_i)/\partial\boldsymbol{\kappa}_j$; $\delta\boldsymbol{\kappa}^{r+1}$ is the vector of estimated increments of parameter $\boldsymbol{\kappa}^r$, and $\boldsymbol{\kappa}^{r+1}$ the vector of estimated parameter values at iteration $r+1$. At the end of the iteration the error is calculated for the set of current parameters, $\boldsymbol{\kappa}^r$, which is represented in the error landscape as a step towards the location of the minimal error. The $r+1$th step is determined by the product of the error vector at the rth step, $\mathbf{e}^r$ and the pseudoinverse matrix

$$\mathbf{J}_r = \left[\mathbf{G}_r^T\mathbf{G}_r + \varepsilon^2\mathbf{I}\right]^{-1}\mathbf{G}_r^T = \mathbf{M}_r^{-1}\mathbf{G}_r^T, \quad (4)$$

which depends on the local topology of the error landscape and provides a search "map" or solution direction in this landscape (see Appendix A for more details). The parameter ε^2 is called the damping parameter and can take an arbitrary small value. This parameter ensures the invertability of $\mathbf{M}_r$ by keeping its smallest eigenvalue equal to, or greater than, ε^2. In other words, ε^2 ensures the stability of the iterative procedure by limiting the length of each iteration step in the error landscape to $1/\varepsilon^2$. As a counterpart, non-zero values of ε may slow convergence and/or degrade the resolution of the parameters because the objective function that is actually minimized for $\varepsilon > 0$ is $\Phi = E(\boldsymbol{\kappa}) + \varepsilon^2\|\Delta\boldsymbol{\kappa}\|$, instead of just $E(\boldsymbol{\kappa})$. A suitable small value of ε should "damp" any disturbance (i.e. experimental data noise) and only slightly affect the model parameter values.

When the parameter values, and therefore the increments $\Delta\boldsymbol{\kappa}$, vary in a wide range it is more efficient to use logarithmic variations (Levenberg, 1944). In this approach we look for the solution of $\log(\mathbf{g}(\boldsymbol{\kappa}, t)) \cong \log(\mathbf{d}^{obs}(t))$ and the model deviation from the observations is $\mathbf{e}(\boldsymbol{\kappa}, t) = \log(\mathbf{d}^{obs}(t)/\mathbf{g}(\boldsymbol{\kappa}, t))$. Also, the partial derivatives that define $\mathbf{G}_r$ are of $\log(\mathbf{g}(\boldsymbol{\kappa}, t))$ with respect to $\log(\boldsymbol{\kappa})$, and the new value of the estimated parameter is $\boldsymbol{\kappa}_{r+1} = \boldsymbol{\kappa}_r \exp(-\Delta\boldsymbol{\kappa}_{r+1}/\boldsymbol{\kappa}_r)$ (see Appendix A for details).

2.2.1.1. The resolution matrix (RM): The sensitivity of the model to each of the estimated parameters is examined by the resolution matrix **R** (an $m \times m$ matrix) at convergence:

$$\mathbf{R}_r = \mathbf{J}_r\mathbf{G}_r = [\mathbf{G}_r^T\mathbf{G}_r + \varepsilon^2\mathbf{I}]^{-1}\mathbf{G}_r^T\mathbf{G}_r \quad (5)$$

When the matrix $\mathbf{G}_r^T\mathbf{G}_r$ is a regular matrix and ε^2 is small, then $\mathbf{R} \sim \mathbf{I}$ and the estimated model parameters are uniquely determined, i.e. the obtained model is close to being correct. However, if the matrix **R** has appreciable non-zero off-diagonal elements, variations of a particular parameter can be compensated by adjustments of other parameters that fit the experimental data equally well. Furthermore, if all the elements in a row of the **RM** are small ($|\mathbf{RM}| \ll 1$), then the corresponding parameter cannot be estimated for the given data set because the error is affected very little, even by large changes in this parameter. In these two situations, the optimal set of parameters occupies a manifold of dimension ≥ 1 and forms a “valley” or “hypervalley” in the error landscape. The **RM** discriminates meaningless estimates and also indicates those parameters that may vary in a wide range without alternation of the system response, as one can move along a valley in the error landscape without modifying the mean-square error.

2.2.2. Quasi-Newton (QN)—This method estimates the parameter set $\boldsymbol{\kappa} = (\boldsymbol{\kappa}_1, \ldots \boldsymbol{\kappa}_m)$ by minimizing the error $E(\boldsymbol{\kappa}, t)$ with the Newton algorithm (Dennis and Schnabel, 1983; Nocedal and Wright, 2006). In contrast with the standard Newton method, which requires numerical calculation of the gradient and Hessian of $E(\boldsymbol{\kappa}, t)$ in $\boldsymbol{\kappa}$ space, the QN method only requires computing the gradient; the Hessian is estimated using the gradients from successive iterations (Press et al., 1986). We used the implementation of the QN method provided by the Microsoft Power Fortran routine, BCONF, which includes an active set strategy (Gill and Murray, 1976). For a given error function, the QN method iteratively computes the search direction by using the step length determined by a local optimization of the function, called a line search, until it reaches the minimum of $E(\boldsymbol{\kappa}, t)$ function values. The Hessian is well-conditioned if the approximation by a quadratic function of the objective function surface $E(\boldsymbol{\kappa}, t)$ is the correct one which will be the case in the neighborhood of a local or global minimum when the convergence rate is quadratic (Dennis and Schnabel, 1989; Nocedal and Wright, 2006). Quite often, however, the objective function is poorly approximated by a quadratic function in the vicinity of the initial guess, especially if the latter is far from the optimum. Consequently, despite fast convergence, many trial and error runs involving different initial guesses may be necessary to determine the set of model parameters precisely.

2.2.3. Simulated annealing (SA)—This algorithm provides a global search method specifically designed for functions with multiple minima over wide range of parameters[2]

[2]The concept of annealing originates in metallurgy, and involves a technique of heating and controlled cooling of a material to increase the size of its crystals and reduce their defects. Heat causes the atoms to depart from a local internal energy minimum through random displacement from one energy state to another. Slow cooling increases the probability of finding configurations with lower internal energy than the initial one.

(Kirkpatrick et al., 1983; Press et al., 1986). This feature is especially important in chemical kinetics systems with many reaction equations and reaction rate constants which may vary considerably. Using a global strategy, each step of the SA algorithm replaces the current solution by a randomly chosen solution that is "nearby" in the parameter space; if the new solution is better it is chosen the new solution; whereas if it is worse the new solution can still be chosen, with a probability that depends on the difference between the corresponding function values and on a global parameter T (called the *temperature)* that is gradually decreased during the iterative process. As a result, the iterating solution changes almost randomly when T is large, but increasingly "down the well" as T goes to zero. The allowance for a random "uphill" enables escaping from local minima. A similar, although less powerful global search can be performed in DLS by gradually increasing the damping parameter ε^2, as the iteration process proceeds. We used the SA implementation provided in the Microsoft Power Fortran routine, SA (Goffe et al., 1994) to search a global minimum of the error function $E(\boldsymbol{\kappa}, t)$. This powerful method is effective in finding an absolute minimum of $E(\boldsymbol{\kappa}, t)$ but requires a large number of model evaluations. Thus its application is limited to parameter estimation in models whose equations can be solved quickly.

2.2.4. Convergence criteria—In DLS, the iterative process is stopped when the error and the norm of the increment of the estimated parameters are less than the prescribed respective tolerances. Also the maximal number of iterations is usually limited (here, the limit was 100 and the convergence is typically reached in less than 20 steps). The proprietary QN and SA implementations have their own exit strategies. There are two stopping criteria for BCONF that occur, either when the norm of the gradient is less than a given gradient tolerance ($<10^{-8}$), or when the scaled distance between the latest two steps is less than the step tolerance ($<10^{-10}$). The program also stops when it reaches prescribed values for the maximum number of function evaluations (default: 100), or of gradient evaluations (default: 400). SA terminates when either one of these two exit conditions is satisfied: (1) when the error is less than default tolerance value of 10^{-6} or (2) when maximum number of function evaluations is reached (currently set to 10^5).

2.2.5. Effective kinetic binding rate constant, k_{obs}—One important determinant of myosin-S1 binding to regulated thin filament via the TmTn complex is the effective kinetic binding rate constant, k_{obs}. When the actin concentration is much larger than the myosin-S1 concentration ([A] $\gg$ [S1]) in the presence of calcium, a single exponential is observed and $k_{obs} = (-d[\mathrm{A}]/[\mathrm{A_o}])k_{+1}[\mathrm{S1}] - k_{-1}[\mathrm{A.M}]$ represents the rate of decrease of unbound myosin concentration (McKillop and Geeves, 1993). Here [A.M] is the concentration of actin sites in weakly bound S1, i.e., A-states. When K_T is large (~200), the population of the A-state is small, and the reverse rate is negligible in absence of nucleotide, and thus the forward rate is approximately proportional to the rate of fluorescence decay in stopped-flow experiment (after subtraction of background fluorescence). Also the observed rate constant, k_{obs}, is independent of the ratio between actin to S1 over the range 5:1 to 20:1. In this case for lower and intermediate concentrations of Ca^{2+}, k_{obs} can be calculated as

$$k_{obs} = \frac{d[\mathrm{A}]}{[\mathrm{A}_O]} = k_{+1}[S1]\left[1 - \frac{1}{1 + K_B(1 + K_T)}\right]. \quad (6)$$

For the family of stopped-flow transient data, we can obtain the fractional k^f_{obs} by fitting the stopped-flow data. The fluorescence data are corrected for the background fluorescence then they are typically normalized by setting the equilibrium value to zero, followed by normalizing to the corrected initial fluorescence. Also for the effective comparison of k^f_{obs} from stopped-flow data with estimated k^f_{obs} from the best fits (6), k_{obs} is further normalizing at any Ca^{2+} concentration to the highest Ca^{2+} concentration (i.e. pCa = 4.6) after subtracting the value of k_{obs} at lowest concentration (i.e. pCa = 4.6):

$$k^f_{obs} = \frac{k_{obs}(\mathrm{pCa}) - k_{obs}(8.9)}{k_{obs}(4.6) - k_{obs}(8.9)}. \quad (7)$$

This procedure allows a comparison of the fractional equivalent rate constant, k^f_{obs}, from the estimated MG model parameters by different estimation methods for every set of experimental kinetic data at a prescribed Ca^{2+} concentration.

