



ELSEVIER

Contents lists available at ScienceDirect

## Data in Brief

journal homepage: [www.elsevier.com/locate/dib](http://www.elsevier.com/locate/dib)



### Data Article

# Compiled data set of exact NOE distance limits, residual dipolar couplings and scalar couplings for the protein GB3



Beat Vögeli<sup>a,\*</sup>, Simon Olsson<sup>a,b</sup>, Roland Riek<sup>a</sup>, Peter Güntert<sup>a,c,d</sup>

<sup>a</sup> *Laboratory of Physical Chemistry, Swiss Federal Institute of Technology, Vladimir-Prelog-Weg 2, ETH-Hönggerberg, CH-8093 Zürich, Switzerland*

<sup>b</sup> *Institute for Research in Biomedicine, Via Vincenzo Vela 6, CH-6500 Bellinzona, Switzerland*

<sup>c</sup> *Institute of Biophysical Chemistry, Center for Biomolecular Magnetic Resonance, and Frankfurt Institute for Advanced Studies, J.W. Goethe-Universität, Max-von-Laue-Str. 9, 60438 Frankfurt am Main, Germany*

<sup>d</sup> *Graduate School of Science, Tokyo Metropolitan University, Hachioji, Tokyo 192-0397, Japan*

#### ARTICLE INFO

##### Article history:

Received 25 August 2015

Accepted 25 August 2015

Available online 4 September 2015

#### ABSTRACT

We compiled an NMR data set consisting of exact nuclear Overhauser enhancement (eNOE) distance limits, residual dipolar couplings (RDCs) and scalar ( $J$ ) couplings for GB3, which forms one of the largest and most diverse data set for structural characterization of a protein to date. All data have small experimental errors, which are carefully estimated. We use the data in the research article Vögeli et al., 2015, Complementarity and congruence between exact NOEs and traditional NMR probes for spatial decoding of protein dynamics, *J. Struct. Biol.*, 191, 3, 306–317, doi:10.1016/j.jsb.2015.07.008 [1] for cross-validation in multiple-state structural ensemble calculation. We advocate this set to be an ideal test case for molecular dynamics simulations and structure calculations.

© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

DOI of original article: <http://dx.doi.org/10.1016/j.jsb.2015.07.008>

\* Corresponding author.

E-mail address: [beat.voegeli@phys.chem.ethz.ch](mailto:beat.voegeli@phys.chem.ethz.ch) (B. Vögeli).

<http://dx.doi.org/10.1016/j.dib.2015.08.020>

2352-3409/© 2015 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Specifications table [please fill in right-hand column of the table below]

Subject area	<i>Physical chemistry, structural biology</i>
More specific sub- ject area	<i>Nuclear magnetic resonance, NMR</i>
Type of data	<i>Exact NOE distance restraints, J couplings, RDCs files, table, figure</i>
How data was acquired	<i>Solution NMR</i>
Data format	<i>CYANA input files</i>
Experimental factors	–
Experimental features	<i>All data has been acquired at 298 K; typically, the GB3 NMR samples contained 350–500 <math>\mu</math>l of 2–4 mM protein solution in 97%/3% or 95%/5% H<sub>2</sub>O/D<sub>2</sub>O, 50 mM potassium phosphate buffer, pH 6.5–7.0, and 0.5 mg/ml sodium azide</i>
Data source location	<i>Zurich, Switzerland; Bethesda, MD, USA</i>
Data accessibility	<i>data is with this article</i>

Value of the data [describe in 3–5 bulleted points why this data is of value to the scientific community]

- one of the largest and most diverse NMR data sets for characterizing the structure and dynamics of a protein to date
- extensive error analysis guarantees high reliability of the data set
- ideal for validation of structure calculation programs and molecular dynamics simulations

