Supporting Information

Using Heavy Mass Isotopomers for Protein Turnover in Heavy Water Metabolic Labeling

Rovshan G. Sadygov*

Department of Biochemistry and Molecular Biology

The University of Texas Medical Branch

301 University of Blvd

Galveston, TX 77555

*To whom the correspondence should be addressed, rgsadygo@utmb.edu, phone: (409)772-3287

Abbreviations: BWE – body water enrichment in deuterium; CPSM - Carbamoyl-phosphate synthase mitochondrial; Eq – equation; FFT – fast-Fourier transform; LC-MS – liquid chromatography and mass spectrometry; MPE – molar percent excess; NEH – number of exchangeable hydrogens; RA – relative abundance.

Content:

Page S13/Figure S8 - The comparison of the accuracy of the rate constants estimations using $I_o(t)$ and $\tilde{I}_1(t)$ under the equal noise condition for peptides with $I_0(0) < I_1(0)$.

Page S14/Figure S9 - The comparison of the accuracy of the rate constants estimations using $I_o(t)$ and $\tilde{I}_1(t)$ under the equal noise condition for peptides with $I_0(0) > I_1(0)$.

Page S15/Figure S10 - The comparison of the accuracy of the rate constants estimations using $I_o(t)$ and $\tilde{I}_1(t)$ under the unequal noise condition for peptides with $I_0(0) < I_1(0)$.

Instructions for running the R codes

The codes consist of two files: Isotopes.R and RateConstantSimulation.R. Each of the files has several functions. Isotopes.R file contains the functions for calculations of isotope distributions of peptides given the amino acid sequence. The isotope distributions are computed using the fft (Fast Fourier Transform) function of R. Other functions in the file compute the number of exchangeable hydrogens and the mass of each peptide sequence. It should be the first file to "source" into R.

RateConstantSimulation.R should be "sourced" into the R environment after Isotopes.R. It contains functions for computing original (I1_t and I2_t) and transformed (I1_t_tilde and I2_t_tilde) RAs of M_1 and M_2. Both original and transformed RAs are computed from formulas presented in the manuscript, Eqs. (4), (5), (8), and (10). RateConstantSimulation.R uses isotope calculation function of Isotopes.R.

To generate the profile of the first four mass isotopomers and transformed RAs of the first and second mass isotopomers, for a peptide sequence (e.g., TVLMNPNIASVQTNEVGLK) the following function is called from inside the R:

```
Profile_Peptide("TVLMNPNIASVQTNEVGLK", 0.35)
```

0.35 is the maximum value of the deuterium enrichment in this example. The sequence and deuterium enrichment can be changed. The function will generate **Figure 2** in the manuscript.

To generate **Figure 4** in the main text of the manuscript (**Figure 3** similar), type the following R function:

Rate_Constant_Simulation("EPLFGISTGNIITGLAAGAK", deg_rate = 0.110, dates=c(0, 3, 5, 7, 14, 21), noise= c(0, exp(-8)), BWE = 0.03)

Deg_rate is the true rate constant used in the simulation, dates – are the time points of labeling, noise is a two component vector, the first component is the location parameter, and the second component is the scale parameter of the Laplace distribution, BWE is the body water enrichment in deuterium.

The function returns the values of the rate constants in the Rates (six component variable) of the results list. The first component is the true rate constant (equal to 0.110 in the example above), the second component is the rate constant computed form $I_0(t)$, the third is computed from $I_1(t)$, the fourth from the transformed $I_1(t)$, the fifth from $I_2(t)$, and the sixth from the transformed $I_2(t)$.

