3D DOSY–TROSY to determine the translational diffusion coefficient of large protein complexes

Tatiana Didenko[1], Rolf Boelens[2] and Stefan G.D. Rüdiger[1,3]

[1]Cellular Protein Chemistry, Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands and [2]Biomolecular NMR Spectroscopy, Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

[3]To whom correspondence should be addressed.
E-mail: s.g.d.rudiger@uu.nl

Edited by Assaf Friedler

The translational diffusion coefficient is a sensitive parameter to probe conformational changes in proteins and protein–protein interactions. Pulsed-field gradient NMR spectroscopy allows one to measure the translational diffusion with high accuracy. Two-dimensional (2D) heteronuclear NMR spectroscopy combined with diffusion-ordered spectroscopy (DOSY) provides improved resolution and therefore selectivity when compared with a conventional 1D readout. Here, we show that a combination of selective isotope labelling, 2D ^{1}H–^{13}C methyl-TROSY (transverse relaxation-optimised spectroscopy) and DOSY allows one to study diffusion properties of large protein complexes. We propose that a 3D DOSY–heteronuclear multiple quantum coherence (HMQC) pulse sequence, that uses the TROSY effect of the HMQC sequence for ^{13}C methyl-labelled proteins, is highly suitable for measuring the diffusion coefficient of large proteins. We used the 20 kDa co-chaperone p23 as model system to test this 3D DOSY–TROSY technique under various conditions. We determined the diffusion coefficient of p23 in viscous solutions, mimicking large complexes of up to 200 kDa. We found the experimental data to be in excellent agreement with theoretical predictions. To demonstrate the use for complex formation, we applied this technique to record the formation of a complex of p23 with the molecular chaperone Hsp90, which is around 200 kDa. We anticipate that 3D DOSY–TROSY will be a useful tool to study conformational changes in large protein complexes.

Keywords: DOSY/NMR spectroscopy/protein folding/protein–protein interactions/TROSY

Introduction

The translational diffusion coefficient of a protein is a function of its molecular weight, size and shape. It is, therefore, a useful parameter to reveal information about a protein's hydrodynamic state in solution, but also to study conformational changes upon addition of ligands and to monitor protein–protein interactions (Serdyuk *et al.*, 2007). NMR spectroscopy is a convenient technique to measure diffusion in solution using a pulse-field gradient (PFG) technique, Diffusion Optimised Spectroscopy DOSY (diffusion-ordered spectroscopy) (Morris and Johnson, 1992; Johnson, 1999).

A conventional DOSY spectrum has a diffusion dimension and one frequency axis. However, for large macromolecules with poorly resolved NMR spectra, the diffusion readout using one frequency dimension will result in considerable spectral overlap when studying macromolecular complexes that contain a mixture of molecules. Barjat *et al.* (1998) showed that the addition of a second frequency axis in a 3D DOSY–heteronuclear multiple quantum coherence (HMQC) experiment is very powerful for a detailed analysis of the diffusion coefficients of complex mixtures of small organic molecules. However, for large protein complexes containing a mixture of different (large) molecules the direct application of this approach would fail. The short T_2 relaxation times of the NMR signals of ^{13}C/^{15}N labelled proteins as commonly used in biomolecular studies, would lead to extremely poor sensitivity for large systems.

