Electronic Supplementary Material (ESI) for RSC Advances.

Electronic supplementary information

High-sensitive NMR protein ligand screening using ^{19}F multiple-quantum coherence spectroscopy

Anna Zawadzka-Kazimierczuk* [a,b], Mate Somlyay [a], Hanspeter Kählig [c], George Iakobson [d], Petr Beier [d] and Robert Konrat * [a]

[a] Department of Structural and Computational Biology, Max F. Perutz Laboratories, University of Vienna,Vienna Biocenter Campus 5, A-1030 Vienna (Austria)
b Biological and Chemical Research Centre, Faculty of Chemistry, University of Warsaw, Żwirki i Wigury 101, 02-089 Warsaw (Poland)
c Department of Organic Chemistry, University of Vienna, Währingerstrasse 27, A-1080 Vienna (Austria)
d Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Flemingovo nam. 2, 160 00 Prague (Czech Republic)

E-mail: anzaw@chem.uw.edu.pl, Robert.Konrat@univie.ac.at

Table of Contents

1. 5Q coherence generation

The simplest way of creating 5Q coherence in an AX_4 spin system (see Fig. S1) starts with excitation of A nucleus followed by a coherence transfer via spin-spin coupling to all X nuclei (realized using 1/2J delay with π pulses acting on both A and X nuclei in the middle of this period). Then a $\pi/2$ pulse applied just on X nuclei will create a coherence, where all X nuclei will have the same phase (x or y), while A nucleus will have also the same or different phase. This coherence will therefore have the form $16A_yX_yX_yX_yX_y$. Such terms consist of 5Q, 3Q and 1Q coherences, and the contribution of 5Q coherence is only 1/16, as:

$A_yX_yX_yX_yX_y$ =((A$^+$-A$^-$)/2i)((X$^+$-X$^-$)/2i)((X$^+$-X$^-$)/2i)((X$^+$-X$^-$)/2i)((X$^+$-X$^-$)/2i) contains:

two 5Q terms: 1 A$^+$X$^+$X$^+$X$^+$X$^+$, 1 A$^-$X$^-$X$^-$X$^-$X$^-$; ten 3Q terms: 4 A$^+$X$^+$X$^+$X$^+$X$^-$, 4 A$^-$X$^-$X$^-$X$^-$X$^+$, 1 A$^+$X$^-$X$^-$X$^-$X$^-$, 1 A$^-$X$^+$X$^+$X$^+$X$^+$

twenty 1Q terms: 6 A$^+$X$^+$X$^+$X$^-$X$^-$, 6 A$^-$X$^-$X$^-$X$^+$X$^+$, 4 A$^+$X$^+$X$^-$X$^-$X$^-$, 4 A$^-$X$^-$X$^+$X$^+$X$^+$

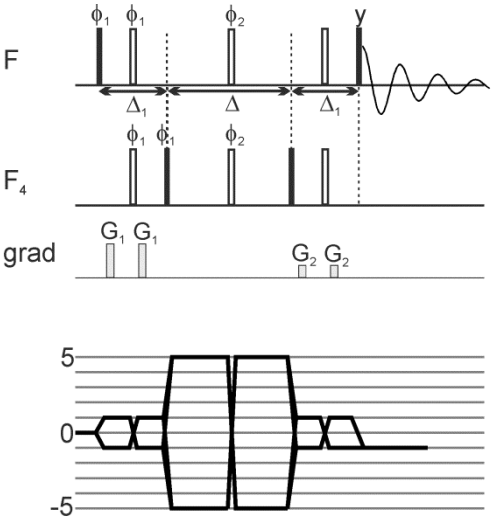

Fig. S1. A scheme of the simplest (but not the most effective) pulse sequence allowing creating 5Q coherence in an AX_4 spin system. All pulses should be selective shaped pulses. The Δ_1 delay should be set to $1/(2J_{A\text{-}X})$. The pulse phases should be set to x, unless shown explicitly. On the phase ϕ_1 a 10-step phase cycle should be performed to select the coherence of +/- 5 order during the multiple-quantum period.

