Functional Protein Structure

DOI: 10.1002/anie.201003495

The Distribution of Fatty Acids Reveals the Functional Structure of Human Serum Albumin**

*Matthias J. N. Junk, Hans Wolfgang Spiess, and Dariush Hinderberger**

Human serum albumin (HSA), the most abundant protein in human blood plasma, serves as a transporting agent for various endogenous compounds and drug molecules.[1,2] Its capability to bind and transport multiple fatty acids (FA), in particular, has been studied extensively in the past.[3,4] Research on HSA was severely hampered by the complexity of the protein and has benefitted tremendously from crystallographic high-resolution structures. Nearly 20 years ago, He and Carter reported the first crystal structure of HSA.[5] To date, a multitude of HSA crystal structures have been deposited in the Protein Data Bank.

Even more important for understanding the binding properties of the protein are the structures of complexes of HSA and transported molecules. Thanks to the pioneering work of Curry et al., crystal structures of various HSA/fatty acid complexes have become accessible.[6–8] They found that fatty acids are distributed highly asymmetrically in the protein crystal although HSA itself exhibits a symmetric primary and secondary structure. Up to seven distinct binding sites were found for long-chain fatty acids, most of which comprised ionic anchoring units and long, hydrophobic pockets.[8,9] The location of two to three high-affinity binding sites[3,10] was assigned by correlation of the X-ray structure with NMR studies on competitive binding of drugs that replaced ^{13}C-labeled fatty acids.[11,12] Sites 2, 4, and 5 bind fatty acids with a high affinity, while sites 1, 3, 6, and 7 exhibit a somewhat lower affinity to fatty acids (see Figure 1 a).

More generally, there is a long-standing debate as to what extent protein crystal structures reflect the dynamic and functional structures of proteins in solution. This debate is often fueled by apparent discrepancies between X-ray crystallographic data and results from solution-state-based techniques (e.g. NMR and other types of spectroscopy as well as neutron scattering) or from molecular dynamics simulations. Moreover, there is an increasing awareness that protein dynamics in solution is connected to biological function. Recent NMR studies revealed that many proteins exhibit pronounced dynamic conformational flexibilities with implications for the function.[13–15]

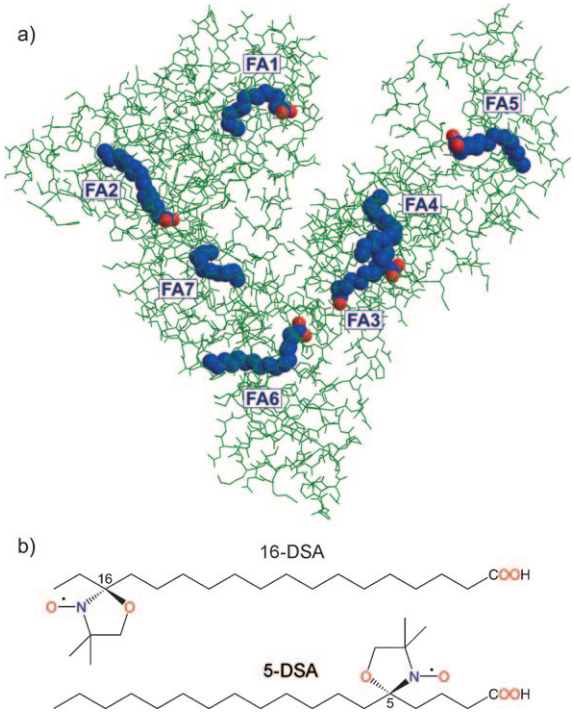

Figure 1. a) Crystal structure (PDB 1e7i) of HSA cocrystallized with seven stearic acid molecules.[8] Blue: C, red: O, green: protein framework. b) Chemical structure of the EPR-active molecules 5-doxylstearic acid (5-DSA) and 16-DSA.

