Science Advances
AAAS

Supplementary Materials for

Functional protein dynamics on uncharted time scales detected by nanoparticle-assisted NMR spin relaxation

Mouzhe Xie, Lei Yu, Lei Bruschweiler-Li, Xinyao Xiang, Alexandar L. Hansen, Rafael Brüschweiler*

*Corresponding author. Email: bruschweiler.1@osu.edu

This PDF file includes:

Supplementary Materials and Methods

Supplementary Materials and Methods

Sample Preparation

The DNA fragment encoding Im7 was PCR-amplified and subcloned into a pTBSG ligation independent cloning vector derivative (pTBSG1) (*49*). The resulting plasmid pTBSG1_Im7 was then transformed into *Escherichia coli* BL21(DE3) strain for protein overexpression. The expressed fusion protein contains a His_6-tag and a TEV protease cleavage site N-terminal to Im7. The overexpression was carried out as following: a single colony was inoculated to 20 mL LB media under vigorous shaking of 250 rpm overnight at 37 °C, the overnight culture was then transferred into 1L M9 minimal media with 1g $^{15}NH_4Cl$ and 5g D-glucose (or 4g D-glucose-$^{13}C_6$ for ^{13}C-labeled samples) as the sole nitrogen/carbon sources and incubated at 37°C under vigorous shaking. When OD_{600} of the culture reached 0.8–1.0, isopropyl *β*-D-1-thiogalactopyranoside (IPTG) was added to it to the final concentration of 0.5mM, and further incubated at 25°C under vigorous shaking for 18 hours. After overexpression, the cells were then pelleted by centrifugation, and lyzed by EmulsiFlex-C5 homogenizer (AVESTIN, Inc.) The cell lysate was subjected to centrifugation at 20,000 xg for 20 minutes. The His_6-tagged Im7 protein in the supernatant was purified by a Ni-NTA agarose (QIAGEN) affinity column and mixed with tobacco etch virus (TEV) protease for His_6-tag cleavage. The final Im7 protein without the His_6-tag can be separated from the His_6-tag and the His_6-tag TEV protease with a second Ni-NTA affinity column, and collected in the flow through, and was concentrated in 50 mM sodium phosphate buffer (pH 7.0) for future use. The resulting Im7 protein contains three non-native residues (SNA) at its N-terminus.

The canine sodium-calcium exchanger NCX1 (CBD1, residues 371–509) was expressed from a pET23b vector in *Escherichia coli* BL21(DE3) (*26*). The CBD1 protein in this study contains eight non-native residues (MSHHHHHH) at its N-terminus. The protein was buffer exchanged to 20 mM HEPES, 100 mM NaCl, 15 mM EDTA, and 20 mM β-mercaptoethanol at pH 7.0.

The primary sequences of Im7 and CBD1 used here are as follows.

Im7:

SNA MELKN SISDY TEAEF VQLLK EIEKE NVAAT DDVLD VLLEH FVKIT EHPDG
    TDLIY YPSDN RDDSP EGIVK EIKEW RAANG KPGFK QG

CBD1:

MSHHH HHHVS KIFFE QGTYQ CLENC GTVAL TIIRR GGDLT NTVFV DFRTE
DGTAN AGSDY EFTEG TVVFK PGETQ KEIRV GIIDD DIFEE DENFL VHLSN
VKVSS EASED GILEA NHVSA LACLG SPSTA TVTIF DDDHA GIFTF EE

(The 9th residue corresponds to V371 according to the conventional residue numbering in full length NCX1 protein.)

Nanoparticles

Bindzil 2040 colloidal silica nanoparticles (SNPs) with a 20 nm diameter were obtained from AkzoNobel and were characterized previously (*50*). They were dialyzed with 50 mM sodium phosphate buffer (pH 7.0) for Im7 and 20 mM HEPES buffer (pH 7.0) for CBD1 and a semi-membrane with 10 kDa molecular weight cut-off purchased from Spectrum Labs. The dialyzed SNPs and freshly purified protein were mixed to prepare the NMR samples. All NMR samples contained 5-10% D_2O as field-frequency lock. The samples were stable over the entire course of NMR data acquisition.

NMR Spectroscopy

NMR experiments were performed on Bruker Ascend spectrometers operating at 850 MHz ^{1}H frequency (19.97 T). A standard suite of triple-resonance correlation spectra was collected for Im7 and CBD1 to obtain chemical shift assignments, including 3D HNCO, HNCA, HN(CO)CA, and CBCA(CO)NH experiments, to confirm previously published assignments (*26, 27*). [15]N spin relaxation rates (R_1 and $R_{1\rho}$) for protein samples both in the absence and presence of SNPs, as

well as a {^{1}H}–^{15}N steady-state nuclear Overhauser effect experiment in the absence of SNPs, were measured using standard ^{15}N relaxation experiments (*43, 44*). Recovery delays were 1.5 s for R_1 and $R_{1\rho}$ experiments and 8 s for the NOE experiment. For $R_{1\rho}$ measurements, ^{15}N magnetization was locked along the effective field using adiabatic half passages (*51*) with a spinlock field strength of approximately 2050 Hz and was calibrated according to (*52*). Examples of relaxation delay intervals are as follows: R_1 = [0.04, 0.16×2, 0.40, 0.52×2, 0.64 s]; $R_{1\rho}$ = [2, 10×2, 18, 26, 34, 42×2 ms] (duplicate delays are indicated with "×2"). The transverse spin relaxation rates (R_2) were then extracted according to