Therefore we decided to use another excitation scheme, inspired by Kay and coworkers [Angew. Chemie – Int. Ed. 2016, 55, 11490-11494]. The coherence transfer in the first part of the proposed pulse sequence (up to creating the 5Q terms) is as follows (see Fig. S2):

$$X_z \xrightarrow{\left(\frac{\pi}{2}\right)_x:F_4} -X_y \xrightarrow{(\pi)_x:F_4,(\pi)_x:F,\pi J\Delta_1} -2X_xA_z \xrightarrow{\left(\frac{\pi}{2}\right)_x:F} 2X_xA_y \xrightarrow{(\pi)_x:F_4,(\pi)_x:F,\pi J\Delta_1} -16X_xA_xX_zX_zX_z \xrightarrow{\left(\frac{\pi}{2}\right)_x:F_4} 16X_xA_xX_yX_yX_y$$

The way back is symmetric to the first part of the pulse sequence, except for the pulses phases. Such scheme results with a coherence of the $A_xX_xX_yX_yX_y$ form, where the contribution of 5Q coherence is 1/6, as:

$A_xX_xX_yX_yX_y$=1/2(A$^+$+A$^-$)1/2(X$^+$+X$^-$)1/(2i)(X$^+$-X$^-$)1/(2i)(X$^+$-X$^-$)1/(2i)(X$^+$-X$^-$) contains:

two 5Q terms: 1 A$^+$X$^+$X$^+$X$^+$X$^+$, 1 A$^-$X$^-$X$^-$X$^-$X$^-$; six 3Q terms: 2 A$^+$X$^+$X$^+$X$^+$X$^-$, 2 A$^-$X$^-$X$^-$X$^-$X$^+$, 1 A$^+$X$^-$X$^-$X$^-$X$^-$, 1 A$^-$X$^+$X$^+$X$^+$X$^+$

four 1Q terms: 2 A$^+$X$^+$X$^-$X$^-$X$^-$, 2 A$^-$X$^-$X$^+$X$^+$X$^+$

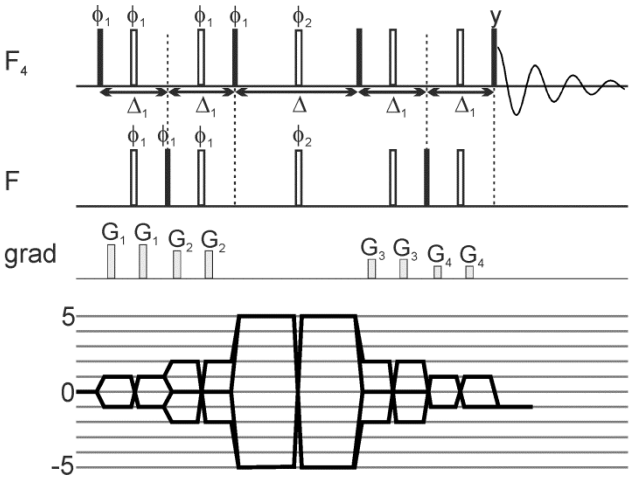

Fig. S2. A scheme of the optimal pulse sequence allowing creating 5Q coherence in an AX_4 spin system. All pulses should be selective shaped pulses. The Δ_1 delay should be set to $1/(2J_{A\text{-}X})$. The pulse phases should be set to x, unless shown explicitly. On the phase ϕ_1 a 10-step phase cycle should be performed to select the coherence of +/- 5 order during the multiple-quantum period.

The latter scheme is therefore more efficient by a factor of 8/3. Using both excitation schemes the pure 5Q coherence can be extracted using a proper phase cycle, but in the latter one the signal is stronger, i.e. the sensitivity of the experiment is higher. The sensitivity gain is even larger than 8/3, due to two facts. Firstly, there are four times more X nuclei than A nuclei in each molecule. Secondly, the signal of the nucleus X is splitted into a doublet (due to scalar coupling with A nucleus), while the signal of the nucleus A is splitted into five components (due to the scalar coupling with four X nuclei). However there is an additional relaxation loss due to the fact that the pulse sequence is longer by ca. 12.5 ms. Assuming transverse relaxation time of a free SF_5 ligand of 70 ms, that gives a loss of a factor of 0.84. Taking all above into account, the sensitivity gain of using the scheme proposed by us is almost twelve-fold.

