# Proteins, Lipids, and Water in the Gas Phase

David van der Spoel,* Erik G. Marklund, Daniel S. D. Larsson, Carl Caleman

Evidence from mass-spectrometry experiments and molecular dynamics simulations suggests that it is possible to transfer proteins, or in general biomolecular aggregates, from solution to the gas-phase without grave impact on the structure. If correct, this allows interpretation of such experiments as a probe of physiological behavior. Here, we survey recent experimental results from mass spectrometry and ion-mobility spectroscopy and combine this with observations based on molecular dynamics simulation, in order to give a comprehensive overview of the state of the art in gas-phase studies. We introduce a new concept in protein structure analysis by determining the fraction of the theoretical possible numbers of hydrogen bonds that are formed in solution and in the gas-phase. In solution on average 43% of the hydrogen bonds is realized, while in vacuo this fraction increases to 56%. The hydrogen bonds stabilizing the secondary structure ($\alpha$-helices, $\beta$-sheets) are maintained to a large degree, with additional hydrogen bonds occurring when side chains make new hydrogen bonds to rest of the protein rather than to solvent. This indicates that proteins that are transported to the gas phase in a native-like manner in many cases will be kinetically trapped in near-physiological structures. Simulation results for lipid- and detergent-aggregates and lipid-coated (membrane) proteins in the gas phase are discussed, which in general point to the conclusion that encapsulating proteins in "something" aids in the conservation of native-like structure. Isolated solvated micelles of cetyl-tetraammonium bromide quickly turn into reverse micelles whereas dodecyl phosphocholine micelles undergo much slower conversions, and do not quite reach a reverse micelle conformation within 100 ns.
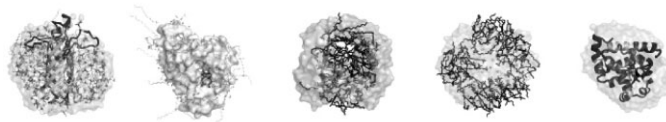
## Introduction

Mass spectrometry (MS) experiments routinely probe biomolecules in the gas-phase in order to—indirectly—gain information about biomolecules under physiological conditions. Experimental probing techniques include regular MS which produces a mass over charge distribution for a given sample[1,2] and ion mobility spectroscopy (IMS), yielding information about the size and shape of biomolecules in the gas phase as a function of charge.[3,4] Fragmentation techniques, like collision-activation detected (CAD) or electron-capture dissociation (ECD), may be used to obtain mass over charge spectra for fragments of molecules,[5] and such techiques are invaluable in proteomics research. A number of reviews cover the basic experimental methodologies and we refer the reader to these papers for more background information.[2,6–11] Here, we focus on biomolecular aggregates in the gas phase—proteins with solvent, proteins with bound lipids or in micelles or solvated micelles—that are prepared to be under near-physiological conditions. A large body of evidence suggests that it is possible to detect intact

D. van der Spoel, E. G. Marklund, D. S. D. Larsson, C. Caleman
Department of Cell and Molecular Biology, Uppsala University, Husargatan 3, Box 596, SE-751 24 Uppsala, Sweden
Fax: 46 18 511755; E-mail: spoel@xray.bmc.uu.se
C. Caleman
Center for Free-Electron Laser Science, DESY, Notkestraße 85, DE-22607 Hamburg, Germany

aggregates largely devoid of solvent. For example, experiments have shown that virus particles that are transferred into the gas phase and collected afterward are still infectious.[12] Similarly, large molecular complexes, like the familiar chaperonin GroEL consisting of 14 proteins, have been detected in the gas phase using time of flight MS,[13,14] and even seemingly fragile structures consisting of a ring of 14 identical proteins.[15] Recent MS work demonstrated that it is possible to dissect virus assembly paths using time of flight MS: Uetrecht et al. found that in particular pentameric assemblies are relatively stable.[16,17] Another interesting example is the study of processes that lead to aggregation of peptides into amyloid plaques.[18–20]

The most gentle method to bring molecules to the gas phase is by electrospray ionization (ESI),[6,21] a technique for which the Nobel prize was awarded in 2002. It works typically by slightly lowering the pH in combination with a buffer solution like ammonium-acetate. By carefully controlling the solvent conditions the charge state distribution of the solute can be tuned and reproducible experiments can be performed. Solvent droplets containing elute ions are then pressed out of a syringe that is charged positively, which will charge up the protein somewhat further, and steer the droplets toward a negative counterelectrode (note that the voltage can be reversed as well). A schematic picture is given Figure 1. A crucial feature is the heating stage: typically the proteins are heated in order to dehydrate them on the way to detection, this in order to increase the resolution of resulting mass over charge spectra. A detailed overview of the instrumentation was given by Jarrold.[22]

Very recently, a new technique based on a µm-wide liquid jet was introduced, from which a solvated (macromolecular) ion can be extracted by a laser pulse.[23,24] This technique is also very gentle on the biomolecule and has the advantage that it is possible to control the amount of remaining solvent on the biomolecule. It can be expected that this technique will grow in importance in the near future as more and more researchers realize the important of solvation for the interpretation of MS experiments.[25]

**David van der Spoel** has an undergraduate degree in physics and a Ph.D. in computational chemistry (1996) from the University of Groningen, The Netherlands. After his Ph.D. he joined Uppsala University where he now is a Professor of Biology with specialization in computational molecular biophysics. Research in his group focuses on protein structure in the gas phase, software development for the GROMACS molecular dynamics (MD) suite and model development for MD simulation.

**Carl Caleman** holds a degree in engineering physics, and a PhD in molecular biophysics (2007), both from Uppsala University. He spent one year as a visiting scholar at UC Berkeley. After a post doc at the Technical University of Munich, he is now employed at the Center for Free Electron Laser, in Hamburg, exploring the possibilities of Xray Free Electron Lasers.

**Daniel Larsson** received his M. Sc. in Molecular Biotechnology at Uppsala University, Sweden in 2007. Currently he is a Ph. D. candidate at the Department of Cell and Molecular Biology, Uppsala University working with computational biology in the group of David van der Spoel. His current interests are in the structure of protein/lipid systems in the gas-phase and the dynamic behavior of virus particles.

