Supplementary Computational methods

The $FRET_R$ forward model. The fluorescence intensities measured in 3 different steady-state experiments are needed to define $FRET_R$: in the first two, CFP and YFP are expressed individually and the spillover factors are calculated, in the third CFP and YFP are co-expressed and $FRET_R$ is measured. Therefore, we need a microscopic interpretation of these 3 experiments to write the $FRET_R$ forward model. A detailed derivation is provided in the following paragraphs.

1) YFP individually expressed

A YFP sample excited by an external radiation can return to the ground state through different independent decay pathways: fluorescence and other non-radiative processes, such as dynamic collisional quenching, near-field dipole-dipole interaction, internal conversion, and intersystem crossing. The quantum yield of fluorescence Q_a is defined as the ratio $Q_a = {k_a^F}/{\sum k_a^i}$, where k_a^F is the acceptor fluorescence rate and the sum is over all the decay pathways, each with rate k_a^i. The *master equation* for this system is:

$$\frac{\partial [A^*]}{\partial t} = -\frac{k_a^F}{Q_a}[A^*] + k_a^X([A] - [A^*]),$$

where $[A^*]$ is the concentration of excited acceptors, $[A]$ is the concentration of acceptors, and k_a^X is the excitation rate. The stationary solution for $[A^*]$ is:

$$[A^*] = \frac{k_a^X \cdot [A]}{\frac{k_a^F}{Q_a} + k_a^X},$$

and the fluorescence intensity is:

$$I_a^F = k_a^F \cdot [A^*] = \frac{k_a^F \cdot k_a^X \cdot [A]}{\frac{k_a^F}{Q_a} + k_a^X}.$$

In the limit of rapid de-excitation and slow excitation, $k_a^F \gg k_a^X$:

$$I_a^F = Q_a \cdot k_a^X \cdot [A].$$

This intensity is spread over a broad emission spectrum:

$$I_a^F = \int d\lambda\, I_a^F(\lambda) = (1 + \Lambda_a) \cdot I_a^F\,(\lambda_{em} = 535\text{nm}),$$

where Λ_a is the fraction of the total fluorescence that is emitted at wavelengths not detected by the emission filter used in the experiment (*lost fluorescence*). The emission and excitation wavelengths come from the center of the bandpass filter sets used in our *in vivo* experiments. The acceptor spillover factor S_a is defined as the ratio between the fluorescence intensities measured at λ_{em}=535nm when the sample is illuminated at λ_{ex}=430nm and λ_{ex}=500nm:

$$S_a = \frac{I_a^F(\lambda_{em} = 535\text{nm},\ \lambda_{ex} = 430\text{nm})}{I_a^F(\lambda_{em} = 535\text{nm},\ \lambda_{ex} = 500\text{nm})} = \frac{k_a^{X,430}}{\frac{k_a^F}{Q_a} + k_a^{X,430}} \cdot \frac{\frac{k_a^F}{Q_a} + k_a^{X,500}}{k_a^{X,500}}, \quad (S1)$$

where $k_a^{X,430}$ and $k_a^{X,500}$ are the acceptor excitation rates with incident radiation at λ_{ex}=430nm and λ_{ex}=500nm, respectively. In the limit of rapid de-excitation and slow excitation $k_a^F \gg k_a^X$, the acceptor spillover factor can be approximated by:

$$S_a \simeq \frac{k_a^{X,430}}{k_a^{X,500}}.$$

2) CFP individually expressed

Similarly to the previous experiment, the donor fluorescence intensity I_d^F can be written as:

$$I_d^F = \frac{k_d^F \cdot k_d^X \cdot [D]}{\frac{k_d^F}{Q_d} + k_d^X},$$

where k_d^F is the donor fluorescence rate, k_d^X is the donor excitation rate, Q_d is the donor quantum yield of fluorescence, and $[D]$ is the donor concentration. This fluorescence is spread over a spectrum of emission. The intensities at λ_{em}=470nm and λ_{em}=535nm are measured, and the donor spillover factor is defined as the ratio:

$$S_d = \frac{I_d^F(\lambda_{em} = 535\text{nm}, \lambda_{ex} = 430\text{nm})}{I_d^F(\lambda_{em} = 470\text{nm}, \lambda_{ex} = 430\text{nm})}.$$