Using the selective isotope labelling instead of uniform labelling makes it possible to study big protein complexes. In recent time new pulse sequences and novel labelling strategies paved the way to close the gap to study large protein systems. Transverse relaxation-optimised spectroscopy (TROSY) allowed studying chemical shift changes in the 900 kDa GroEL/GroES complex (Pervushin *et al.*, 1997; Fiaux *et al.*, 2002). Complementary to those efforts, selective isotope labelling instead of uniform labelling further improved the repertoire to study large systems. Methyl-groups are the very attractive targets for selective labelling because of both relatively easy labelling techniques and their good spectroscopic properties (Tugarinov *et al.*, 2003, 2006). Highly deuterated samples that are ^{1}H–^{13}C labelled at selective methyl positions provide intense and well-dispersed ^{1}H–^{13}C correlation spectra even for large structures including the 20S proteasome, the SecA translocation pore and the Hsp90 chaperone system (Gelis *et al.*, 2007; Sprangers and Kay, 2007; Karagöz *et al.* unpublished results). The relaxation properties of ^{1}H–^{13}C-labelled methyl groups are remarkable because they lead to unusually slow-relaxing density matrix elements. Therefore it is possible to construct TROSY pulse sequences to exploit those slow-relaxing terms for a sensitive observation of the signal.

Here, we combine a DOSY experiment with a ^{1}H–^{13}C methyl-TROSY experiment to monitor diffusion properties of large systems. We applied the 3D DOSY–TROSY technique to study a selectively Ile-specific ^{1}H–^{13}C methyl-labelled but further deuterated protein, p23, in different viscous solutions mimicking large complexes and demonstrated the application to a 200 kDa system, the complex of the molecular chaperone Hsp90 with its co-chaperone p23. We expect this method to be suited to study shape changes of multimeric protein complexes.

Results

The 3D DOSY–TROSY experiment

The 3D DOSY–TROSY experiment (Fig. 1) is essentially a 3D DOSY–HMQC experiment as originally suggested by

Barjat *et al.* (1998), where the HMQC part has been rede-signed as a 3D ^{1}H–^{13}C methyl-TROSY. This enables to measure the translation diffusion coefficients of large methyl-labelled proteins. The pulse sequence consists of two building blocks: a DOSY stimulated-echo (STE) sequence, with a main diffusion delay where the magnetisation is along the *z*-axis, and a ^{1}H–^{13}C methyl-TROSY, which is essentially an HMQC. Both building blocks are highly suitable for large proteins.

It has been shown that the ^{1}H–^{13}C methyl-TROSY spectrum can be best measured via a ^{1}H–^{13}C HMQC experiment (Tugarinov *et al.*, 2003). Interference between relaxation processes in a CH$_3$ spin system causes some density matrix elements to relax more slowly. Therefore it is possible to get a signal enhancement by keeping slow-relaxing coherences non-mixed with fast-relaxing coherences through the experiment. HMQC is known to lead to up to 2.6-fold higher signal intensities than HSQC because it lacks the additional $\pi/2$ ^{1}H pulses that irrevertably mix fast- and slow-relaxing coherences (Ollerenshaw *et al.*, 2003; Tugarinov *et al.*, 2006). Thus, the HMQC pulse sequence is appropriate for methyl-labelled proteins and that is the reason we use it in our experiment.

It is possible to design a 3D DOSY–TROSY experiment using the spin-echo (SE) DOSY with just one ^{1}H $\pi/2$ pulse (Carr and Purcell, 1954). In this case the magnetisation remains in the transverse *xy*-plane during the diffusion time. The fast- and slow-relaxing coherences will not be mixed, but in practice the experiment is not optimal for macromolecules due to fast relaxation of transverse magnetisation in combination with long gradient delays. Therefore, we used an STE pulse sequence (Hahn, 1950) to keep the magnetisation along the *z*-axis during the diffusion time. The STE pulse sequence contains three $\pi/2$ ^{1}H pulses and this is the minimum amount of $\pi/2$ ^{1}H pulses to perform a DOSY experiment on a macromolecule.

