

## **Supporting Information**

# **Accurate interdomain contacts in a mixed folded protein from NMR-guided coarse-grained simulations**

Billy Hobbs<sup>1‡</sup>, Noor Limmer<sup>1‡</sup>, Gwendolyn L. Clenshaw<sup>1</sup>, Felipe Ossa<sup>1</sup>, & Theodoros K.  
Karamanos<sup>1\*</sup>

Department of Life Sciences, Faculty of Natural Sciences, Imperial College London, UK

**Experimental methods, 2 Supplementary Tables and 10 Figures**

## Supplementary methods

**Spectral density mapping.** The measured  $^{15}\text{N}$  relaxation rates can be described as:

$$R_1 = \left(\frac{d^2}{4}\right) (3J(\omega_{\text{N}}) + 6J(\omega_{\text{N}} + \omega_{\text{H}}) + J(\omega_{\text{N}} - \omega_{\text{H}})) + c^2 J(\omega_{\text{N}}) \quad 1$$

$$R_2 = \left(\frac{d^2}{8}\right) (4J(0) + 3J(\omega_{\text{N}}) + 3J(\omega_{\text{N}} + \omega_{\text{H}}) + 6J(\omega_{\text{H}}) + J(\omega_{\text{N}} - \omega_{\text{H}})) \\ + \left(\frac{c^2}{6}\right) (J(0) + 6J(\omega_{\text{N}})) \quad 2$$

$$\text{NOE} = 1 + \left(\frac{d^2}{4R_1}\right) \left(\frac{\gamma_{\text{H}}}{\gamma_{\text{N}}}\right) (J(\omega_{\text{H}} + \omega_{\text{N}}) - J(\omega_{\text{H}} - \omega_{\text{N}})) \quad 3$$

where  $d = (\mu_0 h \gamma_{\text{H}} \gamma_{\text{N}}) \langle r_{\text{NH}}^{-3} \rangle$ ,  $c = \omega_{\text{N}} \Delta\sigma / \sqrt{3}$ ,  $\mu_0$  is the permeability of the vacuum,  $h$  Planck's constant,  $\gamma_{\text{H}}$ ,  $\gamma_{\text{N}}$  are the gyromagnetic ratios of  $^1\text{H}$  and  $^{15}\text{N}$  respectively,  $r_{\text{NH}} = 1.02 \text{ \AA}$  and  $\Delta\sigma = -172 \text{ ppm}$  is the chemical shift anisotropy. Reduced spectral density mapping was performed as in Supplementary reference [1] with values of  $J(0)$ ,  $J(\omega_{\text{N}})$  and  $J(\omega_{\text{H}})$  are obtained as:

$$\sigma_{\text{NH}} = R_1 (\text{NOE} - 1) \gamma_{\text{N}} / \gamma_{\text{H}} \quad 4$$

$$J(\omega_{\text{H}}) = \frac{4\sigma_{\text{NH}}}{5d^2} \quad 5$$

$$J(\omega_{\text{N}}) = \frac{4R_1 - 5\sigma_{\text{NH}}}{3d^2 + 4c^2} \quad 6$$

$$J(0) = \frac{6R_2 - 3R_1 - 2.72\sigma_{\text{NH}}}{3d^2 + 4c^2} \quad 7$$

**Coarse-grained simulations.** The HPS-SS potential has the form<sup>2</sup>:

$$U_{\text{dihed}}(\varphi, \varepsilon_d) = -\ln[U_{\text{dihed},\alpha}(\varphi, \varepsilon_d) + U_{\text{dihed},\beta}(\varphi, \varepsilon_d)] \quad 8$$

Where  $\varphi$  is the dihedral angle of four consecutive residues and  $U_{\text{dihed},\alpha}$ ,  $U_{\text{dihed},\beta}$  are the functions that represent the helical and extended conformations.  $\varepsilon_d$  represents the residue-specific parameter that determines the probability of each residue to be in a helical or extended conformation. The helical part is given by:

$$U_{\text{dihed},\alpha}(\varphi, \varepsilon_d) = e^{-k_{\alpha,1}(\varphi - \varphi_{\alpha,1})^2 - \varepsilon_d} + e^{-k_{\alpha,2}(\varphi - \varphi_{\alpha,2})^4 + \varepsilon_0} + e^{-k_{\alpha,2}(\varphi - \varphi_{\alpha,2} + 2\pi)^4 + \varepsilon_0} \quad 9$$

where  $k_{\alpha,1} = 11.4 \text{ rad}^{-2}$ ,  $k_{\alpha,2} = 0.15 \text{ rad}^{-4}$ ,  $\varphi_{\alpha,1} = 0.90 \text{ rad}$ ,  $\varphi_{\alpha,2} = 1.02 \text{ rad}$  and  $\varepsilon_0 = 0.27$ . The extended part of the function is:

$$U_{dihed,\beta}(\varphi, \varepsilon_d) = e^{-k_{\beta,1}(\varphi-\varphi_{\beta,1})^2+\varepsilon_d} + e^{-k_{\beta,2}(\varphi-\varphi_{\beta,2}-2\pi)^2+\varepsilon_1+\varepsilon_d} \\ + e^{-k_{\beta,2}(\varphi-\varphi_{\beta,2})^4+\varepsilon_2} + e^{-k_{\beta,2}(\varphi-\varphi_{\beta,2}-2\pi)^4+\varepsilon_2} \quad 10$$

where  $k_{\beta,1} = 1.80 \text{ rad}^{-2}$ ,  $k_{\beta,2} = 0.65 \text{ rad}^{-4}$ ,  $\varphi_{\beta,1} = -1.55 \text{ rad}$ ,  $\varphi_{\beta,2} = -2.50 \text{ rad}$ ,  $\varepsilon_1 = 0.14$  and  $\varepsilon_2 = 0.40$ .  $\varepsilon_d$  values were modified from the default values to maximise the agreement with the observed secondary structure propensities calculated by Talos based on the backbone C $\alpha$ , C $\beta$ , C $\gamma$ , N, HN chemical shifts of JD-GF.

## Supplementary Tables

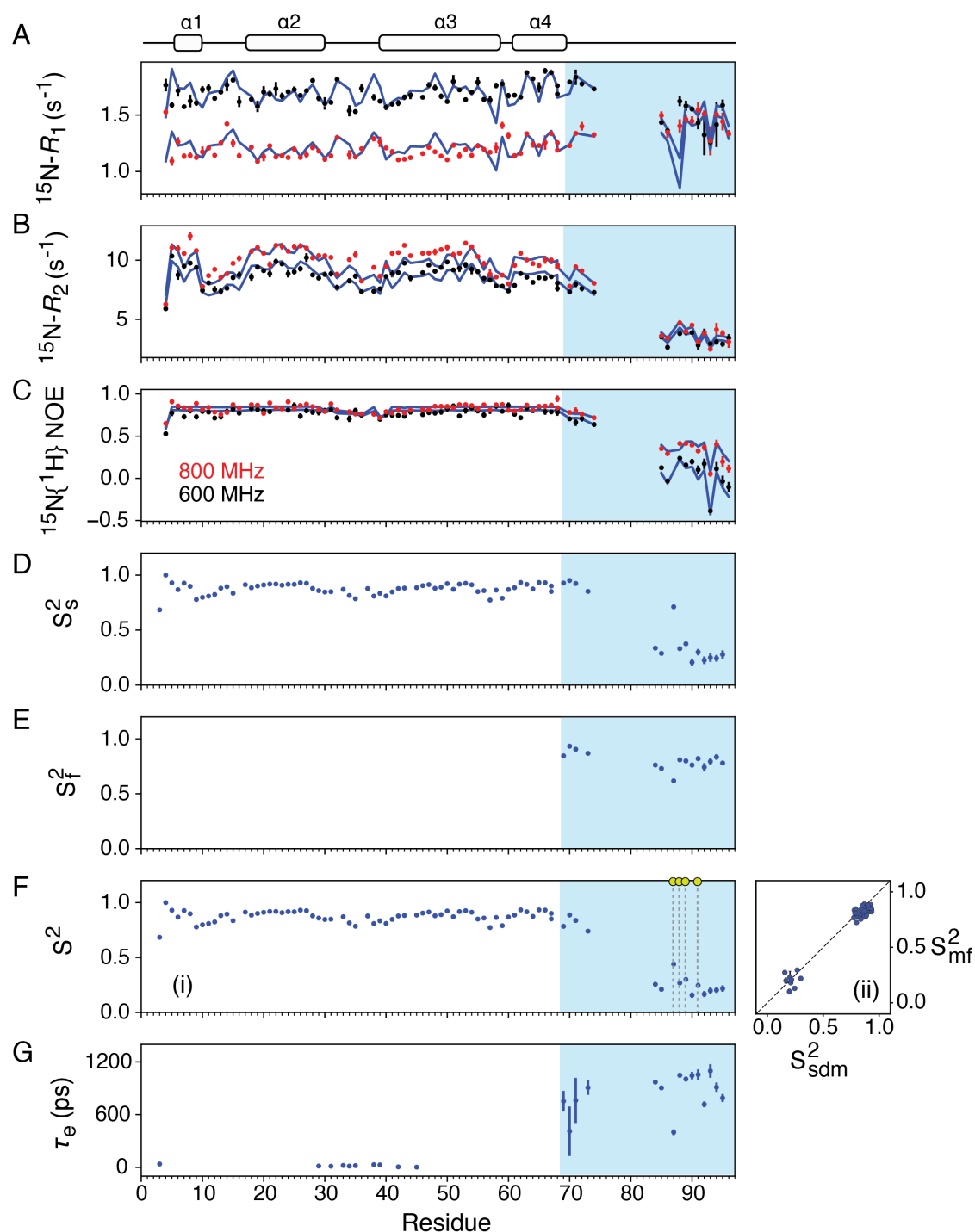
Table 1: Simulation details.