**Erik Marklund** is a Ph. D. candiate with an M. Sc. degree in Molecular biotechnology from Uppsala University (Sweden). He is currently a member of David van der Spoels group at the Department of Cell and Molecular Biology, Computational and Systems Biology, at Uppsala University. His work currently revolves around gas-phase proteins, proton transfer reactions and hydrogen bonding kinetics from a computational point of view.
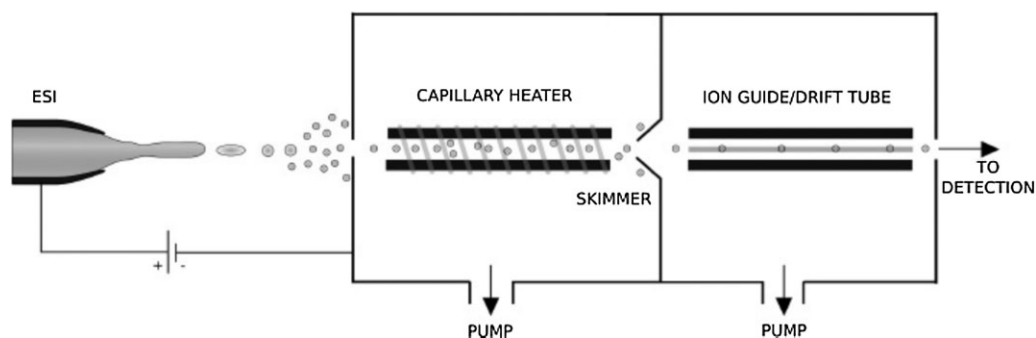
Ion-mobility spectroscopy (IMS) probes the shape and size of proteins by measuring the time of flight through a tube filled with an inert gas, that provides friction. This procedure yields low-resolution structural information, not



■ *Figure 1.* Schematic drawing of an ESI injection system.

unlike small-angle X-ray scattering. By combining IMS with modeling techniques and by varying conditions interesting results can be obtained, for instance relating charge state to molecule size.[2–4,22,26] Although IMS yields very useful results, the link to molecular structure can only be made indirectly through computational methods. It has been suggested that high-resolution structural information may be obtained using X-ray scattering from free-electron lasers.[27,28] Although proof of principle experiments have demonstrated that it is possible to obtain structural information from solid targets at low resolution (60 nm),[29] issues like sample dynamics,[30] and irreproducible scattering from bound solvent molecules[31] will in practice probably limit the resolution of such experiments. Radiation damage to samples at atomic wavelengths is not completely understood either[27,32–36] and this may preclude obtaining high resolution structural information as well. Nevertheless, this technique may provide a powerful complement to IMS in the near future.

An often-heard question is whether the gas-phase structure of a protein, or in general a biomolecular aggregate is similar to the solution structure.[37] The success stories that are referenced above[12–14,16–20] strongly suggest that indeed protein structures in the gas phase must be very similar to the solution structures, since otherwise it is difficult to imagine that non-covalent complexes would stick together. However, this obviously is not a golden rule: ab initio folding simulations of a decapeptide in solution and in vacuo ended up in very different structure, both of which are compact in terms of radius of gyration,[38,39] highlighting that the addition of solvent can modify the energy landscape in a non-trivial manner. It should be noted here that in the simulations in solution the global Gibbs energy minimum corresponded to the experimental structure as determined by NMR spectroscopy,[40] which strongly suggest that the OPLS/AA force field[41] used in that study[38] describes the intramolecular interactions with sufficient accuracy. It is not possible to generalize the differences between solvated and vacuum protein structures, however. For the 20-residue Trp-cage mini-protein reversible folding have been detected using fluorescence decay experiments in an ion trap.[42,43] A melting point of around 420 K, far above the solution melting point of 320 K was determined for this protein.[44] MD simulations of the protein (including attached dye) show however that even the "unfolded" conformations, where the dye molecule can enhance the fluorescence quenching through electron transfer, are compact as well.[44] Further work on the Trp-cage shows, that also mutants are stable in the gas-phase although all mutants have slightly different (3–4 Å root mean square deviation between all atoms) lowest energy structures.[45,46] Difference in RMSD between Trp-Cage in 2+ and 3+ states are similar in magnitude.[44] This example does in no way prove

that refolding to a native like state in the gas phase is possible in general. However, as Breuker and McLafferty have argued, conditions in the ESI process need to be optimized for specific experiments.[47]

In what follows we discuss first the effect of water on protein structure in the gas phase, then the effect of initial temperature. Then, micelles and lipid-coated proteins are discussed followed by conclusion.

## Water-Coated Proteins

Mass spectrometry (MS) experimental setups typically contain a drying stage to make water and or other solvents evaporate before detection (Figure 1). The rationale for this is that water deteriorates the spectrum. Obviously, water is the native environment for most proteins, but it is also well established, for instance based on protein X-ray crystallography, that there are structural water molecules that are very reproducible to the point of being "part of the protein.[48–51]" As such they often contribute to flexibility of proteins.[52–54] In special cases, like for bovine pancreatic trypsin inhibitor in the gas phase a single water molecule has been shown to rigidify the protein.[55] An old crystallography study of subtiliysin in acetonitrile has shown that most water molecules in a protein crystal can be replaced by other solvents, however around 100 water molecules could not be displaced.[56] It has also been shown that most water molecules can be removed from lysozyme crystals without significantly altering the structure.[57] All these results together demonstrate that water often has an impact on protein conformation and flexibility, but they also clearly show that these effects are protein specific.

Water molecules are often involved in enzymatic reactions, both by giving structural "support" [through hydrogen bond (HB) networks] and by partaking in the reaction itself. Hence, it may seem very crude to base conclusions about the physiological properties of proteins on structures from which all waters have been removed. Early IMS experiments on apomyoglobin have shown that it is possible to determine whether or not the protein is compact in size or not,[58,59] where compactness is often taken to suggest foldedness. The authors also attempted to rehydrate the dried proteins by letting in water vapor in the drift tube (Figure 1) and it was found that mildy unfolded proteins pick up fewer water molecules than neatly folded proteins. Experimental water binding studies also indicate that water tends to bind in clusters, i.e., binding one water molecule is more difficult than binding multiple.[60,61] Detailed studies using MS and density functional theory calculations on binding of water to individual amino acids (Phe, Trp, Tyr) in the gas-phase reveal that 2–5 water molecules can be bound with similar affinity.[62] The observations about what happens upon dehydration of a

protein, similar to the drying process following ESI, are completely consistent with the picture emerging from molecular dynamics simulations.[63–68] Some general observations can be made:

(i) Due to absence of the hydrophobic effect in vacuo, electrostatic interactions dominate the intramolecular interactions, leading to more HBs being formed.
(ii) The location of charged sites is inherently different from that solution, it is governed by the gas-phase basicity of amino-acids[69] and the electrostatic potential due to the environment.[65,66,70]
(iii) Solvation by water on the protein surface is to some extent reproducible and follows the patterns of hydrophilic amino-acids.[68]
(iv) Water prefers to cluster on the surface due to strong intermolecular interactions between water. This explains the IMS results mentioned above, that folded proteins bind more water in the gas-phase than unfolded proteins.[58,59,68]

In summary, MD simulations predict limited conformational changes in proteins upon dehydration. The predominant effect is that hydrophilic side-chains fold inwards to form HBs with backbone amide- and carbonyl-groups. Protein function requires, in most cases, protein stability (to some extent) and structural specificity. Stability against unfolding follows predominantly from the packing and burial of hydrophobic groups while the specificity stems from the hydrogen bonding network. It hence seems logical that the number of HBs in proteins (at least those involved in secondary structure formation) is relatively close to the maximum possible number. Table 1 lists the average number of HBs per amino acid based on an analysis of protein databank structures.[71] By simply summing the number of possible HB donors nD and acceptors nA for a given sequence we can determine the