3. Results

To evaluate the effectiveness and robustness of the parameter estimation methods we fitted the model predictions to stopped-flow transients for S1 binding to pyrene-actin in the presence of regulatory proteins for a specified Ca^{2+} concentration (Boussouf et al., 2007a,b; Boussouf and Geeves, 2007). To improve accessibility of the parameter estimation methods for the MG model, we developed Visual Basic GUI, which integrates all computing components into a single shell application.[3]

3.1. Estimation of MG model parameters by DLS, quasi-Newton and SA estimation methods

The McKillop–Geeves (MG) three-state rigid chain model has four free parameters: the equilibrium constants K_B, K_T, K_1, and $K2$ (see Section 2). We estimated these parameters using the DLS, QN and SA methods by fitting the predictions of the MG model to stopped-flow transients for binding S1 to excess actin ([A] = 2.5 μM, [S1] = 0.25 μM) (Boussouf et al., 2007a). The fitted curves for these three methods in the case of high or low Ca^{2+} differ little and agree well with the experimental data (Fig. 3). Table 1 shows the values of the four equilibrium constants obtained by DLS, QN and SA for high or low Ca^{2+} concentrations. It is important to note that different estimation methods estimated somewhat different of values of the parameters K_B, K_T, K_1 and K_2, although the differences in the fitting error are hardly

[3]The program package is available in Srba Mijailovich's Computational Mechanobiology Lab http://www.hsph.harvard.edu/srbamijailovich/. The program can be obtained by clicking on download and then choosing the thin filament regulation V3.20.

noticeable (see Table 1) and the fits are almost identical (Fig. 3). The same behavior is observed at other Ca^{2+} concentrations too, thus a consistent trend of the parameter values with Ca^{2+} concentration is extremely difficult to establish if all four parameters are fitted simultaneously. The parameter values estimated from the different Ca^{2+} concentration by the above three estimation methods showed substantial differences in some cases, or did not show a consistent dependence on Ca^{2+} concentration.

This inconsistency may be attributed to a relatively flat error landscape, which would cause different sets of parameters to yield similar model predictions. When the differences between model predictions are within the experimental error (i.e. fluctuations in measured fluorescence), the parameter estimation methods would not able to discern which solution is better. In the SA paradigm, this effect can be interpreted as a residual temperature increment due to the entropy of the experimental measurements.

We explored the possibility of reducing this effect by filtering the experimental data through averaging pairs of two neighboring points. The difference between the parameters estimated from the original data and the filtered data was minor, the error was slightly reduced and the resolution matrix was improved in some cases. However, this apparent improvement was inconsistent, even when data were filtered a second time. Thus a reduction of experimental fluctuations has not resulted in significantly better convergence and better parameter resolution, pointing out that additional effects affected the goodness of the fit.

These findings indicate that the uniqueness of the sets of parameters estimated from stopped-flow transients should be evaluated carefully. The DLS method enables such an evaluation through analysis of the resolution matrices (see Table 2). Reducing the number of estimated parameters by, for example, fixing some model parameter values by using reliable values from other experiments (e.g. K_1 and K_2) improves the consistency of the remaining parameters with Ca^{2+} concentration. Fig. 4A shows that the best fits of stopped-flow transients obtained by the DLS, QN and SA methods at pCa of 8.9 are almost identical even when the number of estimated parameters is reduced to two. However, these fits are not as good as fits shown in Fig. 3 where all four free parameters are estimated. The reduction of the number of parameters from four to three and then to two also causes some minor differences in fits (Fig. 4B), but these differences are small and similar to the differences observed in Fig. 4A. Despite these differences, the reduction of the number of parameters brings the values of the estimated parameters closer by the different methods (Table 1), and improves the resolution matrix (Table 2), i.e. improves the robustness of the parameter estimation methods, but at the expense of less accurate fits (Fig. 4B, Table 1).

3.2. Calcium concentration dependence of K_B and K_T

Establishing the dependence of the equilibrium constants K_B and K_T on the concentration of Ca^{2+} by estimating the parameters of the MG model from the fits to families of stopped-flow transient experiments is important for the quantitative understanding of thin filament regulation in intact muscle fibers. As discussed above, fixing K_1 and K_2 and estimating only K_B and K_T can improve the uniqueness of the estimated parameters. The rationale for this simplification is based on three observations made in Table 2 and Fig. 5: (i) K_1 is uniquely determined even when fitting all four parameters of the MG model; (ii) K_1 and K_2 vary little

with the Ca^{2+} concentration; and (iii) the best estimates from the best fits of the full set of parameters over eleven calcium concentrations are 0.25 μM^{-1} and 170, respectively. These values vary little with concentration and they are consistent with previous studies. McKillop and Geeves (1993) showed that the binding equilibrium constant, K_1, from the kinetic data in the presence of Ca^{2+} at actin concentrations is between 0.1 and 0.2 μM^{-1}. The same authors reported that the value of K_2 is around 200 in the absence of nucleotide (McKillop and Geeves, 1993).

In order to illustrate how a reduction of the number of estimated parameters affects the values of the parameters and the quality of the fits we compared the values of the estimates of all four free parameters, of three free parameters with prescribed $K_1 = 0.25\ \mu M^{-1}$, and of two free parameters with prescribed $K_2 = 170$ (Table 1,K Fig. 4B). Fixing $_1$ and K_2, i.e. reducing the number of the free parameters to two is advantageous for global methods such as SA because it significantly reduces the number of model calculations and it enhances the uniqueness of the estimated K_B, and K_T over a wider range of Ca^{2+} concentrations (see DLS resolution matrices in Table 2 and Fig. 5). All parameters are estimated from stopped-flow data at actin concentrations in excess of S1 ([A] = 2.5 μM, [S1] = 0.25 μM).

The K_b values estimated by DLS, QN and SA display a sigmoidal Ca^{2+} dependence with the steepest variation at pCa between 6.6 and 5.4 (Fig. 6A). Hill's coefficients obtained from parameters obtained by each of three-parameter estimation methods are similar, averaging of about 2.25 (namely 2.15, 2.28 and 2.34 for DSL, SA and QN, respectively) and the 50% of K_B range is achieved at pCa of 5.57, 5.58 and 5.64, respectively. The K_T-Ca^{2+} dependence is less prominent, showing a mild increase from 0.03 at low Ca^{2+} concentration to ~0.125 at high Ca^{2+} (Fig. 6B). The relatively smooth trend is interrupted by two "spikes" at pCa = 6.0 and pCa = 5.4, which are likely caused by imperfections of the stopped-flow transients at early times, and the imprecise acquisition of the experiment starting time. These large variations of K_T between neighboring concentrations and the variability of the K_T values estimated by different methods are also reflected in the Hill coefficients, which are respectively, 0.47, 1.22 and 0.76 for DSL, SA and quasi-Newton. Half of the change in K_T is achieved at lower Ca^{2+} concentrations (at pCa of 6.54, 6.54 and 6.27, respectively) than estimated from the K_B-pCa relationship. These spikes and variations in the Hill coefficients, however, have almost no effect on the effective kinetic binding rate constant, k_{obs} (Fig. 7).

The dependence of k_{obs} with Ca^{2+} concentration is similar to the rate constant of single exponential decay of fractional fluorescence, k_{obs}, when actin concentration is much larger than the myosin-S1 concentration ([A] ≫ [S1]). The absolute values of k_{obs} calculated from the estimated K_B and K_T and obtained from fitting the observed stopped-flow transient at the range of Ca^{2+} concentrations were similar especially at lower Ca^{2+}, whereas the calculated values overestimated the experimental fits at higher Ca^{2+} concentrations (Fig. 7A). The fractional k^f_{obs} obtained from the DLS, QN and SA methods showed a sigmoidal dependence on pCa (Fig. 7B) and differ little from the common sigmoidal shape even if K_T fluctuate significantly for pCa< 6.0 (see Fig. 6B). Hill coefficients for k_{obs} are also almost identical for the DSL, SA and QN methods, having values of 2.60, 2.61, and 2.71, respectively, and 50% of the change in k_{obs} is achieved at virtually the same Ca^{2+} concentration (i.e. at pCa of

6.0). Interestingly, k_{obs} is closely related to K_B which is further exemplified by the mapping the spike in K_B at pCa = 5.2 to the spike in k_{obs}.

The k_{obs} dependence of pCa derived from the best fits by any of the three-parameter estimation methods is similar to the k_{obs} obtained from single exponential fits of experimental data (Boussouf et al., 2007a). Single exponential decay is observed at high Ca^{2+} concentrations and after a delay, t_{delay}, at lower Ca^{2+} concentrations. Fig. 8A shows an example of curve fitting of a single exponential at low Ca^{2+} concentration (pCa = 7.0). At low concentrations and after an initial delay, the single exponential (a straight line in semi-log plot), fits the data excellently until it reaches the noise floor at long times. These fits provide the single exponential rate constant k_{obs} (s^{-1}) that after appropriate normalization can be compared to the equivalent rate constant, k_{obs} (or k^f_{obs}), calculated from estimates values of K_B and K_T (Eqs. (6) and (7)). Fitting the single exponential to the stopped-flow transient can be achieved by either fitting all available data or just a partial the range of data which excludes the initial delay before rapid decay and the data which reached the noise floor at long times. Fig. 8A demonstrates that more robust fits are obtained when the model is fitted to this partial range of data. Overall differences between the fits of all data vs. the partial range of data are small but significant (Fig. 8B), both fits are similar to data reported in (Boussouf et al., 2007a) and the delay time show the expected sigmoidal trend (Fig. 8C). However, the single exponential rate constant k_{obs} underestimates the values of the equivalent rate constant at higher Ca^{2+} concentrations (Fig. 7A). This underestimation is likely due to the difficulty of separating the initial delay from the single exponential decay. This is also reflected in a lower Hill coefficient of 1.79 comparing to an average value of the coefficient of about 2.6 obtained from the estimated fractional equivalent rate constant k^f_{obs} (Fig. 7B).

We investigated the reason why the values of K_B estimated by different methods are significantly different for pCa = 5.2 (Fig. 6A). We observed that the plot of the experimental data (Fig. 9) has a somewhat irregular shape during the first 60 ms that differs from the initial part of the sigmoidal curve observed at the other concentrations (inset in Fig. 9). For actin concentrations in excess of myosin-S1 ([A] $\gg$[S1]), this early-time irregularity is present at all concentrations but it is much smaller than at pCa = 5.2, only affecting the estimated K_T values for pCa>6.0. It is interesting to note that despite quite different values of K_B estimated by the DSL, QN and SA methods, the fitted curves are almost identical (Fig. 9), and the fit error is roughly the same (3.2603×10^{-5}, 3.3183×10^{-5} and 3.3274×10^{-5}, respectively). A possible reason for the differences observed in K_b is the relatively low value of the diagonal term of the resolution matrix associated to K_B at pCa = 5.2 which is <0.2. Such low value indicates that the error of the fit is not sensitive to a relatively large variation of K_B (Fig. 10A). Thus, the significantly different estimated values of K_B are caused rather by different convergence behaviors of the three-parameter estimation methods than by the particular accuracy of the fit.

Overall, SA provided the most reliable estimates of K_B and K_T because of its ability to search for global minima in a broad parameter domain so that its estimates are virtually independent of the initial guess. This method provided stable convergence in all examined

cases and the smallest fitting error. Interestingly, the DLS estimates agree with the estimates from SA in most of the cases considered in this study, which suggests that DLS can be used equally well as SA to fit data from stopped-flow fast transient experiments. In a few occasions DLS achieved even slightly better fits of the data but the differences in the error were negligible. In addition to an excellent fit, the DLS method provides additional information about the quality of the fit via the resolution matrix (**RM**). Fig. 10 shows the diagonal elements of the **RM** for K_B and K_T. Consistent with Table 2 and Fig. 5, these data indicate that K_B is well resolved, especially for low calcium concentrations. The sharp decrease in the quality of the fit that is observed at high Ca^{2+} concentrations (pCa >6.0, **RM** of $K_B < 0.5$) can be linked to the kinetics of the system (Fig. 10A). In fact, most of TmTn units are in either closed or open state at high Ca^{2+} concentrations (i.e. low pCa), thus further increase of K_B causes a very small increase in the effective binding rate of S1 to regulated actin. In binding kinetic terms this means that at larger values of K_B (>3) the overall rate binding of S1 to regulated actin is affected very little by opening of a few extra units (Fig. 10A). Consequently, large changes in K_B values cause small shifts of the predicted transients which in extreme cases may be within the experimental uncertainty. At lower Ca^{2+} concentrations, however, K_B is better resolved suggesting that the primary regulation mechanism is the transition between blocked and closed sates, i.e.by K_b.

