Supporting information

Automated NMR assignment of protein side chain resonances using APSY

Sebastian Hiller, Rosmarie Joss and Gerhard Wider

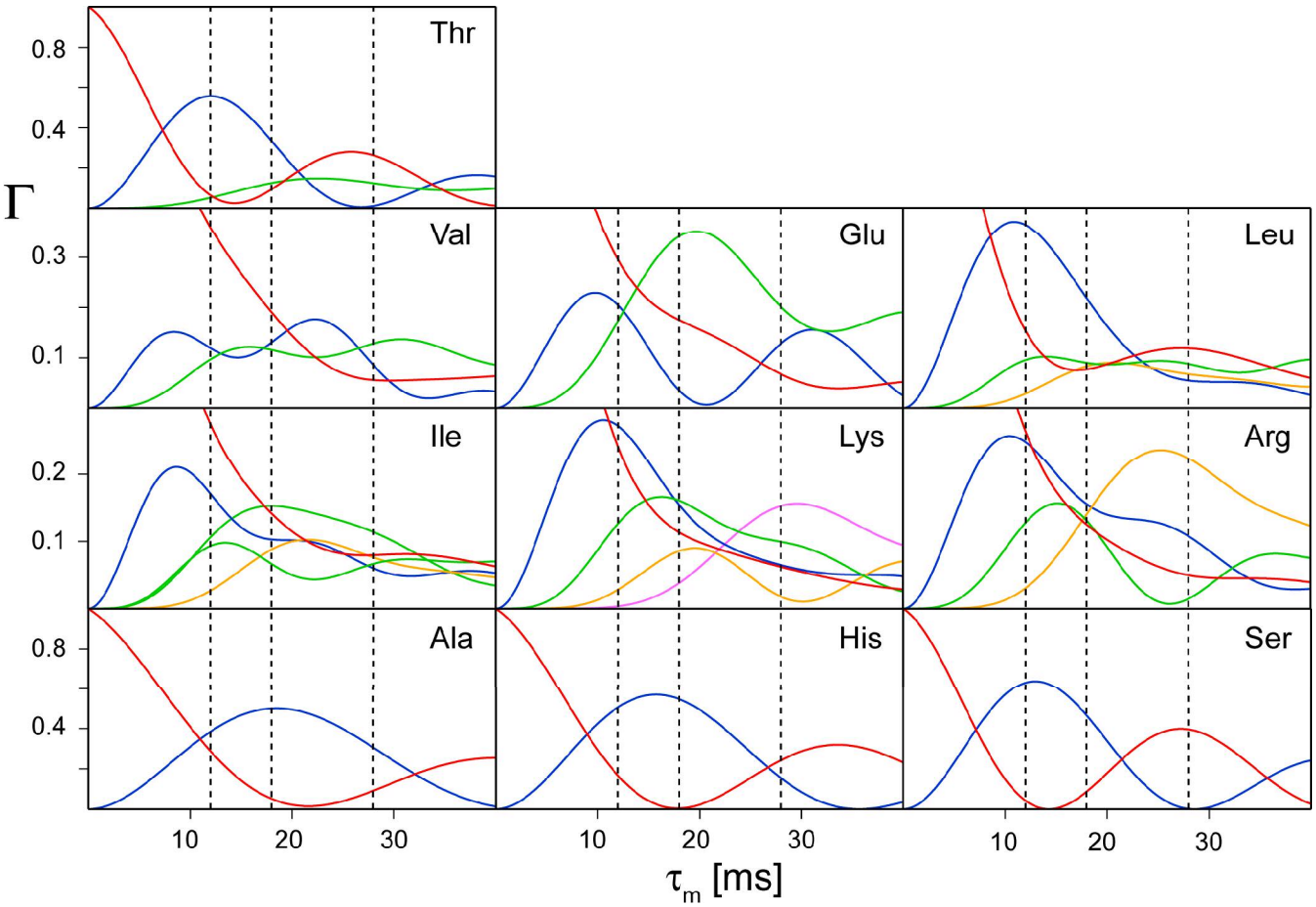

Figure S1. Calculated magnetization transfer amplitudes Γ for the carbon-carbon transfers $^{13}C^{\alpha}$ – $^{13}C^{\alpha}$ (red), $^{13}C^{\beta}$ – $^{13}C^{\alpha}$ (blue), $^{13}C^{\gamma}$ – $^{13}C^{\alpha}$ (green), $^{13}C^{\delta}$ – $^{13}C^{\alpha}$ (yellow) and $^{13}C^{\varepsilon}$ – $^{13}C^{\alpha}$ (magenta) during isotropic mixing in the HC(CC-TOCSY)CONH experiment. The calculations were done as described in Materials and Methods. The vertical dashed lines indicate the three mixing times used in the present work (12ms, 18ms and 28ms). The transfer functions for Ala, His and Ser are representative for the ten