$$R_2 = R_{1\rho} / \sin^2 \theta - R_1 / \tan^2 \theta \qquad \text{(S1)}$$

where θ is the tilt angle in the rotating frame, defined by the arctan of the ratio of the above-mentioned spinlock field, ω_1, and the offset of the corresponding resonance from the carrier frequency, $\Delta\Omega$, i.e. $\theta = \arctan(\omega_1/\Delta\Omega)$. NMR data were processed with NMRPipe (*53*) and visualized with Sparky (*54*). Statistical experimental errors from both the peak-intensity fitting uncertainty and the exponent fitting uncertainty were propagated and displayed as error bars. All Im7 data were collected at 298 K and all CBD1 data at 306 K.

NMR Spin Relaxation Expression for R_2

The following standard expression (*55-58*) was used for the NMR spin relaxation parameters R_2, which are the inverse of the transverse relaxation times T_2 (R_2 = 1/ T_2) for Fig. 1B and model-free analysis.

$$R_2 = d_{00}[4J(0) + 3J(\omega_N) + J(\omega_H - \omega_N) + 6J(\omega_H) + 6J(\omega_H + \omega_N)] + \frac{1}{3}c_{00}\omega_N^2[4J(0) + 3J(\omega_N)] \quad \text{(S2)}$$

where $d_{00} = \dfrac{1}{20}(\dfrac{\mu_0}{4\pi})^2(\dfrac{h}{2\pi})^2\gamma_H^2\gamma_N^2\langle r_{NH}^{-3}\rangle^2$ and $c_{00} = \dfrac{1}{15}\Delta\sigma^2$. $J(\omega)$ is given in Eq. (S3). μ_0 is the permeability of vacuum, h is Planck's constant, γ_H and γ_H are the gyromagnetic ratios of ^{1}H and ^{15}N, and $r_{NH} = 1.02$ Å is the backbone N-H bond length. The ^{15}N chemical shift anisotropy was set to $\Delta\sigma = -160$ ppm. Analogous expressions exist for the longitudinal relaxation rate R_1 and heteronuclear $\{^1$H$\}$-^{15}N hetNOE (see e.g. (28)). Constant c of the main text (Eq. (2)) is

$c = 4d_{00} + \dfrac{4}{3}c_{00}\omega_N^2$. For the illustration of the sensitivity of ΔR_2 on the motional timescale in Fig.

1B, the bound population was set to $p = 0.01$ and the nanoparticle rotational tumbling correlation time was set to $\tau_{NP} = 5$ µs. For the theoretical discussion of the sensitivity of ΔR_2 on the motional timescales (Eqs. (1) – (3) in the main text), Eq. (S2) was simplified by using $J(0) >> J(\omega_N)$, $J(\omega_H + \omega_N)$, $J(\omega_H)$, $J(\omega_H - \omega_N)$, which applies in excellent approximation for globular proteins at high magnetic fields both in the presence and absence of NPs provided that $S^2 > 0$.

Model-free Analysis

Model-free (MF) analysis was performed using an isotropic tumbling model with the following spectral density function (*19, 59, 60*)

$$J(\omega) = S_{fast}^2 S_{slow}^2 \dfrac{\tau_P}{1+\omega^2\tau_P^2} + S_{fast}^2(1-S_{slow}^2)\dfrac{\tau_{i,P}}{1+\omega^2\tau_{i,P}^2}$$

$$\tau_{i,P}^{-1} = \tau_P^{-1} + \tau_i^{-1}$$

(S3)

S_{fast}^2 and S_{slow}^2 are generalized order parameters corresponding to internal motions that occur on fast and slow timescales, respectively. The total order parameter S^2(MF) (see main text) is then the product $S^2(MF) = S_{fast}^2 S_{slow}^2$. The internal correlation time belonging to S_{fast}^2 approaches zero, while S_{slow}^2 has an internal correlation time τ_i. Statistical errors in the model-free parameters

were estimated by Monte Carlo analysis. Consistent with previous studies (*22, 26*), the model-free analysis gives τ_P values of 6.05 ± 0.04 ns for Im7 and 11.16 ± 0.06 ns for CBD1.

Rotational Correlation Time of Silica Nanoparticles

The rotational correlation time of the nanoparticles can be estimated from the Stokes-Einstein-Debye relationship

$$\tau_{NP} = \frac{V\eta}{k_B T} \tag{S4}$$

where $V = 4\pi r_{NP}^3/3$ is the hydrodynamic volume of a spherical nanoparticle of radius r_{NP}, η is the shear viscosity of the water solvent, k_B is the Boltzmann constant, and T is the absolute temperature. With $r_{NP} = 10$ nm, $\tau_{NP} = 0.91$ μs for the experimental conditions of Im7 ($T = 298$ K, $\eta = 0.890$ mPa·s) and $\tau_{NP} = 0.74$ μs for the experimental conditions of CBD1 ($T = 306$ K, $\eta = 0.749$ mPa·s).

Limiting cases of slow and fast exchange

Slow exchange rates k_{ex} between proteins and the nanoparticles do not affect the validity of Eqs. (1) – (3) provided that one does not reach the coalescence regime (because of chemical shift modulations caused by the exchange). If $R_2^{bound} - R_2^{free} > k_{ex}$, relaxation decay is biexponential, although the fast component would be hard to observe (because $p << 1 - p$) leading to $\Delta R_2 = 0$ (at the same time, coalescence effects may also appear) and better nanoparticle conditions would need to be identified.