Simulation #	R <sub>g</sub> (nm)	
	Martini3IDP	CALVADOS
1	1.464 ± 0.025	1.640 ± 0.016
2	1.570 ± 0.021	1.610 ± 0.017
3	1.425 ± 0.030	1.610 ± 0.011
4	1.571 ± 0.015	1.570 ± 0.020

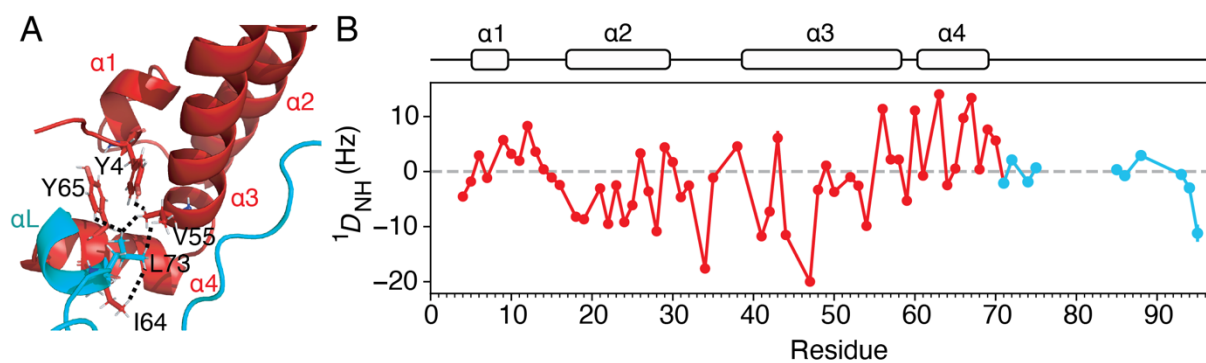
Table 2: Optimised  $\epsilon_d$  values for the HPS-SS backbone dihedral potential.

Residue	$\epsilon_d$
R	3.00
D	0.76
N	1.32
E	-0.58
K	0.54
H	0.11
Q	0.13
S	0.89
C	1.73
G	2.89
T	1.98
A	-1.65
M	-0.38
Y	0.94
V	0.79
W	0.37
L	-0.90
I	-0.12
P	2.65
F	0.94