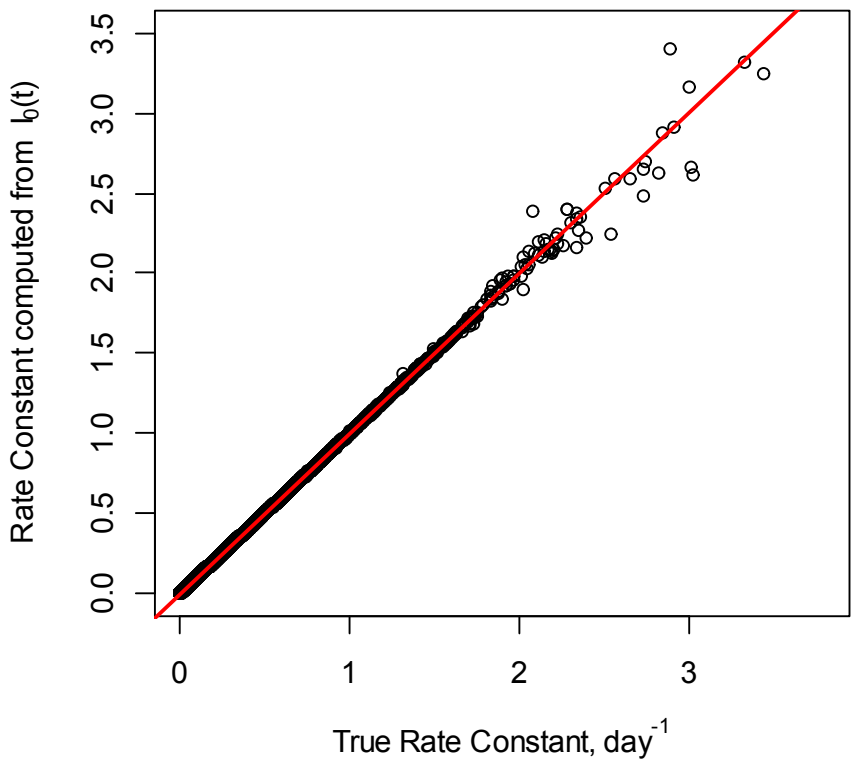
To run the code using the peptides in the Peptides.csv file, first read the peptides into the R environment:

Peptides_Rates = read.csv(“Peptides.csv”)

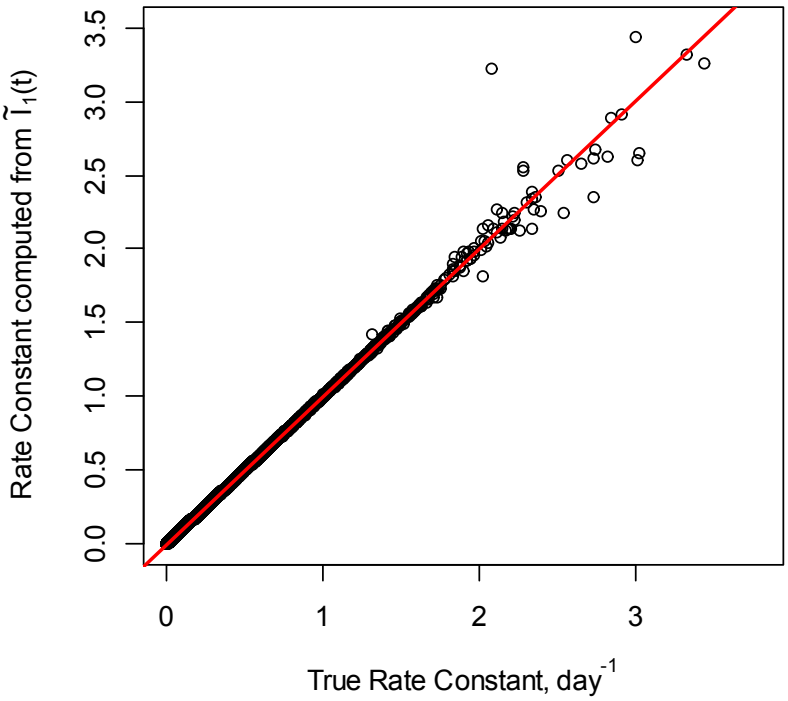
The command:

Rates_Experim_Noise = All_Peptides(Peptides_Rates,c(0, 3, 5, 7, 14, 21), c(-0.0035, 0.0159), pW = 0.03)

