Supporting Online Material for

Recognition Dynamics Up to Microseconds Revealed from an RDC-Derived Ubiquitin Ensemble in Solution

Oliver F. Lange, Nils-Alexander Lakomek, Christophe Farès, Gunnar F. Schröder, Korvin F. A. Walter, Stefan Becker, Jens Meiler, Helmut Grubmüller, Christian Griesinger,* Bert L. de Groot*

*To whom correspondence should be addressed.
E-mail: cigr@nmr.mpibpc.mpg.de (C.G.); bgroot@gwdg.de (B.L.dG.)

Published 13 June 2008, *Science* **320**, 1471 (2008)

DOI: 10.1126/science.1157092

This PDF file includes:

1. METHODS

1.1. **EROS.** Three different types of starting structures for EROS refinement were used in this work, resulting in ensembles termed EROS, EROS2, EROS3 and EROS4. For the ensemble denoted EROS and EROS4, ubiquitin structures were subjected to ensemble treatment from scratch. To this aim, starting from random coordinates, the CONCOORD program[1] was used to generate pairs of ubiquitin structures that together fulfil measured NOE data[2]. The resulting structures are based on a minimal geometrical model of bond lengths and angles and van der Waals radii, together with sum-averaged NOE-based distance bounds. Sum averaging was applied as $\overline{r_{kl}^{-6}} = \frac{1}{2}\sum_{j=1}^{2} r_{kl,j}^{-6}$, corresponding to time scales beyond the supra-τ_c time scale. 500 of such structure pairs were generated. Of these 1000 structures, the 400 best matching the RDC data (in sub-ensembles of 2 structures) were used for further refinement (see below). After refinement, the same selection was repeated once more to select the best 400 sub-ensembles of 2 structures (allowing duplication). Each structure that was selected at least once by the second filter step was added to the final ensemble. Cross-validation showed that R_{free} did not decrease with further refinement/selection cycles. In all selection rounds, only the non-NC couplings were used as working set, allowing cross-validation using NC couplings. From the final ensemble of 132 structures, 16 were removed due to chirality violations, leaving 116 conformers.

To assess the influence of the choice of starting structures, refinement was also carried out with all refinements started from one single configuration, the first entry in the high-resolution NMR ensemble 1d3z. This resulted in the EROS2 ensemble. Similarly, for EROS3, 400 initial structures were generated using CONCOORD that fulfil the NOE restraints in each single configuration.

To assess the impact of NOEs on the refinement, the EROS4 ensemble was started from the same structural pool as EROS with the difference that during refinement against RDC data no NOE restraints were applied.

Refinement was carried out using the GROMACS simulation package[3, 4]. Ensemble refinement was carried out by simultaneously applying NOE and RDC restraints on sub-ensembles of eight structures, as cross-validation indicated eight as the optimal ensemble size. All refinement simulations were carried out using simulated annealing in explicit solvent and periodic boundary conditions. The initial temperature was set to 350K and

slowly cooled to 0K during 100 ps, after which the temperature was raised to 310K during 10 ps. This annealing protocol was chosen as cross-validation indicated that it yields favorable results. Particle Mesh Ewald (PME) was used to describe long-range electrostatic interactions[5, 6]. Unless indicated otherwise, the OPLS/AA-L all-atom force-field[7] and the SPC water model[8] were applied. For the EROS and EROS3 ensembles, 3000 water molecules were used, for the EROS2 ensemble 4107. All simulations were run at constant volume. In addition to the OPLS/AA-L force field, EROS3 ensembles were also refined using the charmm-27[9], the amber03[10, 11], and the gromos96-53a6 force fields[12]. For the simulations carried out in the charmm-27 and amber03 force fields, TIP3P[13] was used as water model instead of SPC. Lincs[14] was used to constrain bond-lengths, allowing a time step of 2 fs. For bonds involving RDCs, bond lengths were chosen as d(H-N)=1.04 Å; d(CA-HA)=1.12 Å; d(N-C)=1.33 Å; d(CA-C)=1.53 Å[15]. Other bond lengths were taken from the respective force fields. Force constants of $1000\,\mathrm{kJ}/\left(\mathrm{mol\ nm}^2\right)$ and $0.25\,\mathrm{kJ}/\left(\mathrm{mol\,Hz}^2\right)$ were chosen for NOEs and RDCs, respectively. For each RDC experiment one alignment tensor was determined from eight structures for every time step during refinement, after superimposing the backbone atoms of residues 1–70 onto each other.

Sum averaging $\overline{r_{kl}^{-6}} = \frac{1}{8}\sum_{j=1}^{8} r_{kl,j}^{-6}$ during EROS refinement ensured that the NOE restraints are solely imposed on structure ensembles rather than individual structures, allowing individual structures to violate the NOE distance bounds. It is well established that r^{-6} averaging allows for substantial flexibility for individual atom pairs involved in NOEs, even in ensembles of only two structures[16, 17]. Such large flexibility, up to local partial unfolding, was indeed observed in the NOE-based ensemble consisting of 2-membered NOE-sub-ensembles that was used as starting point for EROS refinement. It was found that the flexibility of the RDC-refined ensemble (with a mean ensemble backbone RMSD of 1.22 Å, and a maximum of 2.53 Å) is considerably lower than that of the NOE-based ensemble that served as starting ensemble for RDC refinement (mean RMSD: 1.83 Å; max.: 4.85 Å). Hence, the structures and flexibility in the refined ensemble are to a large extent determined by the RDCs. This notion is further corroborated by the fact that the sum of NOE violations (summed over all 2727 NOE's in the 1d3z dataset) is as low as 0.66 nm in the EROS4 ensemble, in which no NOE restraints were applied (compared to 0.11 nm in the EROS ensemble, in which NOE restraints were included).

Order parameters S^2_{EROS} of the EROS ensembles were computed after superimposing the backbone atoms of residues 1-70 onto each other. Expressing the orientation of the NH vectors of the superimposed structures in spherical coordinates (θ, φ) the order parameters were obtained as

$$S^2 = \frac{4\pi}{5} \sum_{M=-2}^{2} \langle Y_{2M}(\theta,\varphi)\rangle \langle Y^*_{2M}(\theta,\varphi)\rangle ,$$

where the Y_{2M} denote the normalized second-order spherical harmonic functions $Y_{20}(\theta,\varphi) = \sqrt{5/(16\pi)}\left(3\cos^2\theta - 1\right)$, $Y_{2\pm1}(\theta,\varphi) = \mp\sqrt{15/(8\pi)}e^{\pm i\varphi}\cos\theta\sin\theta$, and $Y_{2\pm2}(\theta,\varphi) = \sqrt{15/(32\pi)}e^{\pm 2i\varphi}\sin^2\theta$.

As a hybrid protocol between the model-free SCRM method (see below) and EROS refinement, we developed a minimal geometrical model. Structurally, it contains solely bond length and weak bond angle terms, together with a minimal set of van der Waals radii (chosen as 0.13, 0.14, 0.16, 0.16 and 0.05 nm for O, N, C, S, and H, respectively). Compared to the molecular mechanics force fields used in EROS, therefore, torsion angles, dispersion interactions, electrostatic interactions, and solvent interactions are neglected. In addition, this 'geometry filter' shares the method to determine the alignment tensor with the SCRM approach: the NH vectors of the protonated 1UBQ X-ray structure were rotated iteratively to optimally match the NH RDCs. The according alignment tensor was subsequently used in ensemble refinement.

1.2. **Self-Consistent RDC-based Model-free (SCRM).** Order parameters derived from the EROS ensemble method were compared to RDC-based order parameters independently derived by the Self-Consistent RDC-based Model-free (SCRM) method. The SCRM approach allows determination of internal protein dynamics from RDCs largely without structural noise and determines the dynamic average orientation of the inter-nuclear vectors in the protein structure. Removing structural noise from the model-free analysis therefore further increases the accuracy of this analysis and allows more accurate statements about the supra-τ_c motion, i.e. motion beyond the correlation time to be made. The approach relies on the previously introduced model-free approach[18, 19]. Based on a set of dipolar couplings and a static input structure, alignment tensors and dynamically averaged spherical harmonics describing the average orientation of each NH vector as well as its dynamics were derived. The key of the self-consistent analysis is to take output average orientations

as input of the next tensor calculation and repeat until convergence is reached. The iteration is stopped when the order parameters between subsequent cycles agree within 0.01.

The SCRM method yields a relative estimate of RDC-based order parameters, e.g., of NH bond vectors. As the absolute scale of the SCRM order parameters is unknown, they are usually scaled with respect to Lipari-Szabo (LS) relaxation order parameters[19, 20]. This approach rests on the assumption that at least a few residues will show little additional dynamics beyond the overall tumbling time scale, and hence yields a conservative estimate of the overall level of supra-τ_c motion. The optimal relative scaling is non-trivial, and we have therefore developed two independent methods to obtain a scaling estimate.

Scaling method S1 is based on hypothesis tests. In a first step, S1 was applied to test the hypothesis that the LS order parameters and SCRM order parameters (at any scaling) stem from the same distribution (zero supra-τ_c motion hypothesis). This was achieved by generating multiple sets of synthetic order parameters, Gaussian-distributed around the measured SCRM order parameters with a standard deviation given by the respective uncertainties. The zero supra-τ_c motion hypothesis was rejected with a confidence of more than 99.999%, indicating significant supra-τ_c motion. To obtain a scaling estimate, in a second step, the method successively removes residues that display the most significant supra-τ_c motion until the zero supra-τ_c motion hypothesis cannot anymore be rejected for the remaining residues at a confidence level above 95%. The corresponding scale that optimally scales the remaining SCRM order parameters to the respective LS order parameters is the suggested scaling factor. This method is similar in spirit to the step-fitting algorithm recently proposed for microtubule growth[21].

Uncertainties in order parameters were estimated via error propagation both from the experimental error on the RDCs (0.3 Hz), and from the post-SCRM[22] RDC rmsd (0.52 Hz), yielding scaling factors S^2_{Overall} of 0.88 and 0.91, respectively, which were applied to the SCRM curves shown in Fig. 3.

