University of Groningen

Role of charge and hydrophobicity in lipotide formation

Nedergaard Pedersen, Jannik; Wilhelmus Johannes Maria Frederix, Pim; Skov Pedersen, Jan; Marrink, Siewert Jan; Otzen, Daniel

Published in:
 ChemBioChem

DOI:
[10.1002/cbic.201700496](https://doi.org/10.1002/cbic.201700496)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
 Publisher's PDF, also known as Version of record

Publication date:
 2018

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Nedergaard Pedersen, J., Wilhelmus Johannes Maria Frederix, P., Skov Pedersen, J., Marrink, S. J., & Otzen, D. (2018). Role of charge and hydrophobicity in lipotide formation: A molecular dynamics study with experimental constraints. *ChemBioChem*, 19, 263–271. <https://doi.org/10.1002/cbic.201700496>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

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.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Role of Charge and Hydrophobicity in Liprotide Formation: A Molecular Dynamics Study with Experimental Constraints

Jannik Nedergaard Pedersen,^[a] Pim Wilhelmus Johannes Maria Frederix,^[b]
Jan Skov Pedersen,^[c] Siewert Jan Marrink,^[b] and Daniel E. Otzen^{*,[a]}

Bovine α -lactalbumin (aLA) and oleate (OA) form a complex that has been intensively studied for its tumoricidal activity. Small-angle X-ray scattering (SAXS) has revealed that this complex consists of a lipid core surrounded by partially unfolded protein. We call this type of complex a liprotide. Little is known of the molecular interactions between OA and aLA, and no technique has so far provided any high-resolution structure of a liprotide. Here we have used coarse-grained (CG) molecular dynamics (MD) simulations, isothermal titration calorimetry (ITC) and SAXS to investigate the interactions between aLA and OA during the process of liprotide formation. With ITC we found that the strongest enthalpic interactions occurred at a molar ratio of $12.0 \pm 1.4:1$ OA/aLA. Liprotides formed between OA and aLA at several OA/aLA ratios in silico were stable both

in CG and in all-atom simulations. From the simulated structures we calculated SAXS spectra that show good agreement with experimentally measured patterns of matching liprotides. The simulations showed that aLA assumes a molten globular (MG) state, exposing several hydrophobic patches involved in interactions with OA. Initial binding of aLA to OA occurs in an area of aLA in which a high amount of positive charge is located, and only later do hydrophobic interactions become important. The results reveal how unfolding of aLA to expose hydrophobic residues is important for complex formation between aLA and OA. Our findings suggest a general mechanism for liprotide formation and might explain the ability of a large number of proteins to form liprotides with OA.

Introduction

α -Lactalbumin (aLA) is a 123-residue calcium-binding protein (Figure 1). It consists of several α -helices and a β -domain, and the tertiary structure is stabilised by four disulfide bridges.^[1] aLA is known to have several partially unfolded intermediates, known as molten globules (MGs), that can be formed in various ways: by removal of Ca^{2+} , heating or interactions with certain lipids, for example.^[2] A complex with tumoricidal activity was found by Svanborg and co-workers to form between aLA and the fatty acid oleate (OA) when aLA was destabilised to a molten globular state.^[3] Later, several other complexes with different proteins and lipids showing similar tumoricidal activity were formed.^[4] These complexes share the same overall

Figure 1. Crystal structure of α -lactalbumin and structure of oleate. A) α -Lactalbumin (PDB ID: 1F6R). B) Stick representation of oleate with the MARTINI mapping on top.

structure of a lipid core surrounded by partially unfolded protein. We have introduced the term “liprotides” to describe complexes with these structural features.^[4a] Near- and far-UV, together with Trp fluorescence spectroscopy, suggest that the protein in the liprotide retains most of its secondary structure, with a small increase in random coil structure, but completely loses its tertiary structure.^[5]

Liprotides can be formed in several different ways, including by exposure to OA on chromatographic columns or through heating, change of pH or removal of Ca^{2+} .^[6] All of these procedures involve partial unfolding of the protein component. Different procedures can lead to different ratios of lipid and protein in the final complexes. When OA and aLA have been simply mixed together and excess OA removed, aLA:OA molar ratios between 1:9 and 1:11 have been reported.^[4a,7] On the basis of a combination of gel-filtration, dynamic light scattering and electron microscopy, Fontana and co-workers concluded that an aLA–OA liprotide contains four to five protein mole-