In this context we note that the surface-exposed parts of HSA show a high degree of flexibility which constitutes a key to the versatility of this protein in binding various molecules. Already in the 1950s, Karush developed a concept that accounted for this conformational adaptability of the binding sites.[16,17] Further, a model has been proposed that takes into account the conformational entropy arising from the flexibility of the fatty acid alkyl chains.[18]

Our study aims at revealing the functional structure of HSA with respect to its binding of fatty acids directly from the fatty acids' point of view. This is achieved by electron paramagnetic resonance (EPR) spectroscopy, studying spin-labeled fatty acids that have stable nitroxide radicals incorporated and give rise to an EPR signal. Thus, the distribution of the FA binding sites is detected without any contribution from the complex protein itself (Figure 1). Structural information on the binding sites is obtained by determining the distance distributions between the fatty acids in frozen

[*] M. J. N. Junk, Prof. Dr. H. W. Spiess, Dr. D. Hinderberger
Max-Planck-Institut für Polymerforschung
Ackermannweg 10, 55128 Mainz (Germany)
Fax: (+49) 131-379-100
E-mail: dariush.hinderberger@mpip-mainz.mpg.de

[**] We thank Gunnar Jeschke (ETH Zürich) and Marcos Gelos (St. Augustinus-Krankenhaus Düren) for insightful discussions and Christian Bauer for technical support. M.J.N.J. gratefully acknowledges financial support from the Foundation of the German Chemical Industry (FCI) and from the Graduate School of Excellence "Materials Science in Mainz" (MAINZ).

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.201003495.

solution. These distance distributions in the range of several nanometers are retrieved by double electron-electron resonance (DEER), a pulse EPR method that utilizes the inherent distance dependence of the dipolar couplings (acting solely through space) between the unpaired electron spins.[19–21] In recent years DEER has been used increasingly in structural studies on both synthetic[22,23] and biological systems[24] with the focus on (membrane) proteins and nucleic acids.[25–28]

To obtain distance information from different positions along the methylene chain of the fatty acids in the respective binding sites, fatty acids with different labeling positions were applied. In 5-doxylstearic acid (5-DSA), the unpaired electron resides near the anchoring carboxylic acid group; in 16-DSA it is located near the end of the methylene chain (Figure 1b). Thus, information can be retrieved from the anchor positions in the protein as well as from the entry points into the fatty acid binding channel formed by the protein.

All our experiments imply the addition of up to seven fatty acids to one HSA molecule to occupy all binding sites in the protein. To avoid artifacts and allow quantitative interpretation of the data, it is advantageous to limit the amount of EPR-active, spin-labeled fatty acids to two per protein molecule. This avoids complications arising from multispin effects.[29] By simultaneously adding diamagnetic fatty acids, the degree of loading of the fatty acids on HSA can be varied (spin-diluted systems), as the diamagnetic fatty acids occupy the same binding sites as DSA without giving rise to an EPR signal. By adjusting the ratio of diamagnetic FA and DSA, two sites, statistically distributed among all sites, are then occupied by EPR-active molecules. Thus, an artifact-free distance distribution with a complete set of distances from all fatty acid binding sites can be obtained.

The diamagnetic fatty acid was prepared by reduction of the corresponding DSA to the EPR-inactive hydroxylamine (rDSA, see the Supporting Information, Section 0). This molecule is structurally closely related to the paramagnetic DSA. It exhibits the same binding affinities, as checked by continuous-wave (CW) EPR spectroscopy (see the Supporting Information, Section 2). Moreover, similar results are obtained with stearic acid as a spin-diluting diamagnetic species (see the Supporting Information, Section 0).

Before performing nanoscale distance measurements with DEER it is mandatory to validate that the spin-labeled FA, 5-DSA, and 16-DSA and their diamagnetic analogues (rDSA) are actually bound and taken up by HSA. This is shown in detail in the Supporting Information, Section 1. In summary, from the CW EPR signatures of bound and free DSA, we find that at HSA/fatty acid ratios of up to 1:6, more than 99.7% of all fatty acids are complexed by the protein. This proves an essentially quantitative uptake of both types of DSA and confirms that the fatty acid uptake is not disturbed by the spin labeling (see Figure S1 in the Supporting Information). Furthermore, the spin-labeled fatty acids display binding affinities similar to that of stearic acid. Thus, neither steric nor electronic variations severely affect the binding properties of the spin-labeled fatty acid. In view of the wide variety of saturated and unsaturated FAs that can be bound by HSA,[3,8] it is not surprising that the binding of our DSA molecules is comparable to that of other long-chain fatty acids.