An alternative scaling method S2 was developed to allow the comparison with the scaling derived by the S1 method. Scaling method S2 is based on the notion that the true RDC-based order parameters must be strictly smaller than corresponding LS order parameters, due to the longer time scales probed by RDCs. As the SCRM order parameters are determined with a certain experimental uncertainty, the scaling of the SCRM order parameters

	X-ray	heated X-ray	Lipari Szabo	SCRM	EROS
X-ray	x				
heated X-ray	0.71	x			
Lipari Szabo	0.46	0.40	x		
SCRM	0.65	0.50	0.41	x	
EROS	0.76	0.73	0.73	0.74	x
geometry filter	0.7 (0.74)	0.52 (0.56)	0.56 (0.61)	0.82 (0.86)	0.84 (0.87)

TABLE S1. correlation of S^2 (NH) order parameter. The correlation coefficients for the row *geometry filter* reflect the mean over 50 geometry filter ensembles; the mean of the best correlating 10% is given in brackets.

to strictly obey this condition for all residues (with the highest one equal to LS) would underestimate the scaling factor. In order to alleviate this bias, the amount of underestimation thus produced is estimated from a set of synthetic order parameters, Gaussian-distributed around the measured SCRM order parameters, with a standard deviation corresponding to the respective uncertainty. These synthetic data are scaled such that they are strictly below the LS order parameters, with the highest one equal to the corresponding LS value. The distribution of these synthetic scaling factors then allows the estimation of confidence intervals of the originally measured set of SCRM order parameters. Scaling factors derived from 95% confidence intervals (i.e., 95% of the scaling factors are smaller than the suggested scale) were used to derive the best estimate for the scaling factors. The 95% confidence interval scaling factors corresponding to $0.30\,\mathrm{Hz}$ and $0.52\,\mathrm{Hz}$ RDC uncertainties are 0.89 and 0.93, respectively, showing agreement to the estimates from method S1[22].

2. COMPARISON OF SCRM AND EROS ORDER PARAMETERS

As presented in the main text, the order parameters obtained from the RDC data with our model-free approach S^2_{SCRM} agree with S^2_{EROS} within error. This agreement is remarkable considering the many differences between the two approaches. It shows that the dynamics observed in the EROS ensemble is strongly determined by the experimental residual dipolar coupling data (cf. Table S1).

The agreement between S^2_{SCRM} and S^2_{EROS} is observed despite potential force-field bias or sampling problems in the EROS protocol, or due to inclusion of solutions in the model-free approach that are geometrically unfeasible. To evaluate the effect of some of the potential assumptions, we hybridized the two approaches by coupling the alignment procedure

of the model-free approach with a minimal geometrical structural model of ubiquitin that only restricts bond lengths and avoids clashes of atoms (see sec. S1.1). Both, the EROS and model-free results agree well with order parameters obtained by this hybrid approach (r=0.84 and 0.82 respectively). Furthermore, the particular choice of molecular mechanics force-fields affects the obtained EROS order parameters only slightly, as shown by an average mutual correlation of r=0.88 between EROS solutions obtained with four different state of the art force-fields (see also Section S6.2). This indicates that the agreement between EROS and SCRM order parameters is robust independently of how the results are obtained.

3. CROSS-VALIDATION

The EROS ensemble was cross-validated by systematically leaving out data from the refinement. All RDCs between backbone amide N and carbonyl C, as well as all scalar couplings were used for this purpose.

R-values,

$$R_{\mathrm{X}} = \left(\Sigma_k^{n_x} \left(\mathrm{X_{k,calc}} - \mathrm{X_{k,exp}} \right)^2 / \left(2 \sum_k^{n_x} \mathrm{X_{k,exp}^2} \right) \right)^{1/2},$$

for data class X (e.g., an experiment) with n_{x} data points and Pearson correlation coefficients R_{X} were computed individually for every data-class and subsequently averaged. For RDCs, a data-class comprises all couplings between the same type of nuclei determined in the same alignment medium. For scalar couplings the data classes are HNHA, HNCO, HNCB, COHA, COCO, and COCB for the backbone φ-dihedral and the two N- C_γ, C'-C_γ for the side chain χ_1-dihedral, respectively[2]. R-values were averaged as

$$R = \left(\sum_j^N n_j \right)^{-1/2} \left(\sum_i^N n_i R_i^2 \right)^{1/2},$$

where N denotes the number of data-classes. Correlation coefficients were arithmetically averaged. All R-values represented are R_{free}, if not otherwise stated. RDC cross-validation was carried out via amide NC' RDCs obtained in 6 alignment media. The alignment tensor used to compute the NC'-RDCs has been fitted via all non-NC' RDCs.

The ensemble averaged "free" RDC R-factor is 18.5%. This is substantially lower than the values obtained for a single X-ray structure ($25 \pm 4\%$), the NMR ensemble 1d3z (24%) or for the previously published ensembles aimed at describing nanosecond dynamics 1xqq

(27.9%) and 2nr2 (24.1%). The value obtained for the ensemble of X-ray structures is 18.3%. This finding shows that conformational heterogeneity as found in the EROS ensemble or in the ensemble of 46 X-ray structures considerably improves the description of the experimental data with respect to individual structures (the lowest value observed for any of the individual X-ray conformers was 20.4%). Moreover, the conformational heterogeneity found in the nanosecond ensembles (1xqq and 2nr2) does not improve the description of the RDC data over a single X-ray structure. Scalar couplings in the backbone and in the side-chain, as well as scalar couplings across hydrogen bonds are described well by the EROS ensemble. The R_{free} value of 18.5% for the EROS ensemble indicates that potential model bias and over-fitting issues are acceptable. Moreover, it indicates that alignment media have minimal influence on the structure and ensemble of ubiquitin.

(a):

Ensemble	$R_{\text{free}}\left({}^{3h}J_{NC;}\right)$	$R_{\text{free}}\left({}^{3}J_{\varphi}\right)$	$R_{\text{free}}\left({}^{3}J_{\chi_1}\right)$	R_{free} (RDC)
EROS	0.19	0.20	0.10	0.185
EROS2	0.19	0.20	0.10	0.188
EROS4	0.19	0.21	0.10	0.190
X-ray46	0.25[3,5]	0.15	0.10	0.182[3,5]
1xqq	0.21	0.27	0.09	0.279[5,6]
1d3z	0.25 [2]	0.17 [1]	0.14	0.242[1]
2nr2	0.18	0.25	0.10	0.241[5,6]
1ubi	0.35[3]	0.18	0.19	0.24[3]
EROS_avg[4]	0.29	0.23	0.17	0.27

(b):

Ensemble	$r\left({}^{3h}J_{NC;}\right)$	$r\left({}^{3}J_{\varphi}\right)$	$r\left({}^{3}J_{\chi_1}\right)$
EROS	0.81	0.94	0.97
EROS2	0.81	0.93	0.97
EROS4	0.81	0.93	0.97
X-ray46	0.85[3]	0.94	0.98
1xqq	0.77	0.88	0.98
1d3z	0.69	0.93 [1]	0.95
2nr2	0.85	0.88	0.97
1ubi	0.72[3]	0.92	0.88
EROS_avg[4]	0.78	0.94	0.90

TABLE S2. Cross-validation of ubiquitin structures via scalar couplings. (a) RDC R-factors and (b) pearson correlation coefficients. [1]some of the data was used during the refinement[2]. [2]one outlier was removed. [3]Protons were placed with molecular modelling package whatIF[23]. [4]Average structure computed from EROS ensemble. [5]For bonds involving RDCs bond lengths were set to the same lengths chosen for EROS[15] using the Lincs algorithm[14]. [6]R_{free} of ensembles 1xqq and 2nr2 as originally deposited, i.e., without bond length correction, is 0.296 and 0.262, respectively. For 1d3z the value differs from the $16\%/\sqrt{2} = 11\%$ reported in Ref. [2]. The reason for this discrepancy is that here the evaluation is carried out over NC‘ couplings measured in 6 different alignment media, whereas the reported value stems from evaluation against the RDC data (NC', NH, HC') measured in 2 alignment media which was also used to refine the structure. Hence, the much lower reported value of 11% reflects that it is not a *free* R-value and the different choice of couplings (NC' only vs. NC', NH, HC', C_α-H_α,...). Here, for reasons of consistency we evaluated R for 1d3z on the same RDC data as used for all other ensembles. However, two of the six NC'-RDC data sets used here for cross-validation have already been used in Ref. [2] for the refinement of the ensemble 1d3z. The R_{free} taken for the remaining 4 NC'-RDC data sets that were not used to refine 1d3z is 29% and thus comparable to that of 1xqq. R_{free} computed in an analogous manner on 4 sets of HC'-RDC data is 23% for ensemble 1d3z (using all nonHC' RDC to fit the alignment tensor). The consistently higher R-factors and lower correlation coefficients indicate a significant dynamic averaging effect.