In contrast to K_B, K_T is better resolved at higher Ca^{2+} concentrations (pCa>6.0, **RM** of K_T>0.5) than at lower Ca^{2+} concentrations (Fig. 10B). This result suggests that the primary regulation mechanism at high calcium concentrations is the transition between closed and open states. This is because at higher Ca^{2+} concentrations the closed and open states are reasonably well occupied. However, the values of both K_T and **RM** vary significantly from one Ca^{2+} concentration to the next, suggesting that these variations are more likely to be caused by imperfections in experimental methods and data acquisition than by inadequate parameter estimation. This argument is supported by Fig. 6B which displays the same pattern of seesaw variation of K_T and essentially the same estimated K_t values with Ca^{2+} concentration regardless of the parameter estimation methods used. Consistent with this idea, the time courses of fractional fluorescence predicted by three-parameter estimation methods for the same Ca^{2+} concentration are very similar, although they significantly deviate from the observations (Fig. 9). Thus, the observed seesaw variation of K_t with Ca^{2+} concentration is not an artefact of the parameter estimation methods but rather a consequence of imperfections in experimentation to which K_t is very sensitive.

3.3. Estimation errors and convergence of the estimated parameters

There are three desirable features of a good parameter estimation method: (1) ability to escape from local error minima; (2) uniqueness of the estimated parameters; and (3) fast convergence. The first of these features can be achieved with global search methods such as SA. However, as we showed above, some gradient methods, such as DLS, may fit the experimental data equally well. In the circumstances when gradient methods are equally effective in finding an absolute minimum of the estimation error, the gradient methods should be preferred because they converge faster, while at the same time providing information about the uniqueness of the solution via the resolution matrix.

Fig. 11 shows typical iteration histories of three MG model parameters (K_b, K_t and K_2) and of the fit error at pCa = 4.6. This figure shows the estimation of the three parameters of the MG model because the estimation of all four free MG model parameters can only resolve well K_1 having a value of about 0.25 μM^{-1}, while other parameters are poorly resolved (see **RM** in Table 2 and Fig. 5). Thus, we fixed K_1 to 0.25 μM^{-1} and analyzed the convergence of K_B, K_t and K_2 which can now be better resolved. Fig. 11 shows the convergence of K_B, K_T and K_2, estimated by fitting the fluorescence transient for pCa = 4.6 and $\varepsilon = 0.3$. It is important to note that K_2, K_B and K_T showed damped oscillations around their final convergence values. The interdependence between the estimated parameters is reflected in the **RM** (see Table 2, pCa = 4.6, three parameters and Fig. 5B), whose diagonal elements are appreciably smaller than 1 ($K_B \approx 0.47$, $K_T \approx 0.58$ and $K_2 \approx 0.37$), and whose off-diagonal elements corresponding to K_T–K_2 and in lesser degree K_b–K_t are relatively large. Note that off-diagonal of the **RM** elements indicate the degree of correlation between the parameters. For example, at pCa 4.6, K_B can be well resolved if the number of parameters is reduced from three to two by fixing K_2 to 170. In this case diagonal elements of **RM** have values >0.75, whereas offdiagonal elements are much smaller (~0.1, see Fig. 5C). The overall convergence can be improved by a better choice of an initial guess of the estimated parameters and by using a smaller value of the damping coefficient ε.

The effect of the chosen value of s on KB and the convergence and error of the three-parameter estimation (of K_B, K_T and K_2) is shown in Fig. 12. As expected, the convergence is faster and the **RM** improves for lower values of ε. However, decreasing ε below 0.1 causes numerical instabilities that appear as ripples in the error function, and which could lead to divergence of the solutions. The same patterns are observed in the speed of convergence which rapidly increases for smaller values of ε. For example, the mean value of $\bar{K}_B^{(r)}$ converges very quickly for smaller values of ε (Fig. 13A) and the variations between neighboring iterations, $\Delta K_B / \bar{K}_B^{(r)}$ rapidly decrease (Fig. 13B). However, for the very small value of $\varepsilon = 0.05$, the approximate value of the converged $\bar{K}_B^{(r)}$ is achieved in just one or two iterations, but this value is not stable and the estimates kept oscillating around it without converging. In fact in most cases, when $\varepsilon \leq 0.1$ smooth convergence cannot be achieved at all – unlike the case shown in Fig. 13. Fig. 14 shows that in this particular example the resolution matrix does not improve substantially when the damping parameter is decreased below 0.15. Note that the DLS method formally approaches the QN algorithm for very small values of ε, which in turn explains why the QN method is only effective when the initial guess is close to values of fully converged parameters. Thus, the use of the DLS method requires a systematic way of choosing initial guesses for the parameters to be estimated and a careful choice of ε. In our experience, the strategy that proved most successful in fitting the kinetic data is to start with a relatively large damping coefficient, say $\varepsilon = 0.5$, and gradually decrease ε while monitoring the convergence and the stability of the solution. The last estimated parameter values are the best initial guesses for the next parameter estimation with smaller ε values.

4. Discussion and conclusions

Estimating chemical kinetic rate constants from experimental data is complicated due to the nonlinear behavior of protein binding and because the data are prone to procedural imperfections and fluctuations. We developed algorithms to accomplish this goal with the McKillop–Geeves model of thin filament regulation in solution (McKillop and Geeves, 1993), and we suggested strategies of how to successfully extract the best set of values that enable to establish the Ca^{2+} dependence of the model parameters. Establishing relationships between the reaction constants (K_B and KT) and the calcium concentration (pCa) is essential for a better understanding and further the development of quantitative models of thin filament regulation both, in solution and in living muscle fibers.

Three model parameter estimation methods, (i) Damped Least Squares (DLS), (ii) quasi-Newton (QN) which is related to DLS but simpler, and (iii) simulated annealing (SA), were compared by estimating the parameters of the McKillop-Geeves three-state rigid chain model from a set of stopped-flow data collected at different Ca^{2+} concentrations. All three models used different minimization methods to find best fits to a family of stopped-flow transients of S1 binding to regulated actin: (1) the DLS method is based on a gradient method for minimization of the error, (2) the QN method is a powerful optimization method close to DLS; and (3) SA is a global optimization method with an uphill and downhill search.

The MG model has four free parameters, K_B, K_T, K_1, and K_2, but the estimated parameters resulting from these methods have not provided any reasonable dependence of the model parameters on the Ca^{2+} concentration because the model parameters are interdependent and cannot be uniquely resolved (Table 2, and Figs. 5 and 10). Using evidence from the literature and from our own investigations we found that K_1, and K_2 only mildly depend on Ca^{2+} concentration and the trend is quite inconsistent. By setting reasonable values for K_1 and K_2, we found sigmoidal dependence of K_b on Ca^{2+} concentration (Hills coefficient of ~2.25) and a somewhat less prominent and scattered dependence of K_t on Ca^{2+}. The parameters estimated by DLS, SA and QN methods have similar values for all of the studied Ca^{2+} concentrations (Fig. 6). The fractional k_{obs}^{f} which is the effective kinetic binding parameter, derived from model parameter K_B and K_T (Eq. (6)), also displays a sigmoidal dependence on Ca^{2+} concentration, and the difference between the parameters estimated by the different methods was minor (Fig. 6). Furthermore, the Hill coefficients of k_{obs}^{f} were virtually the same (2.6), and k_{obs}^{f} reaches 50% at pCa = 6.0.

Overall, SA appears to be the most reliable of the three estimation methods because it always converges and it is based on a global searching strategy. However, this method converges very slowly and requires a large number of model simulations. Thus, SA is limited for parameter estimation only if the model predictions can be calculated quickly so the converged solution can be achieved over reasonable computational time. In contrast, those models which explicitly include cooperativity between nearest neighbor TnTm units such as the Hill two-state model (Chen et al., 2001; Hill et al., 1980) or the MG cooperative model which can be only be solved by stochastic (i.e. Monte Carlo) simulations will require

an over-extensive amount of computational time. Because these simulations are computationally intensive, only fast converging parameter estimation methods can be effectively used, i.e. the methods which converge in up to 25 iterations. One possibility is the use of the quasi-Newton method which converges rapidly, but its convergence is influenced frequently by the initial guess values and overall it requires a sizeable number of functional evaluations too. This method also imposes a large degree of uncertainty because choosing different initial guess values leads to different estimated model parameter values upon the incomplete convergence. The good fits shown in Figs. 3, 4, 6, 7 and 10, as well the parameter values shown in Table 1 are achieved by guessing initial values close to the values estimated by SA or DSL method; otherwise, the convergence of QN was compromised. These properties limit the application of the QN method to models based on stochastic simulations as well.

The parameter estimation via the sensitivity matrix and the Damped Least Squares (i.e. DLS) method has a number of advantages over the other two methods. Figs. 12 and 13 demonstrate that DLS can converge in less than 20 iterations, when an appropriate value of the damping parameter is used (i.e. ε <0.2), and reasonably well-chosen initial guesses are used. Similar to other iterative least squares methods, DLS can only find solutions that are linearly close to the initial guess and sometimes converge to a local minimum. These problems can be overcome by initially setting ε equal to a relatively large value (e.g. ε = 0.5) for a few iterations (say up to 20) and then progressively decreasing ε until convergence is reached. The initial large value of ε will allow a wider search within the parameter space but it may not converge. Typically, the values of each of the estimated parameters as well as the error fluctuate between some larger and smaller value between two consecutive iterations (Fig. 11). In most cases the estimated parameters will not converge in 20 iterations, but the last few iterations provide sufficient information to choose a new guess which is much closer to the converged solution. The criterion to select new initial parameters should include values with the lowest error. The subsequent estimation with the new guess and ε < 0.2, yields rapid convergence. This methodology was successfully applied to the estimation of the MG model parameters. It provided almost the same parameter estimation values as the global search method SA over all Ca^{2+} concentrations, but in fewer iterations (i.e. <20). Note that the strategy presented above (starting with lager ε and reducing it as approaching a minimum) is a standard optimization procedure for a Levenberg–Marquardt method (Kecman, 2001; Nocedal and Wright, 2006).