It should be noted that the degree of rotational freedom of the nitroxide unit can also be estimated from the CW EPR spectra (see Section S1 in the Supporting Information).[30,31] 5-DSA is rotationally more restricted than 16-DSA, which has a roughly three times faster (i.e. lower) rotational correlation time. This trend is expected since in 5-DSA the nitroxide unit is placed in the tight hydrophobic binding channels near the anchoring point. The nitroxide group of 16-DSA, on the other hand, is located at the end of such a channel or even protrudes from the end of a channel and thus experiences a higher degree of rotational freedom, as reported in detailed CW EPR studies on the binding of spin-labeled fatty acids to different types of serum albumin.[32–34]

To retrieve the desired information about the functional structure of HSA, the HSA/fatty acid complexes were analyzed by DEER spectroscopy. The intramolecular part of the time-domain data and the extracted distance distributions are displayed in Figure 2. For 16-DSA, well-defined dipolar modulations are observed, which originate from narrow distance distributions with a dominating distance at 3.6 nm and two smaller contributions at 2.2 nm and 4.9 nm.

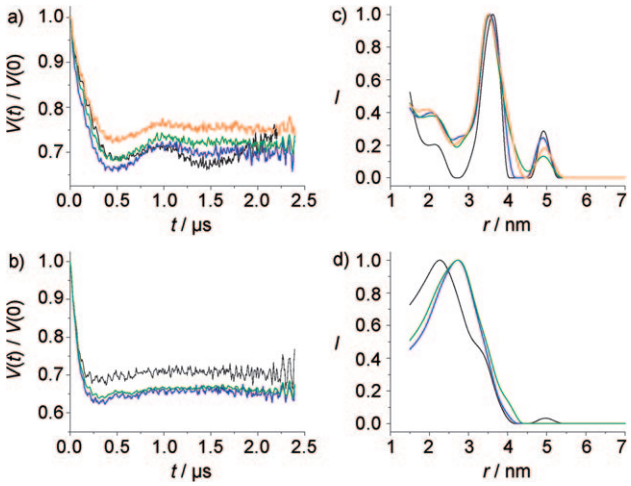

Figure 2. a,b) Intramolecular part of the DEER time-domain data and c,d) extracted distance distributions of spin-labeled stearic acids complexed in HSA with varying numbers of reduced and EPR-active fatty acids per protein molecule: black (0:2), blue (2:2), green (4:2), orange (6:2). The data from 16-DSA are shown on top (a,c), the data from 5-DSA on the bottom (b,d).

For 5-DSA, such a pronounced modulation is missing, which is indicative of a broad distance distribution. In fact, a single, broad distance peak is derived, which covers a range from 1.5 nm (the lower accessible limit with DEER) to 4 nm with a maximum around 2.5 nm.

Remarkably, neither the 5-DSA nor the 16-DSA distance distributions change significantly when the protein is loaded with different amounts of fatty acids. Thus, higher loading does not result in the generation of new distinct distance contributions. It rather results in a slight broadening of the already existing distance peaks. This indicates that although the fatty acids possess different binding affinities to different

protein binding sites, these differences are not very pronounced. Hence, the binding sites are not filled up consecutively. Rather, all binding sites are populated to a certain degree even at low HSA/fatty acid ratios. The high-affinity binding sites are only populated to a larger extent. Variations of the HSA/fatty acid ratio then only leads to changes of the relative populations. This explanation is in full accord with NMR studies with ^{13}C-labeled fatty acids. Only subtle changes of the NMR peak ratios were observed when the fatty acid ratio was increased.[11,12]

While the CW EPR data were collected at 298 K, DEER measurements were conducted at 50 K. For this purpose, the solution containing the HSA/FA complexes was shock-frozen to obtain a vitrified solution. Thus, we study a snapshot of the protein ensemble in solution at the glass transition temperature of around 170 K.[35,36] Note that 20 vol% of glycerol was added to the aqueous buffered solution to obtain such glassy samples. This is a standard procedure in pulse EPR analysis of proteins, and there is plenty of evidence that this small amount of nonaqueous solvent does not alter the structure of proteins but rather stabilizes the native structure in frozen solutions.[25,37] This assumption is further supported by CW EPR measurements that did not show changes of the fatty acid binding dependent on the addition of glycerol.