The total donor fluorescence can thus be written as:

$$I_d^F = \int d\lambda\, I_d^F(\lambda) = (1 + S_d + \Lambda_d) \cdot I_d^F(\lambda_{em} = 470\text{nm}).$$

where Λ_d is the donor lost fluorescence. The intensities $I_d^F(\lambda_{em} = 470\text{nm})$ and $I_d^F(\lambda_{em} = 535\text{nm})$ can thus be written in terms of the total fluorescence as:

$$I_d^F(\lambda_{em} = 470\text{nm}) = \frac{1}{(1 + S_d + \Lambda_d)} \cdot I_d^F,$$

$$I_d^F(\lambda_{em} = 535\text{nm}) = \frac{S_d}{(1 + S_d + \Lambda_d)} \cdot I_d^F.$$

3) YFP and CFP co-expressed

The master equations of a system of fluorophores that are excited by an external radiation and can transfer energy from an excited donor to a non-excited acceptor is:

$$\frac{\partial [D_i^*]}{\partial t} = -\frac{k_d^F}{Q_d}[D_i^*] + k_d^X([D_i] - [D_i^*]) - \sum_j k_{ij}^{ET}[D_i^*]\left([A_j] - [A_j^*]\right), (S2a)$$

$$\frac{\partial [A_j^*]}{\partial t} = -\frac{k_a^F}{Q_a}[A_j^*] + k_a^X([A_j] - [A_j^*]) + \sum_i k_{ij}^{ET}[D_i^*]\left([A_j] - [A_j^*]\right), \quad (S2b)$$

where $[D_i^*]$ is the concentration of the excited donor *i* and k_{ij}^{ET} is the rate constant of energy transfer between donor *i* and acceptor *j*. Following Förster theory, k_{ij}^{ET} can be written in terms of the distance R_{ij} between the two fluorophores:

$$k_{ij}^{ET} = \frac{k_d^F}{Q_d} \cdot \left(\frac{R_0}{R_{ij}}\right)^6,$$

where R_0 is the Förster radius. In general, this radius depends on the orientation factor κ^2 of the interacting dipoles. Here, we adopt the common assumption that donor and acceptor sample their orientations randomly on the time scale of the measurement, so that[1] κ^2=2/3. The equations above form a system of coupled differential equations. In the limit of rapid de-excitation and slow excitation, it is fair to assume that $([A_j] - [A_j^*]) \sim [A_j]$.

In this regime, the stationary solution for $[D_i^*]$ can be written as:

$$[D_i^*] = \frac{k_d^X \cdot [D_i]}{\frac{k_d^F}{Q_d} \cdot (1 + F_j) + k_d^X},$$

where $F_i = \sum_j \left({R_0}/{R_{ij}}\right)^6 [A_j]$. The total intensity in the donor fluorescence channel is:

$$I_d^F = \sum_i k_d^F \cdot [D_i^*] = \sum_i \frac{k_d^F \cdot k_d^X \cdot [D_i]}{\frac{k_d^F}{Q_d} \cdot (1 + F_i) + k_d^X}.$$

In the limit of rapid de-excitation and slow excitation, $k_d^F \gg k_d^X$:

$$I_d^F = Q_d \cdot k_d^X \cdot \sum_i \frac{1}{1 + F_i} [D_i].$$

The stationary solution for $[A_j^*]$ is:

$$[A_j^*] = \frac{k_a^X \cdot [A_j] + \sum_i \frac{\frac{k_d^F}{Q_d} \cdot \left(\frac{R_0}{R_{ij}}\right)^6 \cdot k_d^X \cdot [D_i] \cdot [A_j]}{\frac{k_d^F}{Q_d} \cdot (1 + F_i) + k_d^X}}{\frac{k_a^F}{Q_a} + k_a^X},$$

and the total intensity in the acceptor fluorescence channel is:

$$I_a^F = \sum_j k_a^F \cdot [A_j^*] = \frac{\sum_j k_a^F \cdot \left\{ k_a^X \cdot [A_j] + \sum_i \frac{\frac{k_d^F}{Q_d} \cdot \left(\frac{R_0}{R_{ij}}\right)^6 \cdot k_d^X \cdot [D_i] \cdot [A_j]}{\frac{k_d^F}{Q_d} \cdot (1 + F_i) + k_d^X} \right\}}{\frac{k_a^F}{Q_a} + k_a^X}.$$

In the limit of rapid de-excitation and slow excitation, $k_d^F \gg k_d^X$ and $k_a^F \gg k_a^X$:

$$I_a^F = Q_a \cdot \left\{ k_a^X \cdot [A] + k_d^X \cdot \sum_i \frac{F_i}{1 + F_i} [D_i] \right\},$$

where $[A]$ is the total concentration of acceptors.