In the other extreme, i.e. for very large k_{ex} that approach $1/\tau_P$, one gradually approaches the situation where the system adopts an effective tumbling rate $1/\tau_{eff} = p/\tau_{NP} + (1-p)/\tau_P \cong 1/\tau_P$

again leading to $\Delta R_2 = 0$. This situation only applies for very high k_{ex} that are unrealistic for most protein-nanoparticle interactions.

Molecular Dynamics Simulations

Molecular dynamics (MD) simulations were performed using the GROMACS 5.1.2 package (*45*). The initial structures of Im7 and CBD1 were built based on crystal structures (PDB codes: 1CEI and 2DPK, respectively) and missing residues were reconstructed using MODELLER (*61*). The AMBER ff99SBnmr1 protein force field (*46*) together with the TIP3P explicit water model (*47*) were used. The integration time step was set to 2 fs with all bond lengths restrained involving hydrogen atoms by the LINCS algorithm. A cubic simulation box that extends 8 Å from the protein surface was used, and periodic boundary conditions were applied in all three spatial dimensions. Na^+ ions were added to neutralize the total charge of the system. A cutoff of 10 Å was used for van der Waals and electrostatic interactions. Particle-Mesh Ewald summation with a grid spacing of 1.2 Å was used to calculate long-range electrostatic interactions. After 50,000 steps of steepest descent energy minimization, the system was simulated for 100 ps at a constant temperature of 300 K and constant volume with all protein heavy atoms positionally restrained. The positional restraints were removed for the next 100 ps while the pressure was coupled to 1 atm. The production run was performed in the NPT ensemble at 300 K and 1 atm for 1 μs. Amide order parameters S^2(MD) were back-calculated from MD trajectories using the isotropic reorientational eigenmode dynamics (iRED) method with varying lengths of the time averaging window (*48*).

Principal Component Analysis

Principal component analysis (PCA) was performed on the backbone dihedral angles of Loop I for Im7 (Val27–Asp31) and Loop E-F of CBD1 (Ille445–Glu454) where each dihedral angle was represented in the complex plane (*29, 62*). Specifically, the N dihedral angles φ_n of each loop were represented as points $z_n = e^{i\varphi_n}$ on the unit circle in the complex plane to circumvent the

modulo 2π ambiguity of φ_n. Each MD snapshot at time t was then specified by a complex vector $\mathbf{z}(t)$

$$\mathbf{z}(t) = (z_1, z_2, ..., z_N)^{\mathrm{T}} \quad z_n = e^{i\varphi_n} \quad (n = 1, 2, ..., N) \tag{S5}$$

The (complex) covariance matrix $\mathbf{C}$ was then constructed with elements given by

$$C_{mn} = \left\langle (z_m - \langle z_m \rangle)(z_n^* - \langle z_n^* \rangle) \right\rangle \tag{S6}$$

Next, a principal component analysis was applied to matrix $\mathbf{C}$ by solving the following eigenvalue problem, where $\mathbf{v}_k$ is complex principal component k with real eigenvalue λ_k

$$\mathbf{C}\mathbf{v}_k = \lambda_k \mathbf{v}_k \quad \lambda_1 \geq \lambda_2 \geq \ldots \lambda_N \tag{S7}$$

Finally, each conformer $\mathbf{z}(t)$ was projected along eigenmode $\mathbf{v}_k$, yielding projection coefficients that are generally complex

$$c_k(t) = \langle \mathbf{v}_k | \mathbf{z} \rangle = \mathbf{v}_k^\dagger \mathbf{z}(t) \tag{S8}$$

