Supplementary Material

SVD-aided pseudo principal-component analysis: A new method to speed up and improve determination of the optimum kinetic model from time-resolved data

Key Young Oang,[1,2] Cheolhee Yang,[1,2] Srinivasan Muniyappan,[1,2] Jeongho Kim,[3] and Hyotcherl Ihee[1,2,a)]

[1] Department of Chemistry, KAIST, Daejeon 34141, Republic of Korea

[2] Center for Nanomaterials and Chemical Reactions, Institute for Basic Sciences (IBS), Daejeon 34141, Republic of Korea

[3] Department of Chemistry, Inha University, Incheon 22212, Republic of Korea

a) Author to whom correspondence should be addressed. Electronic mail: hyotcherl.ihee@kaist.ac.kr

Transient absorption (TA) experiment

A homodimeric hemoglobin (HbI) solution was prepared as described previously.[1, 2] The HbI solution in 100-μM sodium phosphate buffer (pH 7) was contained in a rubber-topped airtight quartz cuvette of 2 mm path length. The concentration of heme was determined from the absorbance at 578 nm using the extinction coefficient of heme-oxygenated derivatives (14.3 mM^{-1} cm^{-1}) and adjusted to be 50 μM. The HbI solution was bubbled by N_2 gas for 30 minutes to remove O_2. Then, the HbI solution under nitrogen atmosphere was reduced by adding 10 μL of 100 mM sodium dithionite aqueous solution. The reduced HbI solution was exposed to CO gas for 30 minutes to convert HbI to the CO-bound HbI. Transient absorption (TA) experiment was performed as described previously.[2] Briefly, the second harmonics of Nd:YAG laser (EKSPLA) was used as a pump beam (532 nm, 2 Hz). The fluence of the pump beam per pulse was about 640 μJ mm^{-2}. Output of a Xe lamp (Newport) was used as a probe beam (250 W, continuous wave). The pump beam intersected with the probe beam at a sample position. The angle between the pump and probe beams was adjusted to be as small as possible. The probe beam passing through the sample cell was detected by a combination of a spectrometer (Princeton Instruments) and an ICCD (Andor). The TA experiment was implemented at 296 K.

Supplementary Figures

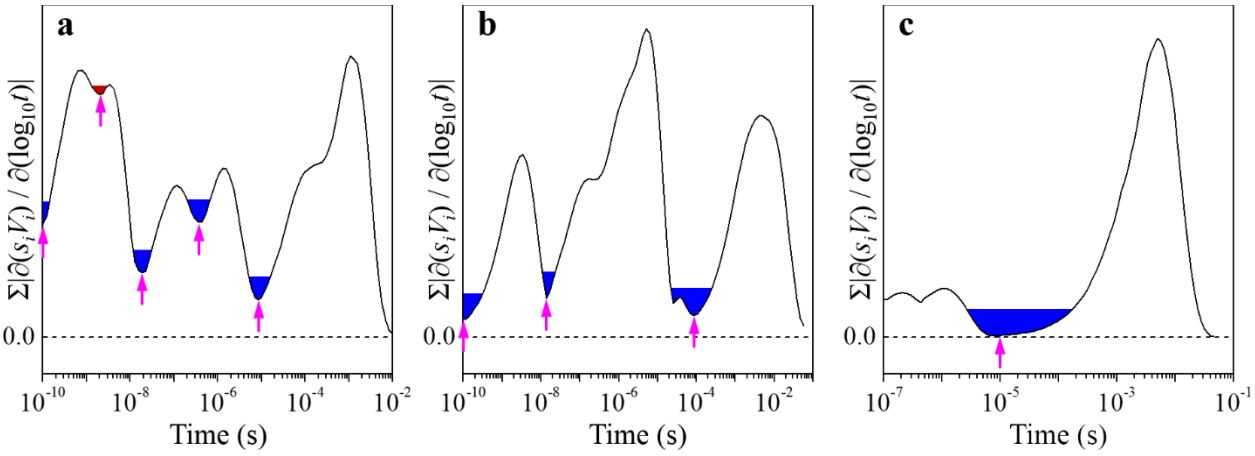
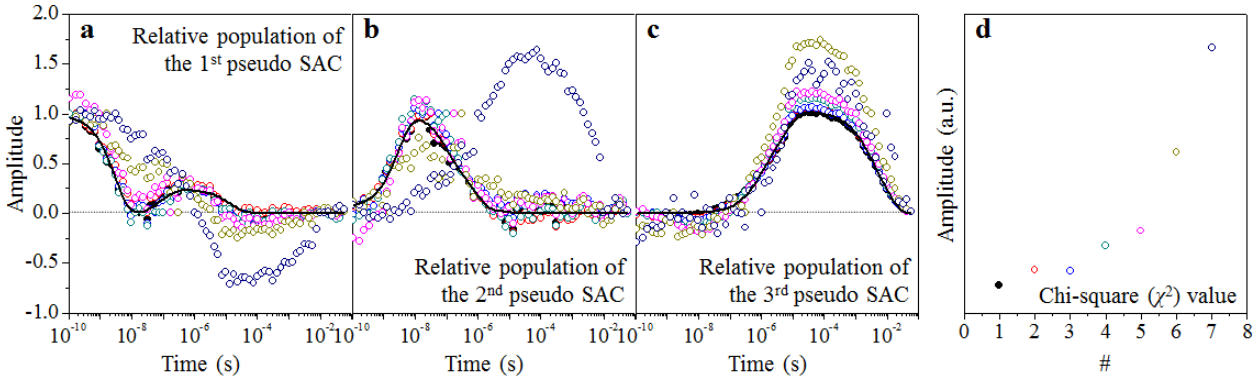