## 1. Data

We compiled 923 exact nuclear eNOEs distance limits (upper distance limits in CYANA format in file 'GB3.upl'; lower distance limits in CYANA format in file 'GB3.lol'), 61 conventional NOEs for aromatics (upper distance limits in CYANA format in file 'GB3.upl'), 1477 RDCs (file 'GB3.rdc'), 225 <sup>3</sup>J scalar couplings ('GB3.cco'), and 52 torsion angle restraints ('GB3.aco') for the third immunoglobulin binding domain of protein G (GB3) (Table 1) [1]. To that purpose, the previously published set consisting of 884 eNOEs [2,3] was supplemented by eNOEs involving methylenes with a degenerate proton pair and Val with a degenerate pair of methyl groups. The RDCs were collected from the literature, in total originating from 8 different alignment conditions. 1335 RDCs were obtained from spin pairs located in the backbone, 129 in side chains, and 13 bridging the backbone and side chains. All J couplings, most of which are taken from the literature, extend over three covalent bonds, of which 147 are from the backbone, and 78 between a backbone and a side-chain spin. In the distance limit, RDC and J coupling CYANA input files, the first six columns describe the two involved atoms. For the distance restraints, the seventh column is the upper or lower distance restraint. In the RDC file, the seventh to eleventh columns contain the measured value, the experimental error, the relative weight (always 1), the alignment tensor number and a constant collecting physical constants of the specific coupling. The magnitudes and rhombicities of the alignment tensors are indicated in the file header. In the J coupling file, the seventh to twelfth columns contain the measured value, the experimental error, the relative weight (always 1), and the coefficients C, B and A of the Karplus equations. In the torsion angle restraint CYANA input file, the first three columns identify the torsion angle, and the fourth and fifth columns list the lower and upper bounds. We also added 41 intrarésidual and 46 sequential  $D_{\text{HNH}\alpha}$  RDCs, which cannot be used in CYANA structure calculations due to the flexible distances between the nuclei (files 'GB3\_rdc\_3Dhnha.f' and 'GB3\_rdc\_4Dhnha.f', respectively).

**Table 1**  
Compiled CYANA restraints from GB3.

Type				#		
eNOEs	upper/lower distance limit	total		984		
		bidirectional	total involving pseudo-atom <sup>a</sup>	355	31	
		unidirectional	total involving pseudo-atom <sup>a</sup>	568	69	
			else		499	
NOEs	upper distance limit	aromatics <sup>b</sup>		61		
scalar couplings	total			225		
	backbone	total		147		
		$^3J_{\text{HN,H}\alpha}$			49	
		$^3J_{\text{HN,C}\beta}$			49	
	$^3J_{\text{HN,C}'}$			49		
	backbone-side chain	total		78		
		$^3J_{\text{H}\alpha,\text{H}\beta(2/3)}$			24	
		$^3J_{\text{N,C}\gamma(1/2)}$			27	
		$^3J_{\text{C}'\text{C}\gamma(1/2)}$			27	
RDCs	total backbone	$^1D_{\text{HN,N}}$	total	1477	372	
			medium 1			50
			medium 2			54
			medium 3			54
			medium 4			52
			medium 5			54
			medium 6			54
	medium 7			54		
		$^1D_{\text{C}\alpha,\text{H}\alpha}$	total		387	
			medium 1			43
			medium 2			50
			medium 3			49
			medium 4			47
			medium 5			50
			medium 6			49
			medium 7			50
		medium 8			49	
		$^1D_{\text{C}\alpha,\text{C}'}$	total		320	
			medium 2			54
	medium 3				54	
	medium 4				50	
	medium 5				54	
	medium 6				54	
	medium 7				54	
	$^1D_{\text{HN,C}'}$	total		256		
		medium 2			52	
		medium 3			51	
		medium 4			47	
		medium 5			54	
		medium 6			52	
		medium 7			54	
	backbone-side chain	$^2D_{\text{C}\beta,\text{H}\alpha}$	medium 1	13	13	
			side chain			
		$^1D_{\text{C}\beta,\text{H}\beta(2/3)}$	total	54		
			medium 1		38	

**Table 1** (continued)

Type		#
		medium 8
	$^1D_{H\beta 2, H\beta 3}$	medium 1
	$^1D_{Cm, Hm3}$	total
		medium 1
		medium 8
Angular restraints	total	52
	phi	26
	psi	26

<sup>a</sup> Pseudo-methylene atom or pseudo-atom for both methyl groups in Val and Leu.

<sup>b</sup> NOEs involving aromatic groups, set to < 8 Å.