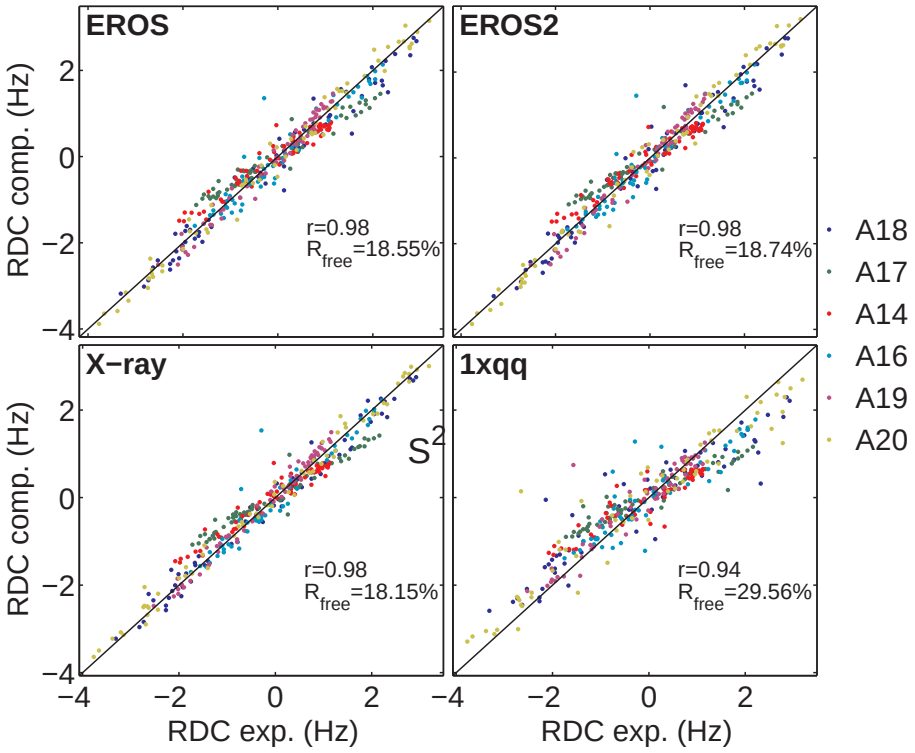

FIGURE S1. Back-calculated amide N- carbonyl C residual dipolar couplings. The colors correspond to different alignment media (cf. legend)

3.1. **Scalar couplings.** Scalar couplings ${}^3J$ were not used in the ensemble refinement and thus are a possible candidate to further cross-validate the structural ensemble. Scalar couplings were computed using the GROMACS Software suite[4] using the parametrization of the respective Karplus curves given in 1d3z[2] for the backbone phi-dihedral and with the parametrization given in Ref. [24] (supporting information therein Table 3, Set 2) for the side-chain dihedral angles. As shown in Table S2 and Figure S2, the experimental data[24] is well reproduced by EROS ensembles. The scalar couplings were computed for each conformer and subsequently averaged. A pronounced effect of dynamical averaging on the scalar couplings can be seen for the side-chain dihedrals by considering the scalar couplings computed for the average structure of EROS (EROS_avg) or a single X-ray structure. This dynamical averaging effect is, expectedly so, much smaller for backbone dihedrals.

3.2. **Scalar couplings across hydrogen bonds.** Scalar couplings across hydrogen bonds ${}^{3h}J_{\mathrm{NC'}}$ are time averages on time-scales similar to the ones probed by RDCs[25]. Because

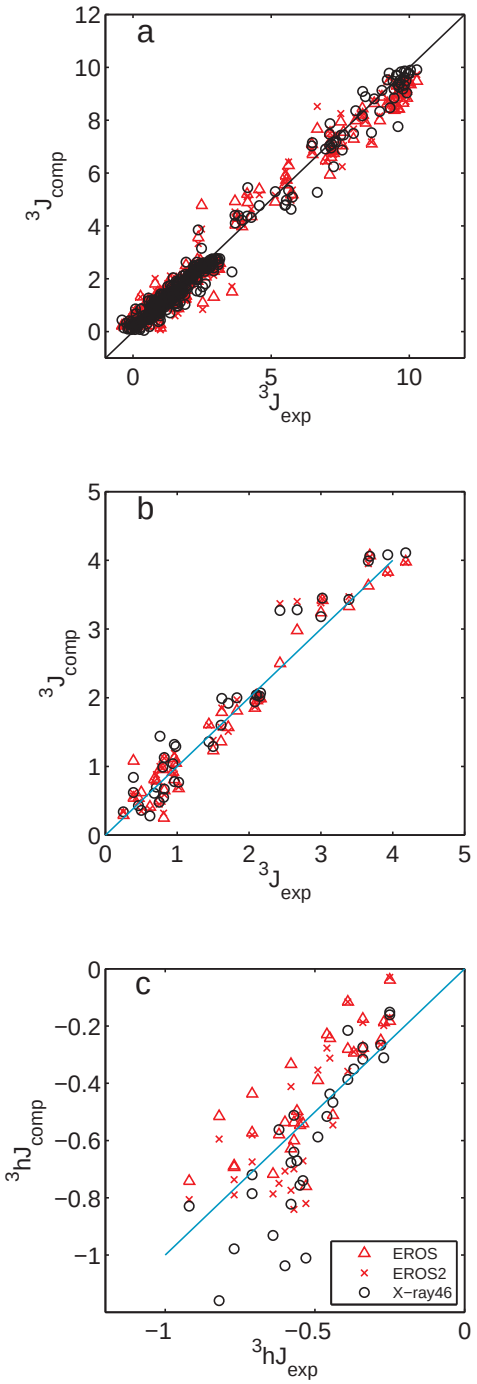

FIGURE S2. Cross-validation of ubiquitin structures by comparison with independently determined NMR data. Calculation of **(a)** backbone scalar couplings ^3J, **(b)** side-chain scalar couplings, and **(c)** scalar couplings across hydrogen bonds $^{3h}J_{NC'}$.

of their strong dependence on H-bond geometries, they are used to cross-validate structural data[26, 17]. We have computed $^{3h}J_{\mathrm{NC'}}$ for several structural ensembles using equation (6) in Ref. [27] which has been parametrized against results obtained with density functional theory. Predicted $^{3h}J_{\mathrm{NC'}}$ are plotted against experimental values[28] in Fig. S2; corresponding correlation coefficients and R_{free}, see Table S2. All ensembles show reasonable agreement, in line with previous findings[29]. As for the side-chain dihedrals χ a strong effect of dynamical averaging is observed.

4. Comparison of RDC-derived order parameters with NMR relaxation data

As explained in Section 1.2, relative, but not absolute order parameters of internuclear vectors can be derived from RDCs. For SCRM, therefore, RDC-based NH order parameters are shown relative to Lipari-Szabo relaxation order parameters. For EROS, absolute order parameters were derived from the RDC-refined ensemble. As the absolute dynamics is not provided by the RDCs, the overall amplitude of order parameters derived from an EROS ensemble is determined by the force field and the simulation protocol and subject to a certain uncertainty, which we estimate below. In particular, the ensemble size and the molecular mechanics force field applied will therefore determine the overall amplitude of the order parameters, whereas the relative distribution is largely determined by the RDCs (see also Section 6). We therefore investigated the role of the ensemble size and molecular mechanics force field applied in EROS on the obtained order parameters. To this end, a systematic study using a synthetic test case was carried out with ensemble sizes of 2,4, and 8 (Fig. S3). This analysis reveals that the used limited ensemble size of eight leads to an overestimation of order parameters by approximately 3%. In addition, it has been found that fast librations of NH bonds reduce the value of the RDC, which can be accounted for by using an effective NH bond length correction from 1.02 to 1.04 Å [15] and translates into an order parameter contribution of approx. 12%. Molecular dynamics simulations of lysozyme with fixed bond length suggest a uniform libration contribution of approx. 8% in the order parameters[30], which implies that this libration effect is partially, but not fully accounted for by molecular mechanics force-fields. We therefore assume that the libration contribution is underestimated in the EROS ensemble by approx. 4%. Based on this 4%

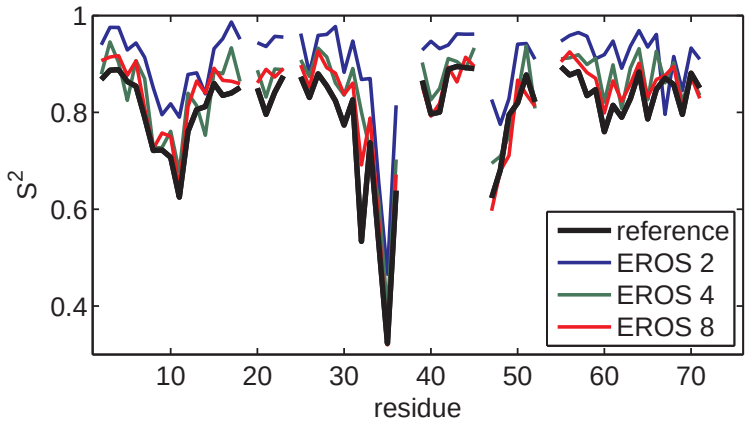

FIGURE S3. Influence of ensemble size on overall scaling of EROS order parameter. The reference shows $S^2\,(\mathrm{NH})$ obtained from a 168ns molecular dynamics simulation in explicit solvent. EROS was carried with ensemble sizes 2, 4 and 8 respectively. For the refinement we used 36 sets of synthetic RDCs generated from the molecular dynamics ensemble, analogous to the experimental data. The alignment tensors were obtained by fitting the trajectory against experimental data sets 1-36, couplings missing in these sets were also left out from the synthetic data set. Not shown are S^2_{EROS} of residues where less than 5 couplings were present in the synthetic data set. Apparently, ensemble sizes 2 and 4 underestimate the dynamics in the reference ensemble substantially. A small dampening of the dynamics of about 3% is observable for ensemble size 8.

and the above effect of limited ensemble size (3%), we thus expect that our EROS derived NH order parameters are too large by approx. 7%. To allow comparison with the Lipari-Szabo relaxation order parameters shown in Fig. 3 of the main text, the EROS order parameters were therefore scaled by 0.93. To properly reflect that this librational correction is necessarily an estimate, the resulting uncertainty in the scaling of the order parameters is depicted as a grey shaded region of $\pm 4\%$ in Fig. 3 of the main text. Note that this correction to the order parameters corresponds to a slight enhancement of the flexibility in the ensemble, leaving the conclusions about the recognition dynamics of ubiquitin unaffected.

5. ANALYSIS OF X-RAY BINDING INTERFACES

Binding interfaces were analyzed for the complexes listed in Table S3. As *contacting residues* we selected all residues of a binding partner (Tab. S3, column 2) with at least one atom in less than 5 Å distance of any atom of the bound ubiquitin structure (Tab. S3, column 1). These contacting residues are shown in Figure 4a as gray spheres together with the unbound protein structure (PDB code 1UBI). Note that although some of the

contacting residues are slightly more than $5\,\text{Å}$ away from the shown structure, they are nevertheless within $5\,\text{Å}$ of the respective bound structure. Biologically irrelevant crystal packing contacts were ignored. Redundancies of binding interfaces, e.g., 2C7M chain A and 2C7N chain C, were removed. However, the redundancy only slightly affects the results, as seen in Figure S4, which shows the number of contacts counting also redundant interfaces.

