A Direct Coupling between Global and Internal Motions in a Single Domain Protein? MD investigation of Extreme Scenarios

Mehdi Bagheri-Hamaneh, Liqun Zhang, and Matthias Buck

SUPPLEMENTARY MATERIALS

Derived S^2 from the four UNR and INT simulations (5-55 ns) compared with experimentally derived S^2 [21]. The correlations coefficients (R-values) comparing the average of the UNR RBD simulations with the experimentally derived NMR parameters are 0.74 for S^2 and 0.63 for τ_e, similar, if not slightly better than the correspondence seen in other studies comparing experiments and simulations (see main text). The pairwise R-value comparisons within the four UNR and the four INT simulations range from 0.58 to 0.74, i.e. the deviation between simulation and experiments is no higher than between duplicates of the simulations.

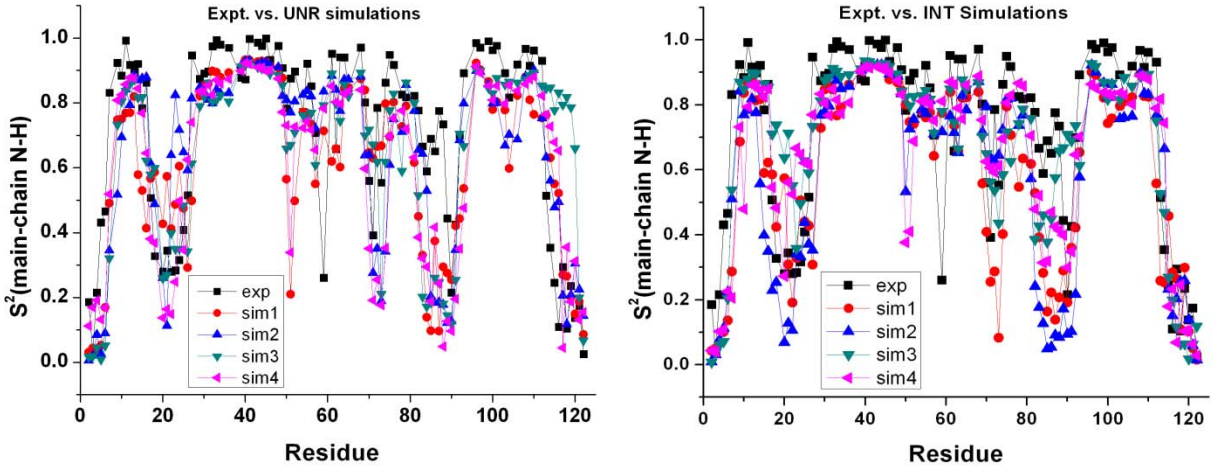

Figure S1: Derived S^2 from the four UNR and INT simulations compared with S^2 from experiment.