The real and imaginary parts of the coefficients from the first principal component (PC1) were used to produce the score plots shown in Fig. 4 of the main text.
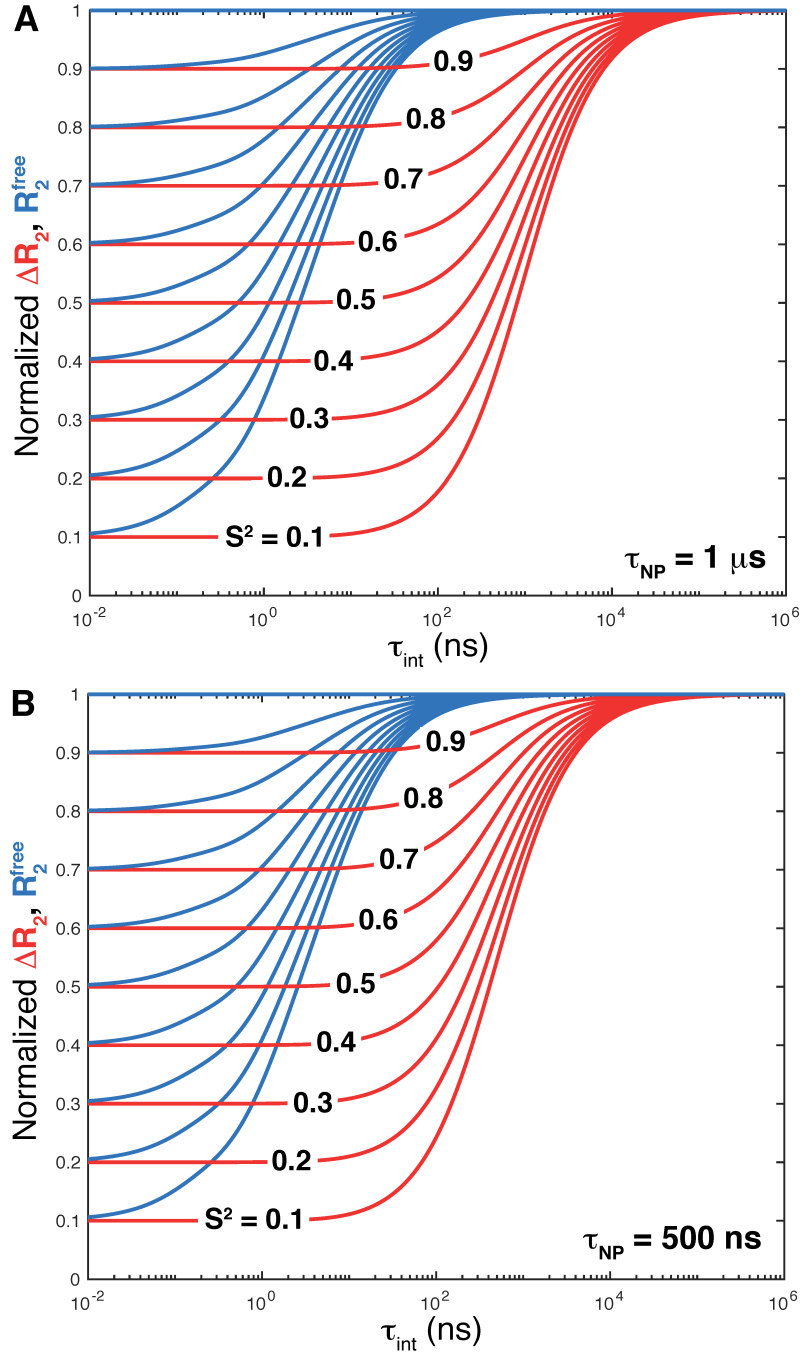
Fig. S1. Simulated dependence of R_2^{free} **and** ΔR_2 **on internal correlation time** τ_i **and** S^2 **order parameter.** R_2^{free} is the transverse relaxation rate in the absence of NPs (blue) and $\Delta R_2 = R_2^{NP} - R_2^{free}$ in the presence and absence of NPs (red). It demonstrates how dynamics on a wide range of timescales is sensitively probed by ΔR_2, which significantly exceeds the range probed by traditional model-free analysis in the absence of NPs. The tumbling correlation time of NPs was set to (A) 1 μs and (B) 500 ns. For better