Table 1. Number of HB donors and acceptors per amino acid (AA) based on the statistical analysis of McDonald and Thornton.[71] Most oxygens can accept two HBs, therefore these are counted as two. For the AA that are charged under physiological conditions (Arg, Asp, Glu, Lys) we also give the corresponding neutral variant.

| AA | Donor | Acceptor |
|---|---|---|
| ALA | 0.95 | 1.16 |
| ARG | 5.40 | 1.20 |
| ASN | 2.24 | 6.01 |
| ASP | 0.89 | 1.24 |
| CYS | 1.36 | 1.26 |
| GLN | 2.27 | 2.41 |
| GLU | 0.93 | 4.43 |
| GLY | 0.96 | 1.18 |
| HIS | 2.59 | 1.94 |
| ILE | 0.93 | 1.12 |
| LEU | 0.98 | 1.20 |
| LYS | 3.59 | 1.12 |
| MET | 1.02 | 1.12 |
| PHE | 0.92 | 1.08 |
| PRO | 0.00 | 1.13 |
| SER | 2.31 | 2.04 |
| THR | 2.29 | 2.01 |
| TRP | 1.83 | 1.21 |
| TYR | 2.00 | 1.16 |
| VAL | 0.93 | 1.11 |

theoretical maximum number of HB $HB_{max}$ for this sequence to be min(nD,nA). Since every donor needs an acceptor to make an HB, $HB_{max}$ yields the maximum number of HB in the absence of solvent. Table 2 shows the

Table 2. Fraction of HBs in proteins studied, in bulk solvent and in vacuo. HBs were defined by geometry: the hydrogen-donor–acceptor angle should be less than 30°, and the donor–acceptor distance less than 0.35 nm. Numbers are averages over the final ns from simulations described in ref.[65,72,87]

| Protein | AA | Don | Acc | HBmax | HBmax/AA | HBsol | HBsol % | HBvac | HBvac % |
|---|---|---|---|---|---|---|---|---|---|
| Ctf | 68 | 99.5 | 115.5 | 99.5 | 1.46 | 39.0 | 39.2 | 57.0 | 57.3 |
| Insulin | 51 | 79.1 | 96.8 | 79.1 | 1.55 | 12.0 | 15.2 | 38.0 | 48.1 |
| Lysozyme | 129 | 243.5 | 244.9 | 243.5 | 1.89 | 94.0 | 38.6 | 122.0 | 50.1 |
| Myoglobin | 152 | 247.8 | 254.5 | 247.8 | 1.63 | 122.0 | 49.2 | 161.0 | 65.0 |
| OmpA | 171 | 274.4 | 303.2 | 274.4 | 1.60 | 142.0 | 51.8 | 154.0 | 56.1 |
| Trp-cage | 20 | 30.9 | 32.0 | 30.9 | 1.55 | 10.0 | 32.3 | 14.0 | 45.3 |
| Ubiquitin | 76 | 132.0 | 134.3 | 132.0 | 1.74 | 56.0 | 42.4 | 71.0 | 53.8 |
| Average | 95.3 | 158.2 | 168.8 | 158.2 | 1.66 | 67.9 | 42.9 | 88.1 | 55.7 |

fraction of HB that are realized in solution and in vacuum, based on simulations. The average fraction increases from 43 to 56% after a few ns upon dehydration indicating that all proteins rapidly reach a local energy minimum. In other words, the proteins are kinetically trapped in a local minimum, the shape of which is determined by residue composition and availability of solvating molecules.

Structures for myoglobin, lysozyme, and ubiquitin are plotted in Figure 2, based on MD simulations in water and in vacuo covered by thin water layers (simulations taken from ref. [68,72]). The structures are colored in order to highlight which part of the structures in vacuo deviate most from the solution structure. The dehydrated parts deviate more than the somewhat solvated parts. Also interesting to note is that

α-helix H in myoglobin becomes more perfect (straight) in vacuo, a feature of α-helices that has been noted previously by Jarrold.[61] In lysozyme the α-helical region seems to be afftected more by the vacuum than the β-sheet, which could be correlated with the fact that this part of the protein is intrinsically less stable in solution as well.[73]

## Temperature Effects on Protein Structure

The structure as well as the dynamics of a protein in vacuum directly depends on the temperature of the system.[68] The charge state is affected by environmental factors as well,[37] and hence protein structure will be
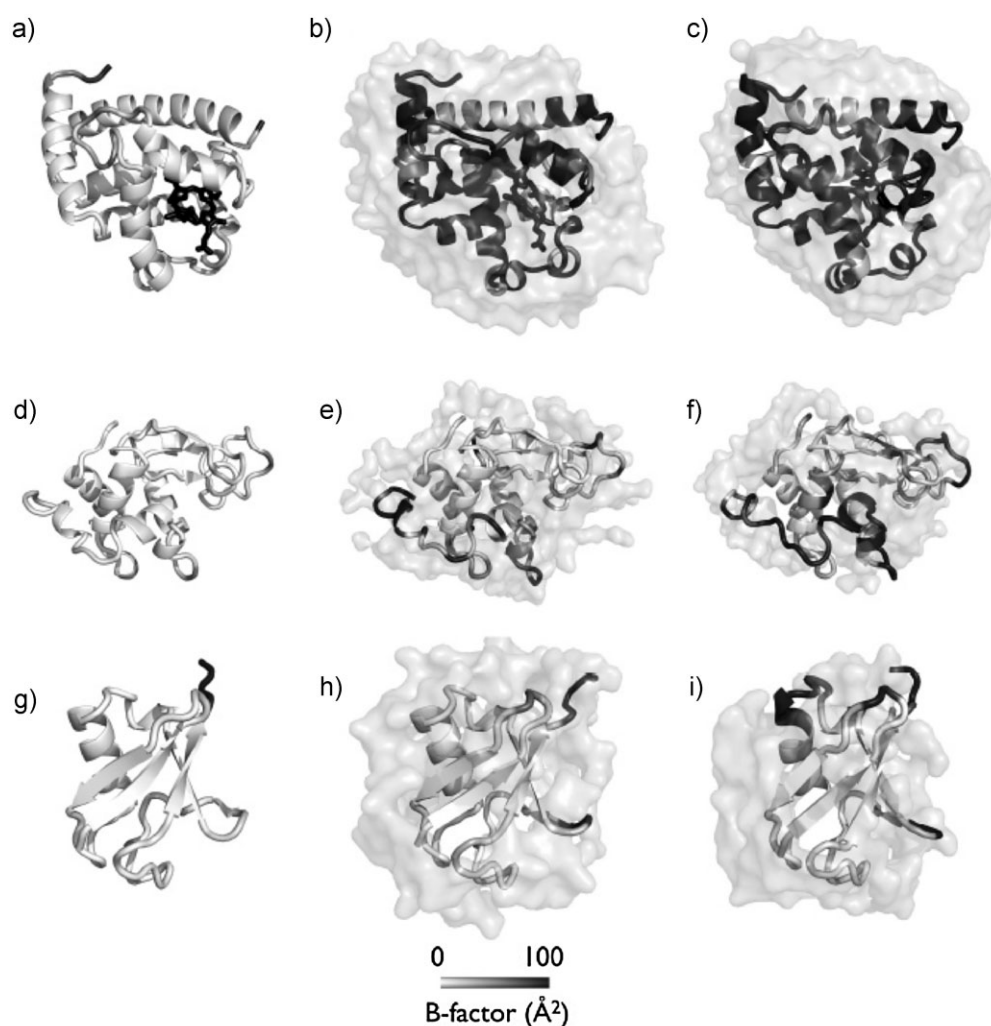