The uniqueness of the estimated parameters is a very important point to consider when dealing with parameter estimation methods. It is well-known that multiple solutions may exist with equal error margins (Tang et al., 2005; Vajda and Rabitz, 1988; Zsely et al., 2004). Various combinations of parameters can fit the experimental data equally well within the experimental uncertainty, or some of the parameters may be interdependent. Among those multiple solutions, the “global” solution might be attractive by offering the smallest residual error. Because we are dealing here with a nonlinear inversion problem, it is desirable to examine the global properties of the fit error in order to be confident that the most accurate set of parameters is estimated. In our case, SA is a convenient method for exploration of the global parameter space searching for the absolute minimum error. However, this global residual error minimum is not guaranteed to be better than other local minima estimated by

other methods such as DLS (Tang et al., 2005). This observation is indeed true in the current instance (see Table 1 and Fig. 5), where K_B–pCa relationships obtained by SA and DLS differ minimally. They only differ appreciably at high Ca^{2+} concentrations where the error landscape is relatively flat and a wide range of estimated K_B values produces model predictions with virtually no change in error. Thus, in a noisy and relatively flat error landscape the meaning of the absolute minimum is somewhat diluted and in this situations there is no a significant benefit in using global search methods, such as SA. In this case, the **RM** obtained from DLS can provide additional information on how to obtain the best solution with a minimum number of iterations by adjusting the appropriate value of ε and monitoring the variation of estimated parameters and the **RM** (Table 3, Figs. 12 and 14). In general the smaller the damping parameter ε, the faster is the convergence and the model parameters are better resolved. However, too small ε may lead to numerical instabilities and divergence of the parameter estimation (Figs. 12 and 13). As pointed out by Menke (1989) choosing an appropriate value for ε is not a straight forward process and typically ε is set by trial and error in usually decreasing sequences of values.

In conclusion, if the complex error landscape has multiple minima, then the solution can be non-unique and *a priori* information must be added to resolve the indeterminacy. Therefore, there is no guarantee that the DLS or any other iterative technique will converge to the solution, even if it is known that the unique solution exists, because it is necessary to explore the global error surface to ensure that a global minimum is reached. Overall the SA method is the most reliable method for the estimation of parameters in complex chemical systems. However, this method is in some cases impractical because it requires an extremely large number of iterations to find absolute minimum of the error. Thus, other methods, such as DLS, or sophisticated combinations of methods should be used. Nevertheless, DLS parameter estimation of the MG model parameters provided almost identical parameter estimates as the SA method and supplies information about the interdependence of the estimated parameters. Furthermore, the main advantage of the DLS vs. other examined methods is its fast convergence which is essential for the solvability of parameters in complex reaction systems which can only by simulated by complex computational algorithms. It should be noted that these conclusions are specific to the single model of myosin binding kinetics to regulated actin. The relative advantage of DSL observed in this study should be further evaluated for other models of mechanochemical systems. It would be especially important to establish the effectiveness of parameter estimation methods when applied to Monte Carlo simulations such as Hill's model of thin filament regulation (Chen et al., 2001; Hill et al., 1980) or Continuous Flexible TnTm Chain model (Smith and Geeves, 2003; Smith et al., 2003). The ultimate parameter estimation method that employs these stochastic simulations would require: (1) fast convergence because the Monte Carlo simulations require extensive computational time for each simulation and (2) sufficiently robust algorithm which can uniquely resolve complex issues related to interference between noise in experimental data and inherent stochastic noise in simulation data.

Acknowledgments

This work was supported by National Institute of Health grant NIH R01 AR048776. We thank Jason Li for participating in the running of curve fitting simulations and the data analysis.

Appendix A

A kinetic system can be described as following general differential equation,

$$\frac{d\mathbf{p}}{dt} = f(\mathbf{p}, \mathbf{k}, t) \quad \text{(A-1)}$$

with the initial condition, $\mathbf{p}(0) = \mathbf{p}^{in}$. Where $\mathbf{p}$ is a vector of the dependent variables, t is the time, $\mathbf{p}^{in}$ is initial value $\mathbf{p}(0)$ at t=0 and $\mathbf{k}$ represents an array of rate transition constants and other relevant model parameters. Once we determine vector of all system states, $\mathbf{p}(\mathbf{k}, t)$, we calculate the function $\mathbf{g}(\boldsymbol{\kappa}, t) = F(\mathbf{p}(\boldsymbol{\kappa}, t))$ which can be compared with corresponding experimental data, $\mathbf{d}^{obs}(t)$. Here $\boldsymbol{\kappa}$ represents the vector containing the m system free input parameters. In this study $\mathbf{g}(\boldsymbol{\kappa}, t)$ is formulated as a sum of contribution of each state to the florescence corresponding to observed pyrene fluorescence intensity during stopped-flow experiments. The details of calculating $\mathbf{g}(\boldsymbol{\kappa}, t)$ from known $\mathbf{p}(\mathbf{k}, t)$ is briefly explained in Section 2 and fully explained in Chen et al. (2001). A sufficient number of measurements $\mathbf{d}^{obs}(t)$ obtained at N discrete time points results in a sufficient number of relations between parameters, $\boldsymbol{\kappa}$, from which the values of these parameters can be obtained as a solution of the inverse problem (Menke, 1989).

Construction of sensitivity matrix

The function $\mathbf{g}(\boldsymbol{\kappa}, t)$ is assumed to be continuous and continuously differentiable. The first-order local sensitivity $\mathbf{s}(\mathbf{g}, \boldsymbol{\kappa}_j)$ of matrix $\mathbf{g}$, with respect to the jth free parameter, $\boldsymbol{\kappa}_j$, is defined as

$$\mathbf{s}(\mathbf{g}; \kappa_j) = \frac{\partial \mathbf{g}(t, \kappa_j)}{\partial \kappa_j}. \quad \text{(A-2)}$$

This sensitivity analysis provides the information about the effect of variation of the model parameters on the system behavior. The local sensitivity approximately calculated for ith element, g_i, which corresponds to observation d_i collected at time t_i (i =1, …, N), by finite difference method (FDM) is

$$s^r_{i,j}(g_i, k_j) = \frac{\partial g_i}{\partial \kappa_j} \approx \frac{g_i(\kappa^r_j) - g_i(\kappa^{r-1}_j)}{\Delta \kappa^r_j}. \quad \text{(A-3)}$$

Here superscript r stands for previous iteration, and $r - 1$ stands for the iteration before r. When $\Delta \boldsymbol{\kappa}_j$ varies in wide range of parameter values it is better represented in log space than in linear space, thus we used here:

$$\Delta\kappa_j^{r+1} = \log\frac{\kappa_j^{r+1}}{\kappa_j^{r}} \quad \text{(A-4)}$$

which gives the local sensitivity

$$s_{i,j}(y_i, \kappa_j^{r+1}) = \frac{\log[g_i(\kappa_j^{r})/g_i(\kappa_j^{r-1})]}{\log\left(\kappa_j^{r}/\kappa_j^{r-1}\right)} \quad \text{(A-5)}$$

The advantage of this setup is that $\Delta\kappa_j^r$ is normalized and it would not be affected by wide range of the parameter values.

The $\mathbf{G}_r$ matrix ($N \times m$)in Eq. (4) is assembled by elements defined by either Eq. (A-3) in linear space, of Eq. (A-5) in logarithmic space as

$$\mathbf{G}_r = \frac{\partial \mathbf{g}}{\partial \boldsymbol{\kappa}} = \begin{pmatrix} s(g_1;\kappa_1) & s(g_1;\kappa_2) & \cdots & s(g_n;\kappa_m) \\ s(g_2;\kappa_1) & s(g_2;\kappa_2) & \cdots & s(g_2;\kappa_m) \\ \vdots & \vdots & \vdots & \vdots \\ s(g_n;\kappa_1) & s(g_n;\kappa_2) & \cdots & s(g_n;\kappa_m) \end{pmatrix} \quad \text{(A-6)}$$

At each iteration the parameter values from previous iteration, $\boldsymbol{\kappa}^r$ are updated by the increments $\Delta\boldsymbol{\kappa}^{r+1}$ obtained as solution of least squares inversion problem (Eq. (4)) as:

$$\boldsymbol{\kappa}^{r+1} = \boldsymbol{\kappa}^{r} - \Delta\boldsymbol{\kappa}^{r+1} \text{ in linear space, or by}$$
$$\boldsymbol{\kappa}^{r+1} = \boldsymbol{\kappa}^{r}\exp\left(-\Delta\boldsymbol{\kappa}^{r+1}/\boldsymbol{\kappa}^{r}\right) \text{in logarithmic space}.$$

The iterative process is stopped when the convergence criteria are met: $\| \mathbf{e} \| < \varepsilon_e$, and $\| \Delta\boldsymbol{\kappa}^{r+1} \| < \varepsilon_e$ where ε_e are prescribed error tolerances for the error and for the parameter estimates, respectively.

References

Arora N, Biegler LT. A trust region SQP algorithm for equality constrained parameter estimation with simple parameter bounds. Comput Optim Appl. 2004; 28:51.

Beck JV, Arnold KJ. Parameter Estimation in Engineering and Science. Bradley RA, Kendall DG, Hunter JS, Watson GS, editorsJohn Wiley & Sons; New York: 1977.

Boussouf SE, Agianian B, Bullard B, Geeves MA. The regulation of myosin binding to actin filaments by Lethocerus troponin. J Mol Biol. 2007a; 373:587–598. [PubMed: 17868693]

Boussouf SE, Geeves MA. Tropomyosin and troponin cooperativity on the thin filament. Adv Exp Med Biol. 2007; 592:99–109. [PubMed: 17278359]

Boussouf SE, Maytum R, Jaquet K, Geeves MA. Role of tropomyosin isoforms in the calcium sensitivity of striated muscle thin filaments. J Muscle Res Cell Motil. 2007b; 28:49–58. [PubMed: 17436057]

Chen Y, Yan B, Chalovich JM, Brenner B. Theoretical kinetic studies of models for binding myosin subfragment-1 to regulated actin: Hill model versus Geeves model. Biophys J. 2001; 80(5):2338–2349. [PubMed: 11325734]

Dennis JE, , JrSchnabel RB. Numerical Methods for Unconstrained Optimization and Nonlinear Equations. Prentice-Hall; Englewood Cliffs, NJ: 1983.

Dennis JE, , JrSchnabel RB. A view of unconstrained optimization. In: Nemhauser GL, Runnooy Kan AHG, Todd MJ, editorsHandbook in Operations Research and Management Science. Elsevier; Amsterdam: 1989.

Eftaxias A, Font J, Fortuny A, Fabregat A, Stuber F. Nonlinear kinetic parameter estimation using simulated annealing. Comput Chem Eng. 2002; 26(12):1725–1733.

Farinha JPS, Martinho JMG, Pogliani L. Non-linear least-squares and chemical kinetics. An improved method to analyse monomer-excimer decay data. J Math Chem. 1997; 21:131–139.

Frank PM. Introduction to System Sensitivity Theory. Academic; New York: 1978.

Gill PE, Murray W. Minimization Subject to Bounds on the Variables. National Physical Laboratory; Teddington, UK: 1976. NPL Report NAC 72

Goffe W, Ferrier GD, Rogers J. Global optimization of statistical functions with simulated annealing. J Econ. 1994; 60:65–99.

Hill TL, Eisenberg E, Greene L. Theoretical model for the cooperative equilibrium binding of myosin subfragment 1 to the actin-troponin-tropomyosin complex. Proc Natl Acad Sci USA. 1980; 77:3186–3190. [PubMed: 10627230]

Hindmarsh AC. GEAR: Ordinary Differential Equation System Solver. 1972

Kecman V. Learning and Soft Computing. The MIT Press; Cambridge, MA: 2001.

