Use of ^1F Differential Labelling for the Simultaneous Detection and Monitoring of Three Individual Proteins in a Serum Environment

Edwards, John M.; Harris, Pernille; Bukrinski, Jens T.; Golovanov, Alexander P.

Published in:
ChemPlusChem

Link to article, DOI:
[10.1002/cplu.201900110](https://doi.org/10.1002/cplu.201900110)

Publication date:
2019

Document Version
Peer reviewed version

[Link back to DTU Orbit](#)

Citation (APA):
Edwards, J. M., Harris, P., Bukrinski, J. T., & Golovanov, A. P. (2019). Use of ^1F Differential Labelling for the Simultaneous Detection and Monitoring of Three Individual Proteins in a Serum Environment. *ChemPlusChem*, 84(5), 443-446. <https://doi.org/10.1002/cplu.201900110>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

A GENUINELY MULTIDISCIPLINARY JOURNAL

CHEMPLUSCHEM

CENTERING ON CHEMISTRY

Accepted Article

Title: Use of ^{19}F Differential Labelling for the Simultaneous Detection and Monitoring of Three Individual Proteins in a Serum Environment

Authors: John M. Edwards, Pernille Harris, Jens T Bukrinski, and Alexander P. Golovanov

This manuscript has been accepted after peer review and appears as an Accepted Article online prior to editing, proofing, and formal publication of the final Version of Record (VoR). This work is currently citable by using the Digital Object Identifier (DOI) given below. The VoR will be published online in Early View as soon as possible and may be different to this Accepted Article as a result of editing. Readers should obtain the VoR from the journal website shown below when it is published to ensure accuracy of information. The authors are responsible for the content of this Accepted Article.

To be cited as: *ChemPlusChem* 10.1002/cplu.201900110

Link to VoR: <http://dx.doi.org/10.1002/cplu.201900110>

WILEY-VCH

www.chempluschem.org

A Journal of

Use of ^{19}F Differential Labelling for the Simultaneous Detection and Monitoring of Three Individual Proteins in a Serum Environment

John M Edwards^[a], Pernille Harris^[c], Jens T Bukrinski^[b], Alexander P Golovanov^{*[a]}

Abstract: Protein behavior in complex mixtures, such as biological fluids, is often modeled by simplified solution buffer systems. Here we have used the recently described differential ^{19}F labelling approach (with NMR detection) to monitor and compare the solution behaviour of three proteins at once: human serum albumin (HSA), transferrin (TrF) and immunoglobulin G (IgG), both in serum and in buffer. We demonstrate that monitoring three proteins simultaneously and independently in biological fluid is possible, and that the presence of other endogenous components greatly changes the association characteristics of these proteins. For example, in the simplified model buffer system all three proteins diffuse at a similar rate, while in serum HSA diffuses around 3 times faster than TrF, and 4 times faster than IgG. This ^{19}F NMR approach opens avenues to characterise the behaviour of complex multi-protein systems in their native environment, e.g., in biological fluids.

The bloodstream and interstitial fluid environment is a complex and crowded mixture of proteins, cells and other molecular components^[1]. These proteins undergo interactions with themselves and with other macromolecules, modulated by a multitude of smaller native molecules and environmental factors. This complexity may not be adequately captured by simplified in vitro model systems consisting of only few selected proteins in a defined buffer. Investigation of this complexity is necessary for molecular understanding of disease states (e.g. protein aggregation in cryoglobulinemia^[2]), or for understanding of the interactions of protein-based therapeutics^[3]. Once injected, biotherapeutics become exposed to this complex environment and it is important to be able to observe their behaviour under these conditions^[4]. Blood serum is a perfect example of a complex environment with high concentrations of endogenous proteins present, including albumin (HSA in humans, at approximately 50 mg/ml^[5]), transferrin (TrF, 3 mg/ml^[6]) and immunoglobulin G (IgG, 7 mg/ml^[5]). The purified or recombinant forms of these proteins are used in their own right as biopharmaceuticals (e.g. IgGs), stabilisers of pharmaceuticals (HSA^[7]), or as drug delivery systems (HSA^[8] and TrF^[9]). Interactions of such biopharmaceuticals with each other and with other native components of interstitial fluids or blood therefore require characterization^[10]. Behaviour of recombinant exogenous proteins in serum may provide a reasonable and yet simple model for the behaviour of biopharmaceuticals injected into a body and their initial interactions with its endogenous components.