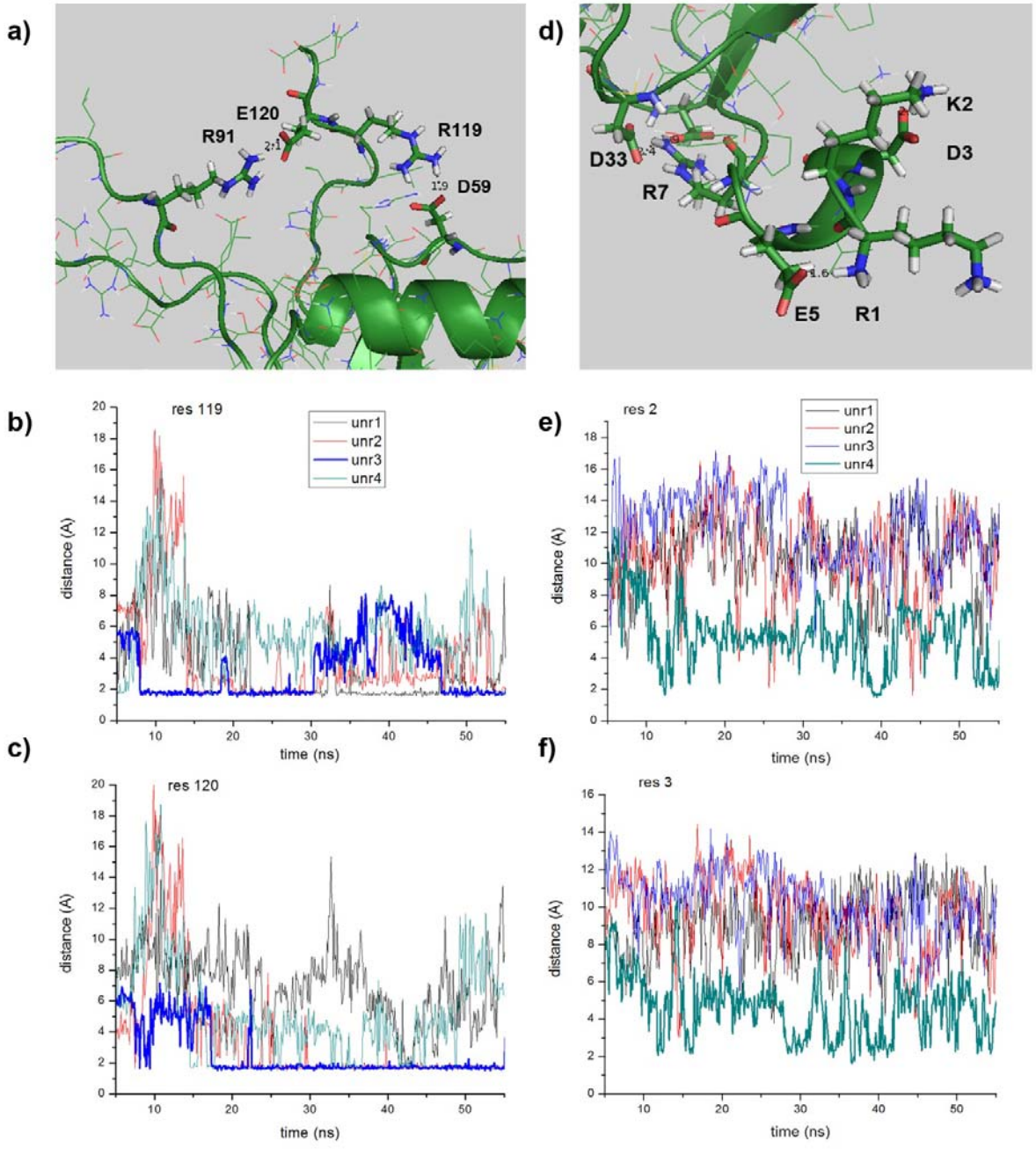

Figure S2: Examples of structure formation in flexible regions a)-c) at the C-terminus in simulation UNR3, e)-f) at the N-terminus in simulation UNR4. The structures at 55ns are shown. Sidechains of res 119 and 120 make contacts with neighboring sidechains (the distance to the

nearest atom is plotted). Sidechains res. 1-3 also make contacts and stabilize a helical turn at the N-terminus, which, however, is less persistent (compared to those at the C-terminus) as does not affect the derived order parameters as much. This latter structure becomes stronger when this and a number of the INT simulations are continued well beyond 55ns. (Thus we aborted the simulations and do not include them in the analysis). The UNR trajectory of simulation3 can be regarded as an outlier for the following reason: Hydrogen bonds form between the C-terminal sidechains and other sidechains in the latter half of this trajectory. Once formed they are not easily disrupted by the surrounding water and general fluctuations (see Suppl. Fig. S2). This formation of structure and a slowing of the dynamics in this region is in disagreement with the experimental NMR relaxation measurements, that show the C-terminal residues to be very dynamic. Deficiencies in the CHARMM22 forcefield that resulted in a higher propensity to form helical or turn-like structures were corrected before by the addition of the CMAP restraint [21]. However, the problem was not removed all-together and recent reports [32,33] point to similar cases to the one observed here.