The *number of contacts* plotted in Figure 4c was calculated by counting for every residue all atoms from the selection described above ($5\,\text{Å}$ from bound structure) that are within $10\,\text{Å}$ of any atom of the respective residue in the unbound structure. This number count is divided by the number of atoms in the residue. The overall shape of the curve is robust against variations in the methodical details. In particular, omitting the normalization by residue size, or changing the cut-off radius from $10\,\text{Å}$ to smaller ($7\,\text{Å}$) or larger values ($12\,\text{Å}$) has no significant influence on the similarity with S^2_{EROS}.

We note that the number of contacts is affected by the chance that some binding motifs might be undiscovered so far. Moreover, contacts and S^2_{EROS} are very unlike observables which have no directly observable common a-priori physical basis. This renders their correlation even more notable, especially when considering that the number of 21 binding interfaces analyzed here bears already considerable statistical relevance.

5.1. **Analysis of interface adaptation from solution ensemble.** PCA was carried out on all (613) heavy atoms of residues 2-70 (and amide protons) after least squares fitting to backbone heavy atoms of residues 2-70 of the X-ray conformer 1UBI. The third principal component shows considerable similarity to the first principal component of the X-ray ensemble, and yields a relatively high correlation of 0.84 for a binding interface analysis analogously to Figure 5a (main text). Nevertheless, an improved collective mode was obtained as that linear combination of the first five principal components that maximizes the overlap with the X-ray ensemble.

To analyze the interface adaptation dynamics along the collective solution mode, $n = 30$ snapshots were generated at equal distances between the extremes obtained by projecting the EROS ensemble to the extracted mode. The snapshots were aligned in PYMOL with the apo ubiquitin structure (1UBI) which was also used to align the interfaces (see above). For each combination of snapshot and binding partner the number of contacts

PDB CODE	Ubiquitin	Binding Partner	# structures used for X-ray ensemble	Comments
1NBF	C	B	2	crystal contacts (chain D,A) ignored
1AAR	A	B	2	polyubiquitin
1AAR	B	A	-	
1CMX	B	A	1	contacts to C ignored
1P3Q	V	R,Q	2	contacts to U ignored
1S1Q	D	C	2	
1TBE	A	B	2	polyubiquitin
1TBE	B	A	-	
1UZX	B	A	1	ignored: redundant to 1S1Q
1XD3	B	A	2	contacts to C ignored
1YD8	U	H	2	
2AYO	B	A	1	
2C7M	B	A	1	ignored: redundant to 2C7N_A
2C7N_A	J	C	5	
2C7N_B	H	I	1	
2D3G	A,B	P	2	ignored: redundant to 2C7N_A
2FCQ	A	B	2	
2FID	A	B	1	ignored: redundant to 2C7N_A
2FIF	A	B	3	ignored: redundant to 2C7N_A
2G45	B	A	2	
1WR6	E	A	4	ignored: redundant to 1YD8
1WRD	B	A	1	ignored: redundant to 1YD8
1OTR, 1WR1	-	-		ignored NMR structures
1YIW			3	ignored: unbound ubiquitin
1UBI			1	
1UBQ			1	
1F9J	B	A	2	

TABLE S3. PDB codes and Chain IDs of crystallographic data used for contact analysis.

T with $3\,\text{Å} < d < 8\,\text{Å}$ (with d the contact distance) was counted to yield a quasi-energy $E = -T$. For each binding partner the snapshot with the lowest quasi-energy was selected. The x-values in Figure 5a were obtained by projecting the respective X-ray structures of ubiquitin to the collective solution mode. The "predicted" value plotted on the y-axis was obtained by interpolating between the minimal and maximal values found for the complex structure, namely $-1.71\,\text{nm}$ and $+1.13\,\text{nm}$, respectively. Then, according to the selected snapshot i, the y value was calculated as $y = -1.71 + (i-1)\,(1.13 + 1.71)\,/n$, where i, is the running number of the snapshot that has the minimal "quasi" contact-energy for the respective binding partner.

Summary of modes that were analyzed in this manner:

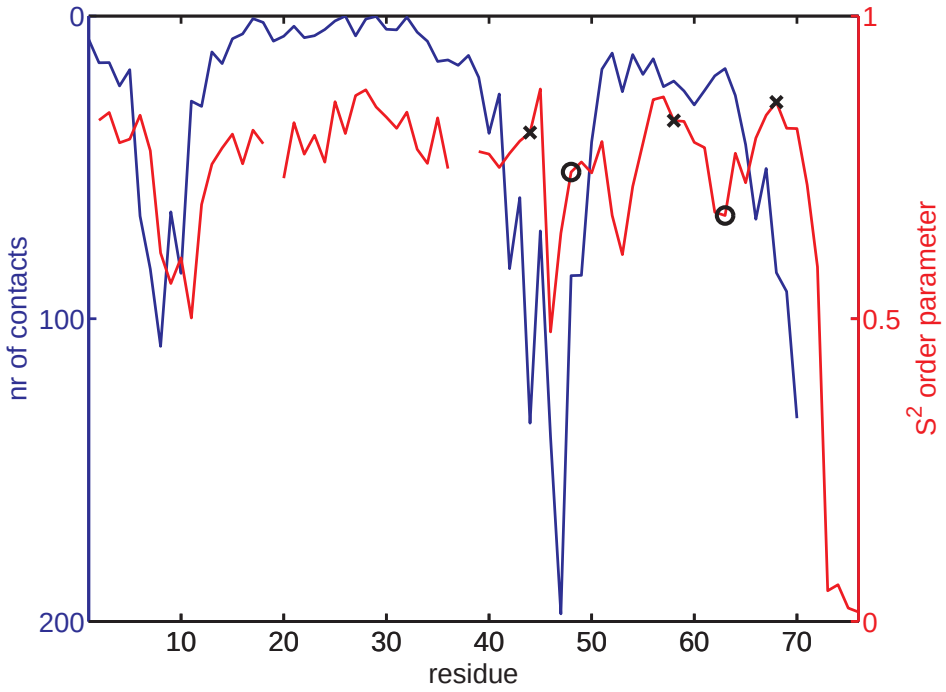

FIGURE S4. This figure shows the same analysis as Figure 4c in the main text but this time considering all (including redundant) ubiquitin complexes. Ubiquitin-binding protein contacts per residue (blue) in comparison to the flexibility in solution (red). Residues known to be crucial for molecular recognition are marked with crosses (ILE44/HIS68, ASP58). Lysines responsible for polyubiquitination are marked with circles (LYS48, LYS63).

mode3 of EROS: 0.84

linear combination of 5 PCA modes of EROS: 0.94

linear combination of 3 PCA modes of EROS: 0.9

linear combination of 5 PCA modes of EROS2: 0.85

linear combination of 5 PCA modes of 2nr2: 0.55

linear combination of 5 PCA modes of 1xqq: 0.64

6. VALIDATION OF EROS APPROACH

Cross-validation shows that the EROS ensembles agree favorably with the available experimental data. For further validation of our approach we tested the robustness of the generalized order parameters S^2_{EROS} (NH) of the ensembles and main principal modes obtained with principal component analysis (PCA), because these are the important observables on which our conclusions drawn in the main text rest. In particular, we systematically

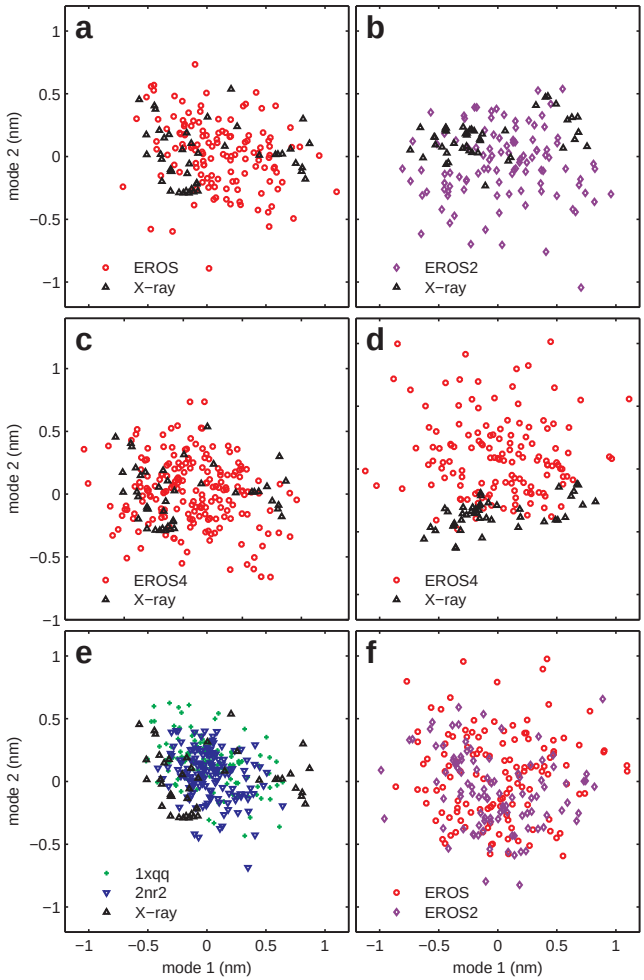

FIGURE S5. Structural comparison of ubiquitin ensembles EROS, EROS2, EROS4, 1xqq, 2nr2, and X-ray. **(a,c,e)** Projection of ensembles onto the principal modes that describe the interface-adaptation motion manifested in X-ray structures. **(b,d,f)** Direct check for systematic structural differences on the order of the overall ensemble size by carrying out PCA on pairs of ensembles that haven't been already shown in Fig. 2 of the main text. For each shown pair of ensembles (cf. legends) the principal components were extracted from the respective ensemble pairs. Note that this is different from panels (a,c,e), where the PCA was carried out over the collection of X-ray conformers, explaining the same appearance of projections of Xray structures in all three panels (a,c,e).

investigated the influence of 1) starting structure, 2) choice of force-fields, and 3) selection of RDC-data (NH only, all non NC) on our results. In the following we consider only those S^2 for which data are presented in Figure 3 of the main text. These are the residues for which model-free results were available (cf. Methods).