*Figure 2.* (a–c) Myoglobin, (a) average solvent structure, (b) starting structure of vacuum simulation in 0.45 nm water, (c) after 20 ns.[72] (d–f) Lysozyme, initially immersed in a 0.3 nm water layer at 300 K. (d) Solution structure, (e) start of vacuum simulation and (f) the end of the 15 ns vacuum simulation.[68] (g–i) Ubiquitin, initially immersed in a 0.3 nm water layer at 300 K. (g) Solution structure, (h) start of vacuum simulation and (i), the end of the 15 ns vacuum simulation. The proteins are colored according to the temperature factor (B-factor) of the $C_\alpha$ atoms (going from white = 0 Å² to dark = 100 Å² or higher). The temperature factors were calculated as $8\pi^2/3$ times the deviation from the average solution structure in order to be comparable with crystallographic temperature factors. For the solution structure the mean-square-fluctuation (converted to B-factors) during the simulation is shown.

influenced by temperature in multiple ways. Systems with lower temperature are naturally more stable and less flexible than those at a higher temperature. For small water droplets carrying proteins or other biomolecules in vacuum, the evaporation of the carrying liquid cools down the system and hence contributes to conservation of the protein structure.[68] Water evaporation is dependent on the initial temperature of the system: a system with higher temperature evaporates faster and therefore cools down faster. As the system cools down evaporation slows down rapidly, as is evident from the decrease of vapor pressure of water with decreasing temperature. Microsecond long simulations of small water droplets have demonstrated that evaporation cools the system to around 225 K, at least for starting temperatures between 250 and 350 K.[74] Once the system reaches these low temperatures the evaporation is almost negligible. Similar behavior have been seen in water droplets containing proteins[68,75] and inorganic ions.[76] Drift tube experiments, backed up by MD simulations, show that the upper temperature limit for rehydration of helices is lower than for other secondary structure types,[60] and corroborates that 225 K is a reasonable temperature limit for hydration of peptides and proteins.[61] For sparingly hydrated globular proteins the structure were found to be rather stable in terms of RMSD, independent of the initial temperature of the system — provided that (i) there are enough surrounding waters to cool down the system by evaporation, and (ii) some water molecules remain bound to the protein to provide hydrogen bonding partners as described above. Heating of completely dry proteins will lead to (partial) unfolding as evidenced by a large body of IMS data.

## Micelles

In order to bring membrane proteins into the gas-phase it is necessary to solubilize them by taking them out of the membrane environment. Micelles containing phospholipids or detergents have proven to be a good carrier for membrane proteins.[77,78] Sharon and Robinson have shown that both regular micelles and reverse micelles can be detected in time of flight MS.[79] What was not clear from their work is how reverse micelles are formed, when starting from regular micelles. Simulations of hydrated cetyl-trimethylammonium bromide (CTAB) micelles in the gas phase clearly demonstrate that such micelles rapidly invert upon dehydration (Figure 3), something that was observed for both 0.45 and 0.6 nm layers of water.[72] As expected the transition takes more time when more water is involved, and remarkably little water evaporates during the process, due to the large amount of charges involved.

In order to test whether this inversion process is general or unique to CTAB micelles we performed similar simulations for zwitterionic dodecyl phosphocholine (DPC) micelles, which we have studied previously in solution[80] and which have been shown to form micelles in simulations on the nanosecond time scale[81] (but note that this time scale is inherently concentration dependent). Snapshots from a simulation of DPC in solution were used as starting points for these simulations where different amounts of water were maintained (0.3, 0.45, and 0.6 nm, respectively). These conformations were simulated in vacuo for 100 ns each, using the GROMACS software[82] and other parameters as described by Wang et al.[72] Figure 4 and 5 show that the amount of change depends on the lipid as well as on the amount of water. The simulations employing the thinnest water layer of 0.3 nm show clustering of water on the surface: after 100 ns the water has separated into three smaller patches. In all clusters we find that the tails move to the surface, but no complete inversion as in the CTAB micelle (Figure 3) is detected. This could be a matter of simulation length, however Sharon and Robinson have demonstrated that solvated CTAB micelles can be detected by MS as well[79] which strongly suggests that micellar structures can be kinetically trapped in the gas-phase into aggregates that do not correspond to the global free energy minimum. In summary, it is possible to obtain both normal and reverse micelles, depending on lipid type, initial conditions, and time spent in the instrument prior to detection.
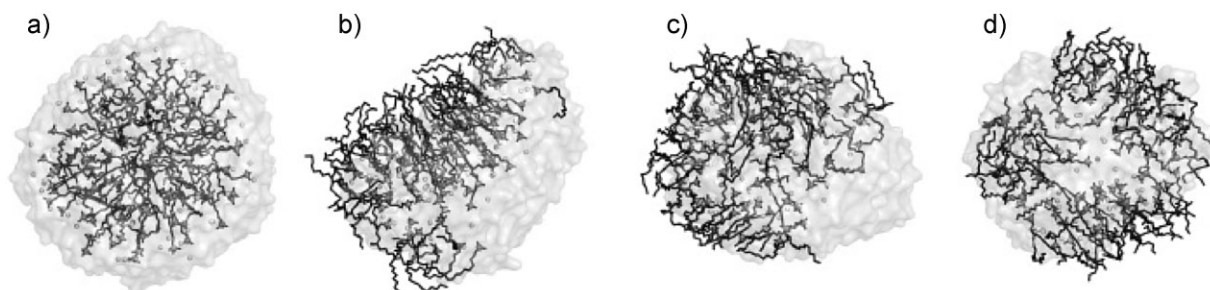