Biophysical investigation of proteins in serum is however problematic due to the difficulty of distinguishing the protein of interest from the many other endogenous proteins present which can mask or overwhelm any useful signal or measurement. Due to its selectivity, NMR spectroscopy has been used for observing specific types of

molecules, but generally requires the protein of interest to be sufficiently small and isotopically labelled with ^{13}C and/or ^{15}N to allow its specific detection^[11]. For therapeutic proteins obtained, for example, from donor blood (such as IgGs) or produced recombinantly in mammalian cells (such as monoclonal antibodies) this isotopic labelling is at best expensive and is often impractical^[11]. An alternative approach is the modification of protein molecules to incorporate non-native ^{19}F nuclei, either using modified amino acids^[12] or extrinsic tags^[13]. A signal in the ^{19}F NMR spectra is then observed and interpreted^[14]. Such approaches have been used to observe individual proteins in buffers, in biological fluids, in cell media or inside cells^[15]. Recently we demonstrated that differential ^{19}F labelling, in which tags with characteristic fingerprint resonance frequency are chemically attached to different proteins before mixing them, provides a convenient method for tracking two proteins at once^[16]. In this approach each protein is represented in a spectrum by a characteristic group of ^{19}F signals, with no overlap between proteins^[16]. Here we extend this approach to three protein components, namely three major constituents of serum, HSA, TrF and IgG, and show that high quality measurements can be obtained in a serum environment as well as in simple buffers. These proteins are individually differentially-labelled in vitro and later reconstituted into different environments, enabling observation of their behaviour and change in apparent size upon changing conditions (temperature), without spectral interference from the endogenous solution components. We demonstrate that the molecular behaviours of HSA, TrF and IgG in model buffer and serum environment are significantly different, highlighting the likely inadequacy of simplified buffer models.

Figure 1. Example of 1D ^{19}F spectrum for a mixture of IgG-TFBPD, HSA-TFCS and TrF-BTFA, each at 20 mg/ml, at 277 K in serum. Structures for each ^{19}F labelling compound are shown above the corresponding NMR signal. Intensity differences are due to variations in labelling efficiency for the three ^{19}F tags.

- [a] Dr. John M Edwards and Dr. Alexander P Golovanov^{*}
Manchester Institute of Biotechnology and School of Chemistry,
University of Manchester,
131 Princess Street, Manchester, M1 7DN, UK.
E-mail: A.Golovanov@manchester.ac.uk
- [b] Dr. Jens T Bukrinski
CMC assist,
ApS, Saxhøjvej,
39A, DK-2500, Valby, DK.
- [c] Dr. Pernille Harris
DTU Chemistry,
Technical University of Denmark,
Building 207, DK-2800 Kgs. Lyngby, DK.

IgG, HSA and TrF were labelled using three orthogonal ^{19}F tags; TFBPD, TFCS and BTFA^[16] respectively, which are shown in Figure 1. TFBPD and BTFA function by modification of endogenous cysteine side chains in the proteins, accompanied by mild reduction with TCEP to ensure availability of the thiol groups. TFCS reacts with lysine side chains and the N-terminus of the protein. To uncover any differences in protein behaviour in buffer and in model biological

fluid (serum) here we have studied the effects of temperature on the molecular motions of ^{19}F -labelled HSA, TrF and IgG, exploring temperature-dependent protein association. These proteins were reconstituted at the same concentration (20 mg/ml each) in two environments: serum and phosphate buffer. The characteristic ^{19}F spectra (with 64 scans) show well-resolved signals representative of these three proteins in solution (Figure 1). Even in the most crowded and viscous environment (in serum at 277 K), the signals of the three differentially labelled proteins are clearly resolved, despite their inherently large size (66.5, 79.5 and 150 kDa, for HSA, TrF and IgG respectively). A control experiment on serum in the absence of labelled proteins showed no detectable background signals in the ^{19}F spectrum. NMR spectra for both samples were then recorded at a range of different temperatures from 277 to 333 K, measuring 1D spectra, translational diffusion coefficients D_L and transverse relaxation rates R_2 for three characteristic ^{19}F signals from the tagged HSA, TrF and IgG.