To demonstrate the efficiency of 5Q coherence creation, we performed the experiment in a two-dimensional version, including the evolution of the 5Q coherence. The corresponding spectra (2Q, 4Q, 5Q) are shown in Fig. S3. It can be appreciated, that the peaks appear at the expected position along the multiple-quantum dimension (2, 4 and 5 times 6.75 kHz away from the spectrum center for 2Q, 4Q and 5Q coherence, respectively). Due to insufficient spectral width in the indirect dimension of the 2Q and 4Q spectra, the peaks are folded by 10 kHz and 30 kHz, respectively. During 2Q and 4Q coherence evolution, the scalar coupling between axial and equatorial fluorine nuclei evolved (which was not the case for 5Q evolution), resulting with doublets along the indirect dimension. Splitting between the doublet components is as expected: 2x and 4x 150 Hz for 2Q and 4Q coherence, respectively, thus confirming the multiple quantum coherence order.

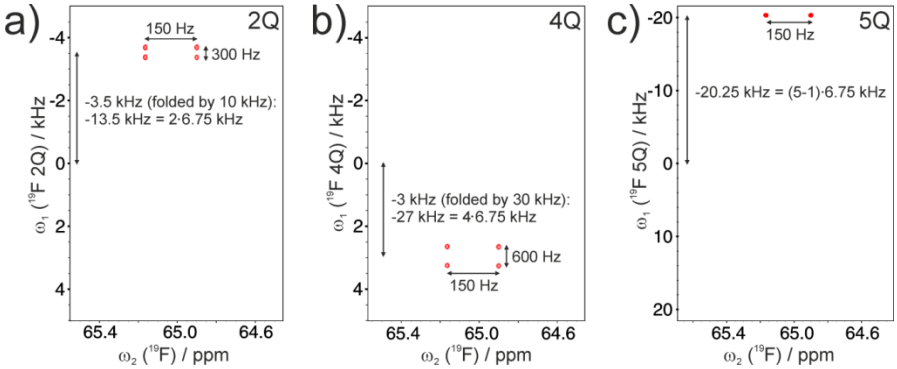

Fig. S3. 2D spectra of the ligand **1** in which 2Q (panel a), 4Q (panel b) and 5Q (panel c) coherence was created.

Table S1. Acquisition parameters for 2D experiments for ligand **1**.

coherence order	number of scans	Direct dimension		Indirect dimension	
		Spectral width, Hz	Number of time-domain points	Spectral width, Hz	Number of time-domain points
5Q	10	56818	16384	43103	256
4Q	16	56818	16384	10000	128
2Q	8	56818	16384	10000	256

2. Doublet components merging

The observed signal of the F_4 group in the fluorine-detected NMR spectra is a doublet, with 150 Hz scalar coupling. Both doublet components relax equally (no differential relaxation due to cross-correlation effects were detectable), therefore one could use information from both of them to decrease the uncertainty of the relaxation rate determination. There are two main ways to do it: (i) to determine the relaxation rate using each component separately, and then calculate their average (weighted with the inverse of the squared uncertainty of each rate), or (ii) to merge the two peaks (by shifting the spectrum by 75 Hz once left and once right and adding the two shifted spectra) and calculate the relaxation rate for the merged (more intensive) peak. The latter procedure is illustrated in Fig. S4. It increases the signal-to-noise ratio of the main peak by a factor of $\sqrt{2}$. We tried both solutions and we found out that merging procedure (ii) results with smaller uncertainties than averaging procedure (i). The results were compared in Fig. S5.