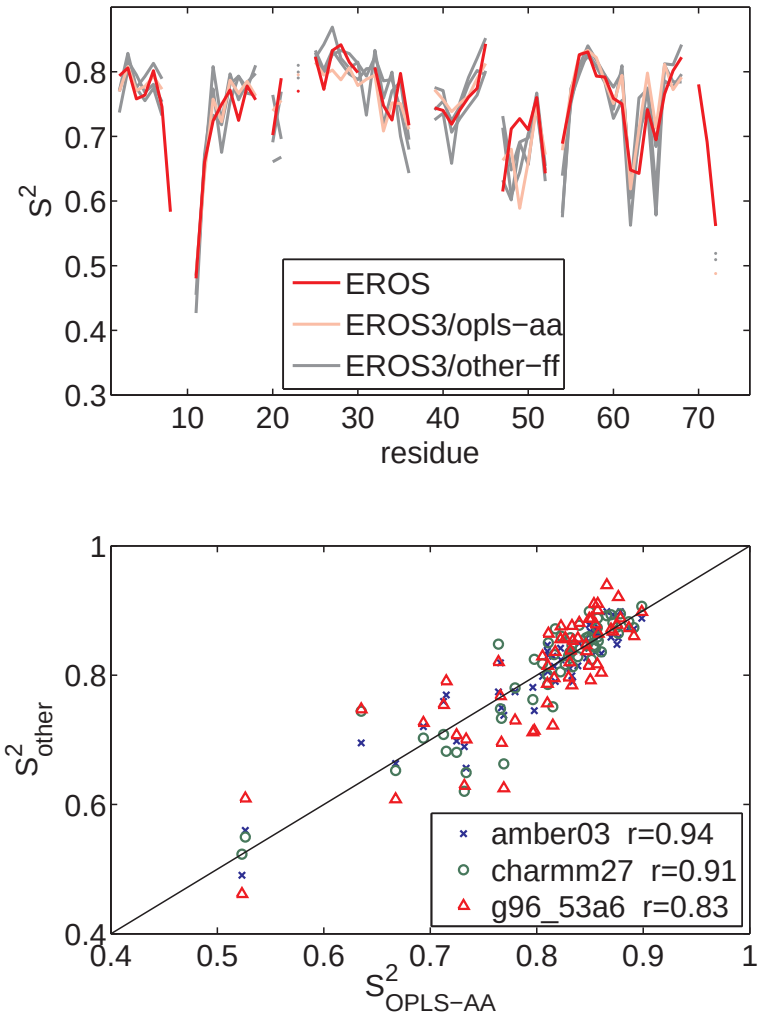

FIGURE S6. Influence of molecular dynamics force-field on NH order parameters. Upper panel: comparison of ensembles that were refined against NH RDC data sets B1-B11 and A1-A30 with different molecular mechanics force-fields. (gray: amber03, charmm27, g96_53a6; pink: opls-aa). For comparison, the order parameters S^2_{EROS} presented in Fig. 3 of the main text (opls-aa, experiments A1-A36) are shown in red. Lower panel: scatter plot of order parameters derived using the amber03, charmm27, and g96_53a6 force-fields compared to the opls-aa force field.

6.1. **Completeness of sampling and overall scaling.** To systematically assess the effect of possibly insufficient sampling in the molecular dynamics refinement step of our protocol, we started refinement from two drastically different situations. EROS refinement was started from a broad pool of structures that were obtained by CONCOORD NOE refinement from extended conformations without ever being individually refined against RDCs

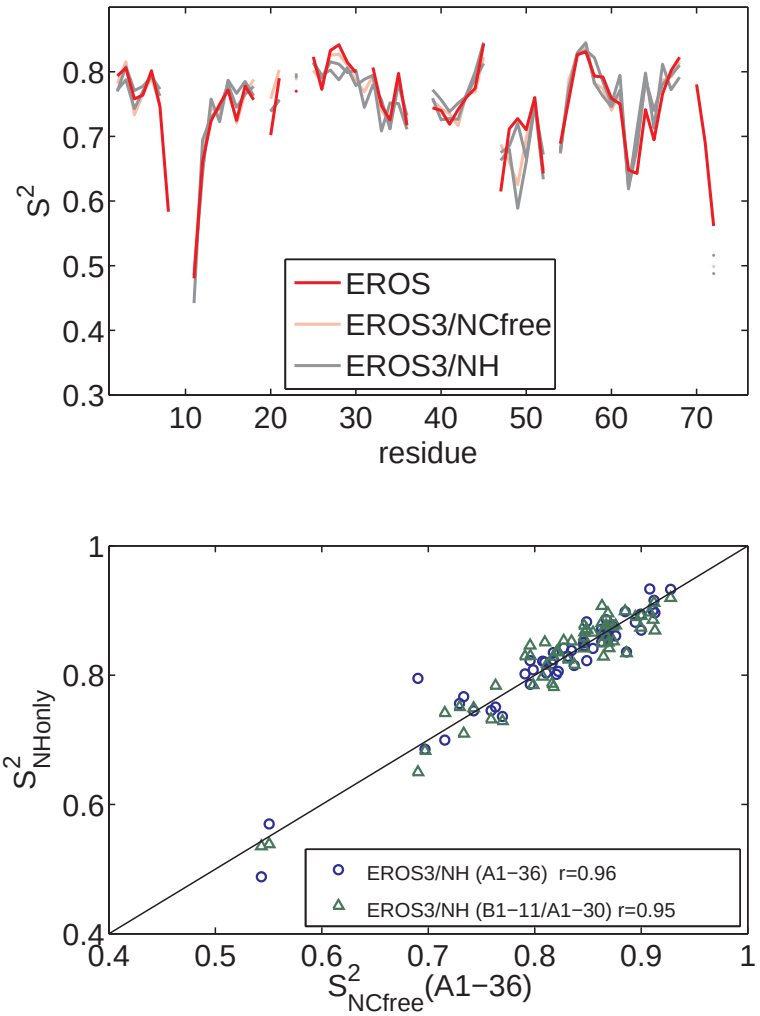

FIGURE S7. Influence of the choice of RDC data on NH order parameter S^2. Upper panel: NH order parameters obtained from refinement carried out against residual dipolar couplings of either solely amide NH (denoted NH; gray curves) or all available couplings, but leaving N-C RDCs for cross-validation (denoted NCfree; pink/red curve). Either alignment media A1-A36 were used (pink/red/gray), or alignment media B1-B11 und A1-A30 (gray). For comparison the order parameters S^2_{EROS} presented in Fig. 3 of the main text (NCfree; A1-A36) are shown in red. These differ from the pink curve in terms of starting structures (see text). Lower panel: scatter plot of NH order parameters obtained from NH-only refinement against datasets B1-B11/A1-A30 and A1-A36 compared to the NCfree refinement against datasets A1-A36.

or NOEs (cf. Section S1). Thus, this pool contained considerable structural diversity (cf. Figure S9). In contrast, EROS2 refinement started from a single structure (cf. Section 1). Despite these radically different starting conditions the resulting ensembles both occupy the same region in the PCA projection shown in Fig. S9. Moreover, for both ensembles

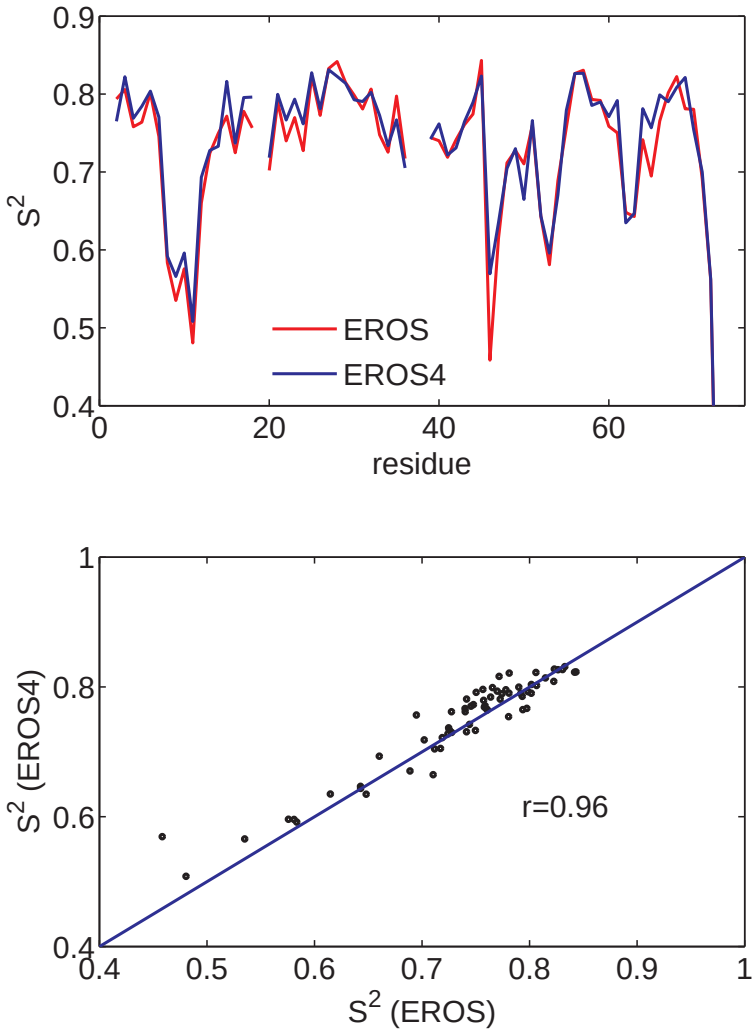

FIGURE S8. Influence of the choice of NOE data on NH order parameter S^2. Upper panel: NH order parameters obtained from ensemble refinement carried out against residual dipolar couplings in the presence of NOE data as addtional restraints (EROS) is compared to NH order parameters obtained without NOE data in the ensemble refinement (EROS4, cf. Methods). Lower panel: scatter plot of order parameters of EROS and EROS4.

the PCA shows a complete coverage of all X-ray structures (cf. Figure S5b, S5e), even the conservative EROS2 ensemble that originated from one single conformer. Finally, as can be seen in Figure S5f, EROS2 and EROS overlap strongly in the PCA projection and hence show no major systematic structural differences. Accordingly, both ensembles are indistinguishable by cross-validation (cf. Section 3).