Derivation of the forward model

Using the microscopic quantities defined above, the forward model for $FRET_R$ can be written as:

$$f(X) = \frac{I_{\mathrm{FRET}}}{S_d \cdot I_{\mathrm{CFP}} + S_a \cdot I_{\mathrm{YFP}}} = \frac{\frac{S_d}{1+S_d+\Lambda_d} \cdot Q_d \cdot k_d^{X,430} \cdot \sum_i \frac{1}{1+F_i}[D_i] + \frac{1}{1+\Lambda_a} \cdot Q_a \cdot \left\{ k_a^{X,430} \cdot [A] + k_d^{X,430} \cdot \sum_i \frac{F_i}{1+F_i}[D_i] \right\}}{\frac{S_d}{1+S_d+\Lambda_d} \cdot Q_d \cdot k_d^{X,430} \cdot \sum_i \frac{1}{1+F_i}[D_i] + S_a \cdot \frac{1}{1+\Lambda_a} \cdot Q_a \cdot k_a^{X,500} \cdot [A]} = 1 + \frac{\frac{Q_a \cdot k_d^{X,430}}{1+\Lambda_a} \cdot \sum_i \frac{F_i}{1+F_i}[D_i]}{\frac{S_d \cdot Q_d \cdot k_d^{X,430}}{1+S_d+\Lambda_d} \cdot \sum_i \frac{1}{1+F_i}[D_i] + \frac{Q_a \cdot k_a^{X,430}}{1+\Lambda_a}[A]}. \quad (S3)$$

Eq. S3 can be simplified by introducing the ratio I_{da} of fluorescence intensities in the FRET channel when CFP and YFP are expressed alone and at equal concentrations:

$$I_{da} = \frac{I_d^F(\lambda_{em}=535\mathrm{nm}, \lambda_{ex}=430\mathrm{nm})}{I_a^F(\lambda_{em}=535\mathrm{nm}, \lambda_{ex}=430\mathrm{nm})} = \frac{S_d \cdot Q_d \cdot k_d^{X,430}}{1+S_d+\Lambda_d} \cdot \frac{1+\Lambda_a}{Q_a \cdot k_a^{X,430}}. \quad (S4)$$

This quantity can be calculated from the data collected in the two experiments measuring the spillover factors S_d and S_a. After substituting I_{da} from Eq. S4 in Eq. S3, the forward model can be written as:

$$f(X, I_{da}, k_{da}) = 1 + \frac{k_{da} \cdot \{[D] - g(X)\}}{I_{da} \cdot g(X) + [A]},$$

where $k_{da} = {k_d^{X,430}}\Big/{k_a^{X,430}}$ and $g(X) = \sum_i \frac{1}{1+F_i}[D_i]$.

Model of linker flexibility. Typically, fluorescent proteins are attached to the N- or C-terminus of a protein by means of a flexible linker. As a consequence, conformational heterogeneity in the complex can be present even when the conformations, positions, and orientations of the two tagged proteins remain relatively fixed. In this situation, the multi-state forward model of Eq. 4 (Online Methods**)** is still applicable, provided that the fluorescent proteins and linkers are explicitly represented. However, in the single-molecule case and when the structures of the tagged proteins are known and kept rigid, the forward model can be extended to account directly for the linker flexibility, without representing the fluorescent proteins explicitly. The probability distribution $p(r_{FT})$ as a function of the distance r_{FT} between a fluorophore and the terminal residue of the linker was first calculated with MD simulations of a GFP and linker alone (below). Each protein of the complex was then taken individually and probes representing possible fluorophore positions were distributed close to either the N or C terminus, according to $p(r_{FT})$. Probes that clashed with the protein structure were not retained. Finally, the probability of having the fluorophore in position x_F when the tagged protein is in state X was estimated by fitting the positions of the probes with a Gaussian Mixture Model (GMM):