[a] Dr. J. Nedergaard Pedersen, Prof. D. E. Otzen
Interdisciplinary Nanoscience Center (iNANO), Department of Molecular
Biology and Genetics, Aarhus University
Gustav Wieds Vej 14, 8000 Aarhus C (Denmark)
E-mail: dao@inano.au.dk

[b] Dr. P. W. J. M. Frederix, Prof. S. J. Marrink
Groningen Biomolecular Sciences and Biotechnology Institute
Zernike Institute for Advanced Materials, University of Groningen
Nijenborgh 7, 9747 AG Groningen (The Netherlands)

[c] Prof. J. Skov Pedersen
Department of Chemistry and Interdisciplinary Nanoscience Center
(iNANO), Aarhus University
Gustav Wieds Vej 14, 8000 Aarhus C (Denmark)

Supporting information and the ORCID identification numbers for the
authors of this article can be found under <https://doi.org/10.1002/cbic.201700496>.

cules.^[6a] We have used small-angle X-ray scattering (SAXS) and reversed-phase HPLC to identify aLA–OA complexes containing between three and four aLA units per 28–36 fatty acid molecules.^[7] Larger variations in the number and ratio of protein and lipid molecules have also been seen when different preparation methods are applied, but all suggest that several protein and lipid units are present in the complexes.^[4a]

Below pH 9–10 and at concentrations above the critical micelle concentration (cmc), OA is protonated and therefore tends to form large visible aggregates.^[8] However, lipotides formed at pH 7.4 are still able to solubilise OA. This is attributed to the ability of lipotides to maintain OA in its deprotonated (ionised) state^[9] when OA is forming micelles.^[8] It is not clear how the carboxylate group's pK_a is lowered in lipotides, but it might occur through electrostatic interactions.

Few theoretical studies have been done to follow unfolding and interaction of proteins with amphiphilic molecules such as OA or surfactants.^[10] Molecular dynamics (MD) simulations of systems containing proteins and surfactants are complex because they involve highly flexible and dynamic surfactants and at the same time require unfolding of the protein. The latter is challenging at the timescales accessible to MD simulations (nano- to microseconds). Docking and simulation studies have already been performed on the lipotides to obtain better understanding of the interaction between protein and fatty acid, but these studies have mainly focused on the interaction of a single or a few lipid molecules with a single protein molecule.^[11] In those studies the tertiary structure of the protein did not change markedly, which is at odds with the known loss of structure of proteins in lipotides. A more representative picture of the interaction between protein and fatty acids requires the full complex with several proteins and fatty acids.

In our study we used the MARTINI coarse-grained (CG) force field to perform MD simulations on the spontaneous self-assembly of aLA–OA lipotides. Using a CG model allows for simulations of large complex systems of multiple proteins and lipids over the range of microseconds, leading to structures that can be compared against experimental data.^[12] To validate our computational results, our MD structures were compared with structures determined experimentally through SAXS measurements. We also used isothermal titration calorimetry

(ITC) to obtain better understanding of the binding of aLA and OA in lipotides and to determine the appropriate ratio of aLA to OA. Reasonable MD simulation structures were only obtained when aLA was allowed to unfold to expose hydrophobic patches. Initial binding of aLA to OA was, however, initiated by electrostatic interactions from an area on aLA with a high local positive charge. This combination of experiments and simulations provided better insight into the interaction between proteins and lipids in lipotides and helps understanding of why lipotides can be formed by a wide range of proteins.

Results and Discussion

aLA can bind more than 20 OA molecules

aLA has been shown to bind different amounts of OA in lipotide complexes, depending on the preparation method.^[13] It was important to determine the appropriate protein/lipid ratio experimentally under conditions equivalent to our simulation conditions, in which the protein is mixed with preformed OA micelles. Accordingly we monitored protein–lipid complex formation by directly mixing protein with a suspension of OA. The process was followed by ITC, which allows us to estimate binding numbers and the ratios giving the highest binding enthalpy (Figure 2 and Figure S1 in the Supporting Information). Our earlier SAXS, far-UV-CD and Trp fluorescence studies have shown that the lipotide structure and OA/aLA binding ratios remain essentially intact between pH 7.4 and 10.^[14] Because OA is only sparingly soluble at pH 7.4, we performed the ITC experiment at pH 10 and 37 °C to increase OA solubility, to avoid OA aggregation during the experiment and to help unfold the aLA. Titration of OA into different concentrations of aLA revealed that the characteristic features of the titration curves were displaced to higher OA concentrations as the concentration of aLA increased (Figure 2A), but the curves more or less collapsed into a single curve when plotted versus the OA/aLA molar ratio (Figure S1), thus indicating that binding largely depended on the ratio between OA and aLA.