Figure S9 shows that the RDCs have a remarkably specific effect on the structural ensembles. Starting from ensembles taken from the broad initial pool (gray), EROS sampling

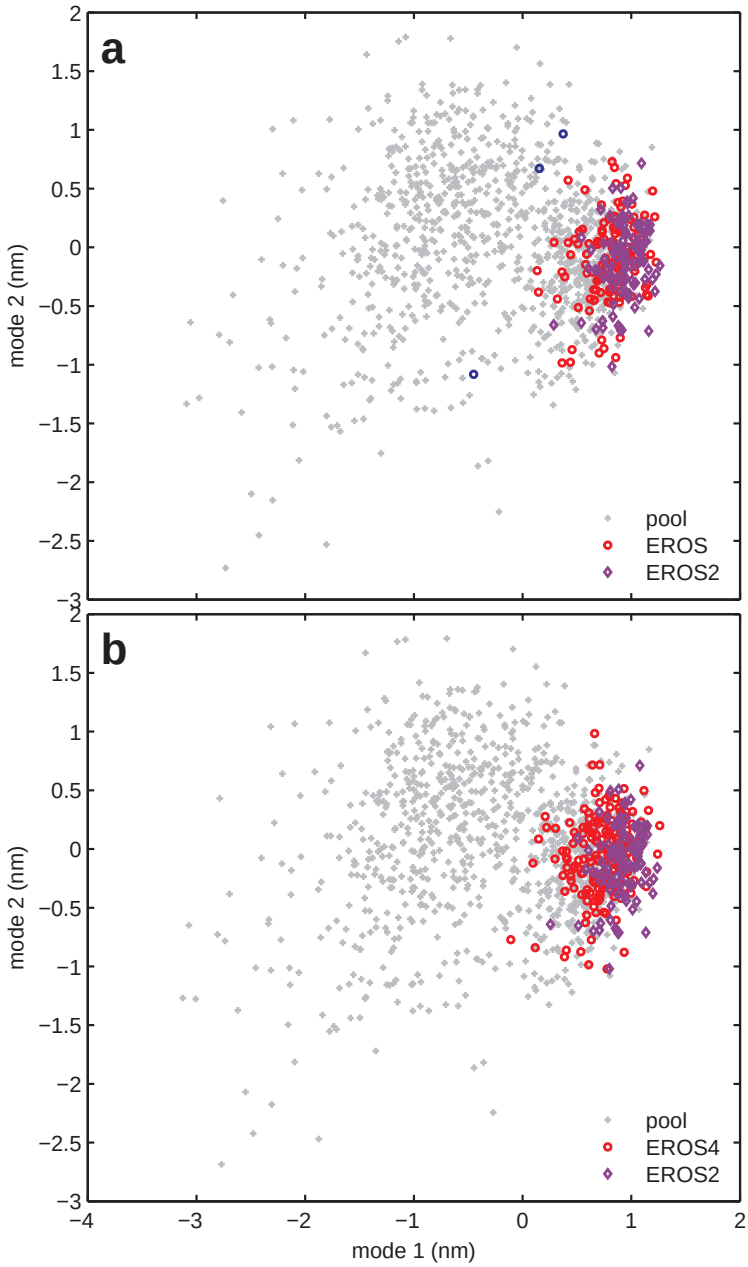

FIGURE S9. PCA projection plot to illustrate the convergence process of EROS ensemble refinement. The starting pool (gray) consists of structural pairs of ubiquitin that fulfill all NOE restraints (generated by CONCOORD, see Methods). The final EROS ensembles are shown in red and purple, respectively. The blue dots in (a) mark three structures that have been manually removed from the EROS ensemble (see text).

under the influence of the RDCs leads to a consistent drift towards the region at the right side of Figure S9 (colored circles) which is also sampled by EROS2 (diamonds). The same result is obtained for the EROS4 refinement that in contrast to the EROS refinement does not apply any NOE restraints in addition to the RDCs. This finding shows that the restraints obtained from the RDCs are sufficient to overcome sampling barriers in most instances and, therefore, the ensemble is well defined up to the remaining small differences in the overall amplitude between EROS/EROS4 and EROS2.

Three structures of the nascent EROS ensemble (Figure S9, blue circles) lie still significantly outside the region that is sampled consistently by EROS, EROS2 and EROS4, i.e., the region sampled independently of the starting conditions. Although the pool has two densely populated regions, all but these three EROS structures are found only within the region on the right side. It seems that a strong energetic barrier separates the two pool regions. Nevertheless, the pull of the RDCs is sufficient to move all but three structures into the region on the right side. Since the three outliers are found exactly at the edge of the region on the left, we concluded that for these three structures frustrated sampling may have occurred due to this barrier and, thus, a significant bias of the starting conditions may remain. Accordingly, these three structures were removed from the final EROS ensemble.

That the EROS refinement is independent from the chosen starting conditions is also supported by a comparison of order parameters S^2_{NH} of EROS and EROS2. Indeed, the correlation between the two sets of order parameters is high (r=0.9) confirming that both ensembles describe the same NH dynamics. The slightly smaller spread of the EROS2 ensemble revealed by PCA is reflected in a 5% increased overall amplitude of the order parameters. Note that the EROS2 ensemble suffers from limited sampling, as refinement started from a single configuration. It therefore can be considered a minimal representation of the dynamics required to fulfil the RDCs. Hence, despite that the overall amplitude of the dynamics is not directly provided by the RDCs, as discussed in section 1.3, the fact that the EROS2 ensemble shows a complete coverage of the X-ray conformers, confirms that the conclusions about the molecular recognition dynamics of ubiquitin are independent of the overall amplitude.

6.2. **Influence of molecular mechanics force-fields.** EROS ensembles have been refined against the RDCs using the state of the art all-atom force-field OPLS-AA. Here we addressed whether the dynamics as reflected by S^2 are sensitive to the applied force-field. Refinement of EROS3 was carried out with the force-fields OPLS-AA/L, AMBER03, Charmm27 and GROMOS96_53A6 (cf. Methods). As seen in Fig. S6, of EROS3 ensembles (pink and gray curves) yield order parameters similar to EROS (red curve). Moreover, correlation coefficients between order parameters derived from ensembles from the different force fields are on average r=0.9, with a slightly lower correspondence to the G96_53A6-EROS3 (r=0.83 to OPLSAA-EROS3). Hence the choice of the molecular dynamics force-field does not strongly affect the ensemble, supporting our previous conclusion that the ensembles are strongly determined by the RDC data.

6.3. **Selection of RDC data for refinement.** Whether our results depend on the choice of RDC data was tested by carrying out refinement with two different collections of data sets. In total, there were 47 data sets available to us for ubiquitin. The 36 data sets used throughout this work are denoted A1-A36. 11 of these data sets replace the oldest available data sets denoted B1-B11 that were obtained with much less concentrated samples than the newer sets. A control data set was constructed by selecting experiments A1-A30 and B1-B11, i.e., it was lacking the most recently acquired data sets A31-A36.

As seen in Figure S7, the change to the more recent data set has no significant effect on the S^2_{EROS}. Moreover, Figure S7b shows that using only NH RDCs instead of all RDCs (except N-C, cf. cross-validation) has no significant effect on the S^2_{EROS} (NH). The ensemble refined against all non-NC work RDCs (EROS/EROS2) has a significant lower (NC) R_{free} value as compared to the case when only NH RDC's are used in refinement. Accordingly, this strategy was used for the ensemble presented in the main text.

7. Experimental Data

7.1. **Experimental data used for refinement of EROS ensembles.** In this section we summarize the used RDCs: In total 47 RDC data sets have been extracted from the literature or measured in our lab. 11 of these data sets were obtained with much less concentrated samples than newer sets, and thus not used for our final refinement. The remaining 36 data sets were labelled A1-A36 (see Table S4). Since the older data the was used for EROS method validation (see Section 6), we list them here with labels B1-B11 (see Table S4).

For some of the 36 alignment conditions also other residual dipolar couplings than the amide NH were measured. A summary of these are given in Table S5 for backbone couplings and in Table S8 for methyl sidechains. The individual non-NH backbone couplings for A14, A15, A17 and A18 are given in Tables S6-S7. non-NH couplings for A19 and A20 are published elsewhere[15]. The NH couplings that are newly measured are given in the supporting information of Ref. [22]. Excluded from refinement were the 373 NC' RDCs obtained in the 6 alignment media A14, A15, and A17-20. These NC' RDCs were used for cross-validation.

Additionally, 2727 H-H NOE restraints from the data set 1d3z[2] were used.

7.2. **HC' and NC' RDCs.** Alignment media have been prepared as described in Ref. [22]; the nomenclature is the same as used in that manuscript. Couplings have been measured using a pulse sequence similar to Ref. [35]. In addition, two further RDC data sets from the literature have been used[15] (cf. Table S5).

7.3. **Side-chain methyl group RDCs.** Methyl group RDCs have been measured using the method described in Ref. [36]. Nomenclature and sample preparation was the same as described in Ref. [22]. The measured RDCs have been scaled to the CC axis along the symmetry axis: RDC(CH3 (axial)) = 0.3155 * RDC(CH3)[37]. Additionally to the side-chain methyl data measured here, we used methyl data corresponding to alignment condition A20 as published here[37].

REFERENCES

[1] B. L. de Groot, *et al.*, *Proteins* **29**, 240 (1997).

[2] G. Cornilescu, J. L. Marquardt, M. Ottiger, A. Bax, *J. Am. Chem. Soc.* **120**, 6836 (1998).

[3] B. Hess, R. M. Scheek, *J. Mag. Res.* **164**, 19 (2003).

[4] D. V. der Spoel, *et al.*, *J. Comput. Chem.* **26**, 1701 (2005).

[5] U. Essmann, *et al.*, *J. Chem. Phys.* **103**, 8577 (1995).

[6] T. Darden, D. York, L. Pedersen, *J. Chem. Phys.* **98**, 10089 (1993).

[7] W. L. Jorgensen, D. S. Maxwell, J. Tirado-Rives, *J. Am. Chem. Soc.* **118**, 11225 (1996).

[8] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, J. Hermans, *Intermolecular Forces*, B. Pullman, ed. (D. Reidel Publishing Company, Dordrecht, 1981), pp. 331–342.

[9] A. D. MacKerell, *et al.*, *J. Phys. Chem. B* **102**, 3586 (1998).