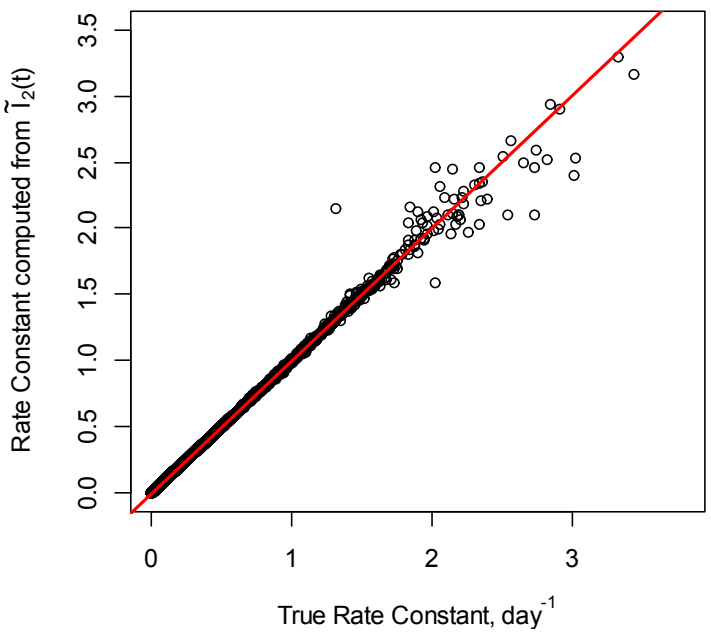
will compute the rate constants of all distinct peptides assuming time points of labeling: 0, 3, 5, 7, 14, 21 days, and the location and scale parameters: 0.0035, 0.0159, respectively. All results will be in Rates_Experim_Noise list variable.



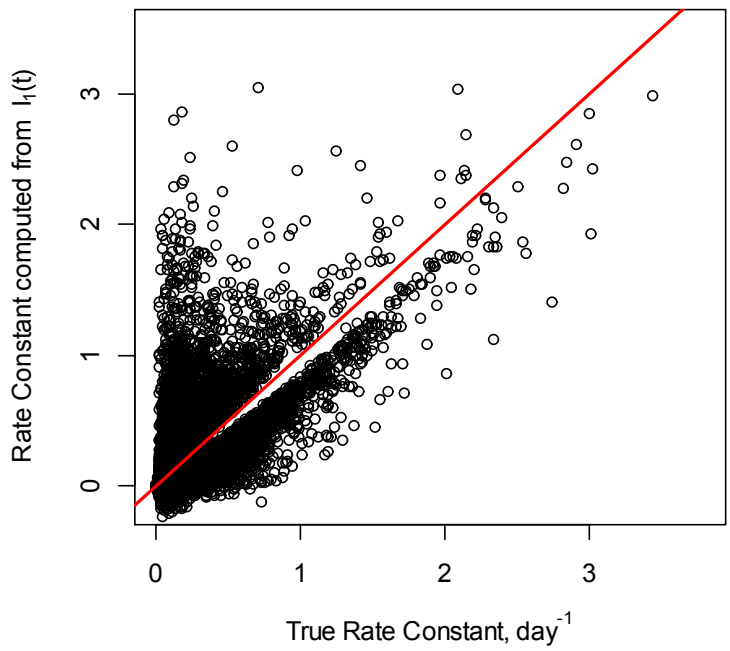
Supplementary Figure S1. For small noise (exp(-10)), the rate constants obtained from the time-course of $I_0(t)$ are identical to the true values, except for large values of the rate constant. The x-axis shows the true values of the rate constant; the y-axis shows the corresponding results obtained via simulations using that time-course of $\tilde{I}_1(t)$. The red line is the identity line. All rate constants reported in Figures S1-S5 are in units of day^{-1}.