Kirkpatrick S, Gelatt CD Jr, Vecchi MP. Optimization by simulated annealing. Science. 1983; 220(4598):671–680. [PubMed: 17813860]

Levenberg K. A method for the solution of certain non-linear problems in least squares. Q Appl Math. 1944; 2:164–168.

Lisy JM, Simon P. Evaluation of parameters in nonlinear models by the least squares method. Comput Chem. 1998; 22(6):509–513.

Maytum R, Lehrer SS, Geeves MA. Cooperativity and switching within the three-state model of muscle regulation. Biochemistry. 1999; 38(3):1102–1110. [PubMed: 9894007]

McKillop DF, Geeves MA. Regulation of the interaction between actin and myosin subfragment 1: evidence for three states of the thin filament. Biophys J. 1993; 65(2):693–701. [PubMed: 8218897]

Menke W. Geophysical Data Analysis: Discrete Inverse Theory. Dmowska R, Holton JR, editorsAcademic Press Inc; San Diego CA: 1989.

Nocedal J, Wright SJ. Numerical Optimization. Springer; New York, NY: 2006.

Press WH, Flannery BP, Teukolsky SA, Vetterling WT. The Art of Scientific Computing. Cambridge University Press; New York, NY: 1986. Numerical Recepies.

Smith DA, Geeves MA. Cooperative regulation of myosin-actin interactions by a continuous flexible chain II: actin-tropomyosin-troponin and regulation by calcium. Biophys J. 2003; 84:3168–3180. [PubMed: 12719246]

Smith DA, Maytum R, Geeves MA. Cooperative regulation of myosin-actin interactions by a continuous flexible chain II: actin-tropomyosin-troponin and regulation by calcium. ystems Biophys J. 2003; 84:3155–3167.

Tadi M, Yetter RA. Evaluation of the rate constants in chemical reactions. Int J Chem Kinet. 1998; 30:151–159.

Tang W, Zhang L, Linninger AA, Tranter RS, Brezinsky K. Solving kinetic inversion problems via a physical bounded Gauss-Newton (PGN) method. Ind Eng Chem Res. 2005; 44:3626–3637.

Terry DB, Messina M. Heuristic search algorithm for determination of rate constants and reaction mechanism from limited concentration data. J Chem Inf Comput Sci. 1998; 38:1232.

Tomovic R, Vukobratovic M. General Sensitivity Theory. Elsevier; New York: 1972.

Vajda S, Rabitz H. Identifiably and distinguishability of first-order reaction system. J Phys Chem. 1988; 92:701.

Varma A, Morbidelli M, Wu H. Parametric Sensitivity in Chemical System. Cambridge University Press; New York: 1999.

Zsely IG, Zador J, Turanyi T. How can very different combustion models produce similar results? Proc Combust Inst. 2004; 30 poster 1F2-11.

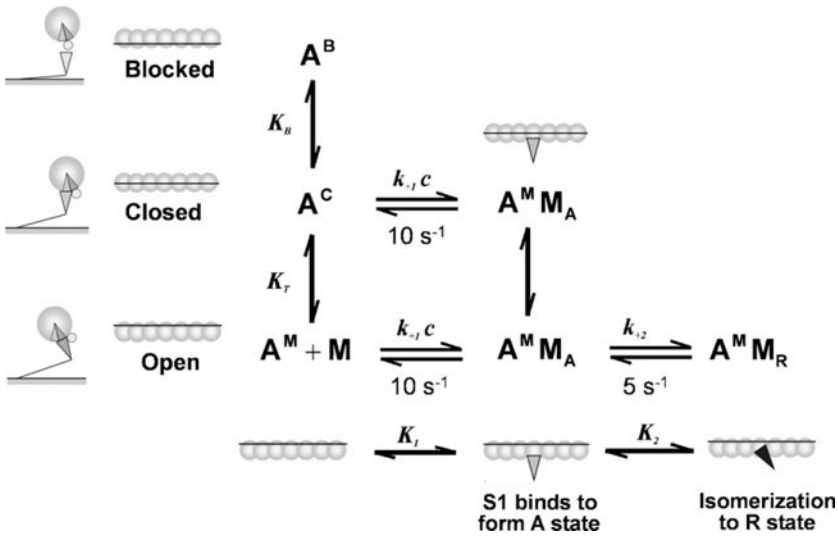

Fig. 1.

McKillop–Geeves (MG) three-state model scheme (McKillop and Geeves, 1993). In the MG model the structural unit, A_7·TmTn, is schematically shown as seven open circles representing the actins connected via a line representing the tropomyosin. This unit exists in a dynamic equilibrium between the three states as represented by the different positions of tropomyosin: the blocked state, A^B, in which no myosin-S1 binding can occur, the closed state, A^C, in which only weak binding of S1 can occur, and the open state, A^M, which allows isomerization of the myosin-S1 to the rigor-like state. The ratio of the three states in the absence of myosin-S1 is defined by the equilibrium constants K_B (between the blocked and closed states) and K_T (between the closed and open states). Weakly bound myosin states are denoted as A-states and rigor-like states are denoted as R-states. Rate of myosin binding is defined by equilibrium constant $K_1 = k_{+1}/k_{-1}$, and the rate of isomeration of S1 into R-state is defined by equilibrium constant $K_2 = k_{+2}/k_{-2}$. Backward rate constants used in all simulations are taken to be $k_{-B} = 100\ s^{-1}$, $k_{-T} = 3000\ s^{-1}$, $k_{-1} = 10 s^{-1}$ and $k_{-2} = 5\ s^{-1}$.

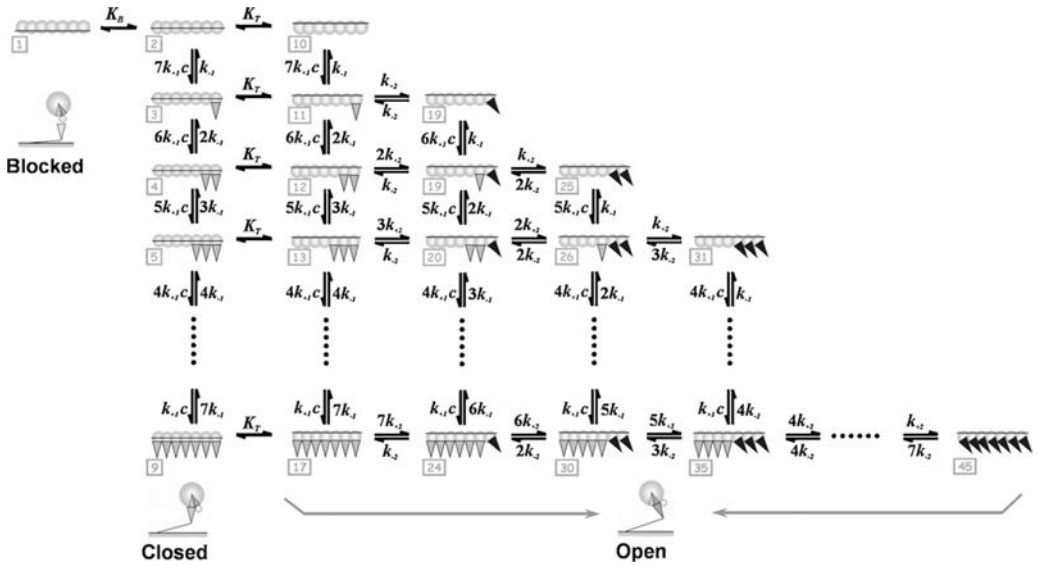

Fig. 2.

Schematic representation of the three-state MG model where each Tm–Tn complex is assumed to cover seven actin sites. The complete kinetic diagram for the binding of S1 to a structural actin$_7$·TmTn unit includes seven actin monomers. The resulting model contains one blocked state, eight closed states and 36 open states. All states are denoted as numbers in the square boxes (gray). The configurations of some states are shown. The fused two-way arrows represent the transitions from the blocked to the closed state with equilibrium rate transition constant, K_b, and the transition from closed to open state with the rate K_t. The two-way arrows represent forward and backward transition rates between myosin states interacting with actin: (i)S1 from solution weakly binding to actin (into A-state) and (ii) transition from A-state to R-state. The forward rate of S1 binding from the solution to the actin in an actin7-TmTn is defined as effective binding rate $k_{+1}\ c\ (s^{-1})$ multiplied by the number of unoccupied actin monomers within the unit, and the backward rate of the S1 unbinding from A-state back to solution by detachment constant k_{-1} multiplied by the number of S1 bound in A-state in the actin$_7$·TmTn. Similarly, transition from A-state to R-state is defined by forward constant $k_{+1}\ c\ (s^{-1})$ multiplied by the number of S1 bound in A-state in the actin$_7$·TmTn, and transition from R-state to A-state is defined by backward rate constant k-1 multiplied by the number of S1 bound in R-state in the unit.

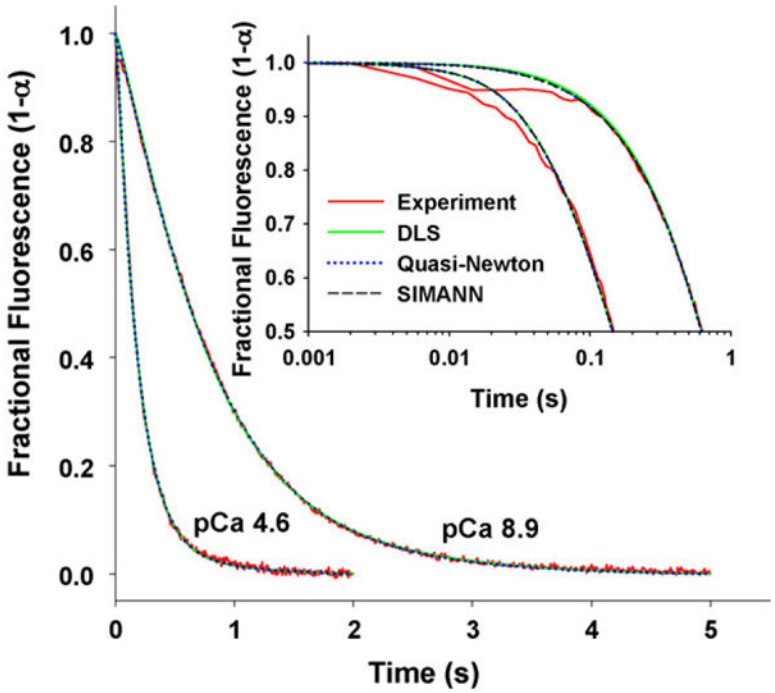

Fig. 3.

The best fits of stopped-flow data for excess actin ([A] = 2.5 μM, [S1] = 0.25 μM) and for high and low Ca^{2+} concentration (pCa of 4.6 and 8.9), by three different methods: DLS, quasi-Newton (QN), and SA. All three fitted curves fit the experimental data excellently. The values of estimated parameters K_B, K_t, K_1 and K_2 have similar but somewhat different values at the same Ca^{2+} concentrations and they are contrasted in Table 1. The backward rate constants are defined in the legend of Fig. 1.