visualization of the difference between S^2(MF) and $S^2(\Delta R_2)$ parameters, the blue and red curves were normalized by linear scaling so that their maximal values are 1.0.
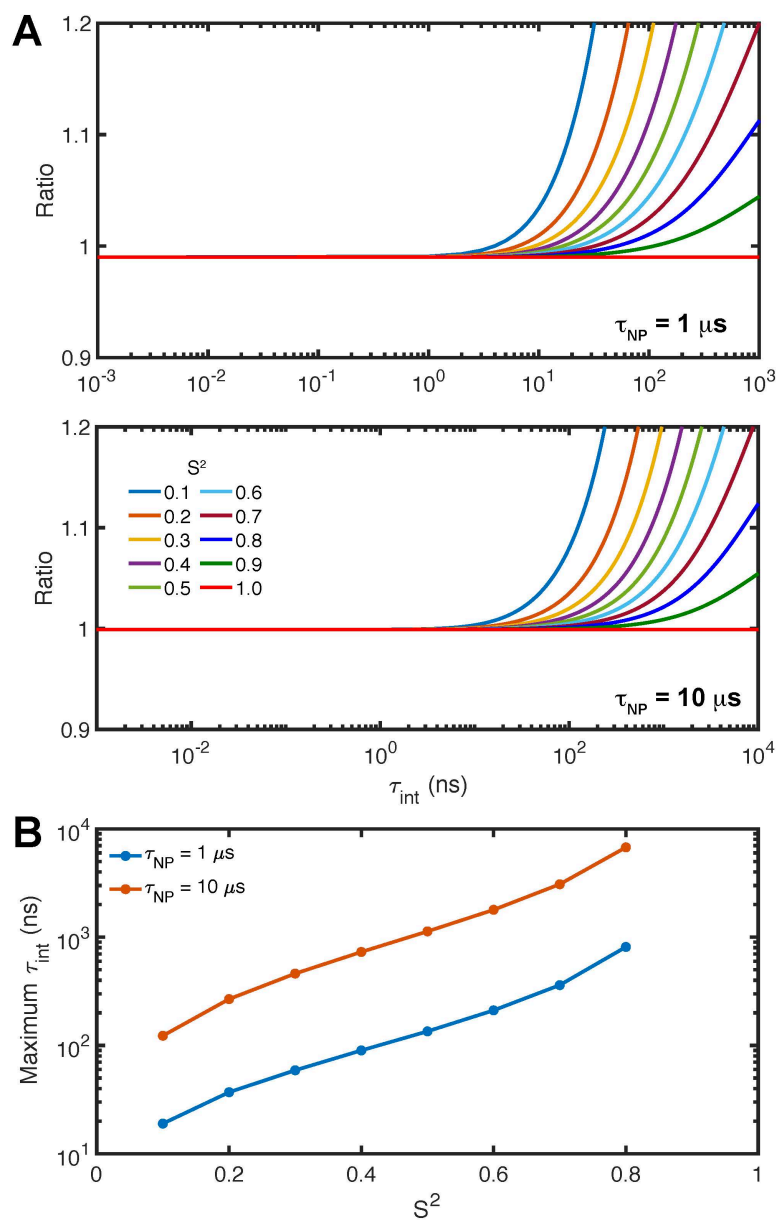
Fig. S2. Range of validity of Eq. 3 for the extraction of S^2 from ΔR_2. (**A**) Simulated dependence of the ratio of the exact over the approximated ΔR_2 (Eq. (3)) on the internal correlation time τ_i and the S^2 order parameter. The tumbling correlation time of the protein was set to 10 ns and the tumbling correlation time of NPs (τ_{NP}) was set to 1 μs and 10 μs. For very long internal correlation times $\tau_i \gg \tau_{NP}$, the ratio approaches $1/S^2$. (**B**) Slowest internal correlation time for which S^2 can be extracted within 10% error. The