Figure S1. Sum of the absolute values of the first derivatives of the principal rSVs weighted by singular values with respect to $\log_{10}$(time) as the selection criteria of the stationary time zones where the amplitude of all the principal lSVs remain constant. Since the original principal rSVs contain noise, the curves fitted to principal rSVs can be used for the calculation of derivatives and the summation of their absolute values. (a) Four stationary time zones determined for TRXSS data of MbCO. Here we note that we only selected the four lowest wells (blue) for the stationary time zones and excluded the fifth lowest well (red) from the stationary time zones. (b) Three stationary time zones determined for TRXSS data of $HbI(CO)_2$. For the cases of (a) and (b), the number of such time zones can match the number of intermediates (n_p) identified by the SVD analysis in the entire time range, and thus the experimental data in each of the stationary time zones can be directly regarded as the pseudo SAC of each intermediate species (see Figures 3d and 4d in the main text). (c) One stationary time zone determined for TA data of $HbI(CO)_2$ due to limited time resolution of our TA measurement. Accordingly, the pseudo SAC of I_2 intermediate had to be selected from the TA data measured at much later time delay (in this case, 100 ns) than the case of the TRXSS data (17.7 ns) (see Figure 5d in the main text).



#	1st SAC	2nd SAC	3rd SAC
●	$\Delta S_{HbI}(q,108\ ps)$	$\Delta S_{HbI}(q,17.7\ ns)$	$\Delta S_{HbI}(q,100\ \mu s)$
○	$\Delta S_{HbI}(q,167\ ps)$	$\Delta S_{HbI}(q,31.5\ ns)$	$\Delta S_{HbI}(q,178\ \mu s)$
○	$\Delta S_{HbI}(q,352\ ps)$	$\Delta S_{HbI}(q,56.1\ ns)$	$\Delta S_{HbI}(q,316\ \mu s)$
○	$\Delta S_{HbI}(q,492\ ps)$	$\Delta S_{HbI}(q,99.9\ ns)$	$\Delta S_{HbI}(q,562\ \mu s)$
○	$\Delta S_{HbI}(q,930\ ps)$	$\Delta S_{HbI}(q,178\ ns)$	$\Delta S_{HbI}(q,1\ ms)$
○	$\Delta S_{HbI}(q,1.71\ ns)$	$\Delta S_{HbI}(q,316\ ns)$	$\Delta S_{HbI}(q,3.16\ ms)$
○	$\Delta S_{HbI}(q,3.09\ ns)$	$\Delta S_{HbI}(q,1\ \mu s)$	$\Delta S_{HbI}(q,10\ ms)$

Figure S2. Time-dependent relative populations of (a) the first, (b) the second and (c) the third pseudo SACs that were extracted from each of seven different combinations of selecting the pseudo SACs for TRXSS data of $HbI(CO)_2$. Black curves in (a) – (c) represent the fit curves of time-dependent relative populations of the $\Delta S_{HbI}(q,108\ ps)$, $\Delta S_{HbI}(q,17.7\ ns)$ and $\Delta S_{HbI}(q,100\ \mu s)$, respectively, as shown in Figure 4e in the main text. (d) Comparison of minimized χ^2 values for all seven cases. As time zones for selecting the first, the second and the third pseudo SACs deviate from 108 ps, 17.7 ns and 100 μs, respectively, the fit to the experimental data becomes worse with increased χ^2 value.