two-carbon spin systems of the amino acids Ala, Asn, Asp, Cys, His, Phe, Ser, Trp and Tyr. The three examples Ala, Ser and His are the systems with the smallest, the largest and an intermediate effective C-C coupling, respectively. The data for Gln and Met are very similar to Glu. The data for Pro (Fig. S2) corresponds to Arg.

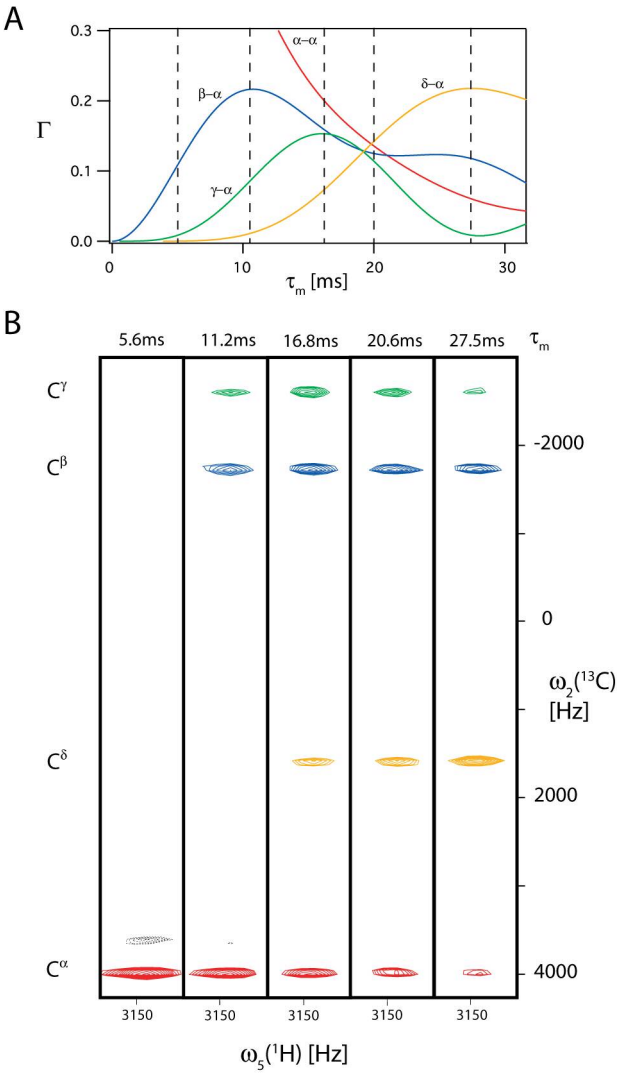

Figure S2. (A) Calculated magnetization transfer amplitudes Γ for the carbon-carbon transfers in proline residues during isotropic mixing in the HC(CC-TOCSY)CONH experiment. (B) Strips from five 2D ($\omega_2(^{13}C)$, $\omega_5(^{1}H)$)-projections of the 5D APSY-HC(CC-TOCSY)CONH experiment for Pro 41 of 434-repressor(1–63). The five experiments were recorded with different mixing times, as indicated above each strip. The cross peaks of C^{α}, C^{β}, C^{γ} and C^{δ} are colored red, blue, green and yellow, respectively.