*Figure 3.* Time laps of the inversion process of a CTAB micelle covered by a 0.6 nm water layer. (a) The initial normal-phase micelle consisted of 151 CTAB in a drop containing 1 643 water molecules. (b,c) Two intermediate stages during the transition at (b) $t = 3.5$ ns and (c) $t = 4$ ns. (d) The reverse phase micelles at $t = 10$ ns.
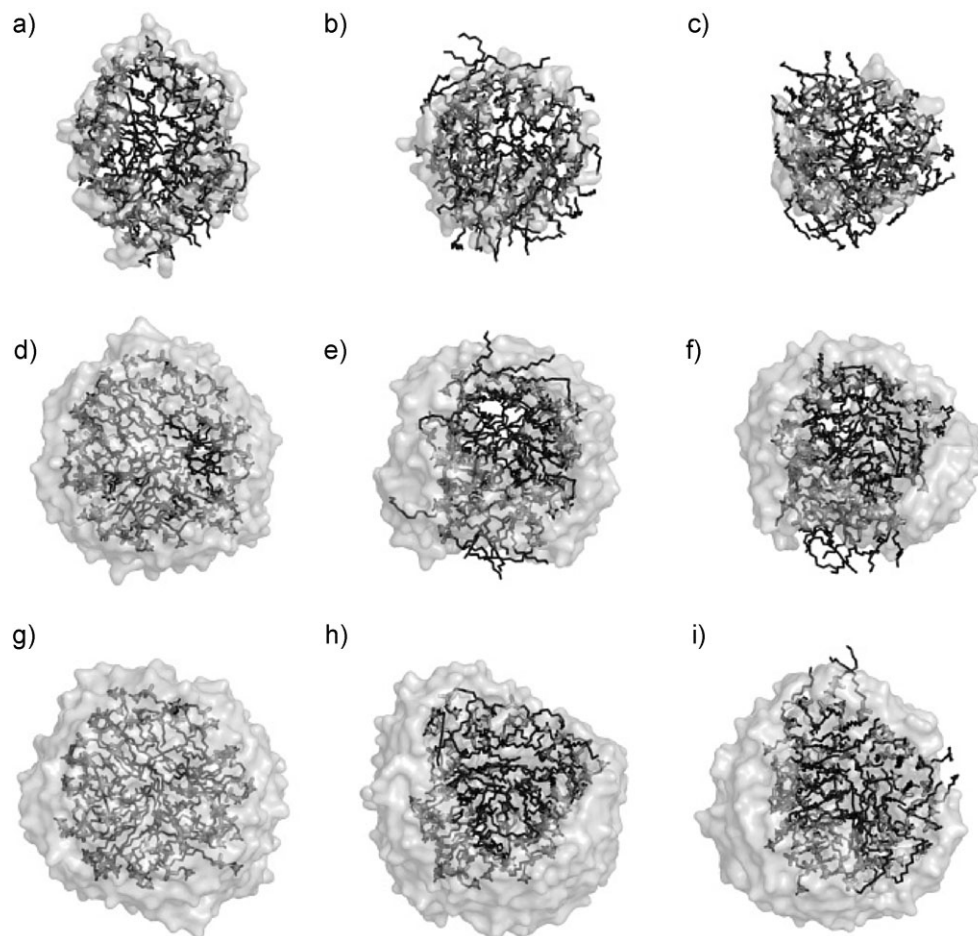
Materials
Views
www.MaterialsViews.com

*Macromol. Biosci.* **2011**, *11*, 50–59
© 2011 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

Macromolecular
Journals

55

*Figure 4.* Snapshots along the trajectory of a DPC micelle (54 lipids) (a–c) covered by a 0.3 nm water layer at (a) $t = 0$, (b) $t = 5$ ns, and (c) $t = 100$ ns; (d–f) covered by a 0.45 nm water layer at (d) $t = 0$, (e) $t = 5$ ns, and (f) $t = 50$ ns, water layer; and (g–i) covered by a 0.6 nm water layer at (g) $t = 0$, (h) $t = 5$ ns, and (i) $t = 50$ ns. Trajectories were generated according to the protocol described by Wang et al.[72]

## Lipid-Coated Proteins

The possibility to carry membrane proteins into the gas-phase in micelles, that is a near physiological environment, yields the prospect of studying drug-binding in a stoichiometric manner.[77] This is crucially important for drug-design as a very large fraction of known drugs binds to membrane proteins like G-protein coupled receptors for which remarkably little structural information is available. Only recently have structures been determined,[83,84] following the making of complex constructs, which induced artifacts as demonstrated by extremely long MD simulations[85] - this just to indicate the complexity involved with direct structural studies of membrane proteins. In many cases information on whether a drug candidate binds to a protein at all is very useful in itself, and hence the demonstration of a micelle–protein complex in the gas-phase by Ilag et al. formed a great breakthrough,[77] even

though other factors such as drug efficacy need to be considered as well.[86] Friemann et al. have studied the structure of the 171 aminoacid N-terminal domain of the outer membrane protein A (OmpA171) in *Escherichia coli* in a DPC micelle,[87] the complex structure of which was generated by MD simulations of lipids condensing onto the protein.[88,89] It was found (Figure 6) that the protein structure hardly changes upon dehydration.[87] In particular the membrane-bound beta-barrel of the protein is maintained very well. This emphasizes the possibility of detailed studies (e.g., using mutations) of membrane–protein complexes with drug molecules.[77] A recent paper by Hebling et al. describes the incorporation of membrane proteins in lipid nanodisks.[90] This may provide an alternative route to membrane protein studies in the gas-phase. Figure 6 shows, based on simulations by Friemann et al.,[87] how lipid-coated proteins may look like in vacuum under native-like conditions. The deviation from solution
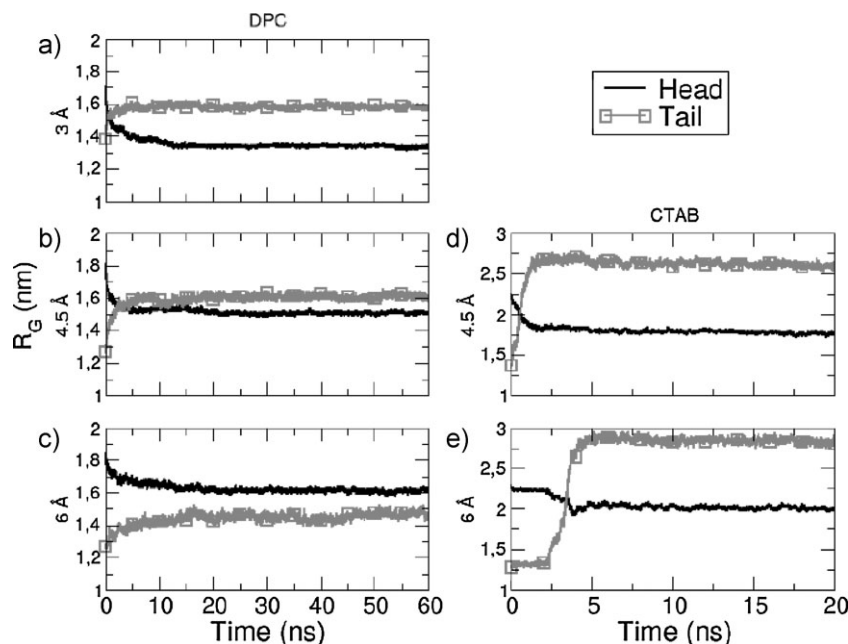
**Figure 5.** The tail groups of the micelles move outwards relative to the head groups when the micelle is exposed to vacuum. DPC micelles with a (a) 3 Å, (b) 4.5 Å, and (c) 6 Å water layers, CTAB micelles with (d) 4.5 Å and (e) 6 Å water layers.