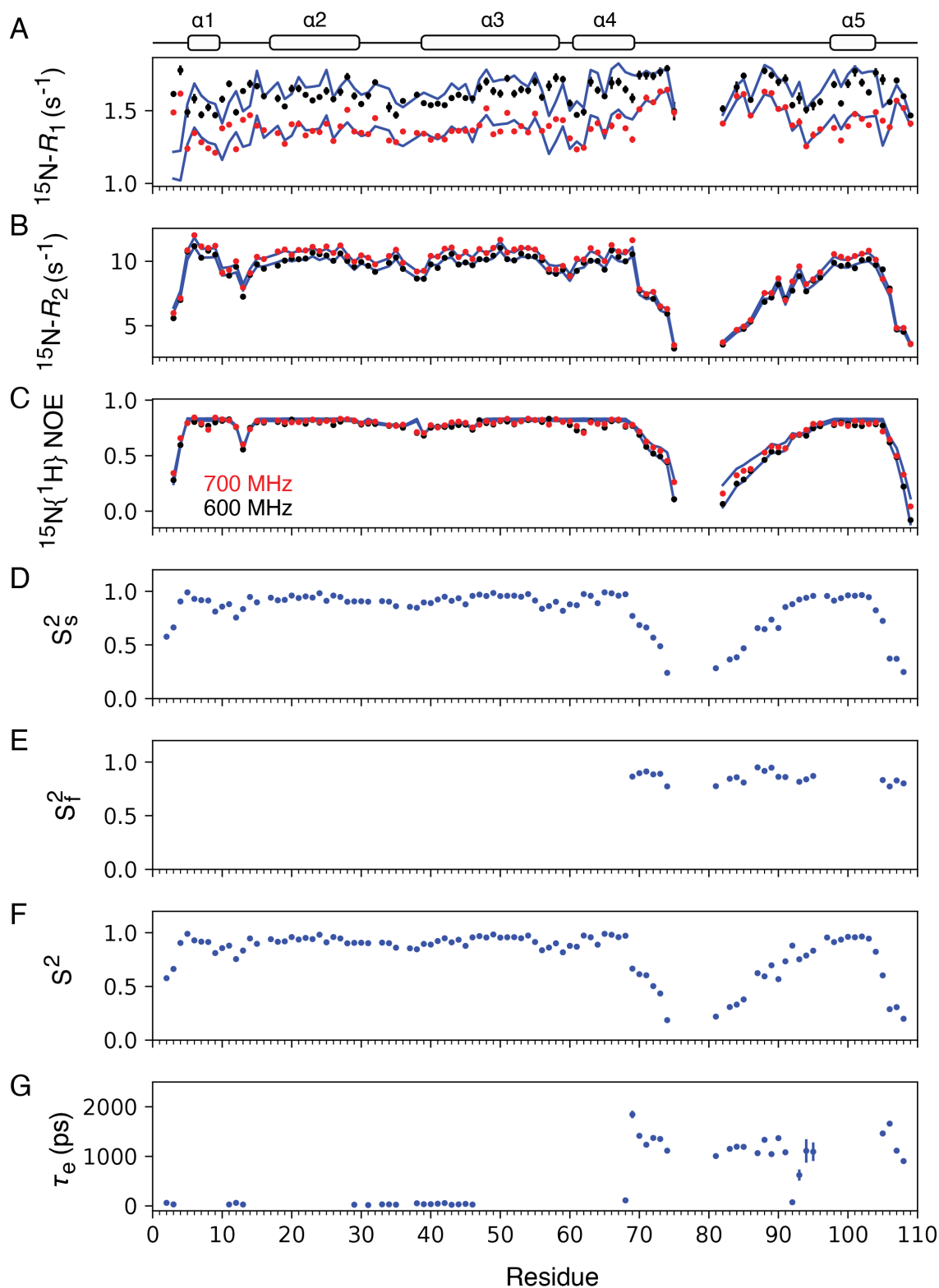
## Supplementary Figures



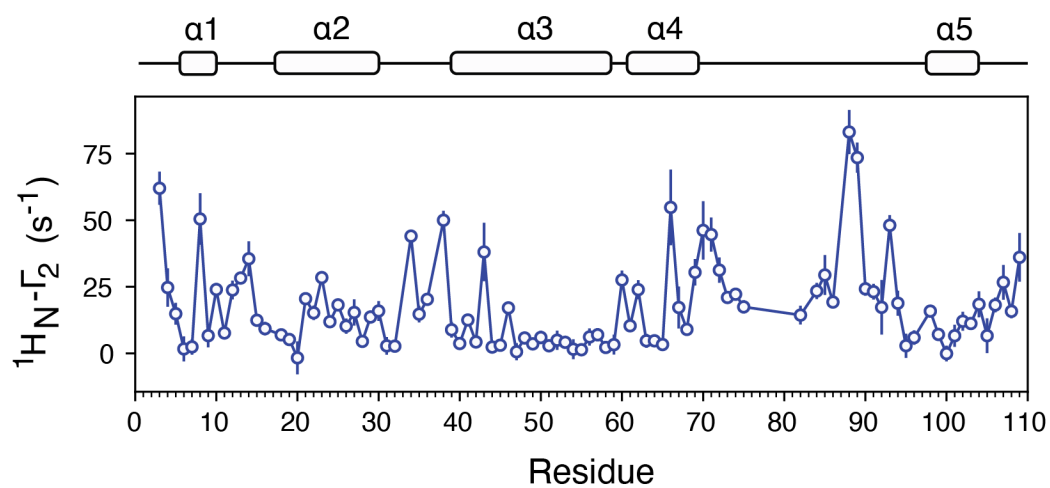
**Figure S1: Model-free analysis of the open JD-GF construct.**  $^{15}\text{N}$   $R_1$  (A),  $R_2$  (B) and  $^{15}\text{N}\{^1\text{H}\}$  NOE data (C) collected at 600 (black dots) or 800 (red dots) MHz with model-free predicted values shown as blue lines. Calculated order parameter squared for the slow ( $S_s^2$ ), fast ( $S_f^2$ ), the generalised order parameter  $S^2 = S_s^2 S_f^2$  and the correlation time for the internal motion ( $\tau_e$ ) are shown in D, E, F (panel i) and G respectively. Panel F(ii) shows the correlation between the spectral density-derived  $S^2$  values with those derived from the model-free analysis (RMSD 0.06, Pearson's correlation coefficient 0.98). An axially symmetric diffusion tensor was used for JD residues and a fully isotropic extended spectral density function for those in GF. When an extended model-free formalism was used  $\tau_e$  specifically refers to the slow motion ( $\tau_e = \tau_s$ ).