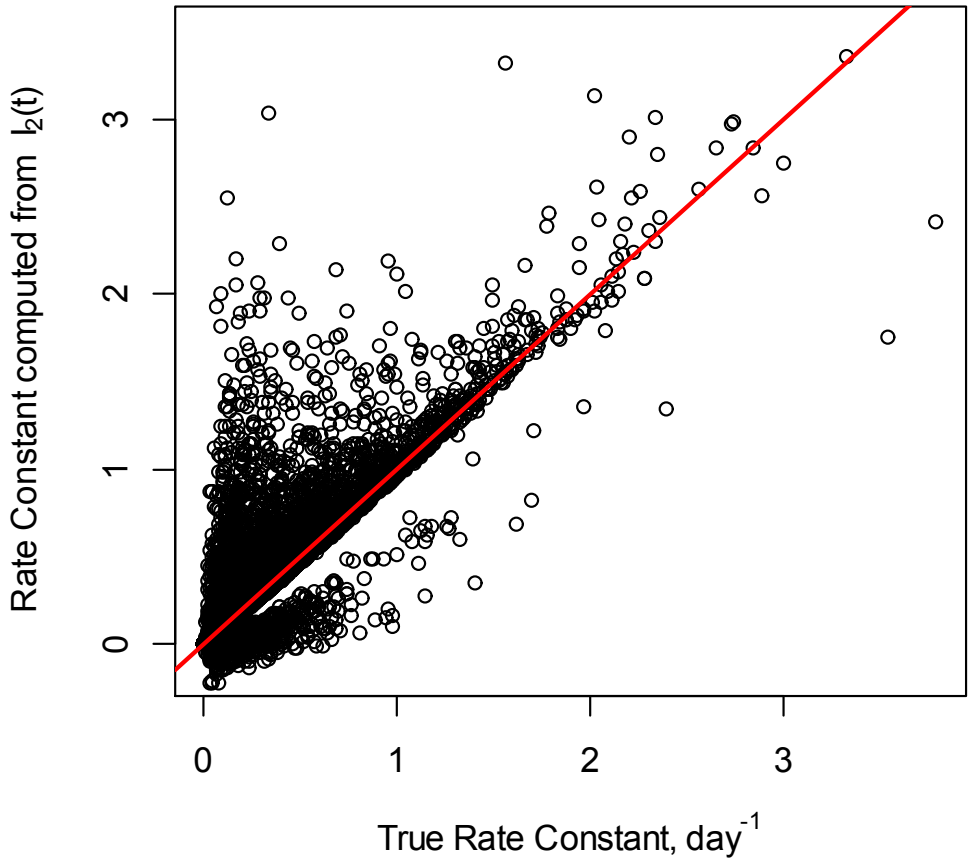
Supplementary Figure S2. For small noise (exp(-10)), the time-course of $\tilde{I}_1(t)$ accurately reproduces the true rate constants, except for large values of the rate constant. The x-axis shows the true values of the rate constant; the y-axis shows the corresponding results obtained via simulations using that time-course of $\tilde{I}_1(t)$. The red line is the identity line.



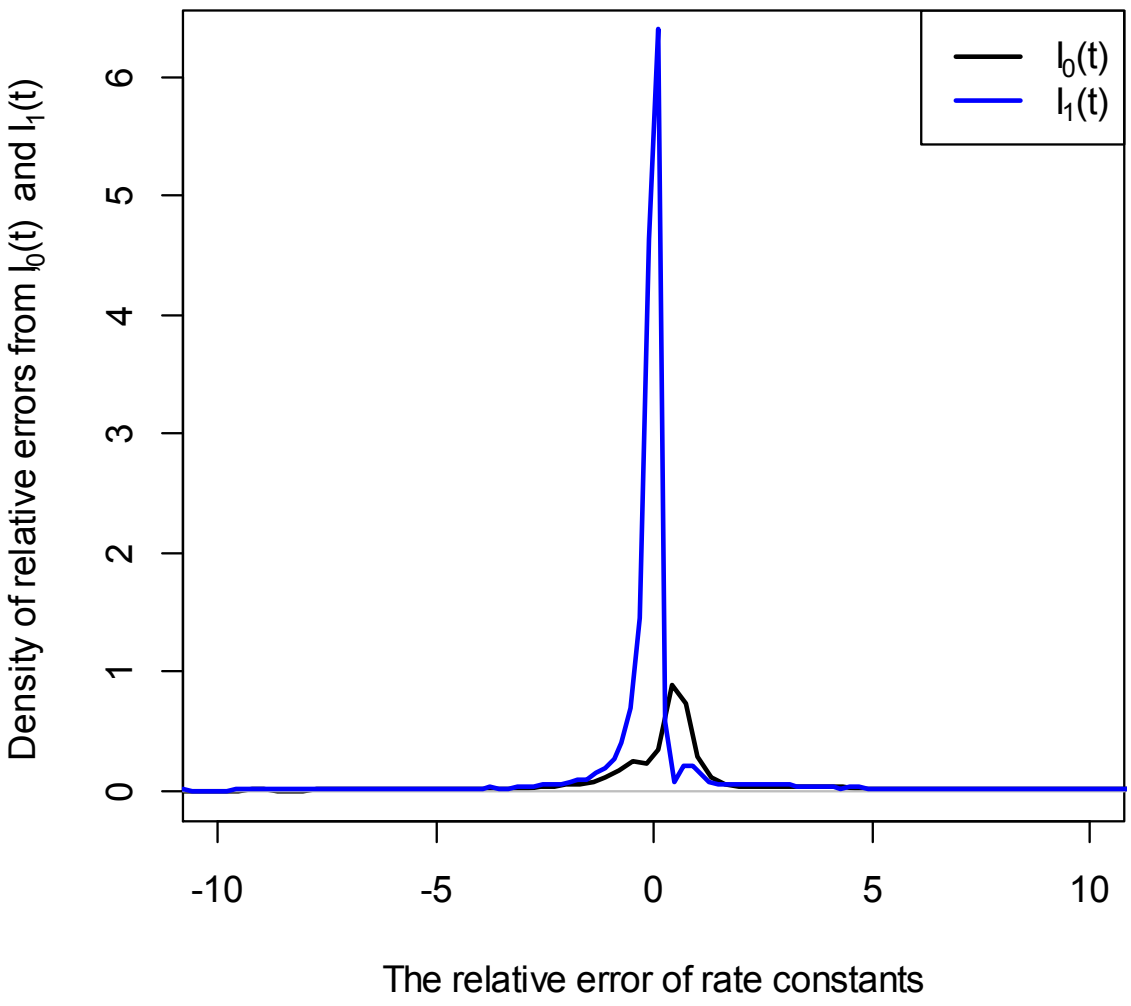
Supplementary Figure S3. For small noise (exp(-10)), the true rate constants are accurately obtained from time-course of $\tilde{I}_2(t)$, except for large values of the rate constant. The x-axis shows the true values of the rate constant; the y-axis shows the corresponding results obtained via simulations using that time-course of $\tilde{I}_2(t)$. The red line is the identity line.