Figure 2. Diffusion data for a mixture of IgG, HSA and TrF in serum (panels A and C), and in phosphate buffer (panels B and D) across a range of temperatures. Diffusion coefficients for all three labelled proteins are shown in panels A and B. Panel B focuses on the slow diffusion range, with the full range displayed in the inset window. Effective radii r_{diff} are shown in panels C and D. Dashed guidelines show the expected changes in diffusion coefficients due solely to the changes in water viscosity with temperature. Confidence intervals are calculated by error propagation.

In the serum sample (Figure 2A) the diffusion coefficients (D_L) of the three proteins are clearly separated, with HSA diffusing the fastest by a large margin, then TrF, then IgG slowest. By contrast when in phosphate buffer these three serum proteins were found to diffuse at a very similar rate to each other across the entire temperature range investigated, to within the margin of error of the experiments (Figure 2B). At high temperatures (above 313K) the apparent diffusion rates of all three proteins in buffer become very rapid, with large error bars. This is likely due to increased contribution from thermal convection in the NMR tube for this less viscous sample, which in turn leads to the apparent decrease in r_{diff} seen in Figure 2D at higher temperature beyond that which would be expected for monomers. The change in observed diffusion coefficients for both samples is dominated by the temperature-dependence of water viscosity, with the relative expected contribution from viscosity shown as guidelines in Figure 2A and 2B. An effective molecular radius (r_{diff}) can be calculated from D_L using the Stokes-Einstein equation, taking into account changing viscosity, and used to assess changes in the apparent association states of these proteins in both samples. In the serum sample (Figure 2C) the three proteins show different behaviour as the temperature is raised, with r_{diff} of HSA marginally increasing but remaining close to the expected monomeric values.

TrF behaves as if it is a significantly larger aggregate ($r_{\text{diff}} \sim 10$ nm vs an expected value of 3.72 nm $^{(17)}$), showing a fairly flat size dependence with respect to temperature. IgG appears to exist as an even larger oligomer at low temperature, and showing a consistent decrease in its effective size above 288 K. At low temperature (below 293 K) the very slow diffusion rates are accompanied by relatively large experimental uncertainty. When the same three proteins are reconstituted in buffer rather than serum their temperature behaviour is much more uniform (Figure 2D), showing effective radii largely consistent with their molecular weight and a monomeric non-associated state throughout. Protein behaviour at the higher end of temperature range in this experimental setup was inconclusive, due to effects of sample convection, and possibility for protein destabilisation in buffer and/or in serum.

Figure 3. Transverse relaxation data for a mixture of IgG, HSA and TrF in serum (panels A and C), and in phosphate buffer (panels B and D) across a range of temperatures. Dashed guidelines indicated expected shape of R_2 dependency due solely to changing solution temperature and viscosity (see Eq. 1). Effective relaxation-derived radii r_{R2} (C and D) were calculated using Eq. 1.

An independently measured parameter, the transverse relaxation rate (R_2) of ^{19}F , also reflects on the relative change in molecular tumbling rates, although additionally it may be modulated by chemical exchange processes. R_2 is therefore dependent on viscosity changes and changes in aggregation state, as well as, potentially, transient interactions. There is a general decrease in the transverse relaxation rate R_2 for all three proteins with increasing temperature but most of this effect is accounted for by the decrease in water viscosity (Figure 3A,B). For each ^{19}F tagged protein the relaxation rates are significantly higher in the serum sample than in the buffer, suggesting larger molecular assemblies and greater contributions from chemical exchange processes brought about by increased transient interactions. Using Equation 1 and the experimentally determined r_{diff} for the three proteins at 298K in buffer, the relationship between the relaxation rate R_2 and the effective radius r_{R2} can be calibrated. The temperature dependences of r_{R2} values can be then compared for the two samples (Figure 3C,D). The r_{R2} dependencies for TrF and HSA in buffer are essentially flat with increasing temperature, with a size larger than would be expected for a monomeric protein state. IgG shows a constant r_{R2} up to 308K, followed by a slight decline at higher temperatures, suggesting disruption of an oligomeric form of the protein. In serum TrF still shows a roughly constant r_{R2} except for a slight increase at 333K (also seen for HSA and IgG), which may reflect onset of melting of the proteins rather than any form of protein clustering. HSA and TrF both display a sigmoidal trend in r_{R2} values, implying cluster formation at low temperatures. Increasing

temperature breaks up these clusters, with the radii reducing towards the monomeric value.