We chose three characteristic points (Figure S1) of binding and used the equation $[OA]_{tot} = [OA]_{free} + n_{OA, bound} \times [aLA]$ to de-

Figure 2. ITC to determine binding stoichiometry of oleate and α -lactalbumin. A) Enthalpogram for titration of oleate into α -lactalbumin as a function of OA concentration. B) OA concentration at points 1, 2 and 3 plotted as a function of aLA concentration. A linear fit to $[OA]_{tot} = [OA]_{free} + n_{OA, bound} \times [aLA]$ provides binding numbers.

termine the amounts of free and bound OA more accurately at these three points.^[15] At low concentrations of OA, there was a small maximum in the signal, most clearly seen when titrating OA into high protein concentrations, in all cases occurring around 5.4 ± 1.6 OA/aLA. We attribute the signal to a combination of exothermic contribution from OA binding and an endothermic contribution from initial unfolding of aLA that happens at low OA concentrations. However, the likely combination of binding and unfolding at these low OA/aLA ratios makes it difficult to determine precisely at what OA/aLA ratio unfolding occurs. After unfolding of aLA at low OA concentrations, the main contribution to the signal is expected to arise from OA binding to aLA. The largest exothermic signal is seen at point 2 (Figure S1) corresponding to 12.0 ± 1.4 OA/aLA. This is in good agreement with binding of 9–11 OA/aLA reported for purified lipotides.^[4a,7] Additional binding occurs at higher [OA], however, and at the third point (halfway between point 2 and the baseline, set to 0 kcal mol^{-1}) we see binding of 21.6 ± 1.8 OA/aLA. Remarkably, this ratio corresponds to that seen with lipotides when large amounts of OA are mixed with aLA and separated by gel filtration (21 ± 5 OA/aLA).^[7b] At all three transitions, the amount of unbound OA was essentially zero within error (y-axis intercept in Figure 2B and Table S2), thus confirming that all of the OA complexes with aLA. Suspension of OA in buffer leads to a very low concentration of monomeric OA because OA has a cmc of $6 \mu\text{M}$ at pH 7.4^[16] and titration of OA into buffer at pH 10 gave no signal (Figure 2A), thus indicating that very little demicellisation takes place upon dilution. Even after the first OA titration in Figure 2A the concentration of OA is already $22 \mu\text{M}$, which is much higher than the cmc, so essentially all OA will be available as micelles to bind aLA.

Oleate forms micelles at high pH

ITC, together with determinations of OA/aLA ratios from earlier studies, suggests that a molar ratio of around nine to 12 OA units per aLA molecule is sufficient to form lipotides. For the pure OA system, we used deprotonated OA, even though OA is protonated at pH 7.4. We did this to obtain OA aggregates that were as similar as possible to the aggregates of the lipotide core, which consist of deprotonated OA.^[9] Increasing pro-

tonation first causes micelle elongation and eventually bilayer structure formation as determined previously with the MARTINI force field,^[17] and this is not likely to be relevant for the lipotides. To determine the amount of OA molecules per micelle to use in this study we performed a CG MD simulation on a solution of 500 OA molecules in water (0.06 wt%) over the course of $1 \mu\text{s}$. The sizes of the observed OA clusters were stable during the simulation period of $0.2\text{--}1 \mu\text{s}$. We observed a relatively broad distribution with an average cluster size of 29 ± 7 OA molecules per micelle (Figure 3). Reassuringly, our earlier SAXS findings suggest that core-shell lipotides often contain 28–36 OA molecules per lipotide,^[7] which is in the same range as in the micelles formed in the simulation. We also note that simulations by King and co-workers of 5–20% OA solutions (a much higher concentration range than we analysed) gave micelles with 27–44 molecules, with the 20% solutions giving the largest micelles, with a prolate shape.^[18] On the basis of these results, we chose an OA/aLA ratio of 36:4 because this gives stable OA micelles and at the same time gives an opportunity for lipotides of 9:1 and 12:1 OA/aLA ratios to form, depending on how many protein molecules bind. Importantly, the conditions are sufficiently similar to the experimental conditions to allow us to calculate SAXS spectra for these simulated complexes and to compare these with experimentally measured SAXS spectra. We chose to use preformed OA micelles for our simulations, because lipotides are known to consist of micellar OA,^[4a] furthermore, lipotides can also form at high pH, at which OA is deprotonated even in the absence of protein.^[19] We cannot rule out the possibility that the lipotides could contain small amounts of protonated OA that could change the structure to some degree.^[17] However, because we do not know the exact degree of protonation, but do know that it is very small, we chose to use completely deprotonated OA. When forming the lipotides experimentally we use the deprotonated form of OA. Deprotonated OA remains soluble for some time even after the pH is changed to 7.4, thus indicating that the protonation of the carboxylate group is slow (despite the extreme rapidity of proton transfers in general). We interpret this as implying that deprotonated OA will have time to be stabilised by aLA before it protonates.