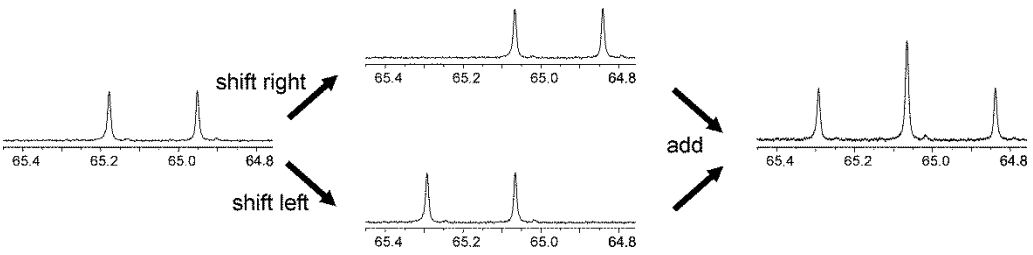

Fig. S4. The idea of the doublet component merging procedure.

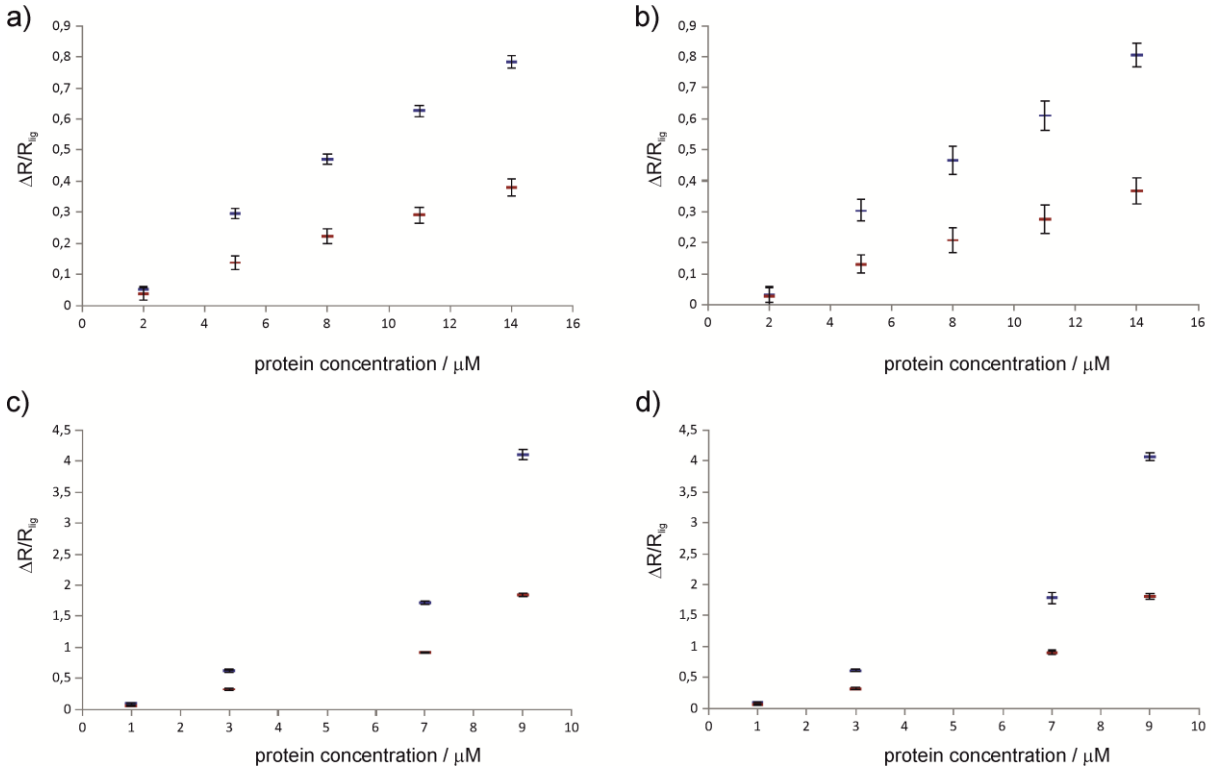

Fig. S5. Comparison of the uncertainties of the results obtained by: merging the doublet components (panels a,c) and calculating the average of the relaxation rates corresponding to the two doublet components (panels b,d). The data are shown for two protein-ligand pairs: NGAL with ligand **1** (panels a,b) and b-catenin with ligand **2** (panels c,d).