$$p(x_F|X) = \sum_{i=1}^{N_G} \omega_i \, G(x_F|x_i, \sigma_i),$$

where G is a normalized Gaussian centered in x_i with standard deviation equal to σ_i, ω_i is the weight of the *i*-th GMM component, N_G is the number of components, and $\sum_{i=1}^{N_G} \omega_i = 1$. When calculating $FRET_R$ for a pair of tagged proteins in a given position (X^d, X^a), we assume that the two attached fluorescent proteins populate all the positions allowed by the linkers. The average over multiple conformations of Eq. 4 (Online Methods**)** can thus be written as:

$$\sum_k w_k g(X_k) = \int dx_F^d \; p\left(x_F^d \middle| X^d\right) \int dx_F^a \; p(x_F^a | X^a) \; \frac{1}{1 + \left({R_0}/{\| \; x_F^d - x_F^a \; \|}\right)^6}$$

$$= \sum_{i=1}^{N_G^d} \omega_i \sum_{j=1}^{N_G^a} \omega_j \int dr_F \; p(r_F | r_{ij}, \sigma_{ij}) \frac{1}{1 + \left({R_0}/{r_F}\right)^6},$$

where $r_F = \| \; x_F^d - x_F^a \; \|$, $r_{ij} = \| \; x_i - x_j \; \|$, $\sigma_{ij} = \sqrt{\sigma_i^2 + \sigma_j^2}$, and [2,3]

$$p\left(r_F \middle| r_{ij}, \sigma_{ij}\right) = \frac{r_F}{\sqrt{2\pi} \; \sigma_{ij} \; r_{ij}} \exp\left(-\frac{r_F^2 \; + r_{ij}^2}{2\sigma_{ij}^2}\right) \sinh\left(\frac{r_F \; r_{ij}}{\sigma_{ij}^2}\right).$$

Provided that the tagged proteins are kept rigid, the positions of the GMM centers $\{x_i\}$ and $\{x_j\}$ can be transformed along with the coordinates X^d and X^a and no further GMM fit is required during sampling. Furthermore, the integral in the above equation was evaluated numerically and tabulated for different values of r_{ij} and σ_{ij} prior to sampling.

Multi-state Bayesian scoring function. The multi-states posterior probability can be written as:

$$p(\{X_k, w_k\}, I_{da}, k_{da}, \{\sigma_n\}|\{d_n\})$$
$$\propto p\left(I_{da}\middle|I_{da}^{exp}, \sigma_{da}^{exp}\right)p(k_{da})\prod_{k=1}^{N}p(X_k)p(w_k)\prod_{n=1}^{N_F}p(\sigma_n|\sigma_0)p(d_n|\{X_k, w_k\}, I_{da}, k_{da}, \sigma_n).$$

Finally, to reduce the number of parameters, it is useful to marginalize all σ_n:

$$p(\{X_k, w_k\}, I_{da}, k_{da}, \sigma_0|\{d_n\})$$
$$\propto p\left(I_{da}\middle|I_{da}^{exp}, \sigma_{da}^{exp}\right)p(k_{da})\prod_{k=1}^{N}p(X_k)p(w_k)\prod_{n=1}^{N_F}\int d\sigma_n\, p(\sigma_n|\sigma_0)p(d_n|\{X_k, w_k\}, I_{da}, k_{da}, \sigma_n)$$
$$= p\left(I_{da}\middle|I_{da}^{exp}, \sigma_{da}^{exp}\right)p(k_{da})\prod_{k=1}^{N}p(X_k)p(w_k)\prod_{n=1}^{N_F}p(d_n|\{X_k, w_k\}, I_{da}, k_{da}, \sigma_0),$$

where the marginal likelihood is:

$$p(d_n|\{X_k, w_k\}, I_{da}, k_{da}, \sigma_0) = \frac{\sqrt{2}\sigma_0}{\pi d_n}\cdot\frac{1}{\log(d_n/f(\{X_k, w_k\}, I_{da}, k_{da}))^2 + 2\sigma_0^2}.$$

Supplementary Experimental Methods

In vivo $FRET_R$ measurements *S. cerevisiae* strain BSY9 (MAT**a**/MATα , ade2-1oc/ade2-1oc, ADE3/ade3Δ100, can1-100/can1-100, CYH2s/cyh2r, his3-11,15/his3-11,15, leu2-3,112/leu2-3,112, trp1-1/ trp1-1, ura3-1/ura3-1) was the host of a series of plasmids that were integrated at either the URA3 or LEU2 locus. Descriptions of the plasmids, along with their Genbank accession numbers, are given in supplemental Table S1. All plasmids are yeast integrative plasmids based on either pRS305 (GenBank: U03437.1) or pRS306 (GenBank: U03438.1).