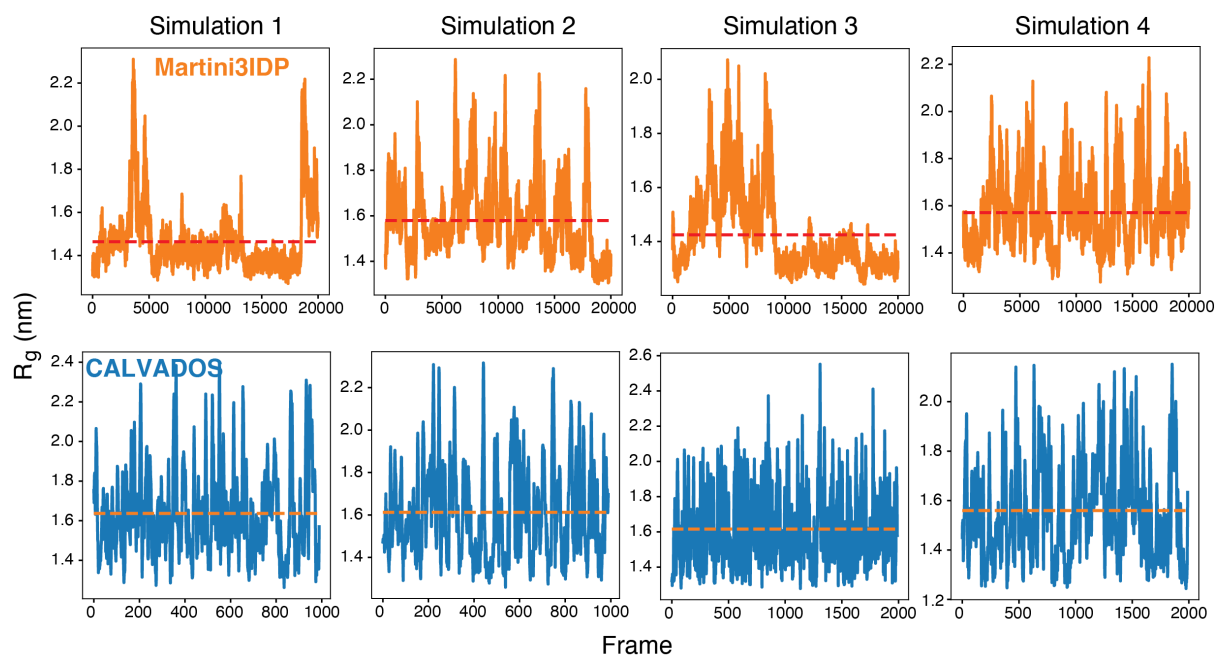
**Figure S2: Structural analysis of the GF region.** A cartoon representation of JD-GF with the J-domain and the GF shown in red and cyan respectively. The observed  $^1\text{H} - ^1\text{H}$  NOEs involving residues K70 – N74 (labelled as  $\alpha\text{L}$ ) are consistent with a helical configuration of this segment (labelled as  $\alpha\text{L}$ ) and are shown as black dashed lines.<sup>3</sup> (B)  $^1D_{\text{NH}}$  RDCs for JD-GF measured in 14 mg/mL Pf1 bacteriophage. Residues in the J-domain are shown in red and those in GF in cyan.



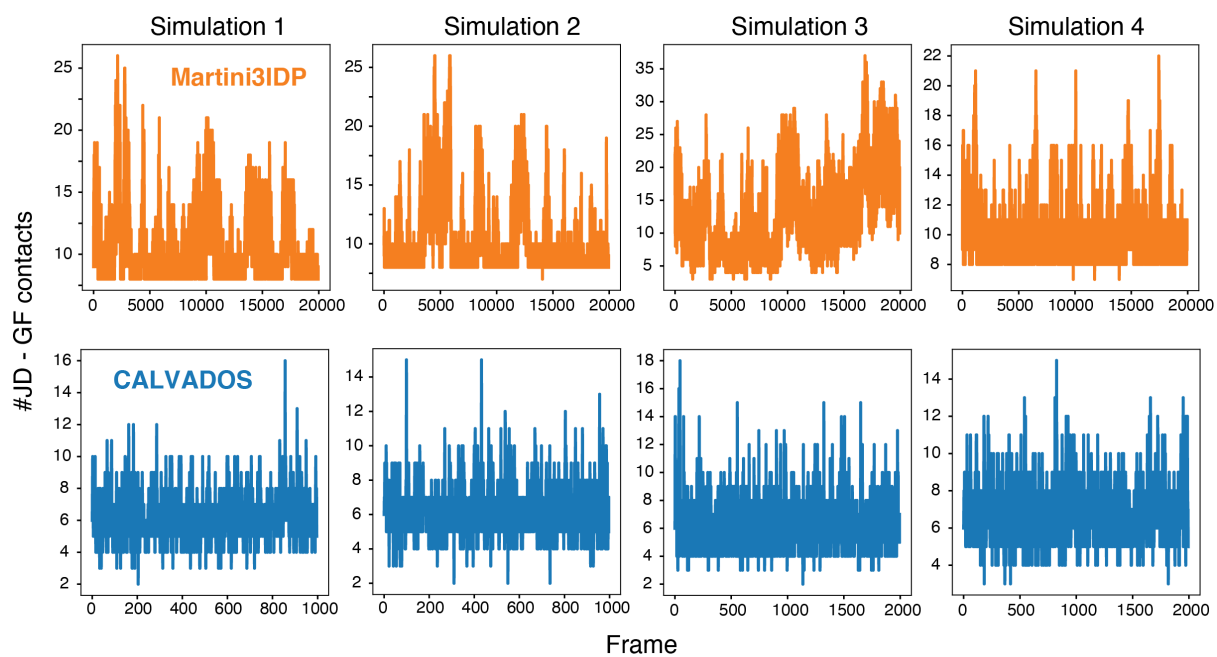
**Figure S3: Model-free analysis of the closed JD-GF- $\alpha 5$  construct.**  $^{15}\text{N}$   $R_1$  (A),  $R_2$  (B) and  $^{15}\text{N}\{^1\text{H}\}$  NOE data (C) collected at 600 (black dots) or 700 (red dots) MHz with model-free predicted values shown as blue lines. Calculated order parameter squared for the slow ( $S_s^2$ ), fast ( $S_f^2$ ), the generalised order parameter  $S^2 = S_s^2 S_f^2$  and the correlation time for the internal motion ( $\tau_e$ ) are shown in D, E, F and G respectively. An axially symmetric diffusion tensor was used for JD residues and a fully isotropic extended spectral density function for those in GF. When an extended model-free formalism was used  $\tau_e$  specifically refers to the slow motion ( $\tau_e = \tau_s$ ).