**Figure 6.** Proteins in micelle enclosure. (a–c) The N-terminal transmembrane domain of OmpA of *E. coli* in a DPC micelle. (a) The crystal structure (missing unresolved loops) color coded according to the crystallographic B-factor. (b) The water equilibration structure color coded according to the fluctuations. (c) The vacuum structure with an initial water layer of 0.3 nm color coded according to the deviation from the water structure. (d–f) Myoglobine-CTAB complex in vacuum. (d) The complex after 10 ns in vacuum. (e) The structure of the protein in (d) rendered in the same style as in Figure 2. (f) The protein in (d) rendered together with CTAB and water molecules within 3 Å, showing surface contacts.

structure is low, as was found for the proteins just coated by a water layer in vacuum (Figure 2).

## Outlook

The packing of proteins strengthened by the many HBs make that they do not unfold under gas-phase conditions, if the total charge is moderate.[3,59] Reversible temperature-driven unfolding experiments of the Trp-cage protein in an ion-trap strongly suggest that the native state in solution is very close to the native state in the gas-phase for this protein,[44] this was also concluded for the particular case of the 1+ charge state of this protein, where the side chain charge locations are the same in solution and gas phase.[91] MD simulations[65,72,87] demonstrate that the fraction of HBs (relative to the theoretical maximum) increases significantly upon dehydration from roughly 43–56% (Table 2). This suggests that in many cases proteins in the gas phase will be a trapped in a low energy state that is structurally close to the native state.

Two important unresolved problems remain in the computational modeling of biological systems in the gas-phase. These are to some extent general modeling issues, but in a "special" case like partially dehydrated proteins they are even more important than for proteins in solution.

(i) Environmental factors such as pH, temperature and solvent composition influence the charge distribution.[37] In cases where proteins have the same charge state in solution as in the gas phase, the charges most likely are on different side chains,[92,93] except in very special cases.[91] A number of algorithms to predict protonation in the gas-phase have been reported.[65,66,94,104] Advances in this field can be expected when MD simulations routinely incorporate protonation and deprotonation reactions in a reliable manner, and some efforts to this end have been published.[95–97]
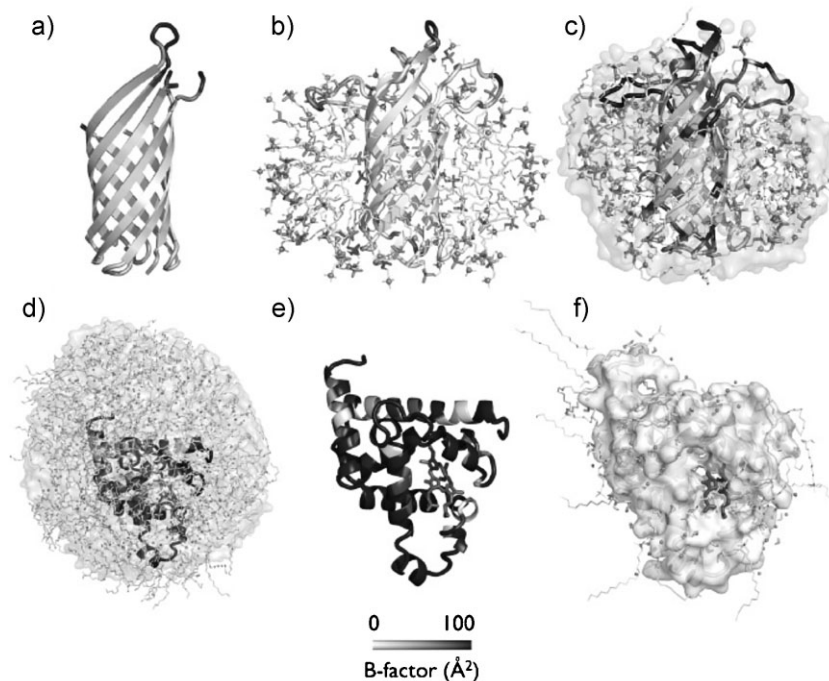
(ii) The force fields employed currently are developed with an effective polarization (by increasing the charges), as described by Berendsen et al.[98] Force field models including explicit polarization are slowly beginning to appear in the literature, and we have previously introduced models for small molecules.[99,100] Presently the only polarizable force field that can be applied to biological systems is AMOEBA[101,102] but this model is rather complex precluding its use for large-scale simulations.

By combining experimental techniques and computational modeling biological systems can be studied in the gas phase in order to directly investigate solvation effects. With the current rate of development of experimental methods, simulation algorithms[82] and computer power[103] it seems likely that the relation between solution- and gas-phase biology can soon be studied in a quantitative manner, and that solution properties may be predicted accurately based on a combination of gas-phase measurements and modeling.

[1] R. van den Heuvel, A. Heck, *Curr. Opin. Chem. Biol.* **2004**, *8*, 519.

[2] C. Uetrecht, R. J. Rose, E. van Duijn, K. Lorenzen, A. J. R. Heck, *Chem. Soc. Rev.* **2010**, *39*, 1633.

[3] D. Clemmer, M. Jarrold, *J. Mass Spectrom.* **1997**, *32*, 577.

[4] T. Wyttenbach, M. Bowers, *Modern Mass Spectrom*, Vol. 225 of Topics Curr. Chem., Springer-Verlag, Berlin 2003, pp. 207–232.

[5] R. Zubarev, *Mass Spectrom. Rev.* **2003**, *22*, 57.

[6] J. Fenn, *Angew. Chem., Int. Ed.* **2003**, *42*, 3871.

[7] A. Heck, R. van den Heuvel, *Mass Spectrom. Rev.* **2004**, *23*, 368.

[8] B. T. Ruotolo, C. V. Robinson, *Curr. Opin. Chem. Biol.* **2006**, *10*, 402.

[9] J. L. P. Benesch, B. T. Ruotolo, D. A. Simmons, C. V. Robinson, *Chem. Rev.* **2007**, *107*, 3544.

[10] R. Zubarev, M. Mann, *Mol. Cell. Proteomics* **2007**, *6*, 377.

[11] T. Taverner, H. Hernandez, M. Sharon, B. T. Ruotolo, D. Matak-Vinkovic, D. Devos, R. B. Russell, C. V. Robinson, *Acc. Chem. Res.* **2008**, *41*, 617.

[12] M. A. Tito, K. Tars, K. Valegård, J. Hajdu, C. V. Robinson, *J. Am. Chem. Soc.* **2000**, *122*, 3550.

[13] C. V. Robinson, M. Gross, S. J. Eyles, J. J. Ewbank, M. Mayhew, F. U. Hartl, C. M. Dobson, S. E. Radford, *Nature* **1994**, *372*, 646.

[14] A. A. Rostom, C. V. Robinson, *J. Am. Chem. Soc.* **1999**, *121*, 4718.

[15] B. T. Ruotolo, K. Giles, I. Campuzano, A. M. Sandercock, R. H. Bateman, C. V. Robinson, *Science* **2005**, *310*, 1658.