## 2. Experimental design, materials and methods

### 2.1. eNOEs

All cross-relaxation rate constants presented in references [2,3] were taken. Upper and lower limits of the distance restraints were set with an allowed distance range of 0% and +15 and –15% for bidirectional and unidirectional eNOEs, respectively. However, NOEs involving methyl groups were processed differently. In this study, the CYANA protocol was executed with an individual treatment of each methyl proton by  $r^{-6}$  summation of the corresponding distances. Therefore, previous input distance restraints were scaled by a factor of  $3^{-1/6} = 0.83268$  per methyl group such that the corresponding cross-relaxation rate constant is a sum over all individual contributions. This would be strictly true if the methyl motion was slow (slower than nanoseconds). Since there is fast rotation present as well, we added an additional tolerance of  $\pm 8.5\%$ .

eNOEs that involve either methylene groups with degenerate chemical shifts or chemically equivalent methyl groups in Val and Leu were added to the previous data set. The apparent cross-relaxation rate constants were fitted to the same formulae as used for single atoms or methyl groups, corrected for spin diffusion [4] and normalized to the equivalent of a superposition of contributions from all pairs of single atoms. Note that this is an approximation because the spins do not undergo fast exchange. Instead, the spectral peaks are superpositions of the individual buildup/decay curves. Therefore, all upper limits and lower limits were given an additional tolerance of 5% in addition to the 0/15% for bidirectional/unidirectional eNOEs.

### 2.2. Residual dipolar couplings

$^1D_{HN,N}$  (43 Hz range) and  $^1D_{C\alpha,H\alpha}$  (92 Hz range) RDCs of wild-type GB3 under alignment induced by Pf1 filamentous phages (tensor 1) were taken from reference [5]. For  $^1D_{HN,N}$ , the errors were uniformly set to 0.5 Hz because half the pairwise r.m.s. deviation from values obtained from a new sample was 0.57 Hz, and singular value decomposition (SVD) fits of the RDCs to a 160-member multi-state ensemble [6], the RDC-refined X-ray structure (PDB 2OED) [7], and RDC-proton optimized 2OED [8–10] yield fitting errors of 0.54 Hz, 0.76 Hz, and 1.29 Hz, respectively. Analogously, the errors for  $^1D_{C\alpha,H\alpha}$  were uniformly set to 1.5 Hz because SVD fits to the same structures yield fitting errors of 1.67 Hz, 2.02 Hz, and 2.51 Hz, respectively, and if the alignment tensor is taken from the  $H^N$ -N RDCs, only slightly larger errors are observed with 2.07 Hz, 2.40 Hz, and 3.52 Hz. An initial tensor estimate was obtained from SVD fit of the  $H^N$ -N RDCs to the RDC-refined X-ray structure [8–10] and omission of the highly mobile residues 12, 40, and 41. This tensor was then used for both  $^1D_{HN,N}$  and  $^1D_{C\alpha,H\alpha}$ . The relative scaling is –2.0327 (corresponding to the bond lengths of 1.02 Å and 1.09 Å).

For the deuterated mutants K19AD47K (tensor 2), K19ED40N (tensor 3), K19EK4A-C-His6 (tensor 4), K19EK4A-N-His6 (tensor 5), and K19AT11K (tensor 6) two slightly different experiments were run to obtain  $H^N$ -N and  $C^\alpha$ -C' RDCs under alignment via Pf1 [10]. For these data sets, the averages are

used and the errors are half the individual pairwise differences. For the 54, 54, 52, 54, 54  $^1D_{\text{HN,N}}$  values, the overall r.m.s. deviations for the sets corresponding to tensors 2, 3, 4, 5, 6 are 0.14, 0.17, 0.16, 0.11, 0.07 Hz for ranges of 28, 30, 35, 27, 30 Hz, respectively. If only one value was available, the error was set to twice the overall r.m.s. deviation. For the 54, 54, 50, 54, 54  $^1D_{\text{C}\alpha,\text{C}'}$  values the errors for the sets corresponding to tensors 2, 3, 4, 5, 6 are 0.041, 0.078, 0.152, 0.073, 0.047 Hz for ranges of 6.2, 6.2, 7.4, 6.7, 7.3 Hz, respectively. If only one value was available, the error was set to twice the overall r.m.s. deviation. The scaling of  $^1D_{\text{C}\alpha,\text{C}'}$  relative to  $^1D_{\text{HN,N}}$  is  $-0.1866$  (corresponding to the bond lengths of 1.525 Å and 1.02 Å). For a seventh alignment condition under Pf1 (tensor 7, mutant K19EK4A), only one data set of  $^1D_{\text{HN,N}}$  (37 Hz range) and  $^1D_{\text{C}\alpha,\text{C}'}$  (8.1 Hz range) is available. The error was set uniformly to twice the largest error of the other five data sets (0.34 Hz and 0.30 Hz).