[10] Y. Duan, *et al.*, *J. Comput. Chem.* **24**, 1999 (2003).

[11] E. J. Sorin, V. S. Pande, *Biophys. J.* **88**, 2472 (2005).

[12] C. Oostenbrink, A. Villa, A. E. Mark, W. F. V. Gunsteren, *J. Comput. Chem.* **25**, 1656 (2004).

[13] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **79**, 926 (1983).

[14] B. Hess, H. Bekker, H. J. C. Berendsen, J. G. E. M. Fraaije, *J. Comput. Chem.* **18**, 1463 (1997).

[15] M. Ottiger, A. Bax, *J. Am. Chem. Soc.* **120**, 12334 (1998).

[16] A. M. J. J. Bonvin, A. T. Brunger, *J. Mol. Biol.* **250**, 80 (1995).

[17] B. Richter, J. Gsponer, P. Varnai, X. Salvatella, M. Vendruscolo, *J. Biomol. NMR* **37**, 117 (2007).

[18] J. Meiler, J. J. Prompers, W. Peti, C. Griesinger, R. Bruschweiler, *J. Am. Chem. Soc.* **123**, 6098 (2001).

[19] N. A. Lakomek, *et al.*, *Angew. Chem.-Int. Ed.* **44**, 7776 (2005).

[20] N. A. Lakomek, T. Carlomagno, S. Becker, C. Griesinger, J. Meiler, *J. Biomol. NMR* **34**, 101 (2006).

[21] J. W. J. Kerssemakers, *et al.*, *Nature* **442**, 709 (2006).

[22] N. A. Lakomek, *et al.*, *J. Biomol. NMR*. DOI: 10.1007/s10858–008–9244–4 (2008).

[23] G. Vriend, *J. Mol. Graph.* **8**, 52 (1990).

[24] J. J. Chou, D. A. Case, A. Bax, *J. Am. Chem. Soc.* **125**, 8959 (2003).

[25] S. Grzesiek, F. Cordier, A. J. Dingley (2001), vol. 338 of *Methods in Enzymology*, pp. 111–133.

[26] G. Bouvignies, *et al.*, *Proc. Natl. Acad. Sci. U. S. A.* **102**, 13885 (2005).

[27] M. Barfield, A. J. Dingley, J. Feigon, S. Grzesiek, *J. Am. Chem. Soc.* **123**, 4014 (2001).

[28] F. Cordier, S. Grzesiek, *J. Am. Chem. Soc.* **121**, 1601 (1999).

[29] H. J. Sass, F. F. F. Schmid, S. Grzesiek, *J. Am. Chem. Soc.* **129**, 5898 (2007).

[30] M. Buck, M. Karplus, *J. Am. Chem. Soc.* **121**, 9645 (1999).

[31] J. R. Tolman, *J. Am. Chem. Soc.* **124**, 12020 (2002).

[32] K. Ruan, J. R. Tolman, *J. Am. Chem. Soc.* **127**, 15032 (2005).

[33] W. Peti, J. Meiler, R. Bruschweiler, C. Griesinger, *J. Am. Chem. Soc.* **124**, 5822 (2002).

[34] H.-J. Sass, G. Musco, S. J. Stahl, P. T. Wingfield, S. Grzesiek, *J. Biomol. NMR* **18**, 303 (2000).

[35] K. Ding, A. M. Gronenborn, *J. Mag. Res.* **158**, 173 (2002).

[36] G. Kontaxis, A. Bax, *J. Biomol. NMR* **20**, 77 (2001).

[37] M. Ottiger, A. Bax, *J. Am. Chem. Soc.* **121**, 4690 (1999).

Source	code	NH
SCRM 2008[22]	A1	60
SCRM 2008	A2	59
SCRM 2008	A3	56
SCRM 2008	A4	62
SCRM 2008	A5	48
SCRM 2008	A6	53
SCRM 2008	A7	62
SCRM 2008	A8	61
SCRM 2008	A9	56
SCRM 2008	A10	61
SCRM 2008	A11	59
SCRM 2008	A12	61
SCRM 2008	A13	59
Lakomek et al. 2006[20]	A14	51
Lakomek et al. 2006	A15	48
Lakomek et al. 2006	A16	53
Lakomek et al. 2006	A17	47
Lakomek et al. 2006	A18	44
Ottiger et al.[15]	A19	63
Ottiger et al.	A20	60
Tolman 2002[31]	A21	44
Tolman 2002	A22	46
Tolman 2002	A23	46
Tolman 2002	A24	45
Tolman 2002	A25	47
Tolman 2002	A26	44
Tolman 2002	A27	48
Tolman 2002	A28	44
Tolman 2002	A29	54
Ruan & Tolman 2005[32]	A30	43
Ruan & Tolman 2005	A31	37
Ruan & Tolman 2005	A32	48
Ruan & Tolman 2005	A33	51
Ruan & Tolman 2005	A34	50
Ruan & Tolman 2005	A35	56
Ruan & Tolman 2005	A36	61

Source	code	NH
Peti et al. 2002[33]	B1	61
Peti et al. 2002	B2	58
Peti et al. 2002	B3	52
Peti et al. 2002	B4	57
Peti et al. 2002	B5	67
Lakomek et al. 2006[20]	B6	59
Lakomek et al. 2006	B7	59
Lakomek et al. 2006	B8	51
Lakomek et al. 2006	B9	63
Lakomek et al. 2006	B10	35
Sass et al. 2000[34]	B11	56

TABLE S4. Experimental sources of NH RDCs used for EROS refinement and SCRM. The second column lists the code assigned to each experimental condition. The last column shows the number of couplings obtained for the amide NH intereaction in the respective alignment medium.

Source	code	NC'	HC'	C_αC'	$C_\alpha H_\alpha$	$C_\alpha C_\beta$
Ottiger et al.[15]	A19	61	61	58	62	39
Ottiger et al.	A20	63	63	54	62	
Lakomek et al. 2006[20]	A14	63	60			
Lakomek et al. 2006	A15	64	64			
Lakomek et al. 2006	A17	63	64			
Lakomek et al. 2006	A18	59	61			

TABLE S5. Experimental sources of the non-NH backbone residual dipolar couplings used for EROS refinement. The second column lists the code assigned to each experimental conditions. The last five columns show the numbers of couplings obtained for the respective types of heteronuclear coupling vectors.

NC’	A14	A15	A17	A18
	E1	E3	E4	E5
1	-2,21	-1,75	0,65	0,06
2	3,31	1,36	-0,91	0,58
3	-2,05	-1,33	1,98	-2,05
4	1,46	0,71	-1,10	0,29
5	-1,92	-0,88	1,66	-2,21
6	0,41	0,63	-1,06	0,34
7	2,27	1,30	-0,75	n
8	n	n	n	n
9	1,59	0,88	-0,33	0,75
10	-1,98	-1,53	0,55	-0,81
11	2,27	1,27	-0,41	n
12	-1,10	-0,65	0,03	0,75
13	n	0,55	0,86	-1,23
14	-0,13	0,58	-0,97	0,97
15	-2,79	-1,85	2,08	-1,88
16	-0,45	0,49	-0,89	-0,10
17	-0,75	-0,10	-0,13	1,43
18	n	n	n	n
19	-2,92	-2,16	1,82	-1,95
20	-0,36	n	n	1,07
21	1,67	0,47	-0,89	0,96
22	-0,03	0,23	1,17	-1,69
23	n	n	n	n
24	2,73	1,27	-0,33	0,10
25	-2,86	-1,85	1,20	-2,05
26	2,21	1,72	-0,81	0,91
27	n	n	1,51	n
28	1,85	0,49	-0,86	0,42
29	-1,01	-0,16	1,12	-1,92
30	-0,55	-0,75	-0,33	1,10
31	1,95	0,71	0,29	-0,45
32	-2,27	-1,72	0,81	-1,49
33	0,91	1,36	-0,03	-0,26
34	-2,32	-1,64	0,67	0,08
35	-1,44	-0,31	-0,31	-1,25

Table S6: NC’ experimental residual dipolar couplings for the alignment media A14, A15, A17, and A18 that have been newly measured for the presented analysis. The second row contains the corresponding nomenclature for the alignment media as used in Lakomek et al. 2006

NC’	A14	A15	A17	A18
36	n	n	n	n
37	n	n	n	n
38	2,63	0,97	-1,10	1,01
39	-0,32	0,73	-0,75	1,56
40	2,08	0,88	-0,03	n
41	-2,47	-2,05	1,46	-1,17
42	1,66	0,65	-0,33	0,29
43	-1,33	-1,36	0,33	0,39
44	0,37	0,57	-1,12	0,21
45	-0,75	-0,29	0,78	n
46	2,14	0,91	-1,04	0,58
47	-1,77	-1,45	1,96	-1,77
48	1,92	0,84	-1,07	1,07
49	-1,59	0,10	0,68	-1,59
50	-0,94	-0,42	0,10	0,78
51	0,03	-0,49	-0,62	-0,29
52	n	n	n	n
53	-2,19	-1,93	0,89	-0,44
54	2,82	1,23	-0,91	0,62
55	0,26	1,56	-0,88	1,14
56	-0,13	-0,42	0,81	0,71
57	2,53	1,07	-0,97	1,36
58	-0,29	0,34	-0,84	1,20
59	2,01	0,81	0,44	-0,32
60	-1,07	-0,52	-0,33	-1,07
61	-0,06	-0,29	-0,42	1,10
62	1,95	1,59	0,08	n
63	0,62	0,62	-0,97	0,36
64	-0,13	0,81	-0,68	1,75
65	-0,36	-0,33	1,56	0,29
66	-0,55	-0,81	-0,13	-0,39
67	1,75	1,43	0,16	-0,16
68	0,52	-0,26	0,03	-0,10
69	1,40	1,46	-1,09	1,36
70	2,16	1,25	-0,46	0,73
71	-1,56	-1,17	0,39	0,36