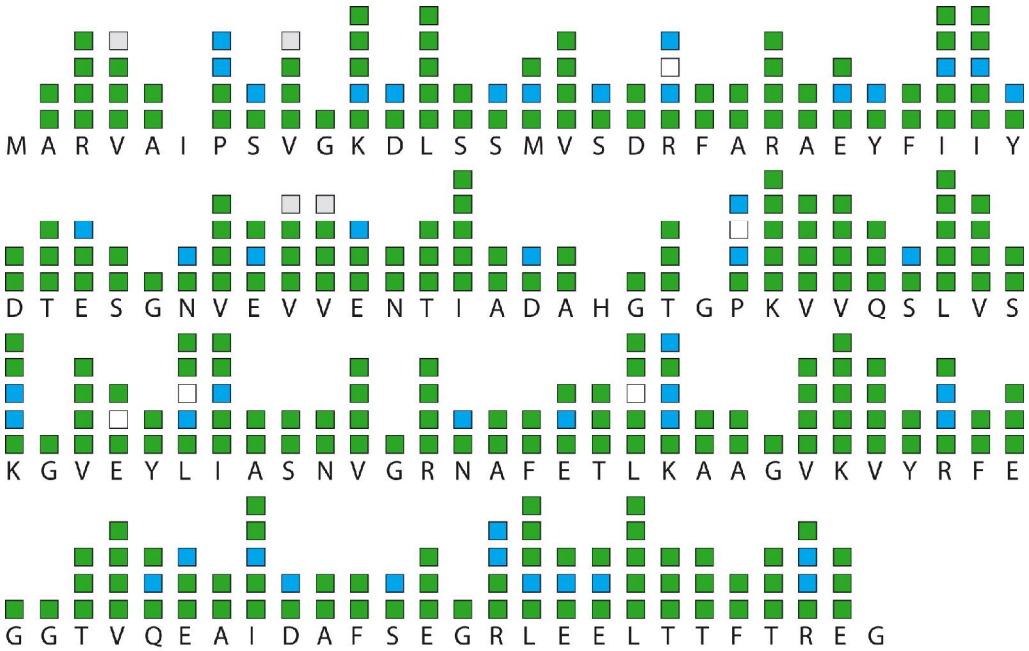

Figure S3. Graphical illustration of the aliphatic side chain assignments obtained for TM1290 based on a 5D APSY-HC(CC-TOCSY)CONH using 2D projection spectra and the backbone $^{1}H^{N}$, ^{15}N and $^{13}C'$ assignments. Squares represent the aliphatic side chain carbons and the C^{α} of each amino acid. Residues without squares can either not be detected by the experiment or are also not represented in the reference assignment established with conventional spectra. Green squares represent carbon atoms, where all expected proton shifts are found, i.e. two proton shifts for CH_2 and one for CH and CH_3 groups. Blue squares indicate CH_2 groups, for which only one proton shift was found. No peak was found for the white squares. Grey squares indicate isopropyl methyl groups that were not detected.

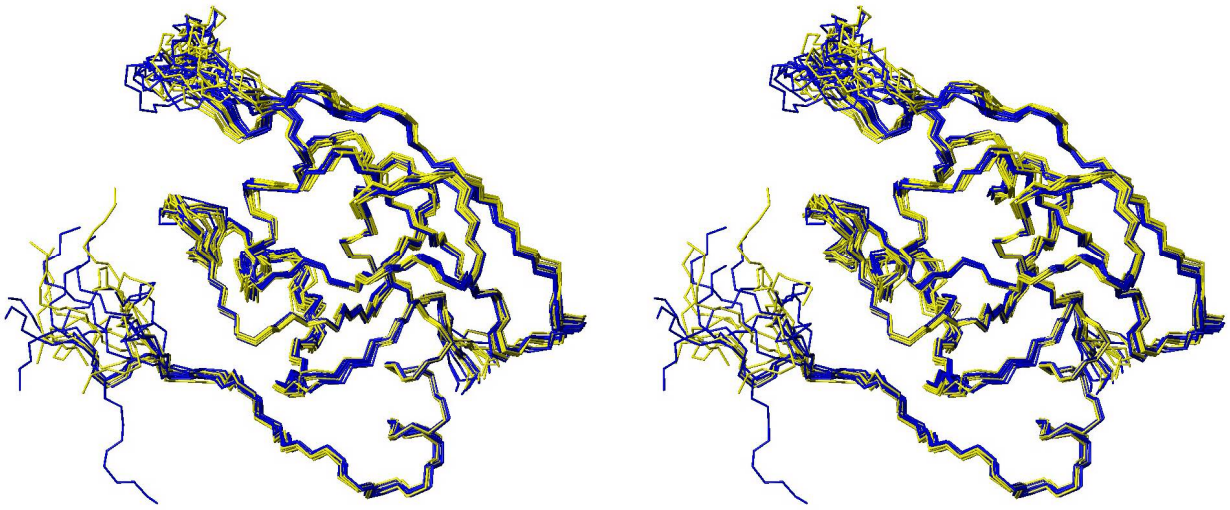

Figure S4. Stereo view of two superimposed backbone bundles of the protein TM1290. Each bundle shows the 10 lowest-energy conformers from 100 calculated CYANA structures. Blue: Calculation with the 2444 upper limit constraints corresponding to the complete manual assignment.[1] Yellow: Calculation with the same input data except the NOEs of all aliphatic atoms that were not assigned in the present work (2159 upper limit constraints). The RMSDs for heavy atoms of residues 3–43 and 52–110 are 0.82 ± 0.07 Å and 0.84 ± 0.07 Å, for the blue and yellow bundle, respectively, and 0.88 ± 0.07 Å for the combined bundle of 20 conformers.