For  $\text{H}^{\text{N}}\text{-C}'$ , one set of RDCs is available for each mutant yielding tensors 2–6 [10]. The relative scaling of the coupling is  $-0.3123$ , assuming an interatomic distance of 2.035 Å corresponding to lengths of 1.020 Å and 1.329 Å for the  $\text{H}^{\text{N}}\text{-N}$  and  $\text{C}'\text{-N}$  bonds and a bond angle of  $119.5^\circ$  [11]. R.m.s. deviations from back-calculated values obtained from SVD are typically 0.4 Hz (using a directly fitted tensor or a tensor obtained from  $^1D_{\text{HN,N}}$ ). Therefore, the errors were uniformly set to 0.2 Hz.

For each of the protonated mutants K19AD47K (tensor 2), K19ED40N (tensor 3), K19EK4A-C-His6 (tensor 4), K19EK4A-N-His6 (tensor 5), K19AT11K (tensor 6), and K19EK4A (tensor 7), one  $^1D_{\text{C}\alpha,\text{H}\alpha}$  data set is available [9]. To account for Pf1 concentration differences, the values were rescaled by the slopes between the  $^1D_{\text{HN,N}}$  values obtained from the protonated samples (not used in this study) and the deuterated samples (see above). The alignment tensors are assumed to be the same as those determined from the  $^1D_{\text{HN,N}}$  sets of the deuterated samples. R.m.s. deviations from back-calculated values obtained by SVD are typically 1.5 Hz (or ca. 2.0–3.5 Hz using the tensors obtained from  $^1D_{\text{HN,N}}$ ). Therefore, the errors are uniformly set to 1.0 Hz.

For the structure calculations, tensors 2–7 were determined from SVD with all measured RDCs in the backbone.

$^2D_{\text{C}\beta,\text{H}\alpha}$ ,  $^1D_{\text{C}\beta,\text{H}\beta 2}$ ,  $^1D_{\text{C}\beta,\text{H}\beta 3}$  and  $^1D_{\text{H}\beta 2,\text{H}\beta 3}$  values were obtained from alignment with Pf1 phage [12]. A 3D HBCBCA type experiment provided four independent values for  $^2D_{\text{C}\beta,\text{H}\alpha}$ , which allows for an estimation of individual errors.  $^1D_{\text{C}\beta,\text{H}\beta 2}$ ,  $^1D_{\text{C}\beta,\text{H}\beta 3}$  and  $^1D_{\text{H}\beta 2,\text{H}\beta 3}$  values are obtained from linear combinations of the effectively measured  $^1D_{\text{C}\beta,\text{H}\beta 2} - ^1D_{\text{H}\beta 2,\text{H}\beta 3}$ ,  $^1D_{\text{C}\beta,\text{H}\beta 3} - ^1D_{\text{H}\beta 2,\text{H}\beta 3}$ , and  $^1D_{\text{C}\beta,\text{H}\beta 2} + ^1D_{\text{C}\beta,\text{H}\beta 3}$  values. Here, the errors of these values were propagated into individual errors of the couplings of interest. Sample conditions were similar to those of the previously mentioned measurement of  $^1D_{\text{HN,N}}$  with Pf1 yielding tensor 1.  $^1D_{\text{C}\alpha,\text{H}\alpha}$  couplings obtained from the same experiment were compared (Pearson's correlation coefficient 0.99) and rescaled to the  $^1D_{\text{C}\alpha,\text{H}\alpha}$  couplings mentioned above in order to estimate the alignment tensor magnitude relative to tensor 1. Then, tensor 1 was used for the structure calculations (the scaling factor was 0.814 and the errors were not scaled).