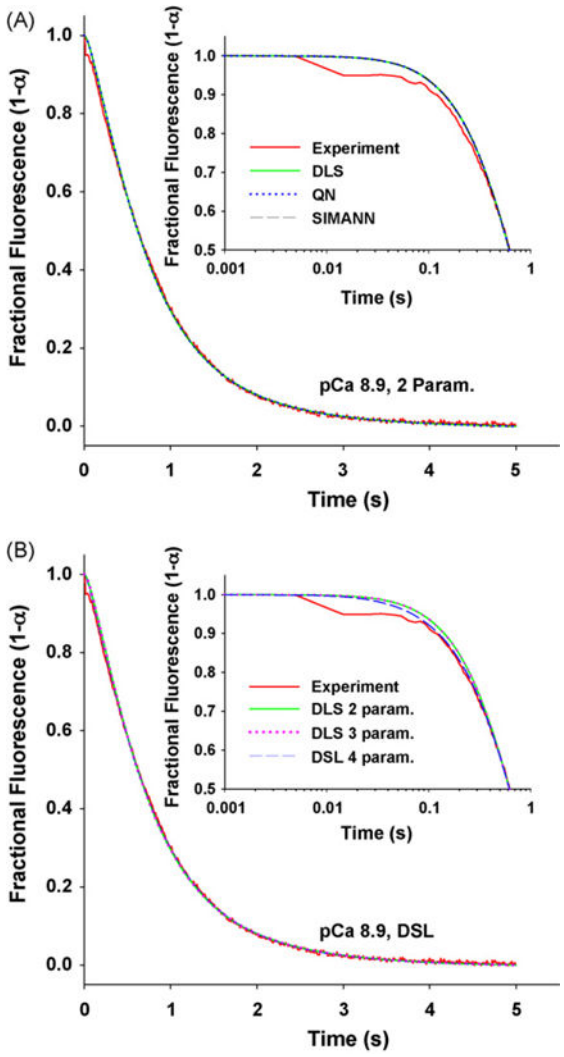

Fig. 4.

Effect of the number of estimated parameters on the parameter robustness and accuracy of the fit. When four parameters are estimated (K_b, K_t, K_1 and K_2) by different parameter methods, the parameter values vary in a wide range, while the differences in the accuracy of the fits are minor (Fig. 3). (A) The reduction of the number of parameters from four to two by fixing K_1 to 0.25 μM^{-1} and K_t to 170 displays some minor differences in quality of the fits, but minimizes the difference of the values of the estimated parameters by the different methods (Table 1); (B) the reduction of the number of parameters from four to three by fixing K_1 to 0.25 μM^{-1} and then to two by fixing K_t to 170 showed minor differences in quality of fit obtained by the DSL method. Despite these differences, the reduction in the number of parameters minimizes the difference in the values of the estimated parameters by the different methods, but at the expense of the accuracy of the fits. Shown are the fits of the stopped-flow transients at pCa= 8.9 for excess actin concentration to S1. The concentrations of actin and S1 are the same as in Fig. 3 and the backward rate constants are denoted in Fig. 1; the other equilibrium constants and error are shown in Table 1.

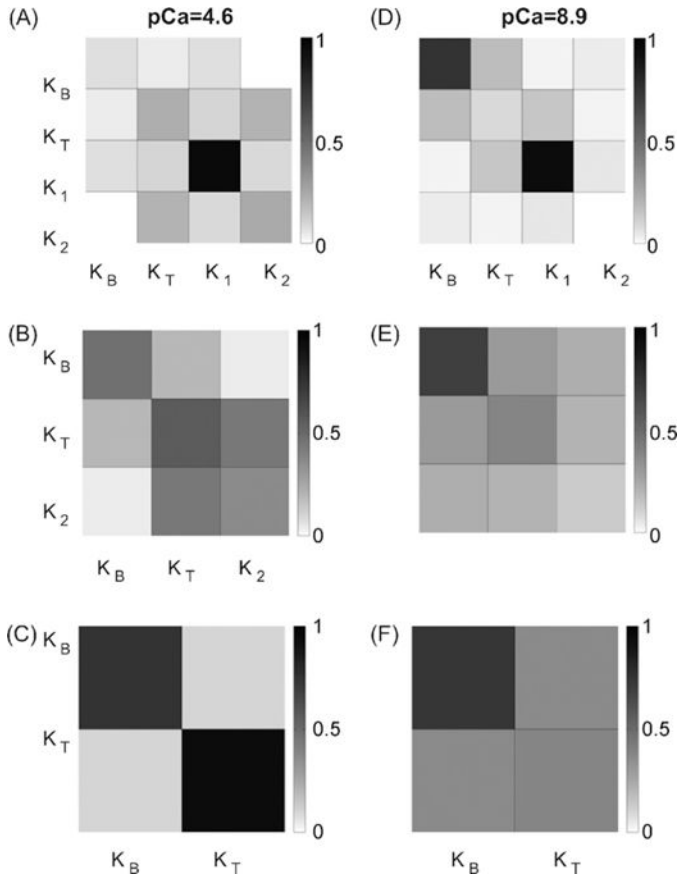

Fig. 5.

Graphical representation of the resolution matrices obtained by DLS (see Table 2 for numerical values). Panels (A)–(C) correspond to pCa = 4.6 and panels (D)–(F) correspond to pCa = 8.9. Panels (A)–(D) are obtained with four free parameters (K_B, K_T, K_1 and K_2). Panels (B)–(E) are obtained by fixing K_1 =0.25 μM^{-1}. Panels (C)–(F) are obtained by fixing K_1 =0.25 μM^{-1} and K_2 = 170.

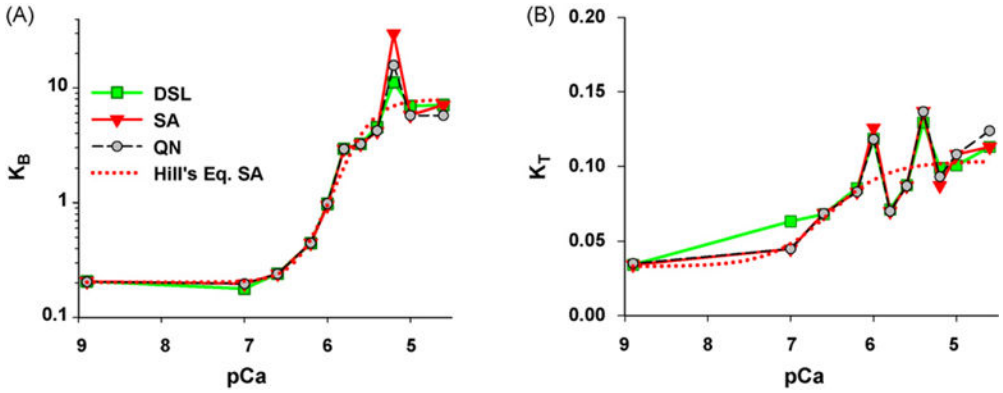

Fig. 6.

Estimated kinetic parameters K_b, K_t obtained by fitting MG model predictions to a family of stopped-flow data for range of Ca^{2+} concentrations by the DLS, quasi-Newton, and simulated annealing (SA) parameter estimation methods. Actin concentration is in excess of S1 ([A] = 2.5 μM, [S1] = 0.25 μM). In all simulations K1 and K2 are taken to be 0.25 μM^{-1} and 170, respectively. Backward rate constants are taken to be the same as denoted in Fig. 1. (A) K_b obtained by all three estimation methods displays sigmoidal dependence with Ca^{2+} concentration and sharply decreases for pCa<5.4 up to pCa = 6.6; (B) sigmoidal dependence K_t on pCa is less prominent having a scattered pattern for 5.4 > pCa>6.0 and differs little between the estimates obtained by different methods.

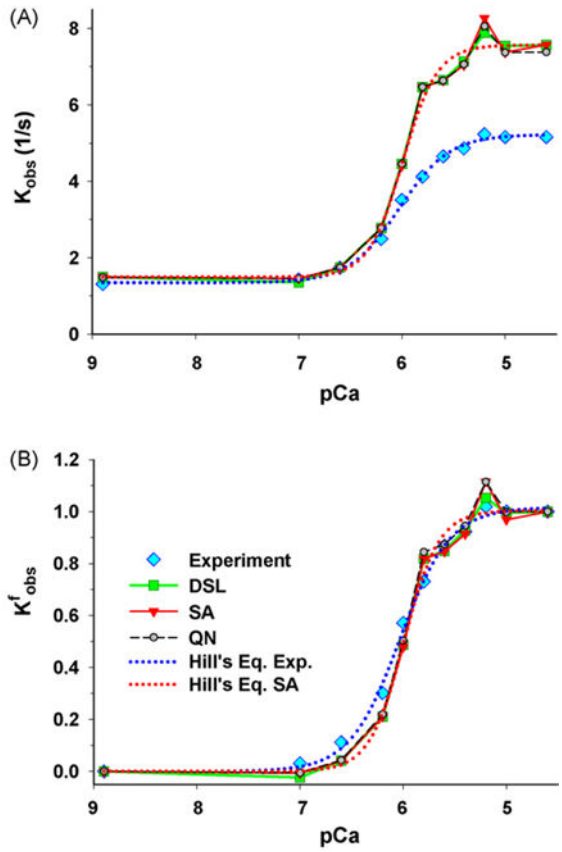

Fig. 7.

The effective kinetic binding rate constant, *kobs* (in s^{-1}), shows sigmoidal dependence on pCa. For all three-parameter estimation methods *kobs* is calculated using Eq. (6) (McKillop and Geeves, 1993) and the values of K_b and K_t are shown in Fig. 6. (A) k_{obs} shows minor differences between different parameter estimation methods regardless of large fluctuations in values K_b and K_b and agrees well with the k_{obs} obtained from single exponential fits of experimental data (Boussouf et al., 2007a) for lower Ca^{2+} concentrations. (B) Estimated fractional equivalent rate constant k^f_{obs} by any of the three-parameter estimation methods agrees well with k^f_{obs} derived from the single exponential fits to the experimental data after appropriate normalization. However, the Hill coefficient for the estimated values somewhat overestimates the Hill coefficient derived from the exponential fits (i.e. 2.6 vs. 1.79, respectively) reflecting some methodological differences in assessing k_{obs} (or k^f_{obs}) from estimated rate constants vs. directly from experimental data.