We analysed the evolution of the CH$_3$ spin system density matrix through the DOSY pulse sequences to estimate the signal attenuation caused by additional $\pi/2$ ^{1}H pulses using the principles described in reference (Ollerenshaw *et al.*, 2003). In case of the SE pulse sequence half of the signal derives from the slow-relaxing coherence terms. After two additional $\pi/2$ ^{1}H pulses corresponding to STE pulse sequence the fast- and slow-relaxing terms do not mix. Nevertheless, the analysis of the density matrix evolution after two successive $\pi/2$ ^{1}H pulses shows that multiple-quantum coherence terms appear. When those terms are eliminated by a spoil-gradient the fast- and slow-relaxing coherences are irrevertably mixed and, therefore, the final signal intensity would be reduced. The signal intensity would decrease 1.3-fold for each additional pair of $\pi/2$ ^{1}H pulses. Thus, it is essential to use not more than three $\pi/2$ ^{1}H pulses in a DOSY pulse sequence for methyl-labelled proteins. In our work, we used STE–DOSY with bipolar pulses (Johnson 1999).

To obtain the translational diffusion constant we recorded a set of DOSY–TROSY spectra with an increasing encoding/decoding gradient strength. We calculated the diffusion coefficient from the decay of signals, which were fitted by using Equation (1):

$$\frac{I}{I_0} = e^{-D\gamma^2 g^2 \delta^2 (\Delta - \delta/3 - \tau/2)}, \tag{1}$$

where *I* is the observed intensity, I_0 the reference intensity, *D* the diffusion coefficient, γ the gyromagnetic ratio of the observed nucleus, *g* the gradient strength, δ the duration of the gradient, Δ the diffusion time and τ is a delay between bipolar gradients.

The protein p23 is a good model to test 3D DOSY–TROSY

To test the 3D DOSY–TROSY method, we chose as model system the protein p23. It is a globular protein of 20 kDa that contains four Ile side chains (Fig. 2A). It forms a complex with the molecular chaperone Hsp90, which is composed of up to two p23 and two Hsp90 molecules (Ali *et al.*, 2006; Karagöz *et al.*, unpublished). Hsp90 is a dimer of 170 kDa, so that resulting complexes with p23 are around 200 kDa. This system is, therefore, excellentially suited to test the potential of 3D DOSY–TROSY for large systems. We applied the method pioneered by Kay and co-workers to selectively label the δ-methyl groups of those side chains with ^{13}C and ^{1}H in an otherwise deuterated background (Tugarinov *et al.*, 2006). We acquired ^{13}C–^{1}H HMQC spectra, which showed four well-resolved peaks, in agreement with previous findings (Fig. 2B; Dyson *et al.*, 2008). The spectra demonstrated that p23 is a suitable system for our study.

We now applied the DOSY–TROSY method to a p23 sample. We measured the decrease of signal intensity with increasing gradient strength. We fitted the experimental values using Equation (1) and calculated the translation diffusion constants, *D* (Fig. 3, Table I). The signals of all four isoleucine decreased at the same rate, resulting in almost identical translational diffusion constants. This was expected since translational diffusion is a property of the entire molecule.

We compared the translation diffusion coefficient obtained in the DOSY experiment with the theoretical constant calculated by HYDRONMR (García de la Torre *et al.*, 2000) using a model-based in the crystal structure of p23. The experimental diffusion coefficient was 9.8×10^{-7} cm/s$^2 \pm 0.2 \times 10^{-7}$ cm/s^2. This value agrees well with the coefficient calculated by using HYDRONMR, $D = 9.6 \times 10^{-7}$ cm/s^2 (Table I). This indicates that p23 is monomeric and has similar shape properties in solution as in the crystal.

Sensitivity of 3D DOSY–TROSY

To show that method is suitable for big proteins and protein complexes, we performed diffusion measurements on p23 in the presence of increasing concentrations of glycerol (Fig. 4). The glycerol increases the viscosity of the solution thus increasing the tumbling time of the protein. This allows one to simulate a much larger protein, e.g. the diffusion coefficient obtained for p23 in 30% glycerol would correspond to the diffusion coefficient of a globular protein of around 200 kDa. We determined the translational diffusion coefficient for each glycerol concentration and compared those to the *D* values predicted by HYDRONMR (Table I). We found experimental and theoretical values to be in excellent agreement. We therefore concluded that 3D DOSY–TROSY is highly suitable to determine the translation diffusion coefficients for large proteins and protein complexes.