The different labeling positions in DSA allow for two different views on the functional structure of the protein. The position of the unpaired electron in 5-DSA is close to the carboxylic acid group of the fatty acid, which interacts with positively charged side groups of the protein. Hence, DEER delivers the characterization of the spatial distribution of the anchoring groups. With 16-DSA, on the other hand, the entry points into the fatty acid binding channels are probed. Information about the spatial distribution of these points is important to gain a better understanding of the uptake and release properties of the protein.

The 5-DSA (and thus the headgroup) distribution in solution is much broader than the distribution of 16-DSA (i.e. the entry points). Despite the increased flexibility of the doxyl moiety of 16-DSA, a very uniform distribution with a well-defined main distance is obtained. This suggests that the entry points of the fatty acid binding sites are distributed rather symmetrically on the surface of the protein molecule.

In Figure 3, the experimental distance distributions in solution are compared with distributions retrieved from the crystal structure. These distributions were calculated assuming a full occupation of all binding sites (for details see Sections 6 and 7 in the Supporting Information). Since the fatty acids in sites 1 and 7 are not resolved up to the C16 atom of the methylene chain, they were extrapolated to this position.

The experimental distribution of 5-DSA nicely fits that of the crystal structure. In both cases, broad distributions centered at a distance of about 3 nm are obtained that coincide remarkably well. Thus, for the anchoring points the highly asymmetric distribution of fatty acid binding sites observed by crystallographic data is supported by the DEER results. First of all, this shows that the distance distributions determined from DEER are reliable. Second, it suggests that the anchoring-point distribution in solution reflects the more

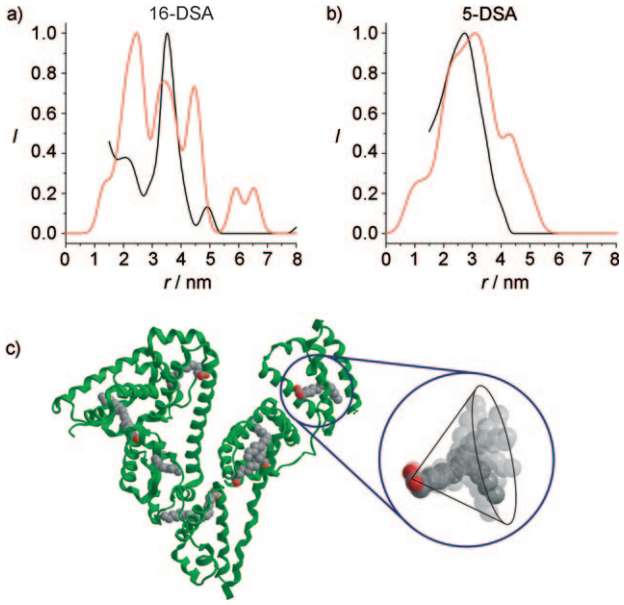

Figure 3. Comparison of experimental distance distributions obtained by DEER (black) with calculated distributions obtained from the crystal structure (red) for the a) C16 position and b) C5 position. The distribution in red is obtained assuming that all seven binding sites are occupied. Note that distances >6 nm cannot be accessed by DEER under the applied conditions. c) Illustration of the found flexibility of the FA binding-site entry points yielding a much more homogeneous and symmetric distribution over the protein surface than expected from the crystal structure. Only one FA binding site is shown for clarity.

rigid inner part of the protein, which is rather similar in the crystal and in solution.