3. Relaxation calculation

In order to simulate the observed relaxation behaviour of the β-catenin / ligand **2** system, we suppose that the non-linearity seen in Fig. 3b is due to the oligomerization of the protein and the concomitantly larger relaxation rate of the bound forms. For sake of simplicity, we considered only dimerization of the protein and the following equilibria:

$PL \leftrightarrow P + L$ K_D

$P_2L \leftrightarrow P_2 + L$ K_D

$P_2L_2 \leftrightarrow P_2L + L$ K_D

$P_2 \leftrightarrow 2\,P$ K'

In our simulation we used the following parameters:

K_D = 1 mM

K'= 100 mM

$[L]^{tot}$ = 0,5 mM

R_L = 10 s^{-1}

R_{PL} = 1000 s^{-1}

$R_{P2L} = R_{P2L2}$ = 21000000 s^{-1}

The simulated and experimentally obtained relative relaxation rates are illustrated in Fig. S6.

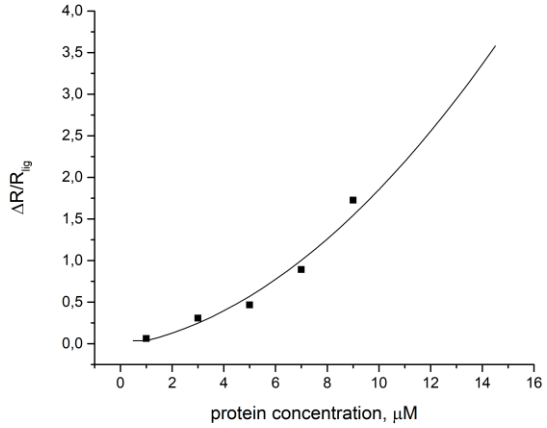

Fig. S6. Comparison of simulated (line) and experimentally measured (squares) relative relaxation rates as a function of β-catenin-concentration in case of concentration-dependent protein oligomerization.

4. Chemical Synthesis

Synthesis of 5-(pentafluoro-λ^6-sulfanyl)-1,3-benzoxazole-2(3*H*)-thione (ligand **1**)

A mixture of 2-amino-4-(pentafluorosulfanyl)phenol (5.00 g, 21.3 mmol), CS_2 (1.93 g, 25.3 mmol), KOH (1.35 g, 24.1 mmol) in EtOH (40 mL) and water (5 mL) was refluxed for 2 h. The mixture was then poured into water (250 mL), extracted with EtOAc (3 × 200 mL), the combined organic phase was dried ($MgSO_4$) and solvent was removed under reduced pressure. Purification by column chromatography (silica gel, hexane/EtOAc, 4:1) gave the desired product as a beige solid (4.56 g, 77%); mp 173-174 °C; ^{1}H NMR (400.1 MHz, acetone-d_6) δ 7.62 (d, J = 8.9 Hz, 1H), 7.77 (d, J = 2.3 Hz, 1H), 7.84 (dd, J = 8.9, 2.3 Hz, 1H), 12.4 (br s, 1H); ^{19}F NMR (376.5 MHz, acetone-d_6) δ 65.2 (d, J = 149.0 Hz, 4F), 85.2 (pent., J = 149.0 Hz, 1F); ^{13}C NMR (100.6 MHz acetone-d_6) δ 109.38 (pent., J = 5.0 Hz, CH), 110.76 (CH), 112.84 (pent., J = 5.0 Hz, CH), 132.59 (C), 150.92 (m, C), 182.75 (C=S); MS (EI) *m/z* 277 (100%) $[M]^+$, 169 (20), 150 (9), 122 (10), 63 (10); HRMS (EI) *m/z* calc for $C_7H_4F_5NOS_2$ $[M]^+$ 276.9654, found 276.9656.

NMR experiments of the synthetic work were performed on a Bruker Avance III HD 400 MHz spectrometer equipped with a broad-band probe (5mm BBO-1H Z-GRD).

2-bromo-4-(pentafluoro-λ^6-sulfanyl)aniline (ligand **2**) is commercially available.