[16] C. Uetrecht, C. Versluis, N. R. Watts, W. H. Roos, G. J. L. Wuite, P. T. Wingfield, A. C. Steven, A. J. R. Heck, *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 9216.

[17] C. Uetrecht, C. Versluis, N. R. Watts, P. T. Wingfield, A. C. Steven, A. J. R. Heck, *Angew. Chem., Int. Ed.* **2008**, *47*, 6247.

[18] D. B. Teplow, N. D. Lazo, S. Bitan, S. Bernstein, T. Wyttenbach, M. T. Bowers, A. Baumketner, J.-E. Shea, B. Urbanc, L. Cruz, J. Borreguero, H. E. Stanley, *Acc. Chem. Res.* **2006**, *39*, 635.

[19] S. L. Bernstein, N. F. Dupuis, N. D. Lazo, T. Wyttenbach, M. M. Condron, G. Bitan, D. B. Teplow, J.-E. Shea, B. T. Ruotolo, C. V. Robinson, M. T. Bowers, *Nat. Chem.* **2009**, *1*, 326.

[20] M. Grabenauer, C. Wu, P. Soto, J.-E. Shea, M. T. Bowers, *J. Am. Chem. Soc.* **2010**, *132*, 532.

[21] J. B. Fenn, M. Mann, C. K. Meng, S. F. Wong, C. M. Whitehouse, *Science* **1989**, *246*, 64.

[22] M. Jarrold, *Annu. Rev. Phys. Chem.* **2000**, *51*, 179.

[23] A. Charvat, B. Abel, *Phys. Chem. Chem. Phys.* **2007**, *9*, 3335.

[24] M. Priebe, S. Kalbfleisch, M. Tolkiehn, S. Koester, B. Abel, R. J. Davies, T. Salditt, *New J. Phys.* **2010**, *12*, 043056.

[25] T. Wyttenbach, M. T. Bowers, *Chem. Phys. Lett.* **2009**, *480*, 1.

[26] A. Scholten, T.-T. Aye, A. J. R. Heck, *Mass Spectrom. Rev.* **2008**, *27*, 331.

[27] R. Neutze, R. Wouts, D. van der Spoel, E. Weckert, J. Hajdu, *Nature* **2000**, *406*, 752.

[28] R. Neutze, G. Huldt, J. Hajdu, D. van der Spoel, *Radiat. Phys. Chem.* **2004**, *71*, 905.

[29] H. N. Chapman, A. Barty, M. J. Bogan, S. Boutet, M. Frank, S. P. Hau-Riege, S. Marchesini, B. W. Woods, S. Bajt, R. A. London, E. Plönjes, M. Kuhlmann, R. Treusch, S. Düsterer, T. Tschentscher, J. R. Schneider, E. Spiller, T. Möller, C. Bostedt, M. Hoener, D. A. Shapiro, K. O. Hodgson, D. van der Spoel, F. Burmeister, M. Bergh, C. Caleman, G. Huldt, M. M. Seibert, F. R. Maia, R. W. Lee, A. Szöke, N. Tîmneanu, J. Hajdu, *Nat. Phys.* **2006**, *2*, 839.

[30] F. R. N. C. Maia, T. Ekeberg, N. Timneanu, D. van der Spoel, J. Hajdu, *Phys. Rev. E* **2009**, *80*, 031905.

[31] F. Wang, E. Weckert, B. Ziaja, D. S. D. Larsson, D. van der Spoel, submitted.

[33] G. O. Z. Jurek, G. Faigel, *Europhys. Lett.* **2004**, *65*, 491.

[34] S. P. Hau-Riege, H. Szoke, H. N. Chapman, A. Szoke, S. Marchesini, A. Noy, H. F. He, M. Howells, U. Weierstall, J. C. H. Spence, *Acta Crystallogr., Sect. A* **2004**, *4*, 294.

[35] C. Caleman, C. Ortiz, E. Marklund, F. Bultmark, M. Gabrysch, F. G. Parak, J. Hajdu, M. Klintenberg, N. Timneanu, *EPL* **2009**, *85*, 18005.

[36] C. Caleman, C. Ortiz, E. Marklund, F. Bultmark, M. Gabrysch, F. G. Parak, J. Hajdu, M. Klintenberg, N. Timneanu, *EPL* **2009**, *88*, 29901.

[37] S. J. Valentine, J. G. Anderson, A. D. Ellington, D. E. Clemmer, *J. Phys. Chem. B* **1997**, *101*, 3891.

[38] M. Seibert, A. Patriksson, B. Hess, D. van der Spoel, *J. Mol. Biol.* **2005**, *354*, 173.

[39] D. van der Spoel, M. M. Seibert, *Phys. Rev. Lett.* **2006**, *96*, 238102.

[40] S. Honda, K. Yamasaki, Y. Sawada, H. Morii, *Struct. Fold. Des.* **2004**, *12*, 1507.

[41] W. L. Jorgensen, J. Tirado-Rives, *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6665.

[42] A. Iavarone, J. Parks, *J. Am. Chem. Soc.* **2005**, *127*, 8606.

[43] A. T. Iavarone, D. Duft, J. H. Parks, *J. Phys. Chem. A* **2006**, *110*, 12714.

[44] A. Iavarone, A. Patriksson, D. van der Spoel, J. Parks, *J. Am. Chem. Soc.* **2007**, *129*, 6726.

[45] A. Patriksson, C. Adams, F. Kjeldsen, J. Raber, D. van der Spoel, R. A. Zubarev, *Int. J. Mass Spectrom.* **2006**, *248*, 124.

[46] C. M. Adams, F. Kjeldsen, A. Patriksson, D. van der Spoel, A. Gräslund, E. Papadopoulos, R. A. Zubarev, *Int. J. Mass Spectrom.* **2006**, *253*, 263.

[47] K. Breuker, F. W. McLafferty, *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 18145.

[48] C. Robert, P. Ho, *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 7600.

[49] V. Denisov, B. Halle, *J. Mol. Biol.* **1995**, *245*, 682.

[50] F. Merzel, J. C. Smith, *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 5378.

[51] J. Smith, F. Merzel, C. Verma, S. Fischer, *J. Mol. Liq.* **2002**, *101*, 27.

[52] L. D. Barron, L. Hecht, G. Wilson, *Biochemistry* **1997**, *36*, 13143.

[53] S. Fischer, C. Verma, *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 9613.

[54] L. Olano, S. Rick, *J. Am. Chem. Soc.* **2004**, *126*, 7991.

[55] Y. Mao, M. A. Ratner, M. F. Jarrold, *J. Am. Chem. Soc.* **2000**, *122*, 2950.

[56] P. Fitzpatrick, A. Steinmetz, D. Ringe, A. Klibanov, *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 8653.

[57] H. Nagendra, N. Sukumar, M. Vijayan, *Proteins* **1998**, *32*, 229.

[58] J. Fye, J. Woenckhaus, M. Jarrold, *J. Am. Chem. Soc.* **1998**, *120*, 1327.

[59] M. F. Jarrold, *Acc. Chem. Res.* **1999**, *32*, 360.