Figure 3. OA cluster size. A) Frequency of cluster sizes from simulation of 500 oleate molecules. B) Snapshot after $1 \mu\text{s}$ simulation of 500 OA molecules. OA is represented as the head group bead (red) with the carbon tail in cyan.

MD of liprotide formation

It is likely that different structures might be formed if different amounts OA are used to form liprotides. Indeed, liprotides are not necessarily the only complexes that can be formed between aLA and OA. Complexes of a few OA molecules with a single aLA unit have been reported but have not yet been shown to have the same membrane interaction properties as liprotides,^[5–6] and the low OA/aLA stoichiometry makes it highly unlikely that they can form the core-shell structure that we consider emblematic of liprotides. Furthermore, the low OA content would explain the lack of membrane effects because a threshold amount of OA is presumably needed to perturb membranes.

We therefore focused on liprotides containing several OA and aLA molecules. To study the formation of liprotides, we set up five independent simulations with one preformed micelle containing 36 OA molecules and four randomly placed aLA molecules. The protein was modelled without the ElnDyn elastic network to allow for conformational changes (see the Experimental Section). Rather than providing a detailed description of each individual binding and unfolding trajectory, we restrict our focus to structures that are consistent with our experimentally acquired SAXS data (see below).

After 1.5 μ s of simulation, one micelle had bound four protein molecules (sim1), three micelles had bound three protein molecules (sim2–4), and one micelle had bound two (sim5, Figure 4). The bound aLA molecules generally covered most of the OA surface, thus impeding binding of additional aLA molecules. For this reason, the micelle that possessed only two bound aLA molecules on the micelle surface (sim5) also had a large part of the surface of aLA close to the OA surface. The aLA molecules that were not bound after 1.5 μ s did not bind during the whole 3 μ s simulation.

We used the Backward script^[20] to obtain all-atom structures of the liprotides and simulated these structures for 20 ns with

Figure 4. Snapshots of liprotides after simulation of α -lactalbumin with oleate and back-mapping. In sim1 the oleate micelle bound four aLA molecules, in sim2–4 three aLA molecules were bound in each case, and in sim5 two aLA molecules were bound. Snapshots were taken 500 ns after the first aLA binding event (at 1200, 850, 2000, 960 and 1500 ns for sim1–5, respectively). OA is shown in the VdW surface representation with the charged head groups in red.

the CHARMM all-atom force field to verify that the structures were still stable and did not alter markedly. During the 20 ns simulation, none of the formed complexes underwent any dissociation of OA, and the protein structures did not change markedly. Instead they became more stable and less fluctuating, with root mean square fluctuations (RMSFs) of (0.28 ± 0.11) nm, in comparison to (0.51 ± 0.17) nm for the CG model. Note that the bound proteins were more flexible than the native protein, which showed an RMSF of (0.11 ± 0.05) nm (20 ns simulation).

The simulated liprotides are consistent with experimentally acquired SAXS data

Simulations of complexes containing several partially unfolded protein units together with micelles have not been investigated in depth before and we therefore used SAXS to validate our computational approach. The all-atom structures obtained from the simulations were converted into SAXS spectra, and these were compared with spectra recorded in the current work and previously.^[14] We can calculate amounts of OA and aLA from experimentally measured SAXS spectra of liprotides with the aid of our SAXS liprotide model, as previously described.^[4a] This allowed us to identify three relevant sets of experimental data that could be compared with our simulated data: liprotides consisting of 35 OA and 4.0 aLA units (lip4), of 31 OA and 3.2 aLA units (lip7) and of 28 OA and 2.8 aLA units (lip8, Table 1). Lip4 spectra were recorded in this study, whereas lip7 and lip8 were obtained from an earlier study^[14] and are based on liprotides made at pH 7.4 and 8, respectively. These experimental data were compared with the structures obtained from MD simulations with OA/aLA ratios of 36:3 and 36:4 (Figures 5 and S2). We did not have any experimental data that resembled the 36:2 ratio obtained in sim5 and data also fitted rather poorly (Figure S2C), effectively providing a negative control.