Both the diffusion and relaxation data indicate substantially different behaviour of HSA, TrF and IgG depending on whether they are reconstituted in a buffer system or in a model serum environment. The likely reason for these differences is the presence of other, as yet unidentified components of the serum. Some small molecules in serum at lower temperature reduce the aggregation state of selected proteins, for example HSA, while other serum molecules, large or small, promote formation of larger protein assemblies, for example for TrF or IgG. These observations highlight the inherent complexity of protein behaviour in biological fluids. Importantly the differential behaviour of these proteins can now be observed in the context of a complex biological fluid environment, despite their relatively large size and increased solution viscosity at low temperature, without interference from any background endogenous signals. Complex fluids, such as defibrinated blood, can be studied at near in-vivo conditions. Taking into account the underlying temperature-dependence of water viscosity, measuring translational diffusion and calibrating the values of observed transverse relaxation rates of ^{19}F signals enables the monitoring of relative changes in protein association states across a range of temperatures, for three proteins at once in this example. Cluster behaviour in such samples can also be assessed by other types of NMR experiments, as we recently demonstrated by applying dark state exchange saturation transfer NMR^[18]. Here we have shown for the first time the usefulness of the differential ^{19}F labelling strategy^[16] to studies of multiple protein behaviour in biological fluids.

Experimental Section

Recombinant HSA (Recombunin®) and recombinant TrF (DeltaFerrin®) were supplied by Albumedix Ltd. The IgG component of our samples is a heterogeneous mixture of wild-type IgGs derived from blood plasma of donors (Subcuvia®), thus having a composition of antibodies typical for an average adult. Here TrF was labelled with BTFA (3-bromo-1,1,1-trifluoroacetone supplied by Sigma-Aldrich, catalogue number 374059). IgG was labelled with TFBPD (1-(4-(trifluoromethyl)benzyl)-1H-pyrrole-2,5-dione), which was custom synthesised by Charnwood Molecular Ltd., Loughborough, UK. HSA was labelled with TFCS (N-(ε-trifluoroacetylcaproyloxy) succinimide ester)^[19] supplied by Fisher Scientific, catalogue number 22299. All reactions were carried out according to previously published protocols, using 5 mg/ml solutions of the proteins of interest, with ^{19}F labelling efficiencies in the 100%-200% range (i.e. 1 to 2 ^{19}F labels per protein molecule on average)^[16].

NMR experiments used a 100 mM sodium phosphate buffer, pH 7.2, with 10% (v/v) D_2O to provide a lock signal. For serum experiments 600 μL aliquots of horse serum (Sigma-Aldrich Cat. No. H0146) were lyophilized and then reconstituted using 600 μL of a 60 mg/mL total concentration equimolar mixture of differentially ^{19}F labelled TrF, IgG and HSA. Reconstitution of lyophilized serum was used to avoid dilution of the other serum components and to maintain the overall high protein concentration. Horse serum was chosen as a model to avoid potential safety issues and limitations associated with usage of human serum. Control experiments comparing lyophilized serum reconstituted in buffer with un-lyophilized serum showed no significant changes in the ^1H NMR spectrum. The buffer experiments used the same mixture of ^{19}F labelled proteins without the addition of lyophilized serum. NMR experiments were carried out on a Bruker Avance III 500 MHz spectrometer using a QCI-F cryoprobe with cooled ^1H and ^{19}F channels and sample temperature control unit. Diffusion coefficients were measured using stimulated echo pulsed-field gradients, by Bruker's standard pulse program stepppgp1s set up for ^{19}F . The diffusion time (Δ) and the gradient length (δ) were set to 200 and 2.0 ms, respectively. Initial gradient

strength calibration was conducted using diffusion of dioxane in water, and verified using known diffusion coefficients and radii of BSA monomers under standard conditions^[16]. Transverse relaxation rates were measured using Carr-Purcell-Meiboom-Gill sequence (CPMG), with total echo times varied between 2.8 ms and 730 ms^[16]. Chemical shift references in ^{19}F spectra were used as provided by standard spectrometer setup. All NMR data was processed and analysed using Topspin 3.5 and Dynamics Centre 2.4.8.