To evaluate the TROSY sensitivity gain of 3D DOSY–TROSY, we compared DOSY–HMQC and DOSY–HSQC experiments on p23 in 20% glycerol (Fig. 5). For this
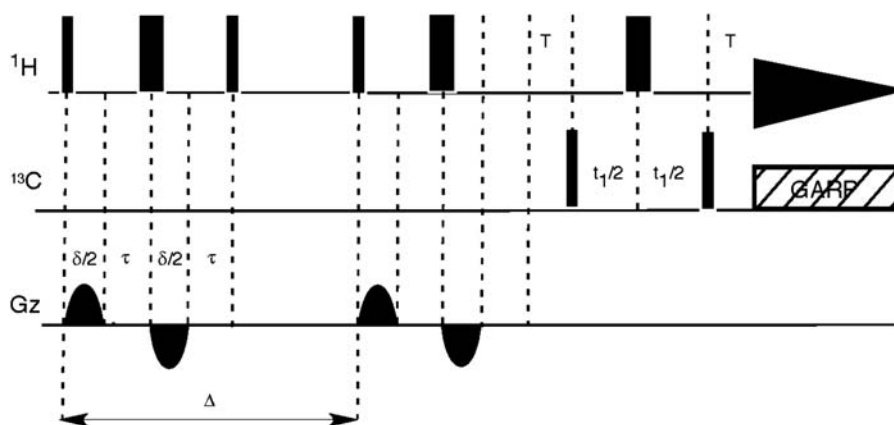
100

Fig. 1. STE–DOSY–HMQC pulse sequence to determine the translation diffusion coefficient of methyl-labelled proteins. The sequence used experimentally contains additional elements for correction of pulse imperfections and suppression of ^{13}C equilibrium magnetisation. Narrow and wide pulses have tip angles of $\pi/2$ and π, respectively. Pulses are applied along the x-axis. The delay T is of duration $1/(2J_{HC})$. ^{13}C decoupling is applied during the ^{1}H acquisition period.
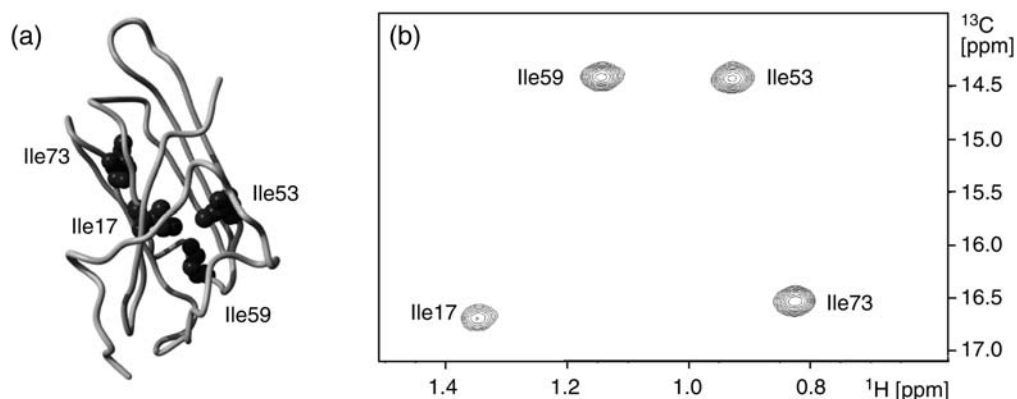

Fig. 2. The four Ile side chains of p23 are well resolved. (**a**) The crystal structure of p23. The Ile side chains are indicated as dark space fillings. (**b**) HMCQ spectra of p23, acquired on an Ileδ1-(^{13}CH$_3$)-labelled p23 sample (D$_2$O).

purpose, we constructed a DOSY–HSQC pulse sequence using the same approach as for the DOSY–HMQC pulse sequence described above. Figure 5 shows the traces from the DOSY–HMQC and DOSY–HSQC recorded under identical conditions. The ratio of intensities $I_{DOSY–HMQC}/I_{DOSY–HSQC}$ was on average 1.8, which demonstrated that DOSY–HMQC profits indeed from a sensitivity improvement that can be ascribed to the TROSY effect.