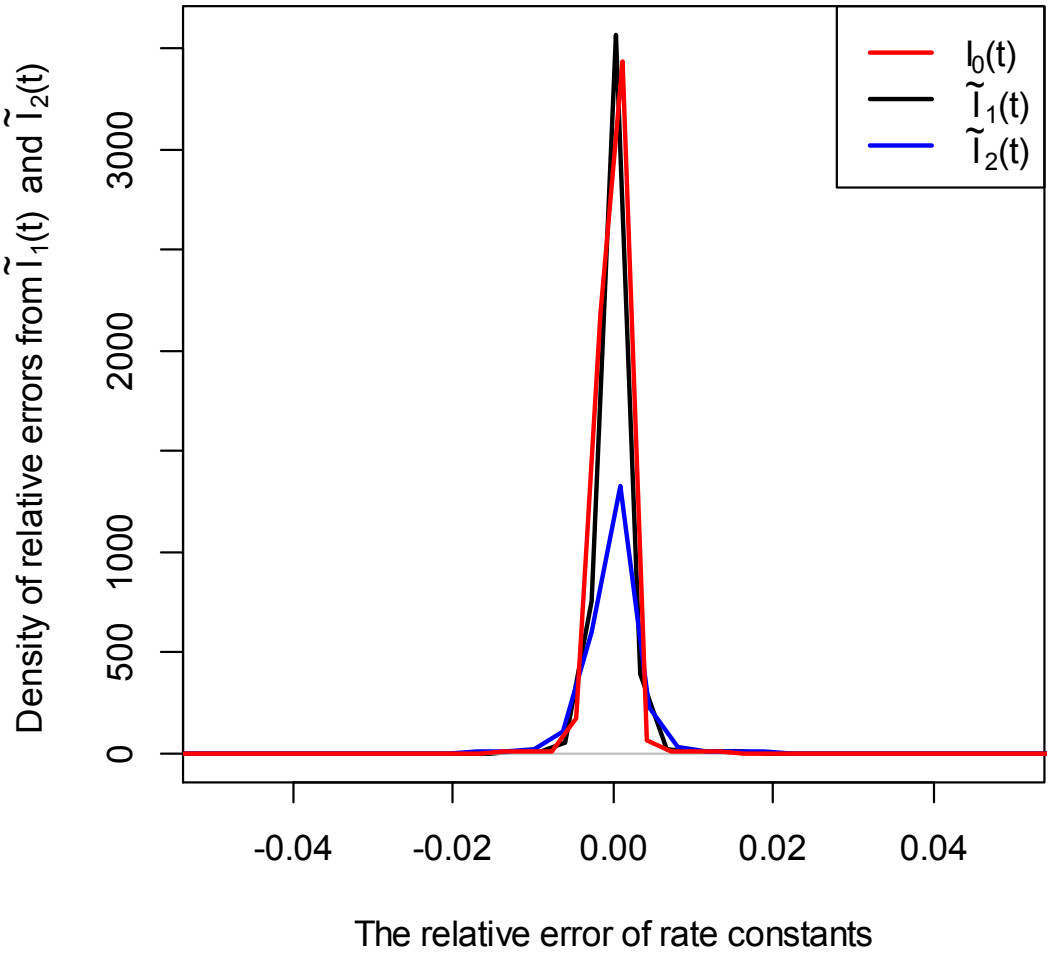
Supplementary Figure S4. The rate constants computed from the time-course of $I_1(t)$ differ from the true rate constants, even for small noise (exp(-10)). The x-axis shows the true values of the rate constant; the y-axis shows the corresponding results obtained via simulations using that time-course of $I_1(t)$. The red line is the identity line.