Transverse relaxation rate R_2 of the tagged protein (in the absence of internal motions and chemical exchange) is proportional to the rotational correlation time (τ_c), which depends on protein size according to Stoke's law:

$$\text{Equation 1: } R_2 \propto \tau_c = \frac{4\pi\eta(r_{R2})^3}{3kT}$$

where r_{R2} is the effective relaxation-derived radius, k is the Boltzmann constant and T is absolute temperature. The viscosity of water (η) at each temperature was taken from literature values^[20]. A calibration factor was calculated to convert R_2 to r_{R2} for each ^{19}F tag by comparison to the experimentally determined radius at 298K from the diffusion experiments. Effective radius r_{diff} can be calculated from the measured diffusion coefficients D_L using the Stokes-Einstein equation adjusted for the effects of molecular crowding^[16].

$$\text{Equation 2: } r_{diff} = \frac{kT}{6\pi\eta D_L} \frac{(1-\phi)^3}{(1+\frac{3}{2}\phi+2\phi^2+3\phi^3)}$$

where ϕ is the total volume fraction of the proteins in solution assuming a protein density factor of 1.25 g/ml^[21].

Acknowledgements

We are grateful to Alexander Fordham for help with the labelling of some proteins and preliminary experiments, and to Matthew Cliff for looking after the NMR Facility. We are grateful to MedImmune and Christopher van der Walle for useful discussions and insight, and to Albumedix Ltd for kindly providing us with Recombunin® and DeltaFerrin®. PH and APG acknowledge the EU ITN PIPPI consortium for funding work in their laboratories.

Keywords: diffusion - ^{19}F NMR Spectroscopy - multiprotein systems - protein-protein interactions - serum analysis

- [1] a) N. L. Anderson, N. G. Anderson, *Mol. Cell Proteomics* **2002**, *1*, 845-867; b) P. Gautam, S. C. Nair, K. Ramamoorthy, C. V. B. Swamy, R. Nagaraj, *Plos One* **2013**, *8*.
- [2] J. Damoiseaux, *Autoimmun. Rev.* **2014**, *13*, 359-362.
- [3] J. G. Elvin, R. G. Couston, C. F. van der Walle, *Int. J. Pharm.* **2013**, *440*, 83-98.
- [4] a) S. Yanaka, T. Yamazaki, R. Yogo, M. Noda, S. Uchiyama, H. Yagi, K. Kato, *Molecules* **2017**, *22*; b) Y. Wang, A. Lomakin, R. F. Latypov, G. B. Benedek, *Proc. Natl. Acad. Sci. U S A* **2011**, *108*, 16606-16611.
- [5] J. R. Marrack, H. Hoch, *J. Clin. Pathol.* **1949**, *2*, 161-192.
- [6] S. W. Stites, M. E. Nelson, L. J. Wesselius, *Chest* **1995**, *107*, 1681-1685.
- [7] W. Wang, *Protein Sci.* **2015**, *24*, 1031-1039.