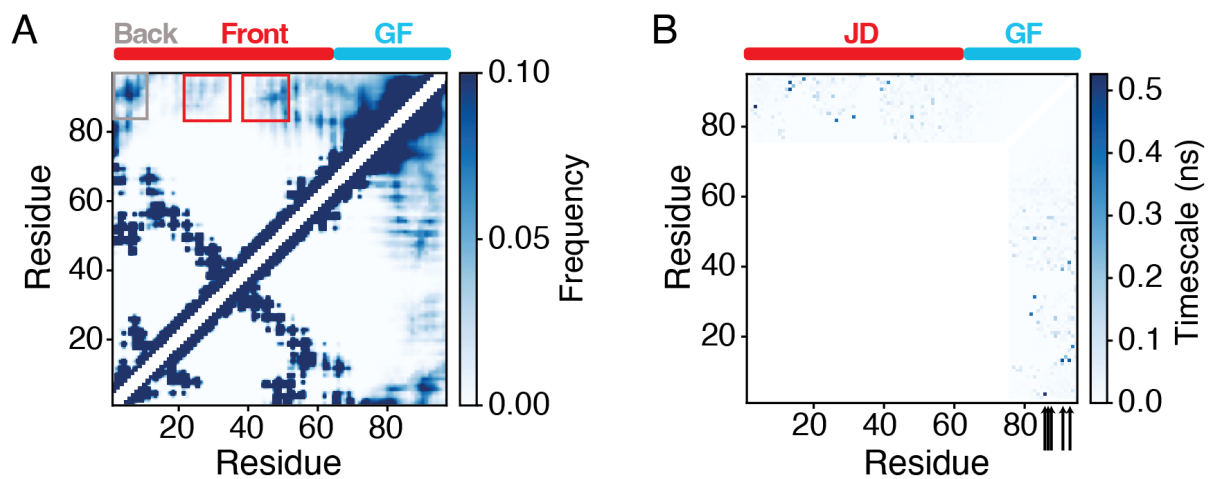
**Figure S4: Solvent PREs for the closed JD-GF- $\alpha 5$  construct.**  $^1\text{H}_\text{N}$  PRE profiles collected in the presence of 5 mM TEMPO. Strong solvent PREs for the hydrophobic parts of the GF linker (residues 70-74 and 88-93) suggest non-specific interactions of the spin label with the GF linker.