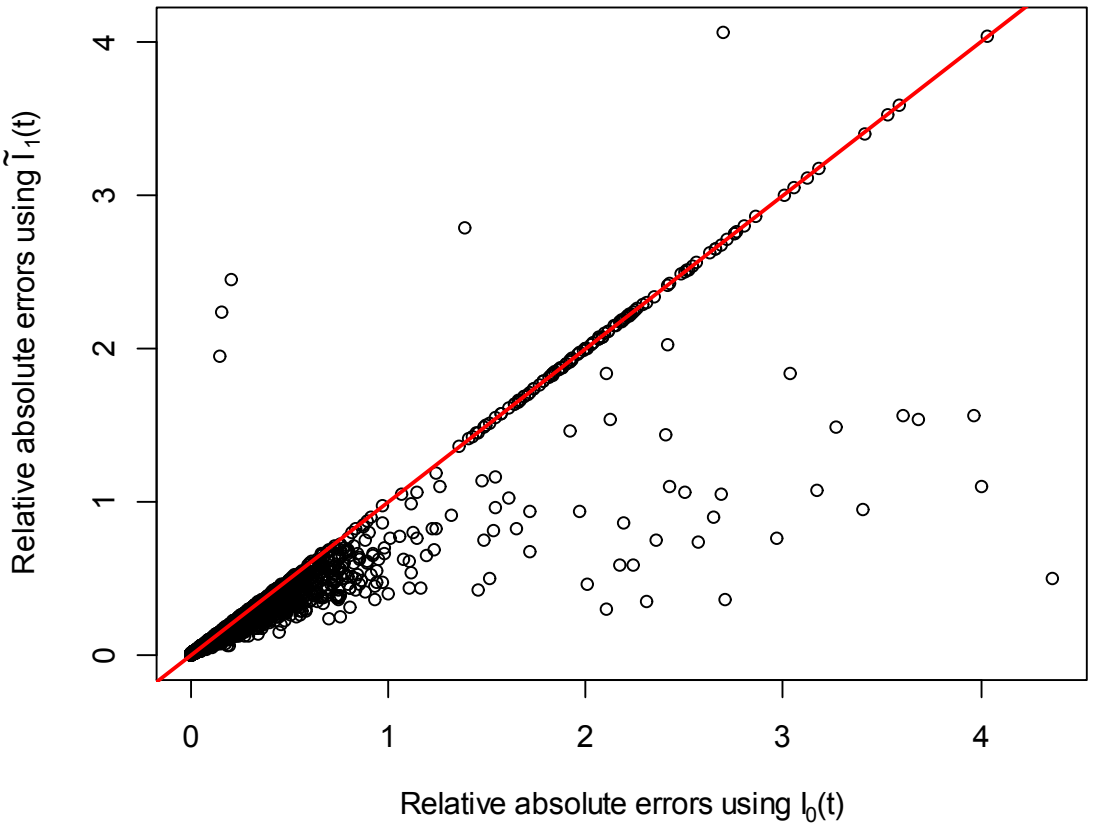
Supplementary Figure S5. Even for small noise (exp(-10)), the rate constants computed from the time-course of $I_2(t)$ do not reproduce the true rate constants. The x-axis shows the true values of the rate constant; the y-axis shows the corresponding results obtained via simulations using that time-course of $I_2(t)$. The red line is the identity line.