$^1D_{\text{C}\beta,\text{H}\beta}$  (Val, Ile, Thr),  $^1D_{\text{C}\beta,\text{H}\beta 3}$  (Ala),  $^1D_{\text{C}\gamma 1/2,\text{H}\beta 3\gamma 1/2}$  (Val, Ile, Thr) and  $^1D_{\text{C}\delta 1/2,\text{H}\beta 3\delta 1/2}$  (Leu, Ile) values were obtained from alignments with Pf1 phage and PEG [6]. For Pf1, a set of  $^1D_{\text{C}\alpha,\text{H}\alpha}$  recorded on the same sample was used to scale the couplings such that the previously mentioned tensor 1 for alignment with Pf1 could be used (scaling factor 0.90 to account for Pf1 concentration difference, Pearson's correlation coefficient 0.99) and also for the estimate of the error (r.m.s. deviation between those two sets is 2.32 Hz, which gives an error of ca.  $2.32/2^{1/2}$  Hz = 1.5 Hz here). The couplings within methyl groups were scaled by  $-1/3.17$  such that they could be used as effective  $^1D_{\text{C}\alpha,\text{C}\beta}$ ,  $^1D_{\text{C}\beta,\text{C}\gamma 1/2}$  and  $^1D_{\text{C}\gamma(1),\text{C}\delta 1/2}$  couplings in the structure calculation [13]. The rescaled errors would be ca. 0.2 Hz, but were uniformly set to 0.5 Hz. For PEG, no  $^1D_{\text{HN,N}}$  couplings were available and the alignment tensor was determined from a set of  $^1D_{\text{C}\alpha,\text{H}\alpha}$  obtained in the same experiment (tensor 8). The measurements were carried out once. The errors were estimated as follows: An SVD from the  $^1D_{\text{C}\alpha,\text{H}\alpha}$  couplings of Pf1 gave an r.m.s. deviation of ca. 2.7 Hz (this tensor was not used in the structure calculations). An analogous SVD from the  $^1D_{\text{C}\alpha,\text{H}\alpha}$  couplings of PEG yielded an r.m.s. deviation of ca. 1 Hz, while the coupling amplitude is half as large. Thus, the RDC data from PEG appear 1.5 times better and if it is assumed that the absolute measuring errors are similar for PF1 and PEG, it would be safe to use the same errors for PEG as for Pf1. However, it is likely that tensor 8 obtained from  $^1D_{\text{C}\alpha,\text{H}\alpha}$  is not as accurate as tensor 1, which is obtained from  $^1D_{\text{HN,N}}$ . Therefore, a uniform error of 2 Hz was chosen for  $^1D_{\text{C}\alpha,\text{H}\alpha}$  and  $^1D_{\text{C}\beta,\text{H}\beta}$ . The errors of the couplings involving methyl groups were uniformly set to 0.5 Hz.

Next, all fitted alignment tensors were corrected for the rescaling due to uniform motion throughout the molecule. Iterative re-determination of the tensors increases them by 4% in the first cycle of a two-state ensemble calculation (CYANA target function value changes by  $+0.91 \text{ \AA}^2$ , or  $-1.08 \text{ \AA}^2$  when using effective bond lengths) and converges to an increase of about 5% after the second cycle. This result is in good agreement with a tensor rescaling by  $1/0.95$  based on SVD of the 160-member ensemble calculated by Schwieters and Clore [6]. Here, we use the re-determined tensor after the initial two-state ensemble calculation.

In addition, intraresidual  ${}^3D_{\text{HN,H}\alpha}$  and sequential  ${}^4D_{\text{HN,H}\alpha}$  RDCs taken from Ref. [5] were used for validation only.

### 2.3. Scalar couplings

${}^3J_{\text{HN,H}\alpha}$  values are averages over couplings derived from CT-MQ( ${}^1\text{H}^{\text{N}}$ ,  ${}^{13}\text{C}^{\alpha}$ ) + SQ( ${}^1\text{H}^{\text{N}}$ )-HNCA spectra and  $J$ -modulated HMQC spectra [14,15]. Based on the pairwise r.m.s. deviation between the values obtained from the two types of measurements, the error of their averaged values equals 0.14 Hz [14]. For each coupling, we use the averages over two sets of each type as CYANA input and the standard deviation as input error (overall 0.15 Hz). If both data sets of one type (in our case only  $J$ -modulated HMQC spectra) are missing the error was set to 0.3 Hz. The averaged measured values were corrected for the residual dipolar couplings between  $\text{H}^{\text{N}}$  and  $\text{H}^{\alpha}$  due to the natural alignment of GB3 in the magnetic field at 600 MHz. The alignment tensor was estimated from sums of  $\text{H}^{\text{N}}$ -N residual dipolar and scalar couplings at 500 and 800 MHz fields. The  $\text{H}^{\text{N}}$ - $\text{H}^{\alpha}$  RDCs were back-predicted from an RDC-refined X-ray structure [8–10].

${}^3J_{\text{HN,C}\beta}$  values are averages over couplings derived from a CT-MQ( ${}^1\text{H}^{\text{N}}$ ,  ${}^{13}\text{C}^{\alpha}$ )-HNCA and a HNCA[CB] E.COSY experiment [14]. The individual errors were based on the pairwise r.m.s. deviation between these two sets of values, with an overall error of their averaged values of 0.07 Hz [14]. If the value of one data set was missing, 0.1 Hz was used.