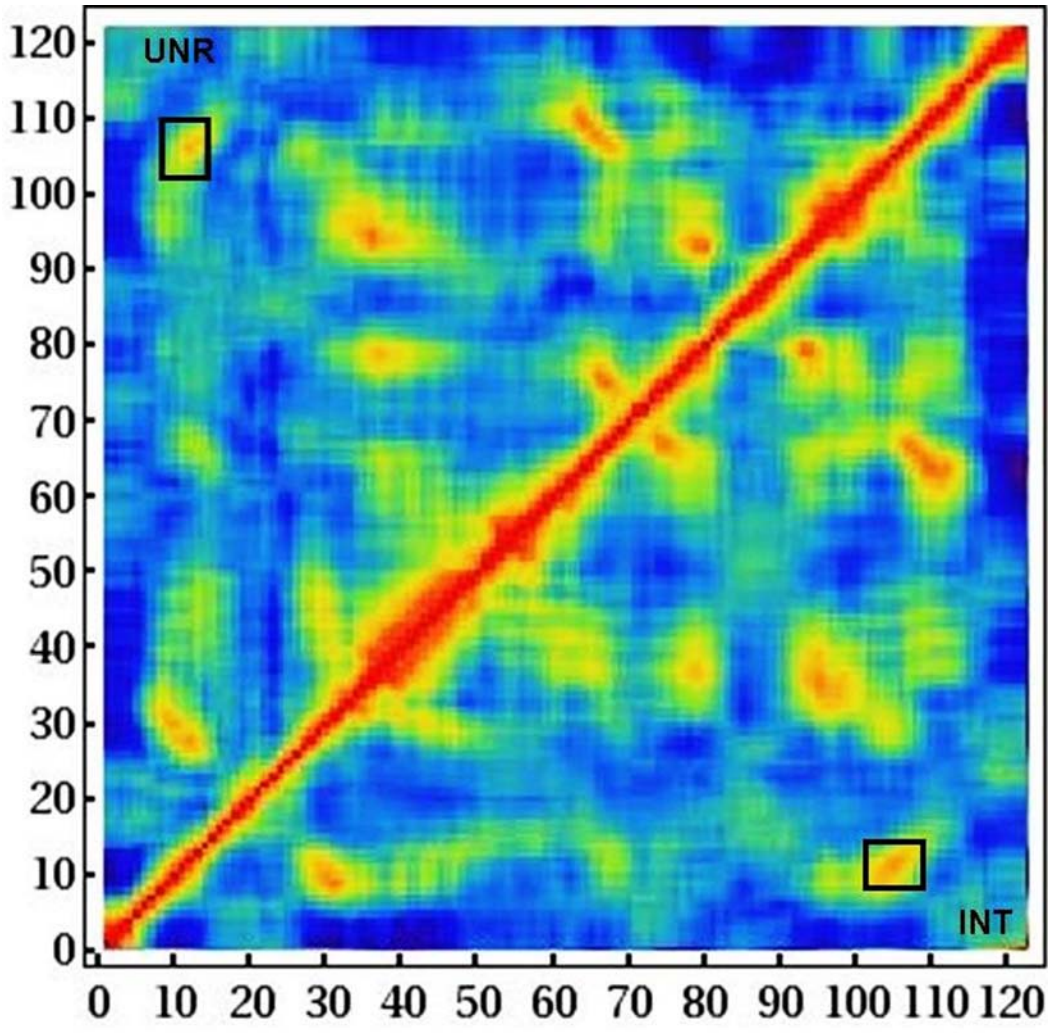

Figure S3. Normalized covariance matrices for UNR and INT simulations.

Normalized covariance matrices for UNR (above -) and INT simulations (below the diagonal). The scale of red to blue coloring show high correlation (+1) to high anticorrelation (-1) of Cα motions, respectively.

The normalized covariance matrix for C_α atoms, with values between 1 (high correlation) and -1 (high anti-correlation), is defined as $C_{ij} = M_{ij} / \sqrt{M_{ii} M_{jj}}$, where $M_{ij} = \left\langle \vec{r}_i \vec{r}_j \right\rangle - \left\langle \vec{r}_i \right\rangle \left\langle \vec{r}_j \right\rangle$ and $\vec{r}_i$ is the position vector of the C_α of the *i*th residue. Normalized covariance matrices were calculated, using CHARMM, for each INT and UNR trajectory (from 5 to 55 ns) and then were averaged over the four simulations. The linear correlation coefficient of two sets a,b of all k C_α pairs is calculated by $R = \Sigma (C_{ak} - \langle C_{ak} \rangle) (C_{bk} - \langle C_{bk} \rangle) / (\Sigma (C_{ak} - \langle C_{ak} \rangle^2)^{1/2} \Sigma (C_{bk} - \langle C_{bk} \rangle^2)^{1/2}))$, where $\langle C_{ak} \rangle$ and $\langle C_{bk} \rangle$ are the averages of the two datasets, a and b, respectively.

The figure shows that these correlations differ slightly between a few protein regions, e.g. the correlation of β-strands 1 and 5 (boxed) is weaker in the unrestrained (UNR) simulations than in those without global motion (INT). However, these small differences are not statistically significant in the context of the extent of sampling seen in the trajectories. Correlation coefficients between all pairwise values are high, with R=0.92 comparing the UNR and INT simulations.