larger the nanoparticles and the higher S^2, the more sensitive and more accurate is the extraction of $S^2(\Delta R_2)$ for slow internal correlation times.
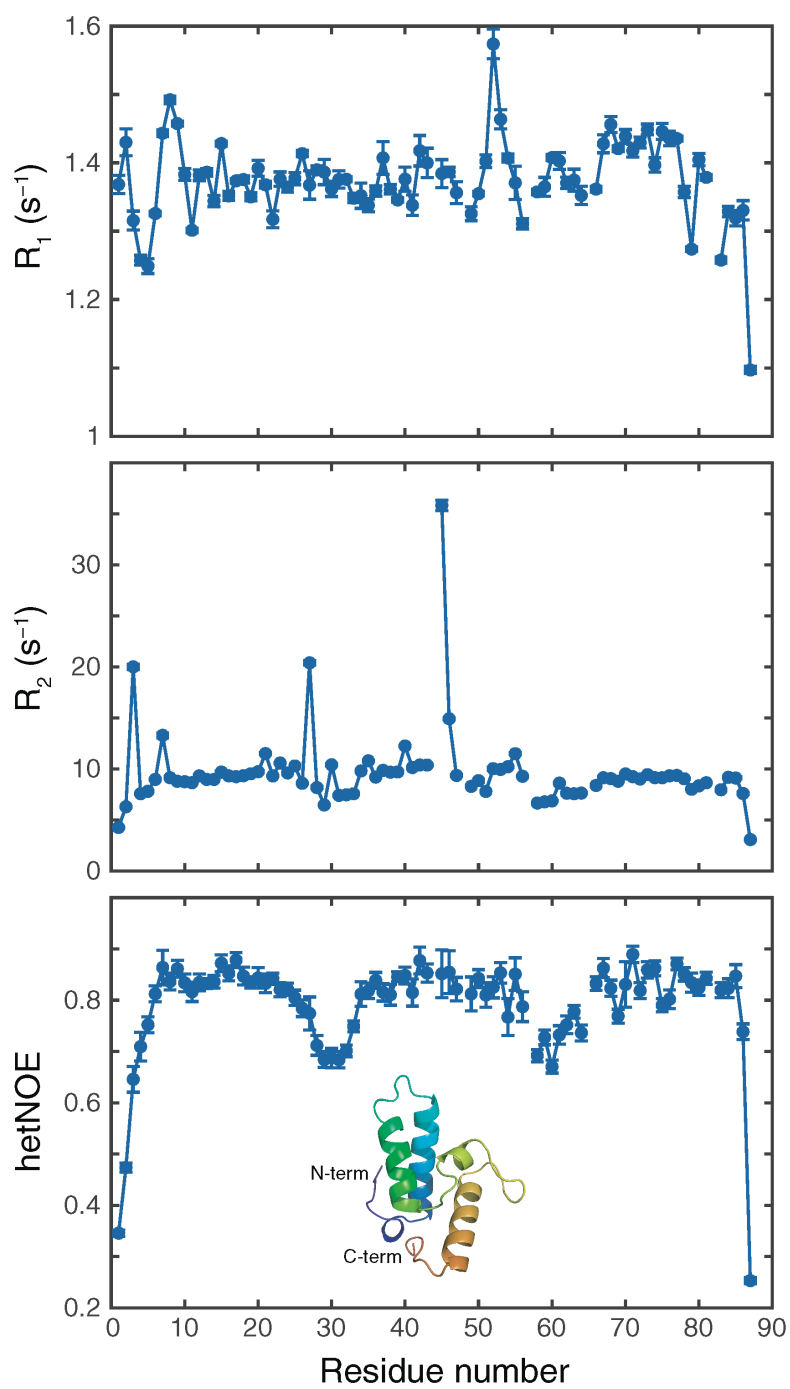
Fig. S3. Experimental ^{15}N spin relaxation parameters of Im7 in the absence of NPs. R_1, R_2, and hetNOE were measured at 850 MHz ^{1}H frequency and 25°C.
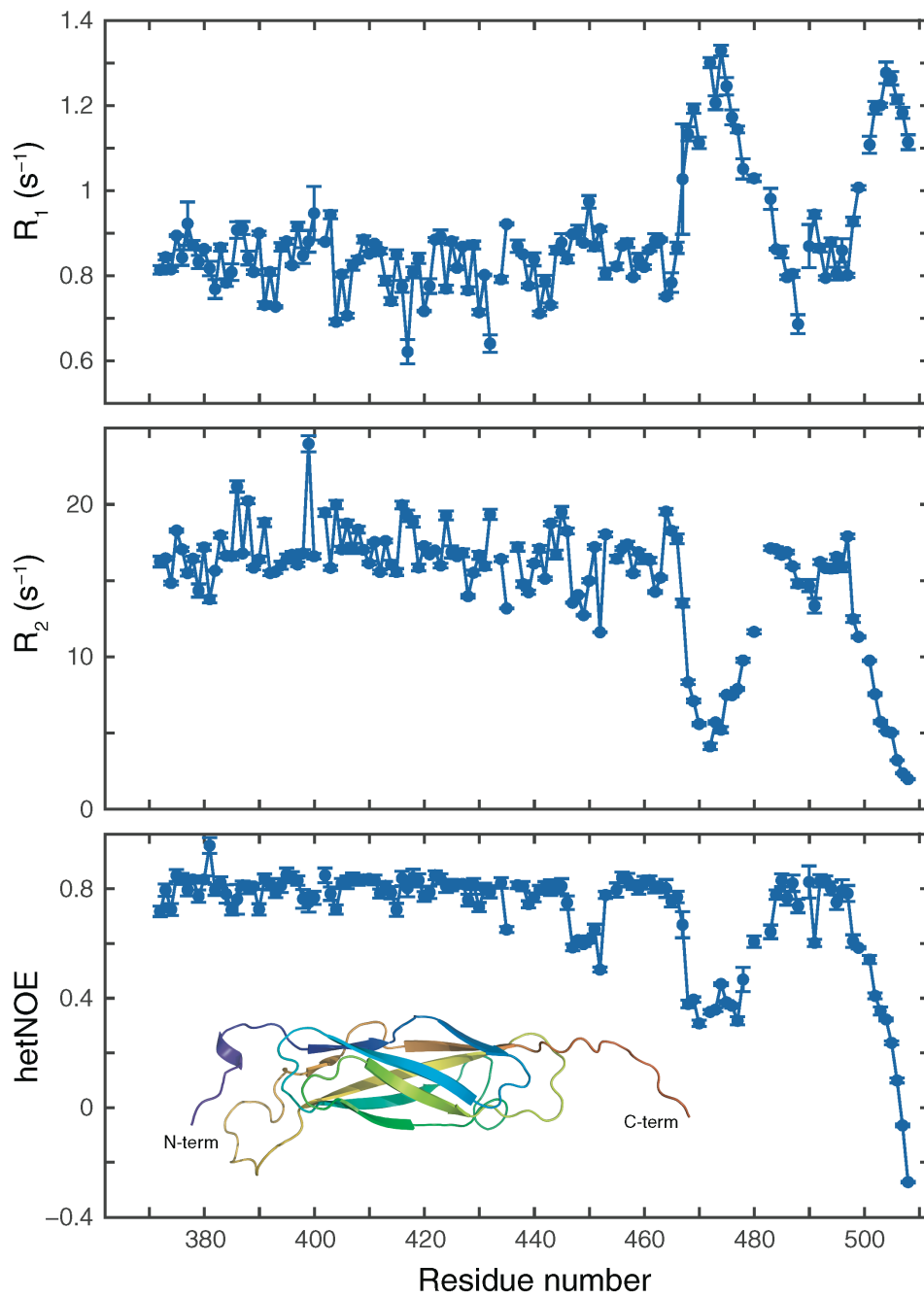
Fig. S4. Experimental ¹⁵N spin relaxation parameters of CBD1 in the absence of NPs. R_1, R_2, and hetNOE were measured at 850 MHz ^{1}H frequency and 33°C.
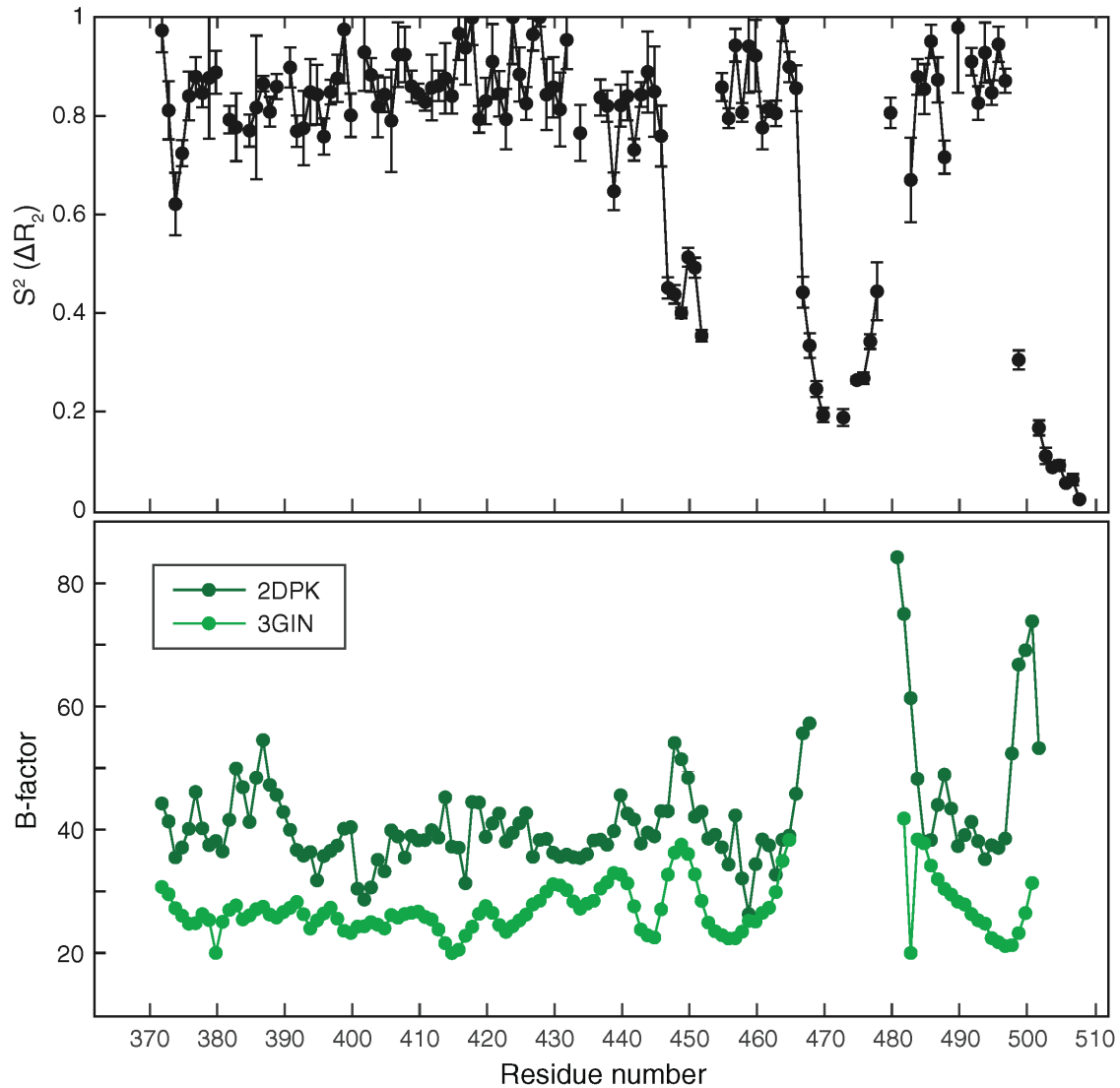
Fig. S5. Comparison between NMR $S^2(\Delta R_2)$ and x-ray B-factors of backbone nitrogen atoms in crystal structures. B-factor profiles of CBD1 WT (PBD code 2DPK) and