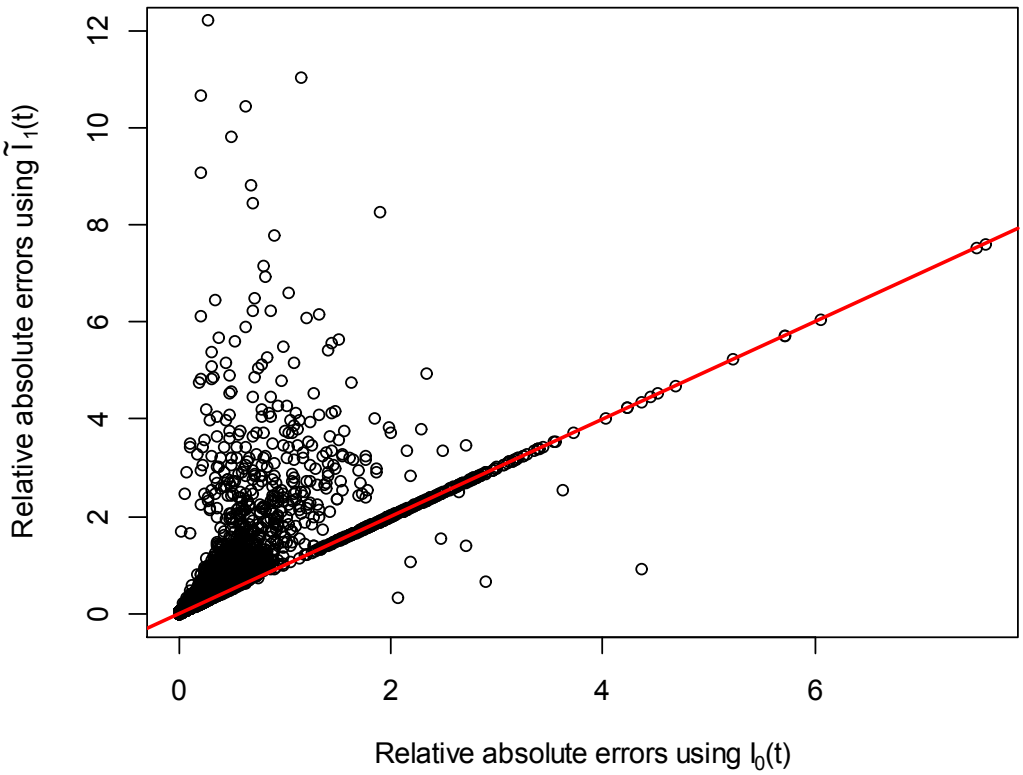
Supplementary Figure S6. The relative errors of rate constant estimations from $I_1(t)$ and $I_2(t)$ are high even from small noise conditions (exp(-10)). The X-axis shows the relative error = (TrueRate – SimulationRate)/TrueRate. The y-axis is the density of the relative errors.



Supplementary Figure S7. The relative errors of rate constant estimations from $\tilde{I}_1(t)$ are similar to those from $I_0(t)$ for small noise conditions (exp(-10)). The errors from $\tilde{I}_2(t)$ are larger. The X-axis shows the Relative Error = (TrueRate – SimulationRate)/TrueRate. The y-axis is the density of the relative errors.



Supplementary Figure S8. For equal noise, the accuracy of the estimation of rate constants is determined by the relative values of RAs. Shown are the relative absolute errors of rate constant estimations from $I_0(t)$ (the x-axis) and $\tilde{I}_1(t)$ (the y-axis) for peptides with $I_0(0) < I_1(0)$. As is seen from the figure, with the exception of five peptides, for the rest of the peptides, the absolute errors of rate constant estimations from $I_0(t)$ are higher than those from $\tilde{I}_1(t)$. The experimentally modeled noise parameters were used: the location parameter was -0.0035, the scale parameter was 0.0159. Shown are the results for more than 4100 peptides. The red line is the identity line.