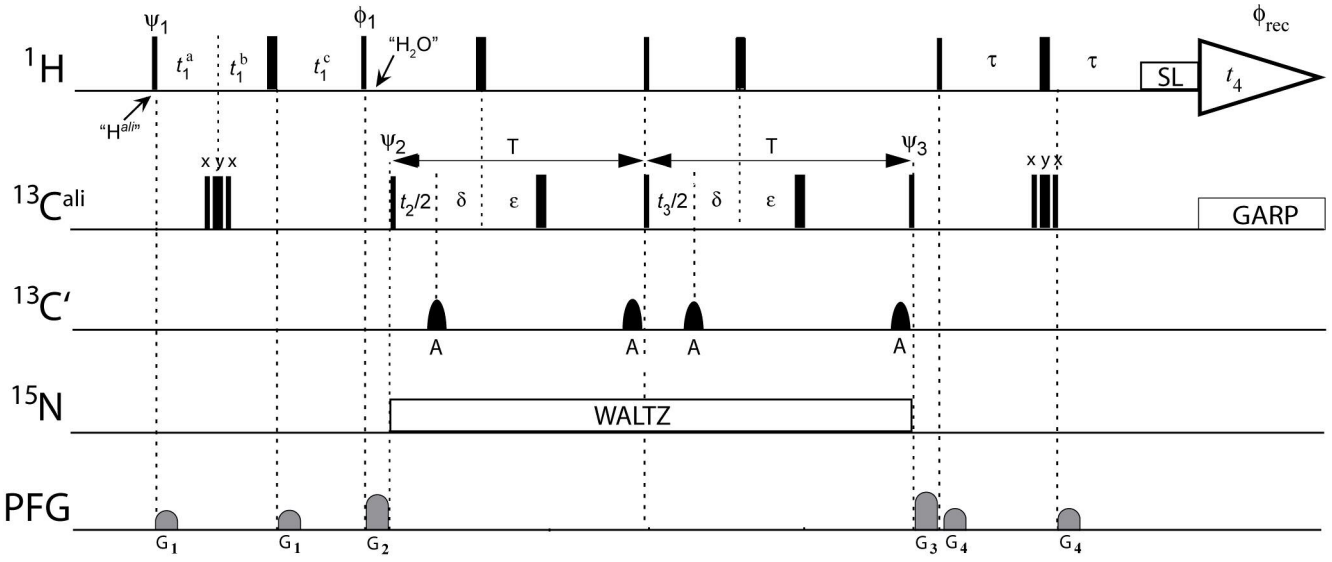

Figure S5. Pulse sequence of the 4D APSY-HCCH-COSY experiment. Radio-frequency (rf) pulses were applied at 39 ppm for aliphatic carbon nuclei, $^{13}C^{ali}$, at 174 ppm for carbonyl carbons, $^{13}C'$, and 119 ppm for ^{15}N. The carrier frequency for protons was set in the aliphatic region at 2.5 ppm at the beginning of the experiment, indicated on the line ^{1}H by "H^{ali}"; at the position "H_2O" the carrier was set to the water resonance (4.7 ppm). Bars stand for rectangular pulses applied at maximum power; the thin bars represent 90° pulses and the wide bars 180° pulses. Shaped rf pulses on the line $^{13}C'$ marked with A are 180° Gauss pulses (5% truncation) with a duration of 120µs on a 750 MHz spectrometer. The decoupling sequences WALTZ-16[2] on ^{15}N and GARP[3] on C^{ali} are indicated with white rectangles. The triangle labeled t_4 represents the acquisition period. The spin lock pulse before the acquisition is shown as a rectangle marked "SL". On the line marked PFG, curved shapes indicate sine bell shaped, pulsed magnetic field gradients along the z-axis with the following durations and strengths: G_1: 400µs, 40%; G_2: 800µs, 50%; G_3: 1000µs, 70%; G_4: 600µs, 45%. The initial delays in the evolution periods were $t_1^a = t_1^c = 1.6$ms, $t_1^b = 5$µs and $t_2/2 = t_3/2 = 10$µs. Further delays were $\delta = 1.1$ms, $\varepsilon = 2.8$ms and $\tau = 1.6$ms. The constant time period T was set to 7.8ms. All pulse were applied with phase x unless indicated otherwise above the pulse. The following phase cycles were used: ϕ_1={y,y,-y,-y}, ψ_3={x,-x} and ϕ_{rec}={x,-x,-x,x}. Quadrature detection for the indirect dimension was achieved with the phases ψ_1, ψ_2