In contrast, the distance distribution of the entry points (16-DSA) strongly deviates from that of the crystal structure. Note that these deviations are most pronounced in the distance range in which DEER is most reliable (2–5 nm). Distances greater 6 nm cannot be accessed by DEER under the applied conditions. The crystal structure distribution exhibits three major peaks at 2.5, 3.5, and 4.5 nm. While the peak positions roughly agree with the DEER data, the relative intensities deviate considerably. In the DEER data, the peak around 3.6 nm by far dominates all others, which remarkably simplifies the distance distribution as compared to that derived from the crystal structure. It suggests that the entry points are distributed much more symmetrically and homogeneously over the protein surface than expected from the crystal structure.

A rather homogeneous distance distribution of six binding sites suggests high symmetry. Considering the six sites to form an octahedron, one expects 12 vertex–vertex distances with edge length r and three diagonal distances with length $\sqrt{2}r$. Assigning the dominating distance of $r = 3.6$ nm to the edge length, a diagonal distance of 5.1 nm is expected, which is in reasonbable agreement with the observed distance of 4.9 nm. Indeed, an octahedron constitutes the most favorable distribution for the entry points of six fatty acid binding sites on a sphere, as they are then easily accessible from every side of the protein and assume the maximum distance with respect to each other.

However, one should not overinterpret these findings. In contrast to such a spherelike, symmetric distribution, the crystal structure (which agrees with the observed 5-DSA distribution) suggests that the main body of the protein is heart-shaped and rather resembles a flattened ellipsoid. This constraint can be accounted for by considering a uniaxially compressed octahedron to describe the entry-point distribution. In this picture, two opposing vertices approach each other. Assuming that in this special case their distance is reduced from $\sqrt{2}r$ to r, one expects five distances with length r, eight distances that are slightly decreased from r to $(\sqrt{3})/2r$, and two distances that remain at $\sqrt{2}r$. The slight (ca. 13%) discrepancy between the two dominating distances r and $(\sqrt{3})/2r$ may simply not be resolved in the experimentally found distance peak centered at $r = 3.6$ nm. which has a full-width at half maximum of approximately 0.7 nm.

Still, such a compressed octahedron can account for the entry points of only six binding sites and not for the seven sites determined by crystallography. However, this seventh site could be responsible for the short distances (ca. 2.2 nm) that cannot be accounted for by the octahedron model at all. It should be noted that the seventh site must be situated in the periphery of the distorted octahedron rather than in its center, as only a small fraction of small distances at around 2.2 nm are observed.

The interpretation derived from comparing DEER and crystal structure distributions is illustrated in Figure 3c. It is tempting to attribute the remarkably symmetric distribution of the entry points to the conformational flexibility of HSA. Furthermore, it may even mirror the optimization of the protein to allow for a fast and facilitated uptake and release of the fatty acids: this would be eased considerably by a homogeneous (average) distribution of FA entry points as well as large conformational flexibility (Figure 3c). A large flexibility on or close to the protein surface may also be entropically favored. The gain in entropy by small conformational variations is much larger at the surface than in the interior of the protein.

A rather symmetric distribution of binding sites on a distorted octahedron also is in accord with the fact that the DEER distance distributions effectively do not change when the fatty acid ratio is varied. Even if the binding sites were filled up consecutively, they would all give rise to similar distances owing to their symmetric distribution on the protein surface. To check the probability of such a symmetric distribution and our tentative interpretation of a distorted octahedron for the FA entry points, a molecular dynamics (MD) simulation on the crystal structure (1e7i) was performed for a period of 6.5 ns (for details, see the Supporting Information, Section 8). Though no substantial deviation from the overall crystal structure is observed, we find that the C16 position of the fatty acids is more strongly affected by small changes of secondary-structure elements than position C5. This is an additional hint that the anchoring region of the protein is rigid while the entry points have a substantial higher degree of freedom and flexibility. One should also note that a potentially lower crystallographic resolution of the conformationally disorderd chain ends of the fatty acid makes the crystallographic data for these sites less reliable. This might also contribute to the deviations between crystal structure and functional structure as derived from DEER.

HSA in solution is optimized for its function as a transporter of fatty acids. Our EPR study suggests that the uptake and release of fatty acids is facilitated by a homogeneous and rather symmetric distribution of the binding sites' entry points, which significantly differs from that expected from the crystal structure (see Figure 3). This leads to a picture of the functional protein structure in which the inner part of the protein is more rigid and asymmetric, while the surface of the protein shows much larger structural flexibility.