E454K mutant (PDB code 3GIN) are very similar. Coordinates of residues 469–480 in Loop F-G are missing from the crystal structures. Loop E-F shows a steep drop in $S^2(\Delta R_2)$ and only slightly increased B-factors.
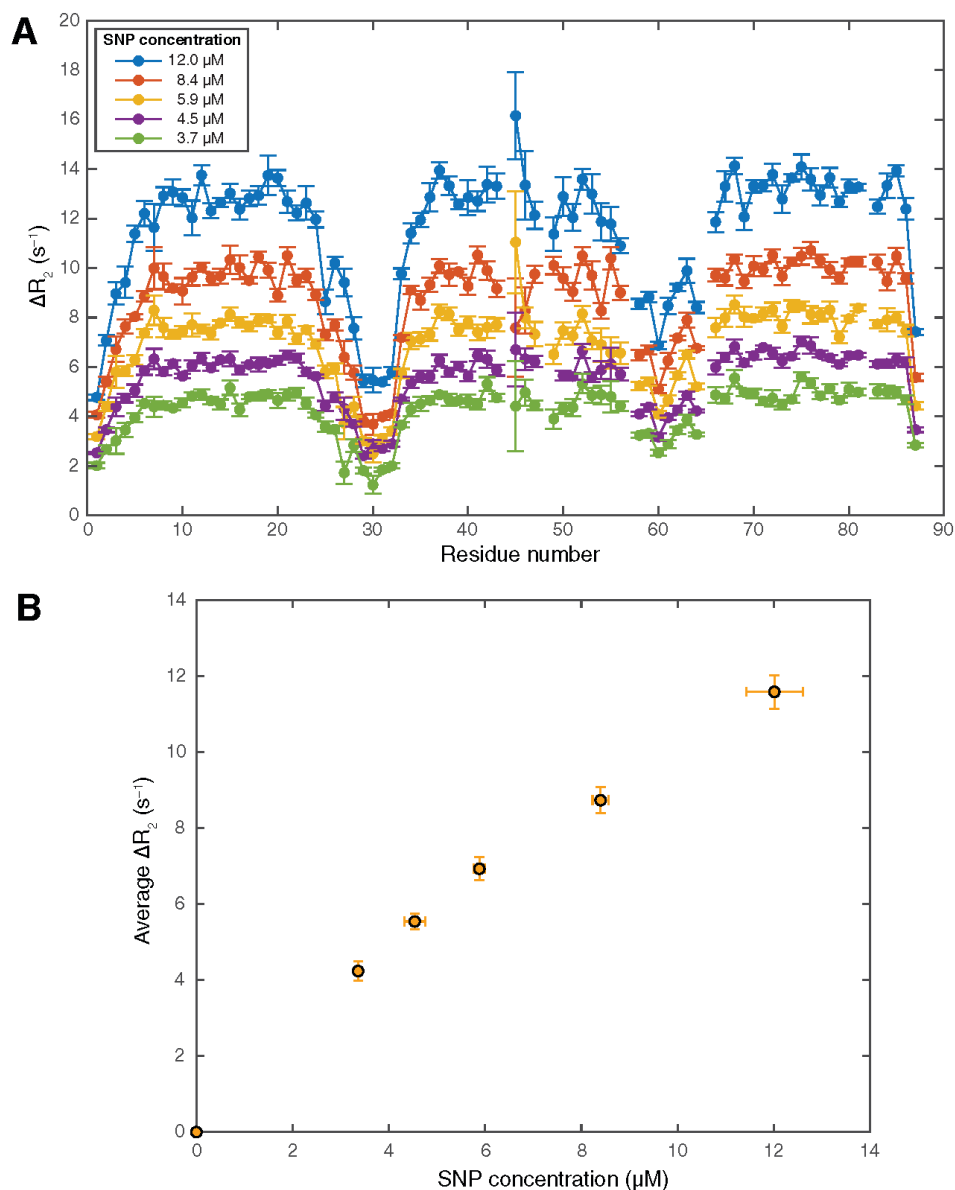
Fig. S6. Dependence of ΔR_2 values on SNP concentration. (**A**) Dependence of experimental ΔR_2 values on SNP concentration while keeping the Im7 concentration fixed at 0.5 mM. For residue Thr45, the large statistical error bars resulted from the propagation of R_2 uncertainties due to large exchange R_{ex} contributions (and, hence, low signal-to-noise), as shown in Fig. 2A of the main text. The five independently determined ΔR_2 profiles of this figure were used to demonstrate the experimental