Supplementary Figure S9. The relative values of RAs determine the accuracy of the rate constant estimations when noises in in $I_0(t)$ and $\tilde{I}_1(t)$ are equal. Shown are the relative absolute errors of rate constant estimations from $I_0(t)$ (the x-axis) and $\tilde{I}_1(t)$ (the y-axis) for peptides with $I_0(0) > I_1(0)$. As is seen from the figure, with the exception of seven peptides, for the rest of the peptides, the absolute errors of rate constant estimations from $I_0(t)$ are higher than those from $\tilde{I}_1(t)$. The experimentally modeled noise parameters were used: the location parameter was -0.0035, the scale parameter was 0.0159. Shown are the results for more than 13000 peptides. The red line is the identity line.

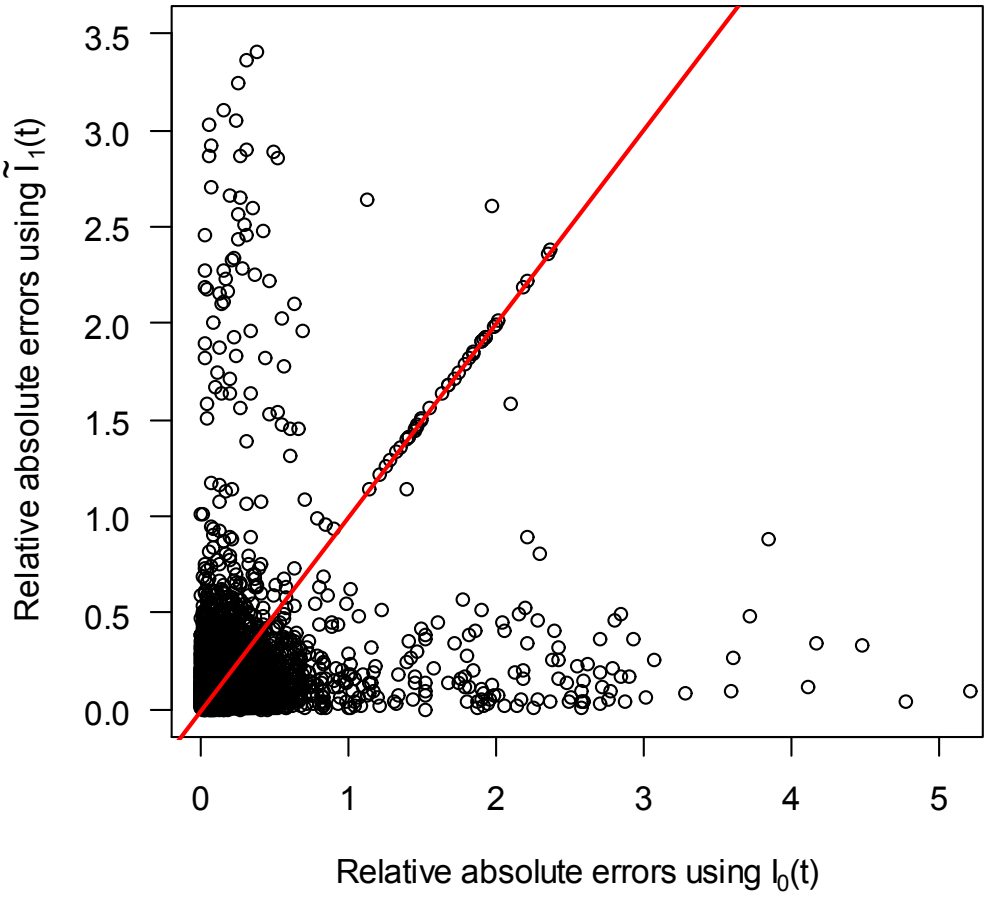


Figure S10. The experimental noise changes the dependency of the accuracy of the rate constant estimations on $I_0(t)$ and $\tilde{I}_1(t)$ time courses. Shown are the relative absolute errors of rate constant estimations from $I_0(t)$ (the x-axis) and $\tilde{I}_1(t)$ (the y-axis) for peptides with $I_0(0) < I_1(0)$. Shown are the results for more than 4100 peptides. The red line is the identity line. The corresponding figure with the equal noise is Supplementary Material Figure 8.