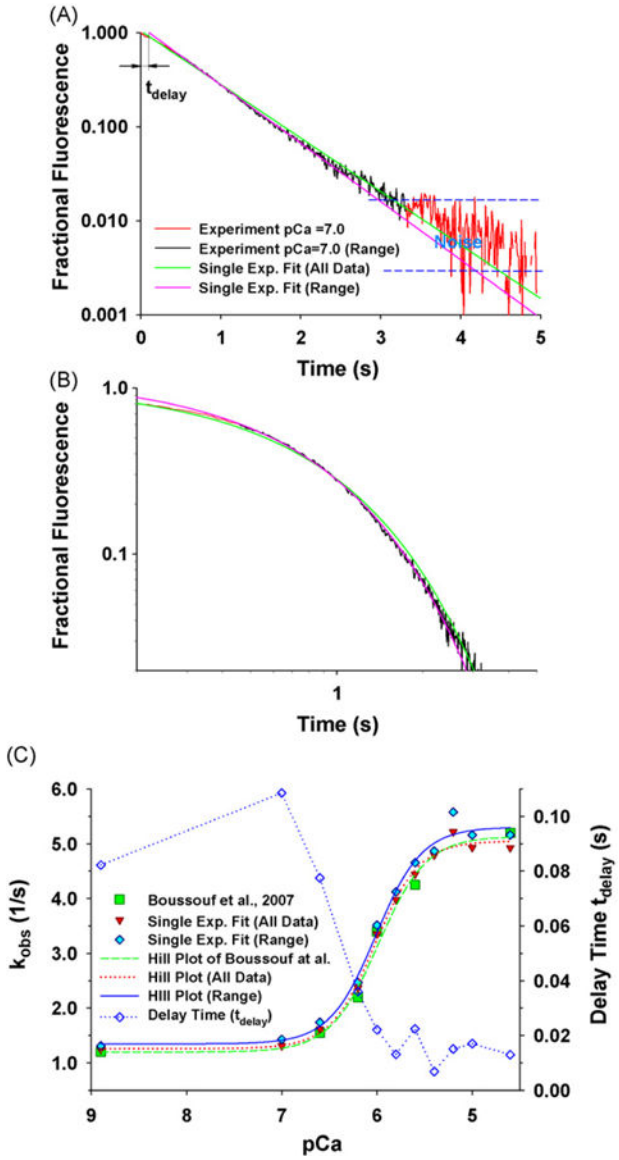

Fig. 8.

Assessment of single exponential rate of decay, k_{obs} (in s^{-1}), obtained from fits of experimental data (Boussouf et al., 2007a). (A) Single exponential fits through whole set of data and through a range of data which are represented by the linear portion in semi-logarithmic plot of the fractional fluorescence vs. time at pCa = 7.0. At these low Ca^{2+} concentrations the fit through the range of data better estimates initial time delay, t_{delay}, and eliminates bias of the fit caused by data noise at long times; (B) the same plot as in A but in log–log plot demonstrates the accuracy of the fits. (C) Calcium dependence of k_{obs} obtained from the fits through all data vs. the range of data. This dependence is compared to k_{obs} reported in Boussouf et al. (2007a). The Hill coefficients for all three sets of k_{obs} are 1.86 for all data, 1.79 for the range data and 1.72 for k_{obs} from Boussouf et al., and the midpoint at pCa ≈ 6.0. The delay time also showed inverse sigmoidal relationship vs. time.

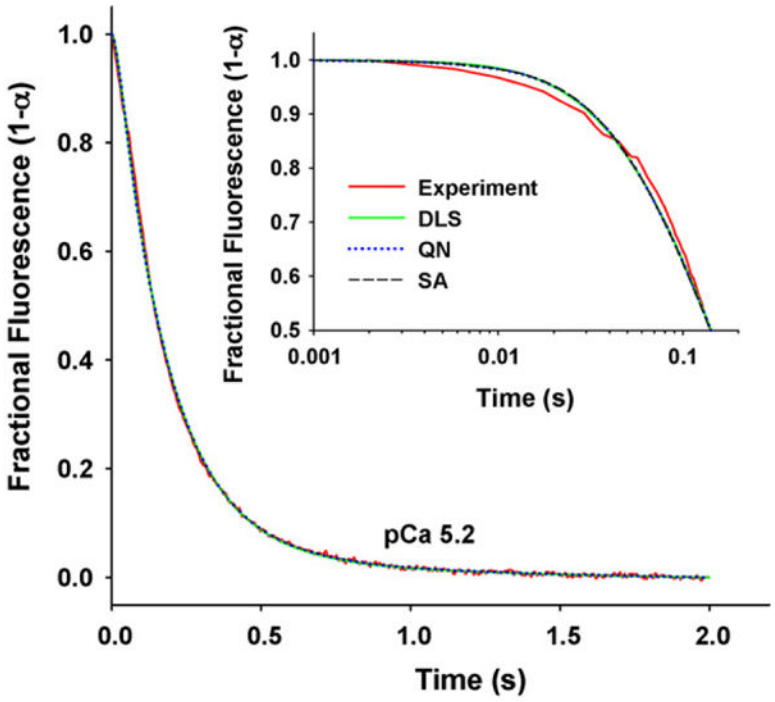

Fig. 9.

Experimental imperfections and quality of the fit. Estimated parameters at pCa = 5.2 shown in Figs. 6 and 7 significantly differ from the values estimated at neighboring Ca concentrations, as well as those estimated by different parameter estimation methods at pCa = 5.2. Parameter values for K_B and K_T vary in wide range (K_b = 11.22, 15.79 and 29.81, K_T = 0.0988, 0.0848 and 0.0870, for DSL, QN and SA, respectively), but all three fits differ a little and differences in errors are minimal (error=3.260×10^{-5}, 3.327×10^{-5} and 3.318×10^{-5}, respectively). Thus the main difference of the estimated parameters at pCa = 5.2 is rather in high sensitivity of estimated fits to minor perturbations in experimental data. In contrast, large deviation of the estimated values at pCa = 5.2 compared to neighboring concentrations is rather caused by imprecise determination of the timing of the beginning of the experiment or some other experimental imperfection (the beginning of any stopped-flow transient is always the most error-prone because of flow or stopping artefacts. Also, at the initial time points where signal changes can be large there can be an effect of the amplifier time constant if it is chosen to be optimal for the full transient). The reverse constants are denoted in Fig. 1 and the concentrations of actin and S1 are the same as in Fig. 3. Only K_B and K_T are estimated while K_1 and K_2 are kept constant at values of 0.25 μM^{-1} and 170, respectively.

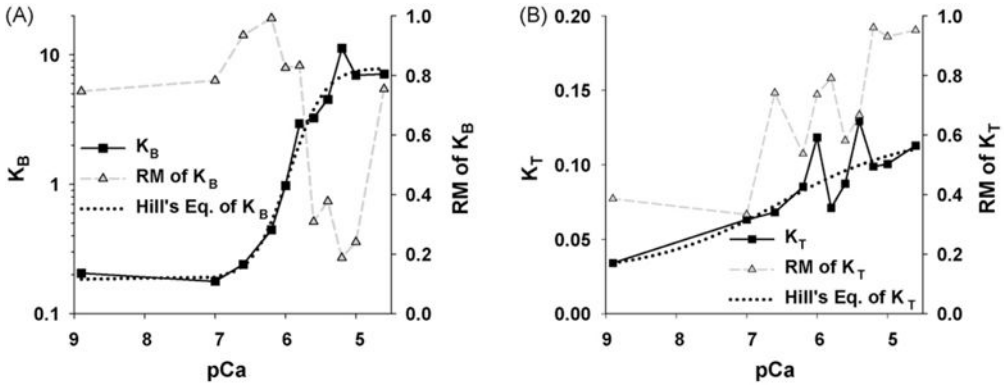

Fig. 10.

Robustness of the estimates of K_B and K_T evaluated by the DLS method. **RM** denotes values of diagonal elements of K_B and K_T in the **RM**. The **RM** value close to one indicates highly resolved and linearly independent estimates of K_B and K_T, while low values of **RM** indicate a poor resolution of the estimated parameters. All fixed model parameters are the same as shown in legend of Fig. 6.

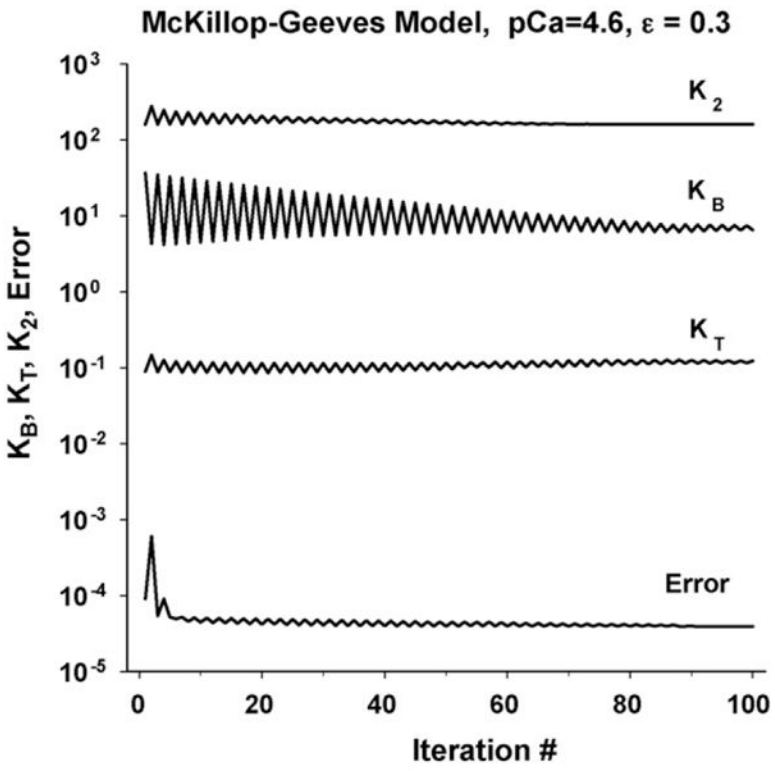

Fig. 11.

Convergence of estimated MG model parameters K_B, K_T, K_2 for excess actin concentration to S1 (i.e. [A] = 2.5mM and [S1] = 0.25mM), and at pCa=4.6. For the damping parameter ε = 0.3 all parameters steadily converge at about 100 iterations, except K_B which requires more iterations to fully converge. The myosin binding equilibrium constant, K_1, is prescribed at 0.25 μM^{-1}. Also all backward rate constants are denoted in Fig. 1. The estimated values of parameters (at iteration 100) are: K_B is between 6.50 and 7.49, K_T is between 0.116 and 0.123, K_2 between 167.6 and 167.8, and error$\cong 3.91 \times 10^{-5}$.

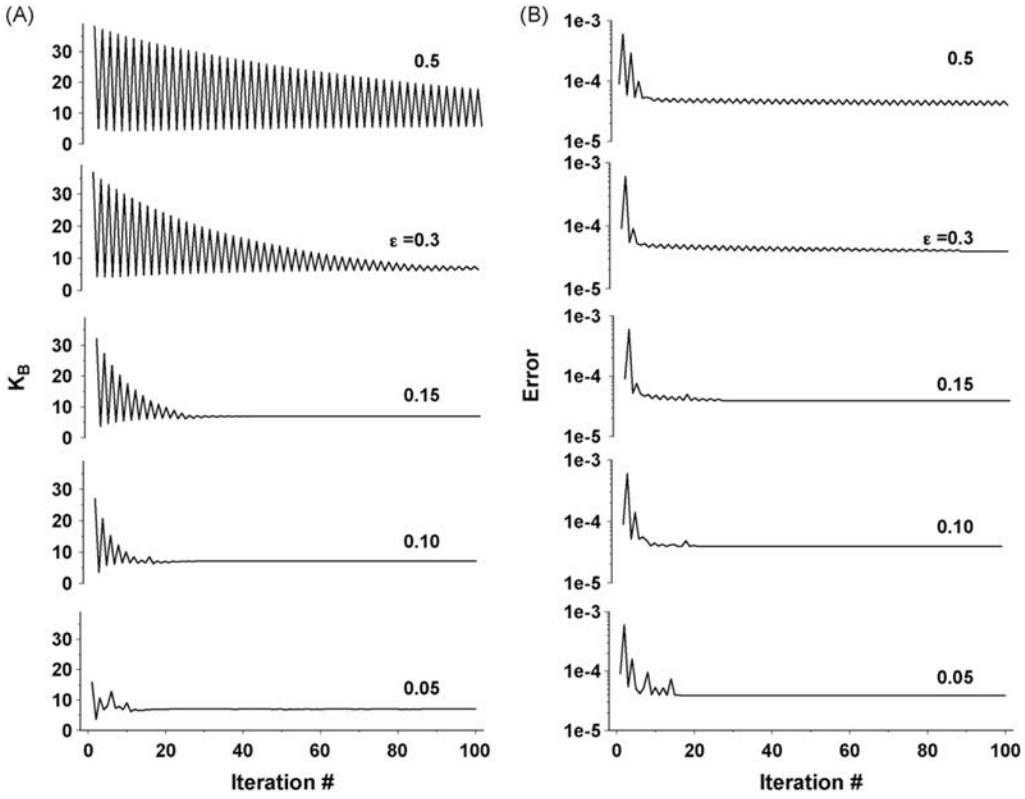

Fig. 12.