reproducibility of $S^2(\Delta R_2)$ measurements in Fig. 2B of the main text. The data were normalized (rescaled) to account for SNP concentration differences. (**B**) The average ΔR_2 values of each profile are plotted against the corresponding SNP concentrations. The relationship is approximately linear for the low SNP concentrations used in this work where at any given time only a small fraction of Im7 molecules are bound to the SNP surface.
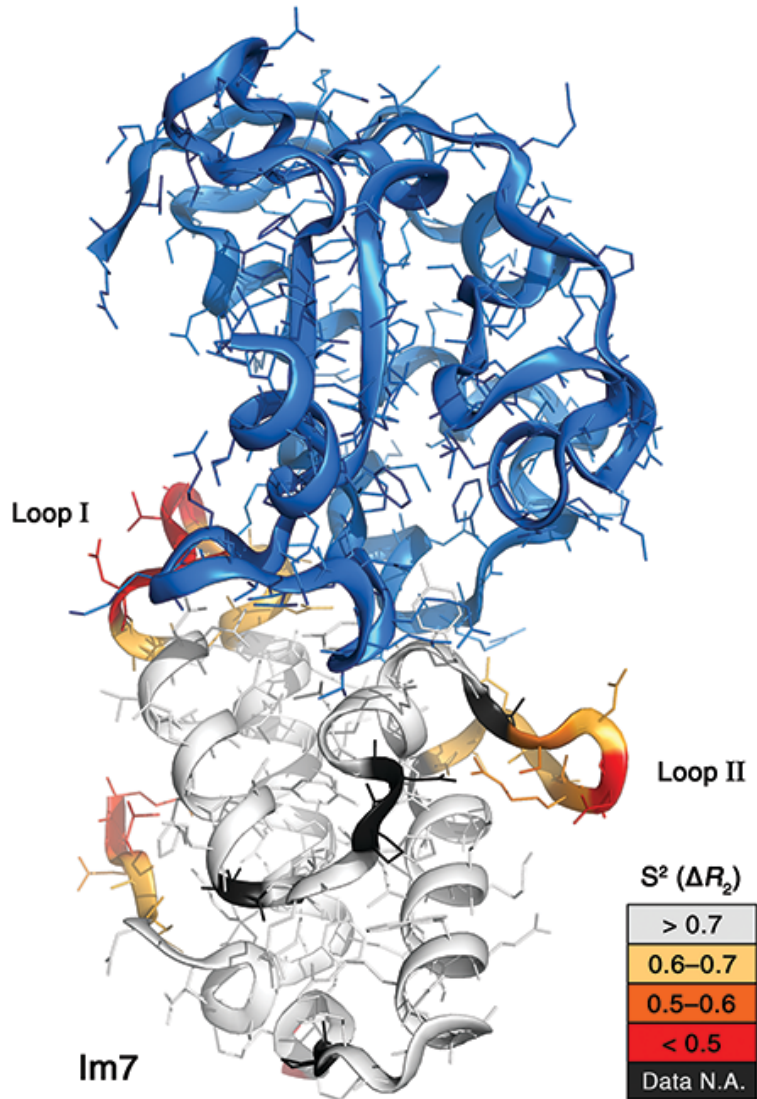
Fig. S7. Mapping of experimental $S^2(\Delta R_2)$ onto the structural model (PDB code 7CEI) of Im7 when bound to the DNase domain of colicin E7. Loop I of Im7, which displays substantially lowered $S^2(\Delta R_2)$ values in the free state, is part of the binding interface (*63*).