${}^3J_{\text{HN,C}'} values are averages over couplings derived from a CT-MQ( ${}^1\text{H}^{\text{N}}$ ,  ${}^{13}\text{C}^{\alpha}$ )-HNCA and a HNCA[C'] E. COSY experiment [14]. The individual errors were based on the pairwise r.m.s. deviation between these two sets of values, with an overall error of their averaged values of 0.1 Hz [14]. If the value of one data set was missing, 0.2 Hz was used.$

The Karplus curve coefficients for  ${}^3J_{\text{HN,H}\alpha}$ ,  ${}^3J_{\text{HN,C}\beta}$  and  ${}^3J_{\text{HN,C}'}$  were determined from fits to the RDC-refined X-ray structure [8–10] under the assumption of uniform fluctuations of  $10^\circ$  of the  $\phi$  angles [16]. The highly dynamic residues 12 and 40 were excluded, and three ubiquitin residues with positive  $\phi$  angles, namely residues 46, 60 and 64, were included in the fits, using angles from an NMR-refined X-ray structure [17]. The Karplus curve coefficients  ${}^3J_{\text{HN,H}\alpha}(\phi) = 8.754 \cos^2\phi - 1.222 \cos\phi + 0.111 \text{ Hz}$ ,  ${}^3J_{\text{HN,C}\beta}(\phi) = 3.693 \cos^2\phi - 0.514 \cos\phi + 0.043 \text{ Hz}$ , and  ${}^3J_{\text{HN,C}'(\phi) = 4.516 \cos^2\phi - 1.166 \cos\phi - 0.038 \text{ Hz}$  were obtained.

It has been shown that the values for  ${}^3J_{\text{HN,H}\alpha}$ ,  ${}^3J_{\text{HN,C}\beta}$  and  ${}^3J_{\text{HN,C}'}$  of GB3 can be predicted somewhat better if density functional theory (DFT) calculations are performed on the structure of the RDC refined X-ray structure (PDB 2OED) than from simple parametrization of the experimental data by Karplus curves [18]. This demonstrates that some discrepancy between the Karplus curves and experimental data is caused by hydrogen bonding, substituent and electrostatic effects rather than fluctuation of the dihedral angles. These errors are very small for  ${}^3J_{\text{HN,H}\alpha}$  and  ${}^3J_{\text{HN,C}'}$  and are clearly dominated by the experimental errors (which are used for the width of the flat bottom CYANA potential). The situation is somewhat different for  ${}^3J_{\text{HN,C}\beta}$ . Here, the error due to the substituent effect is ca. 0.08 Hz. Thus, the errors for  ${}^3J_{\text{HN,C}\beta}$  were set to the propagated errors obtained from the individual random experimental errors plus a uniform error of 0.08 Hz. This increases the overall error from 0.07 Hz to 0.11 Hz.

${}^3J_{\text{H}\alpha,\text{H}\beta 2}$  and  ${}^3J_{\text{H}\alpha,\text{H}\beta 3}$  values were taken from reference [12], where a 3D HBCBCA type experiment provides two independent values allowing for a cross-check. A systematic error arises from the errors in the extracted peak positions caused by  ${}^1\text{H}^{\alpha}$  transverse relaxation during the  $\text{S}^3\text{CT}$  element and is estimated to be 1 Hz (see Fig. 4 in the Supplemental information in Ref. [12]). The systematic and individual random errors were propagated into an overall error. All couplings are in agreement with our previously calculated ensemble [2,3] except for residues 8 and 52. Residue 8 seems to undergo

averaging as indicated by the two nearly identical values of  ${}^3J_{H\alpha,H\beta 2}$  and  ${}^3J_{H\alpha,H\beta 3}$  close to 7 Hz. These values are used here since a potentially wrong stereoassignment would not have an impact. The couplings of residue 52 are also in disagreement with the data set in reference [19], the X-ray structure [8], and our eNOE-based stereospecific assignment [20], all of which suggest a single rotamer state. On the other hand, the  ${}^3J_{C,C\gamma}$  and  ${}^3J_{N,C\gamma}$  couplings for residue 52 in reference [2] appear to be slightly averaged over at least two rotamer states. Due to these inconsistencies, the  ${}^3J_{H\alpha,H\beta 2}$  and  ${}^3J_{H\alpha,H\beta 3}$  couplings of residue 52 are not used here. The substituent-effect-corrected Karplus parametrization (A, B, C)=(7.23, -1.37, 2.40) is used as proposed for Arg, Asx, Glx, His, Leu, Lys, Met, Phe, Pro, Trp and Tyr in Ref. [21].