As there is increasing evidence (e.g. from NMR studies[13–15]) that the conformational flexibility of proteins often is a prerequisite for their function, we expect that the functional solution structures of proteins may differ, even significantly, from the crystal structure also for other transport systems. Using the selectivity of the spin-probing approach, only signals from the transported species, fatty acids in the present case, are obtained which are directly related to the protein's function of interest. Beyond the importance of this specific protein, our spectroscopic approach based on well-established spin labels and nanoscale distance measurements could generally be employed to characterize transport proteins. It results in a tremendous simplification of structural determination, while allowing full characterization of the functionality of the protein in native environment.

Experimental Section

Materials: Non-denaturated human serum albumin (HSA, >95%, Calbiochem), 5- and 16-doxylstearic acid (DSA, Aldrich), and 87 wt% glycerol (Fluka) were used as received. The DSA derivatives were partly reduced to EPR-inactive hydroxylamines (rDSA) by addition of phenylhydrazine (97%, Aldrich) as described in detail in the Supporting Information, Section 0.

Sample preparation: HSA/fatty acid complexes in 100 mM phosphate buffer solution (pH 7.4) were obtained by mixing 0.2 M phosphate buffer solutions (pH 6.4) with and without 2 mM HSA and with 6.7 mM solutions of DSA and rDSA in 0.1 M KOH in appropriate ratios. The combined concentration of DSA and rDSA was kept constant at 2 mM with varying ratios from 2:0 to 2:6 per protein molecule. This method provides for isolated spin pairs in combination with a varying total occupation of the fatty acid binding sites. For DEER measurements, 20 vol% glycerol was added to the final solutions to prevent crystallization upon freezing. The solutions were filled into 3 mm o.d. quartz tubes and shock-frozen in $N_2(l)$ cooled isopentane.

EPR measurements: Details on the CW EPR and DEER measurements as well as data evaluation are presented at the beginning of the Supporting Information.

Received: June 8, 2010
Published online: September 30, 2010

Keywords: albumin · EPR spectroscopy · fatty acids · nanostructures · protein structures

[1] T. Peters, *All About Albumin: Biochemistry, Genetics and Medical Applications*, Academic Press, San Diego, **1995**.
[2] D. C. Carter, J. X. Ho, *Adv. Protein Chem.* **1994**, *45*, 153–203.
[3] A. A. Spector, *J. Lipid Res.* **1975**, *16*, 165–179.