Convergence of K_b (A) and error (B) for different values of the damping parameter ε. All input parameters are the same as in Fig. 11. The decrease of ε significantly increases speed of convergence and typically converges all estimated parameters in less than 20 iterations if $\varepsilon \leq 0.1$. However, for low values of ε some instabilities appear due to interference between experimental error and the change of estimated parameter values between two subsequent iterations. Fully converged parameter values are: $K_b = 7.07$, $K_t = 0.114$, $K_2 = 167.7$, and error$\cong 3.90 \times 10^{-5}$.

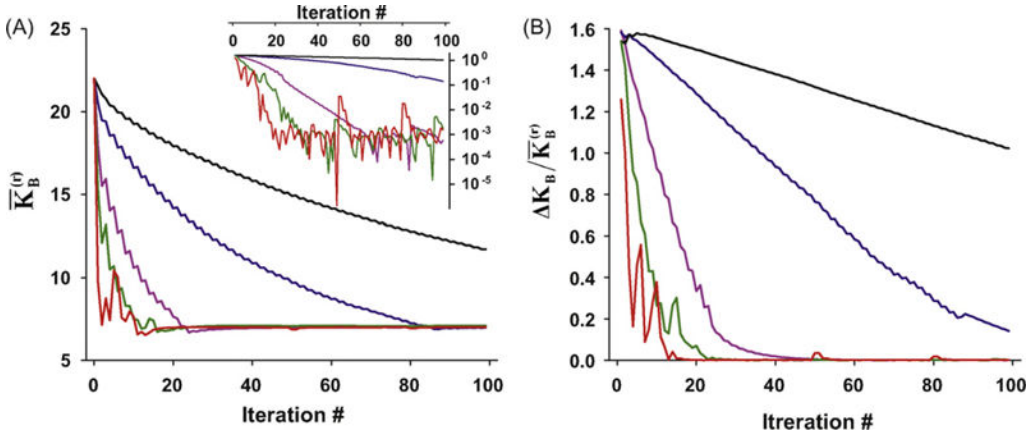

Fig. 13.

Speed of convergence of K_b. (A) Convergence of the mean value of $\bar{K}_B^{(r)} = (K_B^{(r)} + K_B^{(r-1)})/2$, where i is current iteration; (B) convergence of the change K_b value between two consecutive iterations $\Delta K_B/\bar{K}_B^{(r)} = 2(\left|K_B^{(r)} - K_B^{(r-1)}\right|)/(K_B^{(r)} + K_B^{(r-1)})$. Mean value of K_b converges to 7.07. The speed of convergence rapidly increases with decrease of ε and the value of $\Delta K_B/\bar{K}_B^{(r)}$ settles at about 0.1% (see inset in A). The ripples are caused by the numerical instabilities at low values of ε. All fixed parameter values are the same as in Fig. 11.

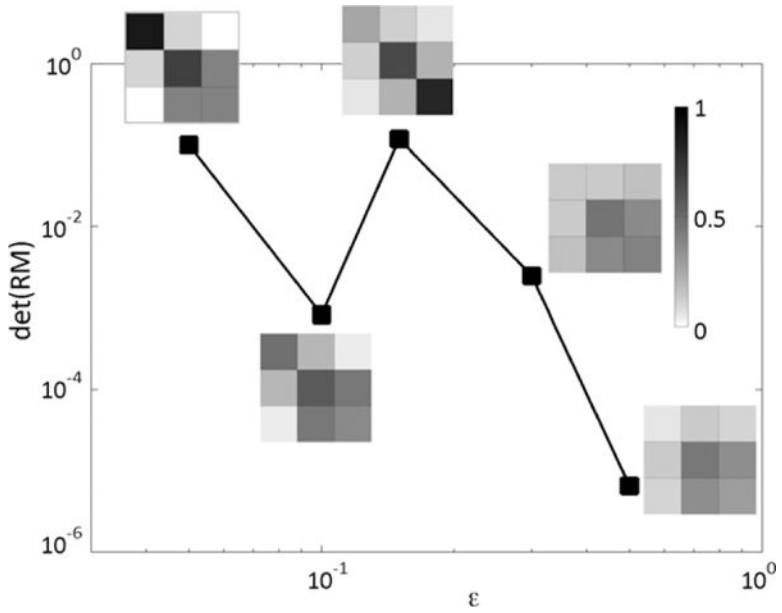

Fig. 14.

Evolution of the resolution matrix (**RM**) with the damping parameter ε. The line plot shows the evolution of the determinant of RM with ε. Note that det(**RM**) = 1 in the ideal case when all the model parameters are independent of each other all the diagonal terms of the **RM** are equal to 1 and the off-diagonal terms of the **RM** are equal to zero. The insets next to each data point in the curve are a graphical representation of the **RM** for the corresponding value of ε. Data are taken from Table 3.

Table 1

The comparison between the estimated parameters and error by the DLS, quasi-Newton and SA methods.

Mode1/pCa	K_b	K_t	K_1	K_2	Error ($\times 10^{-5}$)	Parameters
pCa 4.6						
DLS	5.2717	0.1024	2.592	184.46	3.8896	4
Qnewton	5.7725	0.1002	2.567	190.00	3.8878	4
SIMMAN	5.5881	0.1233	2.561	148.49	3.8903	4
DLS	7.0705	0.1142	2.500	167.83	3.9026	3
Qnewton	7.1043	0.1129	2.500	170.00	3.9026	3
SIMMAN	7.0945	0.1136	2.500	168.62	3.9039	3
DLS	7.1064	0.1129	2.500	170.00	3.9830	2
Qnewton	5.7516	0.1237	2.500	170.00	3.9516	2
SIMMAN	7.1079	0.1129	2.500	170.00	3.9976	2
pCa 8.9						
DLS	0.2738	0.0836	1.6801	204.88	2.8671	4
Qnewton	0.3428	0.0887	1.3884	400.00	2.4059	4
SIMMAN	0.3433	0.0882	1.3876	403.73	2.4057	4
DLS	0.2085	0.0343	2.5000	163.74	6.2195	3
Qnewton	0.2077	0.0351	2.5000	159.68	6.2075	3
SIMMAN	0.2197	0.0392	2.5000	118.80	6.1475	3
DLS	0.2062	0.0341	2.5000	170.00	6.2322	2
Qnewton	0.2038	0.0349	2.5000	170.00	6.2049	2
SIMMAN	0.2061	0.0341	2.5000	170.00	6.3820	2

Table 2

The estimated parameters by the DLS method and corresponding parameter resolution matrix (**RM**).

pCa 4.6		KB	KT	K1	K2
KB	**5.2717**	**0.068646**	0.043227	0.074243	−0.009111
KT	**0.1024**	0.043227	**0.231423**	0.104071	0.211774
K1	**0.2592**	0.074243	0.104071 [	**0.953922**	0.089077
K2	**184.4574**	−0.009111	0.211774	0.089077	**0.237632**

		KB	KT	K2
KB	**7.0705**	**0.471279**	0.18856	0.036523
KT	**0.1142**	0.18856	**0.577765**	0.441546
K2	**167.8280**	0.036523	0.441546	**0.369114**
K1	0.2500			

		KB	KT
KB	**7.1064**	**0.754661**	**0.101425**
KT	**0.1129**	**0.101425**	**0.951412**
K1	0.2500		
K2	170,0000		

pCa 8.9		KB	KT	K1	K2
KB	**0.2738**	**0.750572**	0.184956	0.017104	0.045338
KT	**0.0836**	0.184956	**0.084657**	0.14246	0.025661
K1	**0.1680**	0.017104	0.14246	**0.951724**	0.05895
K2	**204.8835**	0.045338	0.025661	0.05895	**0.0083121**

		KB	KT	K2
KB	**0.2085**	**0.697916**	0.305689	0.222463
KT	**0.0343**	0.305689	**0.378839**	0.204682
K2	**163.7400**	0.222463	0.204682	**0.127652**
K1	0.2500			

		KB	KT
KB	**0.2062**	**0.746649**	0.371836
KT	**0.0341**	0.371836	**0.385705**
K1	0.2500		
K2	170.0000		

Table 3

The effect of the damping parameter, ε, on the estimated model parameters by the DLS method and on the fit error.

ε		Estimated parameters			Resolution matrix				Error ($\times 10^{-5}$)	
		Last	Last-1		K_b	K_T	K_{2+}		Last	Last-1
0.50	K_b	5.7307	17.7194	K_B	0.0549	0.1286	0.1047	Error	3.9609	4.6782
	K_T	0.1140	0.0879	K_T	0.1286	0.4517	0.3585			
	K_{2+}	188.1002	169.5407	K_{2+}	0.1047	0.3585	0.2857			
0.30	K_B	6.8095	7.2387	K_B	0.1318	0.1396	0.1703	Error	3.9650	3.9652
	K_T	0.1104	0.1092	K_T	0.1396	0.4626	0.3715			
	K_{2+}	178.4369	173.9432	K_{2+}	0.1703	0.3715	0.3955			
0.15	K_B	7.1265	6.9980	K_B	0.2653	0.1113	0.0509	Error	3.9037	3.9030
	K_T	0.1059	0.1187	K_T	0.1113	0.6483	0.2170			
	K_{2+}	173.0033	159.1455	K_{2+}	0.0509	0.2170	0.8207			
0.10	K_B	7.0705	7.0774	K_B	0.4713	0.1886	0.0365	Error	3.9026	3.9026
	K_T	0.1142	0.1143	K_T	0.1886	0.5778	0.4415			
	K_{2+}	167.8288	167.5711	K_{2+}	0.0365	0.4415	0.3691			
0.05	K_B	6.9512	7.0276	K_B	0.87994	0.09587	−0.04318	Error	3.9031	3.9031
	K_T	0.1111	0.1145	K_T	0.09587	0.69256	0.39667			
	K_{2+}	175.0200	176.4080	K_{2+}	−0.04318	0.39667	0.40598			