We also did the same experiments with an STE–longitudinal eddy-current delay (LED) pulse sequence as a DOSY part. The STE–LED pulse sequence contains two additional $\pi/2$ ^{1}H pulses (Johnson, 1999). We compared the signal intensity of STE–LED–DOSY–HMQC with that of the STE–DOSY–HMQC (Fig. 5). The signal intensity of STE–DOSY–HMQC is 1.2-fold more intense than of STE–LED–DOSY–HMQC. This correlates with the predicted intensity decrease upon addition of two $\pi/2$ ^{1}H pulses to STE–DOSY. Thus, the STE–DOSY–HMQC pulse sequence is preferable for diffusion measurements of large protein complexes.

3D DOSY–TROSY monitors the formation of a 200 kDa complex

Two of the possible applications of DOSY–TROSY are the visualisation of the formation of large protein complexes and the characterisation of its hydrodynamic properties. Here, we used the interaction of p23 with its natural interaction

partner, the molecular chaperone Hsp90. Hsp90 is a 170 kDa dimer that can bind two p23 molecules in a positively co-operative mechanism (Karagöz et al., unpublished). We set out to test whether the 3D DOSY–TROSY method was able to monitor the formation of such a large complex.

We used ^{13}C-methyl-labelled p23 and unlabelled Hsp90, which will be invisible in a ^{1}H–^{13}C HMQC spectrum. We formed the Hsp90–p23 complex by using a 1.5-fold excess of Hsp90 to ensure that most p23 was bound to the chaperone. We added ATPγS as a non-hydrolysable ATP analogue, which is required for efficient formation of the Hsp90–p23 complex (Karagöz et al., unpublished). We performed a series of DOSY–TROSY diffusion measurements on the p23–Hsp90 mixture and compared the result to the coefficients obtained in the same experiment in the absence of Hsp90 (Fig. 6). The presence of Hsp90 reduced the translational diffusion rate of p23 1.7-fold ($D = 6.4 \times 10^{-7}$ cm/s$^2 \pm 0.2 \times 10^{-7}$ cm/s^2), which indicated that p23 was now part of a large complex.

Discussion

We introduced a 3D DOSY–TROSY technique to study diffusion properties of large proteins. We demonstrated, using p23 as a model, that this pulse sequence allows the precise
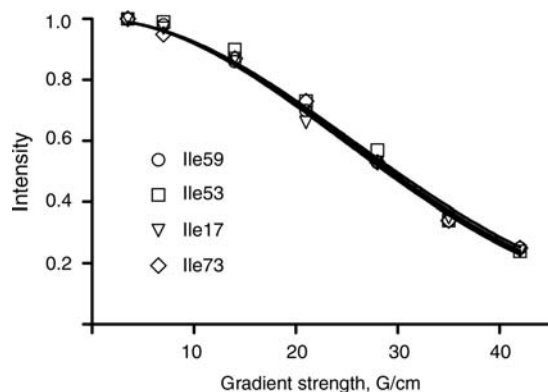
Fig. 3. The DOSY–HMQC experiments reveals identical diffusion coefficients for all four Ile side chains. Pulsed-field gradient diffusion experiment on an Ileδ1-(^{13}CH$_3$)-labelled p23 sample (D$_2$O). The circle, diamond, square, inverted triangle symbols represent experimental values for each Ile with increasing gradient strength. The data were fitted to Equation (1) (black lines).

Table I. Translation diffusion constants of Hsp90 complexes.