[60] M. Kohtani, M. F. Jarrold, *J. Am. Chem. Soc.* **2002**, *124*, 11148.

[61] M. F. Jarrold, *Phys. Chem. Chem. Phys.* **2007**, *9*, 1659.

[62] B. Gao, T. Wyttenbach, M. T. Bowers, *J. Am. Chem. Soc.* **2009**, *131*, 4695.

[63] Y. Mao, M. A. Ratner, M. F. Jarrold, *J. Phys. Chem. B* **1999**, *103*, 10017.

[64] Y. Mao, M. A. Ratner, M. F. Jarrold, *J. Am. Chem. Soc.* **2001**, *123*, 6503.

[65] A. Patriksson, E. Marklund, D. van der Spoel, *Biochemistry* **2007**, *46*, 933.

[66] L. Konermann, *J. Phys. Chem. B* **2007**, *111*, 6534.

[67] T. Meyer, X. de la Cruz, M. Orozco, *Structure* **2009**, *17*, 88.

[68] E. G. Marklund, D. S. D. Larsson, D. van der Spoel, A. Patriksson, C. Caleman, *Phys. Chem. Chem. Phys.* **2009**, *11*, 8069.

[69] A. Harrison, *Mass Spectrom. Rev.* **1997**, *16*, 201.

[70] J. D. Carbeck, J. C. Severs, J. Gao, Q. Wu, R. D. Smith, G. M. Whitesides, *J. Phys. Chem. B* **1998**, *102*, 10596.

[71] I. McDonald, J. Thornton, *J. Mol. Biol.* **1994**, *238*, 777.

[72] Y. Wang, D. S. D. Larsson, D. van der Spoel, *Biochemistry* **2009**, *48*, 1006.

[73] C. M. Dobson, P. A. Evans, S. E. Radford, *Trends Biochem. Sci.* **1994**, *19*, 31.

[74] C. Caleman, D. van der Spoel, *J. Chem. Phys.* **2006**, *125*, 154508.

[75] M. Z. Steinberg, K. Breuker, R. Elber, R. B. Gerber, *Phys. Chem. Chem. Phys.* **2007**, *9*, 4690.

[76] C. Caleman, D. van der Spoel, *Phys. Chem. Chem. Phys.* **2007**, *9*, 5105.

[77] L. L. Ilag, I. Ubarretxena-Belandia, C. G. Tate, C. V. Robinson, *J. Am. Chem. Soc.* **2004**, *126*, 14362.

[78] N. P. Barrera, N. Bartolo, P. J. Booth, C. V. Robinson, *Science* **2008**, *321*, 243.

[79] M. Sharon, C. V. Robinson, *Annu. Rev. Biochem.* **2007**, *76*, 167.

[80] D. P. Tieleman, D. van der Spoel, H. J. C. Berendsen, *J. Phys. Chem. B* **2000**, *104*, 6380.

[81] S. J. Marrink, D. P. Tieleman, A. E. Mark, *J. Phys. Chem. B* **2000**, *104*, 12165.

[82] B. Hess, C. Kutzner, D. van der Spoel, E. Lindahl, *J. Chem. Theory Comput.* **2008**, *4*, 435.

[83] S. G. F. Rasmussen, H.-J. Choi, D. M. Rosenbaum, T. S. Kobilka, F. S. Thian, P. C. Edwards, M. Burghammer, V. R. P. Ratnala, R. Sanishvili, R. F. Fischetti, G. F. X. Schertler, W. I. Weis, B. K. Kobilka, *Nature* **2007**, *450*, 383.

[84] V. Cherezov, D. M. Rosenbaum, M. A. Hanson, S. G. F. Rasmussen, F. S. Thian, T. S. Kobilka, H.-J. Choi, P. Kuhn, W. I. Weis, B. K. Kobilka, R. C. Stevens, *Science* **2007**, *318*, 1258.

[85] R. O. Dror, D. H. Arlow, D. W. Borhani, M. O. Jensen, S. Piana, D. E. Shaw, *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 4689.

[86] D. Mackay, *J. Theor. Biol.* **1990**.

[87] R. Friemann, D. S. D. Larsson, Y. Wang, D. van der Spoel, *J. Am. Chem. Soc.* **2009**, *131*, 16606.

[88] P. J. Bond, M. S. P. Sansom, *J. Mol. Biol.* **2003**, *329*, 1035.

[89] P. J. Bond, J. M. Cuthbertson, S. Deol, M. S. P. Sansom, *J. Am. Chem. Soc.* **2004**, *126*, 15948.

[90] C. M. Hebling, C. R. Morgan, D. W. Stafford, J. W. Jorgenson, K. D. Rand, J. R. Engen, *Acc. Chem. Res.* **2010**, *82*, 5415.

[91] A. Patriksson, C. Adams, F. Kjeldsen, R. A. Zubarev, D. van der Spoel, *J. Phys. Chem. B* **2007**, *111*, 13147.

[92] K. Breuker, H. Oh, D. M. Horn, B. A. Cerda, F. W. McLafferty, *J. Am. Chem. Soc.* **2002**, *124*, 6407.

[93] C. M. Adams, F. Kjeldsen, R. A. Zubarev, *Am. Soc. Mass Spectrom.* **2004**, *15*, 1087.

[94] M. Miteva, P. A. Demirev, A. D. Karshikoff, *J. Phys. Chem. B* **1997**, *101*, 9645.

[95] M. A. Lill, V. Helms, *J. Chem. Phys.* **2001**, *115*, 7993.

[96] R. Burgi, P. A. Kollman, W. F. van Gunsteren, *Proteins* **2002**, *47*, 469.

[97] J. Khandogin, C. L. Brooks, *Biochemistry* **2006**, *45*, 9363.

[98] H. J. C. Berendsen, J. R. Grigera, T. P. Straatsma, *J. Phys. Chem.* **1987**, *91*, 6269.

[99] P. C. Jordan, P. J. van Maaren, J. Mavri, D. van der Spoel, H. J. C. Berendsen, *J. Chem. Phys.* **1995**, *103*, 2272.

[100] P. J. van Maaren, D. van der Spoel, *J. Phys. Chem. B* **2001**, *105*, 2618.

[101] P. Ren, J. W. Ponder, *J. Phys. Chem. B* **2003**, *107*, 5933.

[102] J. W. Ponder, C. Wu, P. Ren, V. S. Pande, J. D. Chodera, M. J. Schnieders, I. Haque, D. L. Mobley, D. S. Lambrecht, R. A. DiStasio, Jr., M. Head-Gordon, G. N. I. Clark, M. E. Johnson, T. Head-Gordon, *J. Phys. Chem. B* **2010**, *114*, 2549.

[103] R. Schulz, B. Lindner, L. Petridis, J. C. Smith, *J. Chem. Theory Comput.* **2009**, *5*, 2798.

[104] R. Marchese, R. Grandori, P. Carloni, S. Raugei, *PLoS. Comp. Biol.* **2010**, *6*, e1000775.