[4] J. A. Hamilton, D. P. Cistola, J. D. Morrisett, J. T. Sparrow, D. M. Small, *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 3718 – 3722.
[5] X. M. He, D. C. Carter, *Nature* **1992**, *358*, 209 – 215.
[6] S. Curry, H. Mandelkow, P. Brick, N. P. Franks, *Nat. Struct. Biol.* **1998**, *5*, 827 – 835.
[7] S. Curry, P. Brick, N. P. Franks, *Biochim. Biophys. Acta Mol. Cell Biol. Lipids* **1999**, *1441*, 131 – 140.
[8] A. A. Bhattacharya, T. Grüne, S. Curry, *J. Mol. Biol.* **2000**, *303*, 721 – 732.
[9] M. Fasano, S. Curry, E. Terreno, M. Galliano, G. Fanali, P. Narciso, S. Notari, P. Ascenzi, *IUBMB Life* **2005**, *57*, 787 – 796.
[10] J. A. Hamilton, S. Era, S. P. Bhamidipati, R. G. Reed, *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 2051 – 2054.
[11] J. R. Simard, P. A. Zunszain, C. E. Ha, J. S. Yang, N. V. Bhagavan, I. Petitpas, S. Curry, J. A. Hamilton, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 17958 – 17963.
[12] J. R. Simard, P. A. Zunszain, J. A. Hamilton, S. Curry, *J. Mol. Biol.* **2006**, *361*, 336 – 351.
[13] K. Henzler-Wildman, D. Kern, *Nature* **2007**, *450*, 964 – 972.
[14] O. F. Lange, N.-A. Lakomek, C. Farès, G. F. Schröder, K. F. A. Walter, S. Becker, J. Meiler, H. Grubmüller, C. Griesinger, B. L. de Groot, *Science* **2008**, *320*, 1471 – 1475.
[15] L. Salmon, G. Bouvignies, P. Markwick, N. Lakomek, S. Showalter, D.-W. Li, K. Walter, C. Griesinger, R. Brüschweiler, M. Blackledge, *Angew. Chem.* **2009**, *121*, 4218 – 4221; *Angew. Chem. Int. Ed.* **2009**, *48*, 4154 – 4157.
[16] F. Karush, *J. Am. Chem. Soc.* **1950**, *72*, 2705 – 2713.
[17] F. Karush, *J. Am. Chem. Soc.* **1954**, *76*, 5536 – 5542.
[18] N. Laiken, G. Nemethy, *Biochemistry* **1971**, *10*, 2101 – 2106.
[19] A. D. Milov, K. M. Salikhov, M. D. Shirov, *Fiz. Tverd. Tela* **1981**, *23*, 975 – 982.
[20] M. Pannier, S. Veit, A. Godt, G. Jeschke, H. W. Spiess, *J. Magn. Reson.* **2000**, *142*, 331 – 340.
[21] R. E. Martin, M. Pannier, F. Diederich, V. Gramlich, M. Hubrich, H. W. Spiess, *Angew. Chem.* **1998**, *110*, 2994 – 2998; *Angew. Chem. Int. Ed.* **1998**, *37*, 2834 – 2837.
[22] G. Jeschke, *Macromol. Rapid Commun.* **2002**, *23*, 227 – 246.
[23] D. Hinderberger, O. Schmelz, M. Rehahn, G. Jeschke, *Angew. Chem.* **2004**, *116*, 4716 – 4721; *Angew. Chem. Int. Ed.* **2004**, *43*, 4616 – 4621.
[24] O. Schiemann, T. F. Prisner, *Q. Rev. Biophys.* **2007**, *40*, 1 – 53.
[25] C. Dockter, A. Volkov, C. Bauer, Y. Polyhach, Z. Joly-Lopez, G. Jeschke, H. Paulsen, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 18485 – 18490.
[26] D. Hilger, H. Jung, E. Padan, C. Wegener, K. P. Vogel, H.-J. Steinhoff, G. Jeschke, *Biophys. J.* **2005**, *89*, 1328 – 1338.
[27] O. Schiemann, N. Piton, J. Plackmeyer, B. E. Bode, T. F. Prisner, J. W. Engels, *Nat. Protoc.* **2007**, *2*, 904 – 923.
[28] O. Schiemann, P. Cekan, D. Margraf, T. F. Prisner, S. T. Sigurdsson, *Angew. Chem.* **2009**, *121*, 3342 – 3345; *Angew. Chem. Int. Ed.* **2009**, *48*, 3292 – 3295.
[29] J. D. Morrisett, H. J. Pownall, A. M. Gotto, *J. Biol. Chem.* **1975**, *250*, 2487 – 6591.
[30] S. Stoll, A. Schweiger, *J. Magn. Reson.* **2006**, *178*, 42 – 55.
[31] D. J. Schneider, J. H. Freed in *Biological Magnetic Resonance*, *Vol. 8* (Eds.: L. J. Berliner, J. Reuben), Plenum, New York, **1989**.
[32] J. D. Morrisett, H. J. Pownall, A. M. Gotto, *J. Biol. Chem.* **1975**, *250*, 2487 – 2494.
[33] S. J. Rehfeld, D. J. Eatough, W. Z. Plachy, *J. Lipid Res.* **1978**, *19*, 841 – 849.
[34] V. A. Livshits, D. Marsh, *Biochim. Biophys. Acta Biomembr.* **2000**, *1466*, 350 – 360.
[35] K. Kawai, T. Suzuki, M. Oguni, *Biophys. J.* **2006**, *90*, 3732 – 3738.
[36] C. Inoue, M. Ishikawa, *J. Food Sci.* **2000**, *65*, 1187 – 1193.
[37] R. V. Rariy, A. M. Klibanov, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 13520 – 13523.