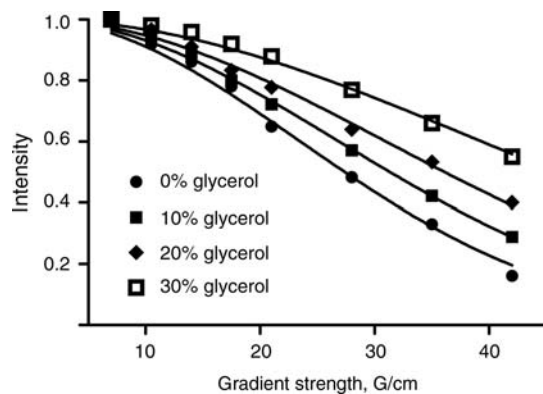|	Experimental diffusion coefficient, D	Prediction by HYDRONMR
Ile59	9.94×10^{-7} cm/s$^2 \pm 0.05 \times 10^{-7}$ cm/s^2	
Ile53	9.45×10^{-7} cm/s$^2 \pm 0.10 \times 10^{-7}$ cm/s^2	
Ile17	1.00×10^{-8} cm/s$^2 \pm 0.07 \times 10^{-7}$ cm/s^2	
Ile73	9.77×10^{-7} cm/s$^2 \pm 0.07 \times 10^{-7}$ cm/s^2	
p23 in 0% glycerol	9.9×10^{-7} cm/s$^2 \pm 0.2 \times 10^{-7}$ cm/s^2	9.86×10^{-7} cm/s^2
p23 in 10% glycerol	8.54×10^{-7} cm/s$^2 \pm 0.2 \times 10^{-7}$ cm/s^2	8.17×10^{-7} cm/s^2
p23 in 20% glycerol	6.43×10^{-7} cm/s$^2 \pm 0.2 \times 10^{-7}$ cm/s^2	6.23×10^{-7} cm/s^2
p23 in 30% glycerol	3.99×10^{-7} cm/s$^2 \pm 0.2 \times 10^{-7}$ cm/s^2	4.11×10^{-7} cm/s^2
p23 + Hsp90	6.70×10^{-7} cm/s$^2 \pm 0.2 \times 10^{-7}$ cm/s^2	

Fig. 4. The DOSY–HMQC experiment allows to monitor diffusion in viscous solutions. Pulsed-field gradient diffusion experiment on an Ileδ1-(^{13}CH$_3$)-labelled p23 sample (D$_2$O) in the presence of increasing concentrations of glycerol. The closed circle, closed square, closed diamond, square symbols represent experimental values for each glycerol concentration (averaged values of all four Ile side chains). The data were fitted to Equation (1) (black lines).

determination of the translational diffusion coefficient of 200 kDa systems.

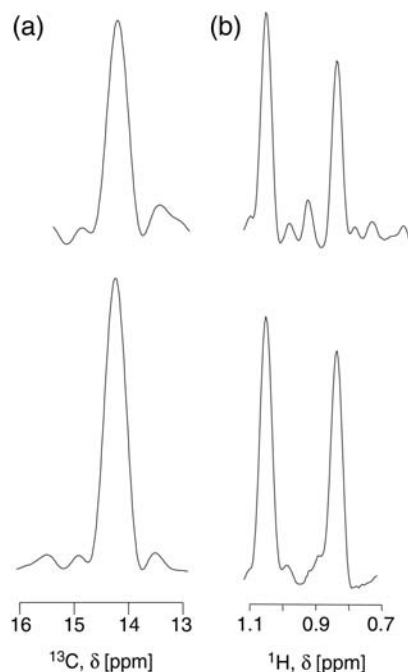Protein machines often undergo conformational changes during their active cycle, the Hsp90 system is a prime