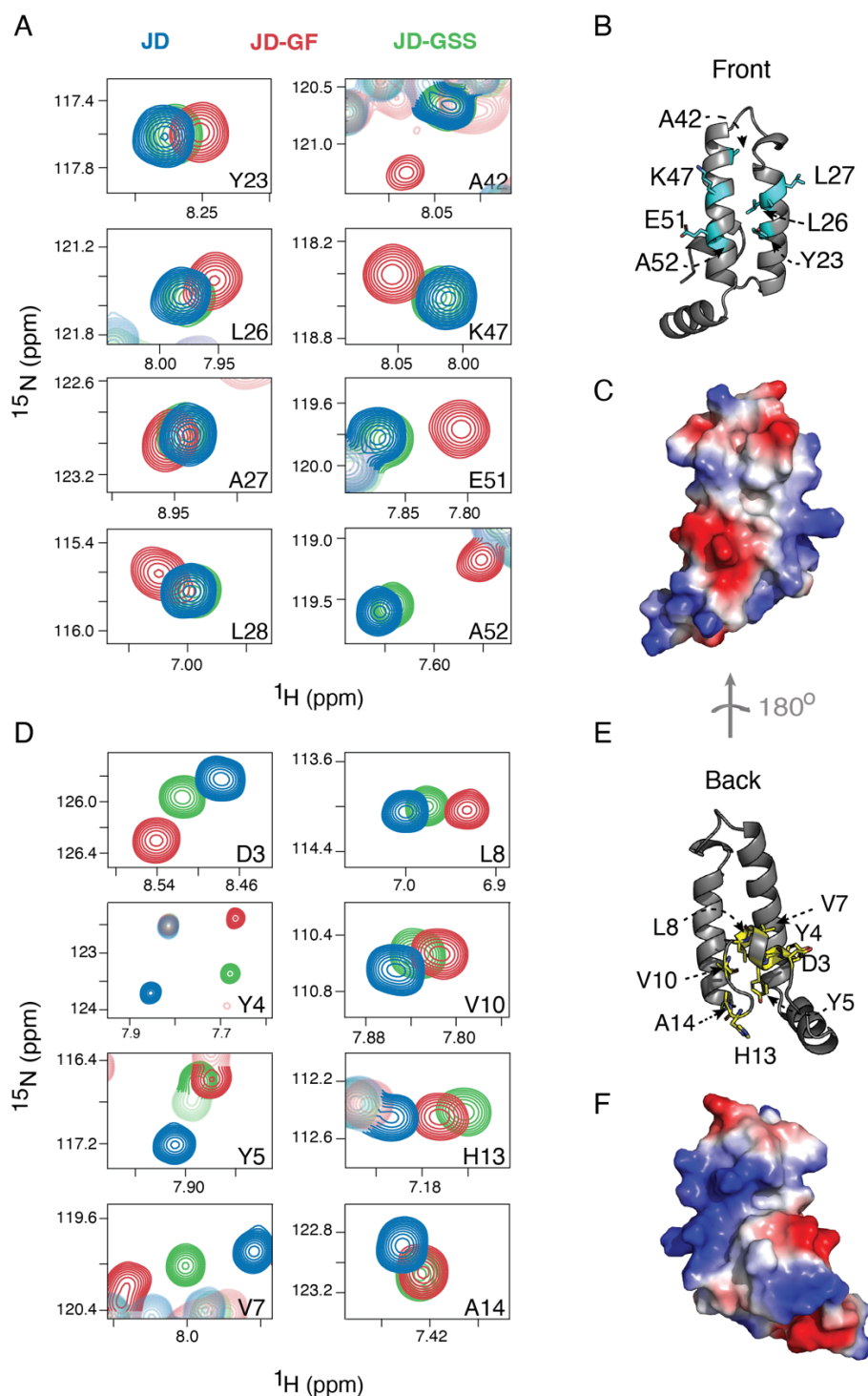
**Figure S5: Time evolution of  $R_g$ .**  $R_g$  as a function of frame number for the Martini3IDP (top) and CALVADOS simulations (bottom). The average  $R_g$  value for each simulation is shown as a dashed line.



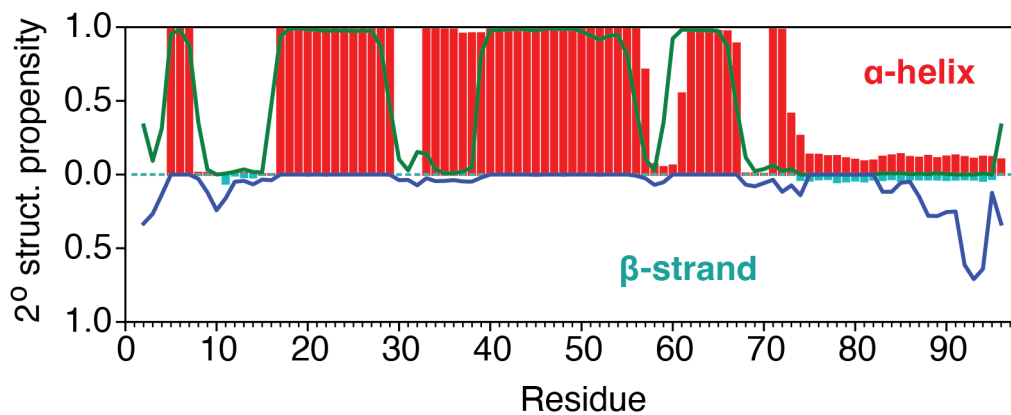
**Figure S6: Time evolution of interdomain contacts.** The number of JD – GF contacts as a function of frame number for the Martini3IDP (top) and CALVADOS simulations (bottom). A contact is defined when two backbone beads are within 10 Å of each other.