Microscopy and FRET analysis were performed as described in [4]. In brief, images were captured using a Deltavision microscope equipped with CFP/YFP filter set 89002-ET from Chroma Technology Corp and a mercury HBO 100W light source. Independent emission and excitation filter wheels were used to position filters. Images were captured on a Coolsnap HQ camera with 0.08 or 0.1 second exposure times. The order of image acquisition is important and all images were captured in the sequence YFP, then FRET and finally the CFP channel. A DIC image was captured at the end of the experiment to provide cellular context to the fluorescence images. Images were then analyzed for FRET using Fretscal. Fretscal is an integrated set of custom Matlab scripts and is available online at no cost at the MATLAB Central file exchange, http://www.mathworks.com/matlabcentral/.

References

1 Stryer, L. Fluorescence Energy-Transfer as a Spectroscopic Ruler. *Annu Rev Biochem* **47**, 819-846, (1978).

2 Churchman, L. S., Flyvbjerg, H. & Spudich, J. A. A non-Gaussian distribution quantifies distances measured with fluorescence localization techniques. *Biophys J* **90**, 668-671, (2006).

3 Kohler, S. D., Spitzbarth, M., Diederichs, K., Exner, T. E. & Drescher, M. A short note on the analysis of distance measurements by electron paramagnetic resonance. *J Magn Reson* **208**, 167-170, (2011).

4 Kollman, J. M. *et al.* The structure of the gamma-tubulin small complex: implications of its architecture and flexibility for microtubule nucleation. *Mol Biol Cell* **19**, 207-215, (2008).

5 Laio, A. & Parrinello, M. Escaping free-energy minima. *Proc Natl Acad Sci U S A* **99**, 12562-12566, (2002).

6 Barducci, A., Bonomi, M. & Parrinello, M. Metadynamics. *Wires Comput Mol Sci* **1**, 826-843, (2011).

7 Bussi, G., Gervasio, F. L., Laio, A. & Parrinello, M. Free-energy landscape for beta hairpin folding from combined parallel tempering and metadynamics. *J Am Chem Soc* **128**, 13435-13441, (2006).

8 Sugita, Y. & Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett* **314**, 141-151, (1999).

9 Bonomi, M., Barducci, A. & Parrinello, M. Reconstructing the Equilibrium Boltzmann Distribution from Well-Tempered Metadynamics. *J Comput Chem* **30**, 1615-1621, (2009).

Figure Legends

Fig. S1. Bayesian model of $FRET_R$ data. (A) $FRET_R$ calculated from the forward model of Eq. 1 (Online Methods) as a function of the distance between donor and acceptor. $FRET_R$ is calculated for different values of the ratio between donor and acceptor excitation rates $\boldsymbol{k_{da}}$ and the ratio of donor and acceptor intensities in the FRET channel $\boldsymbol{I_{da}}$, measured when CFP and YFP are expressed separately. (B) Negative logarithm of the marginal likelihood function (Eq. 7 Online Methods) for 3 values of the uncertainty $\boldsymbol{\sigma_0}$, plotted against the $FRET_R$ value of data point d_n and given a prediction by the forward model of a $FRET_R$ value of 2.0. (C) The unimodal distribution used as the prior for uncertainty $\boldsymbol{\sigma_n}$ (Eq. 6 Online Methods) for 3 values of $\boldsymbol{\sigma_0}$.

Fig. S2. Accuracy of the $FRET_R$ forward model. The relative deviation of the $FRET_R$ forward model values from KMC simulations is calculated as a function of the parameters k_d^X/k_d^F and k_{da}, for systems of (A) one CFP-one YFP, (B) two CFP-one YFP, and (C) one CFP-two YFP. The relative deviation is averaged over a distance range from 3 to 10 nm. The YFP photobleaching ratio during the acquisition of the CFP channel is set to 0.3.