Fig. 5. The TROSY effect improves sensitivity in the DOSY–HMQC experiment. (**a**) ^{1}H traces from HMQC (bottom) and HSQC (top) correlation maps of p23. (**b**) ^{13}C traces from HMQC (bottom) and HSQC (top) correlation maps of p23.
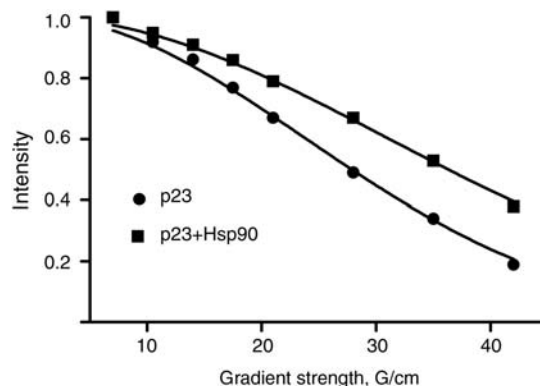

Fig. 6. Monitoring the formation of a Hsp90–p23 complex by DOSY–HMQC. Pulsed-field gradient diffusion experiment on an Ileδ1-(^{13}CH$_3$)-labelled p23 sample (D$_2$O) alone and in the presence of 1.5-fold excess of Hsp90. Experimental values in the presence (closed square) and absence (closed circle) of Hsp90 are indicated (averaged values of all four Ile side chains). The data were fitted to Equation (1) (black lines).

example (Pearl and Prodromou, 2006; Southworth and Agard, 2008; Mayer, 2010). The diffusion constant is an important parameter to monitor such changes (Dehner and Kessler, 2005). NMR spectroscopy itself has significant advantages compared with other techniques for diffusion measurement such as dynamic light scattering, size exclusion chromatography and analytical ultracentrifugation. Those advantages are selective monitoring of a labelled component only, a possibility to use big amounts of fluorescent ligands such as nucleotides and the wide range of experimental conditions, which may also be very close to physiological if necessary.

The application of 3D DOSY–TROSY to ^{1}H–^{13}C methyl-labelled proteins combines high sensitivity with high

resolution. This enables high accuracy with the possibility to monitor processes that occur during formation of macromolecular complexes with high selectivity. So far, diffusion NMR was often applied to small organic molecules, peptides or disordered proteins such as the N-terminal domain of the transcription factor p53 (Dehner and Kessler, 2005), all with relatively long transversal relaxation times. The sensitivity of our DOSY–TROSY approach in combination with the selective labelling makes it rather suitable for large systems. We found the TROSY effect in the DOSY–HMQC pulse sequence was comparable with that observed earlier for HMQC experiments on methyl-labelled proteins (Fig. 5; Ollerenshaw *et al.*, 2003). Experimental values obtained using this pulse sequence correlated very well with theoretical constants for a system that corresponded to a 200 kDa protein, and we expect that the method could be used for even larger systems.

The monitoring of complex formation is a rather useful application of DOSY-NMR. We showed that the 3D DOSY–TROSY technique applied on a ^{1}H–^{13}C methyl-labelled protein made it possible to visualise the complex formation of a 20 kDa protein with a 170 kDa non-labelled protein. In addition, as the ^{1}H–^{13}C correlation spectra of ^{1}H–^{13}C methyl-labelled proteins are usually well spread it may be possible to differentiate in 3D DOSY–TROSY separate components in complex mixtures of labelled proteins. We anticipate that our technique may be applied to study conformational changes and complex formation of large, multiprotein systems in the future.

Methods

Expression and purification

The human Hsp90β and p23 proteins were produced with a N-terminal his tag. *E. coli* cells carrying an expression plasmid for p23 were grown in 99% D_2O minimal medium containing NH_4Cl as the sole nitrogen source and ^{1}H/^{12}C-glucose as the carbon source. To achieve methyl labelling, the media were supplemented with 100 mg/l 2-keto-3,3-d2-4-^{13}C-butyrate (for Ile-[δ1^{13}CH$_3$]) 1 h prior to induction of p23 expression, according to a protocol, modified from Tugarinov *et al.* (2006). Hsp90 was produced in LB media without isotope labelling. Both proteins were purified on Poros 20MC metal chelate media followed by Poros 20HQ anion exchange media. For the NMR experiments, p23 sample conditions were as follows: 100 μM protein, 99.9% D_2O, 150 mM sodium phosphate pD 7.2.