- [8] J. T. Bukrinski, P. Sonderby, F. Antunes, B. Andersen, E. G. W. Schmidt, G. H. J. Peters, P. Harris, *Biochemistry* **2017**, *56*, 4860-4870.
- [9] a) M. E. Brandsma, A. M. Jevnikar, S. W. Ma, *Biotechnol. Adv.* **2011**, *29*, 230-238; b) H. Y. Li, H. Z. Sun, Z. M. Qian, *Trends Pharmacol. Sci.* **2002**, *23*, 206-209; c) Z. M. Qian, H. Y. Li, H. Z. Sun, K. Ho, *Pharmacol. Rev.* **2002**, *54*, 561-587.
- [10] a) S. Mitragotri, P. A. Burke, R. Langer, *Nat. Rev. Drug Discov.* **2014**, *13*, 655-672; b) I. R. Correia, *Mabs* **2010**, *2*, 221-232.
- [11] A. Dutta, K. Saxena, H. Schwalbe, J. Klein-Seetharaman, *Methods Mol. Biol.* **2012**, *831*, 55-69.
- [12] K. E. Arntson, W. C. Pomerantz, *J. Med. Chem.* **2016**, *59*, 5158-5171.
- [13] a) G. F. Wang, C. Li, G. J. Pielak, *Chembiochem* **2010**, *11*, 1993-1996; b) N. G. Sharaf, A. M. Gronenborn, *Isotope Labeling of Biomolecules - Labeling Methods* **2015**, 565, 67-95; c) J. T. Gerig, *Prog. Nucl. Mag. Res. Sp.* **1994**, *26*, 293-370; d) C. M. Tressler, N. J. Zondlo, *Biochemistry* **2017**, *56*, 1062-1074.
- [14] G. F. Wang, C. Li, G. J. Pielak, *Protein Sci.* **2010**, *19*, 1686-1691.
- [15] a) Y. Ye, X. Liu, Z. Zhang, Q. Wu, B. Jiang, L. Jiang, X. Zhang, M. Liu, G. J. Pielak, C. Li, *Chemistry* **2013**, *19*, 12705-12710; b) Y. Ye, Q. Wu, W. Zheng, B. Jiang, G. J. Pielak, M. Liu, C. Li, *Anal. Bioanal. Chem.* **2018**, *410*, 869-874; c) C. Frieden, S. D. Hoeltzli, J. G. Bann, *Energetics of Biological Macromolecules, Pt E* **2004**, *380*, 400-415; d) C. Li, G. F. Wang, Y. Wang, R. Creager-Allen, E. A. Lutz, H. Scronce, K. M. Slade, R. A. Ruf, R. A. Mehl, G. J. Pielak, *J. Am. Chem. Soc.* **2010**, *132*, 321-327.
- [16] J. M. Edwards, J. P. Derrick, C. F. van der Walle, A. P. Golovanov, *Mol. Pharm.* **2018**, *15*, 2785-2796.
- [17] J. K. Armstrong, R. B. Wenby, H. J. Meiselman, T. C. Fisher, *Biophys. J.* **2004**, *87*, 4259-4270.
- [18] J. M. Edwards, J. E. Bramham, A. Podmore, S. M. Bishop, C. F. van der Walle, A. P. Golovanov, *Anal. Chem.* **2019**, Just Accepted, DOI 10.1021/acs.analchem.1029b00143.
- [19] a) A. Ranganathan, S. J. Gee, B. D. Hammock, *Anal. Bioanal. Chem.* **2015**, *407*, 7263-7273; b) Y. H. Ding, S. B. Fan, S. Li, B. Y. Feng, N. Gao, K. Ye, S. M. He, M. Q. Dong, *Anal. Chem.* **2016**, *88*, 4461-4469.
- [20] L. Korson, Drosthan.W, F. J. Millero, *J. Phys. Chem.-Us* **1969**, *73*, 34-39.
- [21] a) K. M. Andersson, S. Hovmoller, *Z. Kristallogr.* **1998**, *213*, 369-373; b) M. Medina-Noyola, *Phys. Rev. Lett.* **1988**, *60*, 2705-2708.

Entry for the Table of Contents (Please choose one layout)

Layout 1:

COMMUNICATION

Complex mixture analysis: Complex mixtures in their native biological environments are difficult to analyse. Here we demonstrate that monitoring three proteins simultaneously and independently in biological fluid is possible using a ^{19}F differential labelling approach, and that the presence of other endogenous components greatly changes the association characteristics of these proteins compared to a simplified model buffer system.

John M Edwards, Pernille Harris, Jens T Bukrinski, Alexander P Golovanov *

Page No. – Page No.

Use of ^{19}F Differential Labelling for the Simultaneous Detection and Monitoring of Three Individual Proteins in a Serum Environment