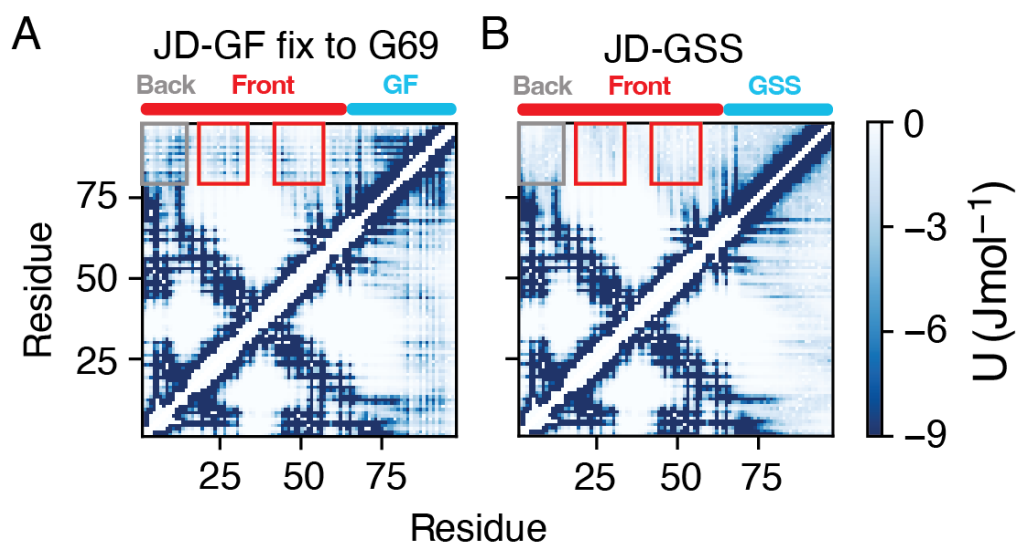
**Figure S7: Contacts in Martini3IDP simulations.** Maps showing the frequency (A) and the lifetime (B) of contact for each pair of residues in the concatenated 4  $\mu$ s Martini3IDP simulation. The map in (A) is capped at a frequency value of 0.1 and only contacts for GF residues are shown in (B). Residues with reduced motions as identified by NMR are highlighted with black arrows in (B).



**Figure S8: Contacts between the GF-linker and JD.** (A) Regions of the  $^1\text{H}$ - $^{15}\text{N}$  of JD-GF (red), JD alone (green) or JD-GSS (blue) for various hydrophobic, front-facing JD residues. The CSPs caused by GF essentially disappear in the presence of a disordered Gly-Ser-Ser repeats of the same length in JD-GSS. The residues highlighted in (A) are shown in cyan on the structure of JD in (B). The electrostatic surface potential of the front of JD is shown in (C) using the same pose as in (B) (red – negatively charged, white – hydrophobic, blue – positively charged). (D) As in (A) but focusing on residues 1 – 15 at the back of JD. CSPs for these residues are still present in the JD-GSS construct. Residues highlighted in (D) are shown in yellow on the structure of JD in (E) and the electrostatic surface potential of the back of JD is shown in (F).



**Figure S9: Conformational analysis of the JD-GF CALVADOS ensemble with no backbone dihedral angles included.** Secondary structure propensities calculated in the CG-simulation are shown as bars ( $\alpha$ -helix, red/ $\beta$ -strand, cyan). Talos-derived secondary structure<sup>4</sup> based on the JD-GF backbone shifts are shown as lines ( $\alpha$ -helix, green/ $\beta$ -strand, blue).



**Figure S10: Modulating interdomain contacts.** Energy maps of the Ashbaugh-Hatch term for a simulation where harmonic restraints were applied only up to residue G69 in JD-GF (A) or for the JD-GSS construct in which the GF is swapped with Gly-Ser-Ser repeats of the same length as GF (B). Grey/red boxes highlight the back (residues 1-15)/or the front (residues 25 - 30) of the JD.

## Supplementary References

1. Farrow, N.A., Zhang, O., Szabo, A., Torchia, D.A. & Kay, L.E. Spectral density function mapping using  $^{15}\text{N}$  relaxation data exclusively. *J. Biomol. NMR* **6**, 153-62 (1995).
2. Rizuan, A., Jovic, N., Phan, T.M., Kim, Y.C. & Mittal, J. Developing bonded potentials for a coarse-grained model of intrinsically disordered proteins. *J. Chem. Inf. Model.* **62**, 4474-4485 (2022).
3. Hobbs, B. et al. A low-complexity linker as a driver of intra- and intermolecular interactions in DNAJB chaperones. *Nat. Commun.* **16**, 5070 (2025).
4. Shen, Y., Delaglio, F., Cornilescu, G. & Bax, A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR* **44**, 213-223 (2009).