Structures obtained after 3 μ s simulation showed significant differences from structures obtained during the 500 ns after aLA had bound OA (Figure 5A). The local maxima seen in the SAXS spectra around $0.15 \text{ \AA}^{-1}$ had shifted to higher q values and the intensity of the peak decreased with time (Figure 5A). These changes reflect aLA rearrangements during simulation that caused tighter packing of aLA around OA, giving less flexi-

Table 1. Overview of simulation results and SAXS experiments.

Simulation	Final OA/aLA ratio	$t_{\text{bound}} [\mu\text{s}]^{\text{[a]}}$	Experiment ^[b]	Experimentally determined ratio ^[c]
1	36:4	0.70	lip4 ^[d]	35:4.0
2	36:3	0.35	lip7/lip8 ^[14]	31:3.2/28:2.8
3	36:3	1.50	lip7/lip8 ^[14]	31:3.2/28:2.8
4	36:3	0.46	lip7/lip8 ^[14]	31:3.2/28:2.8
5	36:2	0.50	not found	–

[a] Simulation time before the final aLA has bound to OA. [b] Name of SAXS sample used for comparison with simulated data. [c] Ratio between OA and aLA in the experimental SAXS data used for comparison. [d] This work.

Figure 5. Comparison of experimentally measured SAXS data with SAXS spectra calculated from simulations. A) SAXS spectra of lip4 compared with structures from sim1. B) SAXS spectra of lip7 and lip8 compared with structures from sim2.

bility and more penetration of aLA chains into the OA micelle. For all experiments, better fitting was obtained with use of MD structures in which all protein units had been bound for up to 500 ns than in the case of structures in which the protein molecules had had time to wrap fully around the micelle (Figure 5; also compare sim1–4 with sim5 in Figure 4). It is likely that tertiary structural interactions in the MARTINI force field are too weak when EINEDyn is not used and that the protein therefore “over-unfolds” (i.e., becomes too extended in structure) over larger timescales. Another problem might be the general overestimation of interactions in the MARTINI force field.^[21] Gratifyingly, MD simulations with an OA/aLA ratio of 36:3 gave good agreement with experimentally acquired SAXS data for lip7 and lip8 with ≈ 3 aLA moieties per ≈ 30 OA units (Figure 5B for sim2 and Figure S2 A–B for sim3 and sim4).

On the basis of the above observations, we investigated whether inclusion of the elastic network obtained from EINEDyn would also give rise to lipotides with predicted spectra resembling the experimentally measured SAXS spectra (Figure S2D). Four simulations of a 36-molecule OA micelle together with three aLA molecules were used for comparison. In two cases (sim1E and sim4E), the protein mainly bound on one side of the micelle, giving spectra very different from those obtained experimentally (Figure S2D). When the protein was more evenly distributed around the micelle (sim2E and sim3E, Figure S2D), spectra that showed a maximum at high q values, characteristic for lipotides, were obtained. However, the subsidiary maximum was markedly shifted towards a higher intensity and towards smaller q values, showing that unfolding of the protein is necessary to reproduce the experimentally obtained SAXS spectra. We conclude that simulations including EINEDyn were not able to reproduce the experimental data.

aLA assumes a molten globule (MG) state in lipotides

The MG state of aLA is considered emblematic for MG states in general.^[22] In this state, the protein becomes more flexible, losing its tertiary structure, and the radius of gyration (R_g , measured by SAXS) increases from 1.57 nm (native state) to 1.72 nm for the MG state.^[23] In our simulations we saw that aLA alone in solution over 500 ns had an R_g of $(1.47 \pm$