We determined  ${}^3J_{C,C\gamma}$  and  ${}^3J_{N,C\gamma}$  couplings for aromatic residues using the pulse sequences proposed in Ref. [22].  ${}^3J_{C,C\gamma(1/2)}$  and  ${}^3J_{N,C\gamma(1/2)}$  couplings for Val, Ile and Thr residues were taken from Ref. [23]. The Karplus coefficients for the aromatic residues were taken from Ref. [21] proposing (A, B, C) to be (2.31, -0.87, 0.49/1.29, -0.49, 0.34), and for the methyl bearing residues (2.76, -0.67, 0.19/2.01, 0.21, -0.12) for Thr and (3.42, -0.59, 0.17/2.64, 0.26, -0.22) for Val and Ile as proposed in Ref. [23].

#### 2.4. Torsion angle restraints from $C^\alpha$ chemical shifts

Restraints for the  $\phi$  and  $\psi$  backbone torsion angles were generated from  ${}^{13}C^\alpha$  chemical shifts with CYANA [24,25]. The allowed ranges were chosen conservatively [26] and are either  $-200^\circ$  to  $-80^\circ$  for  $\phi$  and  $40$ – $220^\circ$  for  $\psi$  if the  ${}^{13}C^\alpha$  secondary chemical shift was larger than 2 ppm,  $-120^\circ$  to  $-20^\circ$  for  $\phi$  and  $-100^\circ$  to  $0^\circ$  for  $\psi$  if the  ${}^{13}C^\alpha$  secondary chemical shift was less than  $-1.5$  ppm, or  $-120^\circ$  to  $80^\circ$  for  $\phi$  and  $-100^\circ$  to  $60^\circ$  for  $\psi$  if the  ${}^{13}C^\alpha$  secondary chemical shift was between 1.5 and 2 ppm. No torsion angle restraints were generated if the  ${}^{13}C^\alpha$  secondary chemical shift was between  $-1.5$  and  $+1.5$  ppm.

## Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at <http://dx.doi.org/10.1016/j.dib.2015.08.020>.

## References

- [1] B. Vögeli, S. Olsson, R. Riek, P. Güntert, Complementarity and congruence between exact NOEs and traditional NMR probes for spatial decoding of protein dynamics, *J. Struct. Biol.* 191 (3) (2015) 306–317. <http://dx.doi.org/10.1016/j.jsb.2015.07.008>.
- [2] B. Vögeli, P. Güntert, R. Riek, Multiple-state ensemble structure determination from eNOE spectroscopy, *Mol. Phys.* 111 (2013) 437–454.
- [3] B. Vögeli, S. Kazemi, P. Güntert, R. Riek, Spatial elucidation of motion in proteins by ensemble-based structure calculation using exact NOEs, *Nat. Struct. Mol. Biol.* 19 (2012) 1053–1057.
- [4] J. Orts, B. Vögeli, R. Riek, Relaxation matrix analysis of spin diffusion for the NMR structure calculation with eNOEs, *J. Chem. Theory Comput.* 8 (2012) 3483–3492.
- [5] B. Vögeli, L. Yao, A. Bax, Protein backbone motion viewed by intraresidue and sequential HN–Ha residual dipolar couplings, *J. Biomol. NMR* 41 (2008) 17–28.
- [6] G.M. Clore, C. Schwieters, Concordance of residual dipolar couplings, backbone order parameters and crystallographic B-factors for a small alpha/beta protein: a unified picture of high probability, fast atomic motions in proteins, *J. Mol. Biol.* 355 (2006) 879–886.
- [7] T.S. Ulmer, B.E. Ramirez, F. Delaglio, A. Bax, Evaluation of backbone proton positions and dynamics in a small protein by liquid crystal NMR spectroscopy, *J. Am. Chem. Soc.* 125 (2003) 9179–9191.
- [8] J.P. Derrick, D.B. Wigley, The third IgG-binding domain from streptococcal protein G. An analysis by X-ray crystallography of the structure alone and in a complex with Fab, *J. Mol. Biol.* 243 (1994) 906–918.
- [9] L. Yao, B. Vögeli, D.A. Torchia, A. Bax, Simultaneous NMR study of protein structure and dynamics using conservative mutagenesis, *J. Phys. Chem. B* 112 (2008) 6045–6056.
- [10] L.S. Yao, B. Vögeli, J.F. Ying, A. Bax, NMR determination of amide N–H equilibrium bond length from concerted dipolar coupling measurements, *J. Am. Chem. Soc.* 130 (2008) 16518–16520.
- [11] B. Vögeli, How uniform is the peptide plane geometry? A high-accuracy NMR study of dipolar C $\alpha$ –C'/H N–N cross-correlated relaxation, *J. Biomol. NMR* 50 (2011) 315–329.
- [12] E. Miclet, J. Boisbouvier, A. Bax, Measurement of eight scalar and dipolar couplings for methine–methylene pairs in proteins and nucleic acids, *J. Biomol. NMR* 31 (2005) 201–216.
- [13] M. Ottiger, A. Bax, How tetrahedral are methyl groups in proteins? A liquid crystal NMR study, *J. Am. Chem. Soc.* 121 (1999) 4690–4695.