Fig. S3. $FRET_R$ from KMC simulations of systems of (A) two CFP-one YFP and (B) one CFP-two YFP, as a function of the distances between fluorophores. k_{da} is set to 8.0, k_d^X/k_d^F to 0.001. The YFP photobleaching ratio during the acquisition of the CFP channel is set to 0.3. (C) The ratio of $FRET_R$ of the two CFP-one YFP to the one CFP-two YFP systems, calculated from the forward model with donor-acceptor distances set both to R_0, as a function of k_{da} and I_{da}.

Fig. S4. Effect of YFP photobleaching on $FRET_R$. $FRET_R$ calculated from KMC simulations of a system of one CFP-two YFP as a function of the distances between fluorophores, when (A) YFP are not photobleached during the acquisition of the CFP channel, and (B) the YFP photobleaching ratio is equal to 0.3. The values of k_{da} and

k_d^X/k_d^F are the same as in Fig. S3. (C) $FRET_R$ of the photobleached system calculated from the multi-state forward model of Eq. 4 (Online Methods**)**. The average relative deviation from the KMC simulations of panel (B) is 0.8%**.**

Fig. S5. Free energies of a series of CFP-YFP pairs separated by polyproline peptides of length 5 (A), 10 (B), 15 (C), and 20 (D). The free energies are calculated by all-atom molecular dynamics simulations as a function of $\Omega = \sum_{i=1}^{n-1} \cos \omega_i$, where ω is the torsional angle formed by the proline quadruplet Cα-C-N-Cα. ω is equal to 0˚ and 180˚ for the *cis* and *trans* proline isomers, respectively. For a peptide of length *n*, Ω=-(*n*-1) indicates an ideal left-handed polyproline II helix, Ω=*n*-1 an ideal right-handed polyproline I helix. Sampling is accelerated by a combination of Parallel Tempering and metadynamics[5-7].

Fig. S6. Free energy of a CFP-YFP pair separated by a flexible linker as a function of the distance between the centers of the two fluorophores (x axis) and the angle formed by the major axes of inertia of the two fluorescent proteins (y axis). The free energy is calculated by all-atom molecular dynamics simulations, accelerated by Parallel Tempering[8]. Isoenergy lines are drawn every 1 k_BT.

Fig. S7. Assessment of sampling convergence. (AB) Convergence of the MD simulations of YFP–CFP pairs separated by (A) polyproline peptides of different length and (B) a flexible linker or other proteins of known structure. The weighted average $\langle g(X) \rangle$ needed by the multi-state forward model of Eq. 4 (Online Methods**)** was calculated as the ensemble average of $g(X)$ over the MD conformers, with R_0=4.9nm, after discarding the first 10% of the trajectory. To assess the convergence, $\langle g(X) \rangle$ is calculated and plotted as a function of the simulation time. A reweighting algorithm[9] was used to obtain from the biased PTMetaD simulations canonical averages. (CD) Convergence of the benchmark accuracy as a function of the number of tests. Accuracy is defined as the average Cα dRMS between the crystallographic structure and the most

probable model, calculated (C) on the entire complex and (D) on the N- and C-terminal residues. Results are with low-noise data and using 100% of the data points.

Fig. S8. *In vivo* experimental validation of the forward model. $FRET_R$ values measured on 5 systems of defined structure (black bars) expressed in *S. cerevisiae* are compared to the values predicted by the forward model (white bars), using the model of linker flexibility. The other 4 data points of Fig. 3 (polyproline series) were not used for this comparison, because in these systems heterogeneity was dominated by the presence of conformations with different patterns of proline isomers, rather than the linker flexibility. Red lines indicate experimental and model errors.

Fig. S9. Benchmark flowchart. (A) A flowchart for benchmarking the accuracy of our Bayesian approach to determining the molecular architecture of a complex from synthetic $FRET_R$ data. The benchmark contains 16 protein complexes. The stages include target selection, system representation, synthetic data generation, sampling, and analysis. (B) Schematics of the sampling algorithm based on a Gibbs sampling Monte Carlo scheme coupled with Simulated Annealing. The flowchart depicts a typical step of our sampling scheme in which random changes of the position and orientation of each subunit, the values of the forward model, and likelihood parameters are proposed, and either rejected or accepted based on the Metropolis criterion. The temperature of the system is cyclically varied from 1.0 to 5.0 $k_B T$ to avoid trapping the system in local energy minima, thus enhancing the sampling.