and ψ_3 for t_1, t_2 and t_3, respectively. These phases were simultaneously incremented in 90° steps for consecutive FID's. Quadrature detection for the indirect dimensions was achieved using the trigonometric addition theorem to obtain pure cosine and sine terms for a subsequent hypercomplex Fourier transformation.[4, 5] The pulse phases ψ_1, ψ_2 and ψ_3 were incremented in 90°-steps for t_1, t_2 and t_3, respectively; only the pulse phases of the evolution periods which are part of the given projection are incremented.

Table T1. List of the 42 BMRB entries used for statistical chemical shift data analysis in this work.

Entry No.	Protein name	Year	Residues
bmr4050	Rubredoxin	1997	54
bmr4068	Turkey ovomucoid third domain	1998	56
bmr4070	FimC	1998	205
bmr4081	Interferon-alpha-2a	1997	165
bmr4092	Core binding factor b subunit	1998	143
bmr4117	Human Elongation Factor-1beta	1998	91
bmr4126	Single chain three helix bundle	1999	73
bmr4140	EH1 domain of mouse Eps15	1998	120
bmr4146	F1Fo ATP Synthase Subunit c	1998	79
bmr4154	Oxidized putidaredoxin	1999	106
bmr4156	Protein Disulfide Isomerase	1999	110
bmr4184	Eps15 homology domain	1998	95
bmr4205	Ets-1 Pointed	1998	110
bmr4223	TATA box binding protein associated factor II 230	1998	67
bmr4237	MinE topological specificity domain	1999	58
bmr4249	Human NER factor XPA	1998	122
bmr4284	Calmodulin-Ca^{2+}Pump-Peptide Complex	1999	148
bmr4296	Major cold shock protein from E. coli	1998	70
bmr4302	Protein disulfide isomerase a' domain	-	115
bmr4311	Human T-cell leukemia virus type I capsid monomer	1999	214
bmr4313	ARD	2002	179
bmr4317	NS1(1-73) dimmer	1997	73
bmr4318	Glutaredoxin 2	1999	215
bmr4326	N-Terminal Domain of DNA Polymerase B	1994	87
bmr4327	N-Terminal inhibitory domain of metalloproteinases-1 inhibitor	1999	126
bmr4334	ARID domain of dead-ringer protein	1999	139
bmr4371	Recombinant Onconase/P30 protein	1999	105

bmr4395	Ribosomal protein L25	1998	94
bmr4401	Skeletal N-troponin C	1998	90
bmr4437	Merozoite surface protein 1	1999	96
bmr4455	CD58 adhesion domain, 1dCD58	1999	95
bmr6655	Zinc finger domains 1 and 2 of dsRBP-ZFa	2005	127
bmr6751	Asl1650	2006	88
bmr6868	Protein ydhR precursor	2005	123
bmr6955	Hypothetical protein ydhA	2006	102
bmr7014	Nsp1	2006	116
bmr7080	NTD-CTD complex	2006	33
bmr7106	RGS18 monomer	2003	151
bmr7119	32324 monomer	2006	189
bmr7170	Sr482 monomer	-	117
bmr7191	Pat90 monomer	-	97
bmr7229	Small inducible cytokine B14	2006	78

References

(1) Etezady-Esfarjani, T.; Herrmann, T.; Peti, W.; Klock, H. E.; Lesley, S. A.; Wüthrich, K. *J. Biomol. NMR* **2004**, *29*, 403–406.

(2) Shaka, A. J.; Keeler, J.; Frenkiel, T.; Freeman, R. *J. Magn. Reson.* **1983**, *52*, 335–338.

(3) Shaka, A. J.; Barker, P. B.; Freeman, R. *J. Magn. Reson.* **1985**, *64*, 547-552.

(4) Brutscher, B.; Morelle, N.; Cordier, F.; Marion, D. *J. Magn. Reson. B* **1995**, *109*, 238–242.

(5) Kupce, E.; Freeman, R. *J. Am. Chem. Soc.* **2004**, *126*, 6429–6440.