NMR spectroscopy

All the spectra were recorded on an 750 MHz Bruker Avance2 spectrometer with a TXI probe with triple-axis gradients at 300 K with (80, 1024) complex points in (^{13}C–^{1}H) frequency dimensions. Thirty-two scans were recorded for each FID, with total acquisition time 2 h for each spectrum. A recycle delay of 1.5 s was used in all experiments. The parameters for DOSY diffusion time $\Delta = 400$ ms, $\delta = 3.4$ ms, $\tau = 0.5$ ms. The gradient strength varied from 5 to 42.5 G/cm.

Data analysis

Experimental values were plotted and fitted using a GraphPad Prism software. Hydrodynamic calculations were done by using the shell model with hydration layer 3.2 Å of HYDRONMR (García de la Torre *et al.*, 2000) with parameters corrected according to experimental conditions.

References

Ali,M.M., Roe,S.M., Vaughan,C.K., Meyer,P., Panaretou,B., Piper,P.W., Prodromou,C. and Pearl,L.H. (2006) *Nature*, **440**, 1013–1017.
Barjat,H., Morris,G.A. and Swanson,A.G. (1998) *J. Magn. Reson.*, **131**, 131–138.
Carr,H.Y. and Purcell,E.M. (1954) *Phys. Rev.*, **94**, 630–638.
Dehner,A. and Kessler,H. (2005) *Chembiochem*, **6**, 1550–1565.
Dyson,H.J., Kostic,M., Liu,J. and Martinez-Yamout,M.A. (2008) *FEBS Lett.*, **582**, 1495–1500.
Fiaux,J., Bertelsen,E.B., Horwich,A.L. and Wuthrich,K. (2002) *Nature*, **418**, 207–211.
García de la Torre,J., Huertas,M.L. and Carrasco,B. (2000) *J. Magn. Reson.*, **147**, 138–146.
Gelis,I., Bonvin,A.M., Keramisanou,D., Koukaki,M., Gouridis,G., Karamanou,S., Economou,A. and Kalodimos,C.G. (2007) *Cell*, **131**, 756–769.
Hahn,E.L. (1950) *Phys. Rev.*, **80**, 580–594.
Johnson,C.S., Jr. (1999) *Prog. Nucl. Magn. Reson. Spectrosc.*, **34**, 203–256.
Mayer,M.P. (2010) *Mol. Cell.*, **39**, 321–331.
Morris,K.F. and Johnson,C.S. (1992) *J. Am. Chem. Soc.*, **114**, 3139–3141.
Ollerenshaw,J.E., Tugarinov,V. and Kay,L.E. (2003) *Magn. Reson. Chem.*, **41**, 843–852.
Pearl,L.H. and Prodromou,C. (2006) *Annu. Rev. Biochem.*, **75**, 271–94.
Pervushin,K., Riek,R., Wider,G. and Wüthrich,K. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 12366–71.
Serdyuk,I.N., Zaccai,N.R., Zaccai,J. and Zaccai,G. (2007) *Methods in molecular biophysics: structure, dynamics, function*. Cambridge University Press.
Southworth,D.R. and Agard,D.A. (2008) *Mol. Cell.*, **32**, 631–640.
Sprangers,R. and Kay,L.E. (2007) *Nature*, **445**, 618–622.
Tugarinov,V., Hwang,P.M., Ollerenshaw,J.E. and Kay,L.E. (2003) *J. Am. Chem. Soc.*, **125**, 10420–10428.
Tugarinov,V., Kanelis,V. and Kay,L.E. (2006) *Nat. Protoc.*, **1**, 749–754.