HC’	A14	A15	A17	A18
(2006)	E1	E3	E4	E5
1	n	n	n	n
2	2,48	2,73	0,48	-1,04
3	-4,41	-1,20	2,57	-1,20
4	7,45	2,89	-2,48	6,01
5	n	-0,72	2,32	n
6	5,21	0,48	-1,00	4,57
7	1,04	-1,52	2,32	-2,00
8	-4,57	-0,96	2,00	-2,00
9	n	n	n	n
10	-0,72	-0,96	2,16	-2,32
11	0,80	1,12	0,96	-1,12
12	-4,17	-3,73	0,40	-0,96
13	0,88	-0,80	1,68	-3,77
14	3,01	-1,00	-0,84	3,17
15	0,72	-0,96	2,57	0,08
16	7,93	4,81	-2,32	4,89
17	0,40	0,40	1,36	1,52
18	-0,96	-0,40	0,88	-3,37
19	n	n	n	n
20	4,97	2,57	-3,21	5,77
21	2,00	0,56	-0,80	-3,29
22	2,97	1,76	0,12	1,04
23	n	-1,36	-0,88	2,08
24	n	n	n	n
25	-5,61	-2,00	2,16	-1,12
26	5,53	1,04	-0,32	4,09
27	-2,89	-3,13	1,44	-2,72
28	2,24	1,52	-1,04	-0,64
29	-2,40	-0,56	2,20	0,16
30	2,48	-1,12	-0,16	2,16
31	0,32	0,48	0,72	-3,53
32	-4,49	-1,52	0,56	-0,96
33	4,65	1,92	0,00	3,53
34	-0,64	-2,65	1,44	-1,76
35	3,21	3,45	-1,44	0,96
36	-1,36	0,32	1,20	0,56

Table S7: HC’ experimental residual dipolar couplings for alignment conditions A14, A15, A17, and A18 that have been newly measured for the presented analysis.

HC’	A14	A15	A17	A18
37	n	n	N	n
38	n	n	N	n
39	-1,52	0,72	1,36	0,08
40	n	-2,61	1,68	-1,28
41	1,76	1,20	-1,84	1,60
42	5,77	4,25	-2,48	n
43	-0,24	0,24	0,16	-1,84
44	2,08	1,04	0,88	-2,72
45	1,20	-1,76	1,12	-1,68
46	n	n	N	n
47	-2,16	-1,48	2,73	-1,68
48	-0,64	-0,16	-2,40	3,37
49	-5,57	-3,41	2,61	-3,17
50	2,08	-1,60	0,48	0,16
51	-0,68	-0,76	1,32	-1,32
52	-7,13	-1,68	2,00	-1,68
53	n	n	N	n
54	4,49	3,69	-2,57	0,64
55	-4,17	-1,60	1,52	-0,64
56	1,84	-1,76	0,24	0,88
57	4,57	3,37	-3,77	3,45
58	-2,32	0,08	1,52	-0,88
59	0,80	-2,00	2,00	-0,80
60	0,32	1,52	-1,76	2,08
61	4,41	3,29	-1,20	2,48
62	4,65	3,29	-2,08	0,16
63	-6,73	-2,48	-0,24	0,96
64	n	n	1,44	1,68
65	2,89	1,36	0,08	-2,40
66	6,01	2,52	-3,69	6,01
67	4,49	2,16	-0,64	3,69
68	-1,68	-2,89	1,20	-0,72
69	2,40	0,48	-1,36	n
70	0,24	-0,96	1,60	-4,57
71	-0,96	0,24	-0,40	-0,96

methyl		Spin1	Spin2	A1	A2	A3	A4	A7	A8	A9	A10	A12	A13
AA	Res												
MET	1	SD	CE	n	n	n	n	N	n	n	n	n	n
ILE	3	CG1	CD1	0,61	0,80	2,46	-1,40	-1,10	-0,55	1,06	-0,20	-0,98	-1,87
ILE	3	CB	CG2	0,90	1,43	2,54	n	-1,13	-1,34	-4,24	-0,81	-1,76	-2,36
VAL	5	CB	CG1	n	n	n	n	N	n	n	n	n	n
VAL	5	CB	CG2	-0,16	-0,15	-0,70	1,49	0,65	-0,21	-2,01	-0,04	0,14	0,95
THR	7	CB	CG2	-0,77	-0,86	-5,62	1,35	0,90	1,06	-0,18	0,09	0,84	1,89
LEU	8	CG	CD1	0,53	0,66	1,22	-1,01	-1,20	-0,67	-1,08	-0,73	-1,44	-1,95
LEU	8	CG	CD2	n	n	n	n	N	n	n	n	n	n
THR	9	CB	CG2	0,57	1,05	2,25	-2,36	-3,17	-1,28	-2,60	-1,79	-3,64	-3,94
THR	12	CB	CG2	0,37	1,19	1,54	0,14	-1,09	-1,22	-4,59	-0,82	-1,84	-1,90
ILE	13	CG1	CD1	-0,01	-0,05	1,42	-1,20	-0,55	-0,17	1,24	-0,25	-0,64	-0,79
ILE	13	CB	CG2	-0,72	-0,95	-2,76	1,40	1,43	0,86	0,23	0,53	1,31	2,30
THR	14	CB	CG2	0,42	1,29	1,14	-0,53	-2,21	-1,24	-4,99	-1,52	-2,96	-3,35
LEU	15	CG	CD1	n	n	n	n	n	n	n	n	n	n
LEU	15	CG	CD2	-0,13	-0,48	0,38	0,53	1,43	0,35	1,73	0,86	0,31	1,25
VAL	17	CB	CG1	n	n	n	n	n	n	n	n	n	n
VAL	17	CB	CG2	-1,13	-2,10	-1,55	1,20	2,52	1,23	3,74	1,47	2,91	4,19
THR	22	CB	CG2	-0,19	-0,24	0,11	1,68	0,94	-0,52	-1,93	0,14	0,37	0,82
ILE	23	CG1	CD1	0,26	0,33	2,59	-1,16	-1,39	-0,66	0,47	-0,37	-1,15	-1,54
ILE	23	CB	CG2	0,31	1,00	-0,17	-0,05	-1,66	-1,08	-5,52	-1,45	-2,41	-2,32
VAL	26	CB	CG1	-0,93	-1,72	-0,84	1,93	3,05	0,90	2,87	1,22	2,81	4,24
VAL	26	CB	CG2	n	n	n	n	n	n	n	n	n	n
ALA	28	CA	CB	n	n	n	n	n	n	n	n	n	n
ILE	30	CG1	CD1	0,11	-0,48	-2,55	0,48	1,75	1,05	3,13	1,27	2,14	1,93
ILE	30	CB	CG2	0,04	0,43	-1,45	0,38	-1,51	-0,69	-4,89	-1,06	-1,66	-2,47
ILE	36	CG1	CD1	-0,54	-1,00	-0,04	0,34	1,25	0,45	1,13	0,44	0,92	1,52
ILE	36	CB	CG2	0,53	0,81	-1,32	-1,16	-1,44	0,03	0,41	-0,33	-0,78	-1,68
LEU	43	CG	CD1	n	n	n	n	n	n	n	n	n	n
LEU	43	CG	CD2	-0,04	-0,20	1,66	0,39	1,33	0,07	2,02	0,80	1,51	1,24

Table S8: Experimental residual dipolar couplings for methyl groups, that have been newly measured for the presented analysis. Data set A20 has been taken from Ref. [36].

methyl		Spin1	Spin2	A7	A2	A1	A3	A4	A9	A8	A10	A13	A12	A20
AA	Res													
ILE	44	CG1	CD1	-0,11	0,28	0,10	0,01	0,10	-1,12	-0,27	-0,32	-0,32	-0,41	-0.25
ILE	44	CB	CG2	2,57	-1,00	-0,30	-1,85	1,11	3,86	1,24	1,59	3,44	2,81	3.15
ALA	46	CA	CB	-2,82	1,87	1,01	2,89	-1,06	-7,63	-2,42	-2,04	-4,85	-4,31	-4.56
LEU	50	CG	CD1	-1,30	0,90	0,52	-0,76	-1,11	1,03	0,21	-0,17	-1,63	-1,65	-1.52
LEU	50	CG	CD2	-0,11	0,57	0,37	2,05	0,14	-0,83	-0,52	0,37	-0,47	0,19	-0.52
THR	55	CB	CG2	4,12	-2,19	-1,41	-2,79	2,12	4,57	1,46	1,55	5,85	4,00	5.04
LEU	56	CG	CD1	n	n	n	n	n	n	n	n	n	n	
LEU	56	CG	CD2	-0,77	0,28	-0,06	-2,48	0,05	-3,02	-0,09	-1,10	-1,04	-1,40	-0.70
ILE	61	CG1	CD1	-1,47	0,37	0,12	2,09	-1,30	-1,88	-0,88	-1,02	-2,30	-2,06	-2.24
ILE	61	CB	CG2	1,35	0,42	0,37	2,28	0,58	2,37	-0,02	1,32	1,14	1,89	0.91
THR	66	CB	CG2	n	n	n	n	n	n	n	n	n	n	
LEU	67	CG	CD1	-0,71	-0,20	-0,05	0,25	-1,11	-0,56	-0,19	-0,71	-0,95	-1,07	-0.80
LEU	67	CG	CD2	2,50	-1,24	-0,73	-0,61	1,16	3,01	0,94	1,33	3,19	2,44	3.12
LEU	69	CG	CD1	n	n	n	n	n	n	n	n	n	n	
LEU	69	CG	CD2	n	n	n	n	n	n	n	n	n	n	
VAL	70	CB	CG1	n	n	n	n	n	n	n	n	n	n	
VAL	70	CB	CG2	0,99	-0,81	-0,45	1,38	0,34	2,10	0,19	0,72	1,47	1,14	1.33
LEU	71	CG	CD1	n	n	n	n	n	n	n	n	n	n	
LEU	71	CG	CD2	0,53	0,37	0,20	0,94	0,34	-0,79	-0,45	0,20	-0,07	0,26	N
LEU	73	CG	CD1	1,13	-0,48	-0,19	-0,64	0,87	0,14	0,10	0,44	1,10	0,83	1.04
LEU	73	CG	CD2	0,97	-0,20	-0,10	-1,57	0,43	1,78	0,67	0,71	1,03	1,12	1.09

Table S8 (continued): Experimental residual dipolar couplings for methyl groups