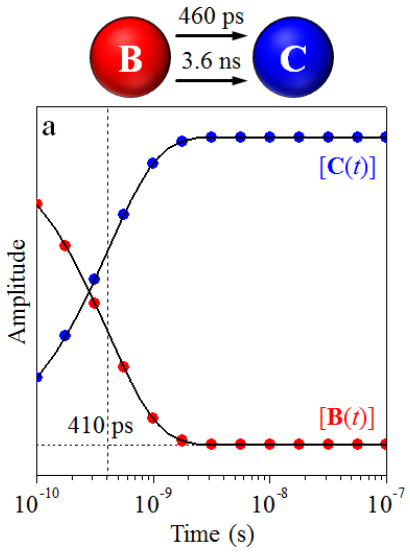


Figure S3. The result of a simulation where an intermediate (termed **B**) transforms to another intermediate (termed **C**) with two relaxation times, Λ_1 (460 ps in this case) and Λ_2 (3.6 ns in this case). The population changes of **B** (red points) and **C** (blue points) are well described by single-exponential fits (black curves) with a relaxation time, $1/(1/\Lambda_1 + 1/\Lambda_2)$ (= 410 ps).

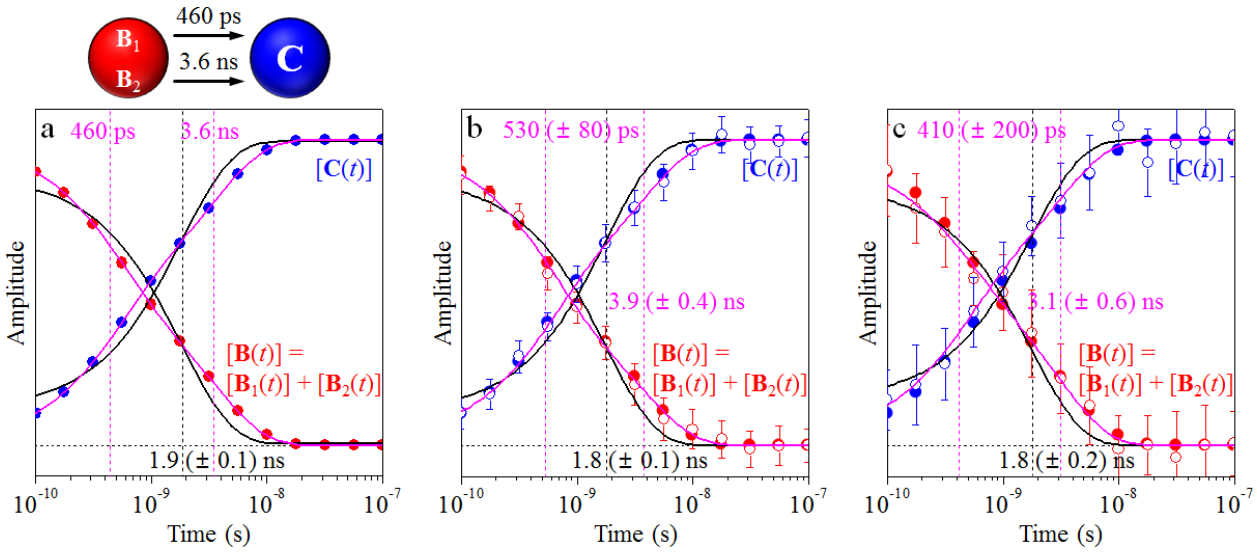


Figure S4. (a) The result of a simulation where an intermediate (termed **B**), which have two conformational substates ($\mathbf{B}_1$ and $\mathbf{B}_2$), transforms to another intermediate (termed **C**) with two relaxation times, Λ_1 (460 ps in this case) for the $\mathbf{B}_1 \rightarrow \mathbf{C}$ transition and Λ_2 (3.6 ns in this case) for the $\mathbf{B}_2 \rightarrow \mathbf{C}$ transition. We assumed that the relative population ratio of $\mathbf{B}_1$ and $\mathbf{B}_2$ is 46:54. The population changes of **B** (red solid circles) and **C** (blue solid circles) are well described by two exponentials (magenta lines). When a single exponential (black line) is used for the fits, the fit quality is much worse. (b and c) The results of simulations identical to the one shown in (a) except the addition of random noise of (b) ±10 % and (c) ±20 % to the population changes of **B** and **C**. The population changes of **B** (red open circles with error bars) and **C** (blue open circles with error bars) are well described by two exponentials (magenta lines), even with the noise included. When a single exponential (black line) is used for the fits, the fit quality is much worse. For comparison, we superimposed the data points for the populations of **B** and **C** without any noise shown in (a).

References

1. K. H. Kim, S. Muniyappan, K. Y. Oang, J. G. Kim, S. Nozawa, T. Sato, S. Y. Koshihara, R. Henning, I. Kosheleva, H. Ki, Y. Kim, T. W. Kim, J. Kim, S. Adachi and H. Ihee, J. Am. Chem. Soc. **134**, 7001 (2012).

2. J. Choi, S. Muniyappan, J. T. Wallis, W. E. Royer, Jr. and H. Ihee, ChemPhysChem **11**, 109 (2010).