- [14] B. Vögeli, J. Ying, A. Grishaev, A. Bax, Limits on variations in protein backbone dynamics from precise measurements of scalar couplings, *J. Am. Chem. Soc.* 129 (2007) 9377–9385.
- [15] H. Kuboniwa, S. Grzesiek, F. Delaglio, A. Bax, Measurement of HN–H alpha J couplings in calcium-free calmodulin using new 2D and 3D water-flip-back methods, *J. Biomol. NMR* 4 (1994) 871–878.
- [16] R. Brüschweiler, D.A. Case, Adding harmonic motion to the Karplus relation for spin–spin coupling, *J. Am. Chem. Soc.* 116 (1994) 11199–11200.
- [17] J.S. Hu, A. Bax, Measurement of three-bond,  $^{13}\text{C}'$ – $^{13}\text{C}$  beta J couplings in human ubiquitin by a triple resonance, E. COSY-type NMR technique, *J. Am. Chem. Soc.* 119 (1997) 6360–6368.
- [18] B. Wang, X. He, K.M. Merz, Quantum mechanical study of vicinal j spin–spin coupling constants for the protein backbone, *J. Chem. Theory Comput.* 9 (2013) 4653–4659.
- [19] L.Y. Lian, J.P. Derrick, M.J. Sutcliffe, J.C. Yang, G.C.K. Roberts, , Determination of the solution structures of domains II and III of protein G from streptococcus by  $^1\text{H}$  nuclear magnetic resonance, *J. Mol. Biol.* 228 (1992) 1219–1234.
- [20] J. Orts, B. Vögeli, R. Riek, P. Güntert, Stereospecific assignments in proteins using exact NOE, *J. Biomol. NMR* 57 (2013) 211–218.
- [21] C. Perez, F. Löhr, H. Ruterjans, J.M. Schmidt, Self-consistent Karplus parametrization of  $^3\text{J}$  couplings depending on the polypeptide side-chain torsion  $\chi_1$ , *J. Am. Chem. Soc.* 123 (2001) 7081–7093.
- [22] J.S. Hu, S. Grzesiek, A. Bax, N.M.R. Two-Dimensional, Methods for determining  $\chi_1$  angles of aromatic residues in proteins from three-bond  $\text{J}\text{C}'\text{C}\gamma$  and  $\text{J}\text{N}\text{C}\gamma$  couplings, *J. Am. Chem. Soc.* 119 (1997) 1803–1804.
- [23] J.J. Chou, D.A. Case, A. Bax, Insights into the mobility of methyl-bearing side chains in proteins from  $(3\text{J})\text{(CC)}$  and  $(3\text{J})\text{(CN)}$  couplings, *J. Am. Chem. Soc.* 125 (2003) 8959–8966.
- [24] P. Güntert, Automated NMR protein structure calculation, *Prog. Nucl. Magn. Reson. Spectrosc.* 43 (2003) 105–125.
- [25] P. Güntert, Automated structure determination from NMR spectra, *Eur. Biophys. J.* 38 (2009) 129–143.
- [26] P. Luginbühl, T. Szyperki, K. Wüthrich, Statistical basis for the use of  $^{13}\text{C}\alpha$  chemical shifts in protein structure determination, *J. Magn. Reson. B* 109 (1995) 229–233.