0.03) nm. One part of the reason for this discrepancy is probably the hydration layer around the protein, which contributes to experimentally measured SAXS data but is not included when determining R_g with the aid of g_gyrate from the GRO-MACS standard tools. For example, the R_g of lysozyme has been determined by SAXS to be 1.52 nm, whereas the PDB structure (PDB 1DPX) gives R_g values of 1.43 nm when hydration is not accounted for and of 1.48 nm when a hydration layer around the protein is used.^[24] Nevertheless, the R_g value for aLA from the simulation is still low in comparison with the experimentally measured value and suggests that the protein might be too compact in its native state in the simulation. However, aLA bound to micelles in the last 500 ns of the five simulations gave an average R_g of (1.67 ± 0.05) nm. This is close to that of the MG state and suggests an appropriate level of unfolding. The R_g at intermediate times was not higher than at the end of the simulation. We also used SAXS to determine the R_g values of lip7 and lip8 to be (3.08 ± 0.22) nm and (2.76 ± 0.45) nm, respectively. For comparison, the lipotides with 36:3 OA/aLA ratio obtained in the simulations have R_g values of (2.29 ± 0.03) , (2.29 ± 0.03) and (2.39 ± 0.04) nm. Some of the difference might be accounted for by hydration but this cannot account for the whole difference. Taken together, we believe that differences between MD simulations and experimental data may also be caused by tight packing of aLA and OA rather than by the degree of unfolding of aLA, because the protein seems to be unfolded to a reasonable degree. The RMSF of atomic positions in aLA showed an increase from unbound aLA $((0.11 \pm 0.05)$ nm) to micelle-bound aLA $((0.28 \pm 0.11)$ nm), showing that some gain of flexibility is induced when aLA binds to OA micelles in the simulation. The changes in the R_g values of the OA micelles in our case are similar for the five simulations and go from (1.45 ± 0.01) nm before aLA binding to (1.53 ± 0.03) nm after lipotide formation.

Hydrophobic patches on aLA are exposed in the presence of OA

We used the simulated lipotide structures to identify residues important for interaction with OA by investigating interactions formed from 2–3 μ s. To investigate the initial protein–micelle

contacts we considered the four simulations in which the EIneDyn prevented the protein from unfolding, whereas final contacts were calculated from the five simulations without the tertiary structure stabilisation. We quantified contacts in terms of the average surface area of residues of aLA that was buried by OA (Figure 6A; note that the y-axis is positive in both directions to allow direct comparison of data with and without EIneDyn). At pH 7.4 aLA is overall negatively charged, like the OA micelle. However, upon closer inspection it has a small area with a local positive charge and it is here that binding between aLA and OA takes place (Figure 6B, area around Lys58, Lys94 and Lys98). A closer look at the ten residues mostly involved in binding also reveals that the three most important residues are Lys98, -58, and -94 (Figure S3). Previously, NMR studies of aLA's interaction with negatively charged phospholipids had identified residues 10, 12, 77, 94, 98, 99 and 104 located in α -helix C as particularly important for the interaction.^[25] Four of these residues are also on the list of the most important residues involved in the interaction from our MD simulation (Figure S3) and the rest of the residues are generally located around the same area. It is reasonable to expect similar residues on aLA to be involved in interactions with negatively charged phospholipids as with negatively charged OA micelles.

When the protein was stabilised by EIneDyn only a small area of the protein was involved in the interaction with OA (Figure 6B). Removal of the EIneDyn network allowed changes in aLA's tertiary structure to take place, allowing many more residues to become involved in the interaction with OA (Figure 6A). The residues were no longer located in a single area but were spread out over the protein (Figure 6C). Residues located in the positively charged area around α -helix C were still important, likely because the first interaction between aLA and the micelle still took place in this area when EIneDyn was not used and the charge still played a dominant role in the interaction. Nevertheless, the most strongly binding residues were hydrophobic residues normally buried inside the core of aLA (Figure S3). Interestingly, in the native structure, the three main

helices of aLA form a hydrophobic pocket that keeps the helices together. The helices are amphipathic with hydrophobic residues forming the helices pocket and more hydrophilic residues pointing away from the pocket. These buried hydrophobic residues are often involved in the interactions with OA in the final lipotide structures and further demonstrate why unfolding of the protein is important. Earlier MD simulations by Chaudhuri and co-workers suggested that residues from all over aLA interact with OA when aLA is folded.^[11a] However, Chaudhuri and co-workers only used a small number of OA molecules, leading to full exposure of individual OA molecules and facilitating interactions between the hydrophobic tail of OA and aLA. In an OA micelle, the tails are to a certain extent sequestered in the micelle interior, leaving the carboxylate heads more exposed to the surroundings. In vitro, OA may form mixtures of protonated and deprotonated states and exist in different aggregated forms. However, only a very small fraction of OA is likely to be monomeric (6 μM at pH 7.4^[16]) and hydrophobic patches of OA are generally not exposed. Furthermore, several earlier MD simulations on the protein-lipid complexes only involved single aLA molecules,^[11] even though it is known that several proteins and lipids are present in the complexes. Overall, as previous studies have shown,^[4b,6d] the interaction between OA and aLA is a combination of hydrophobic and electrostatic interactions and both are necessary for strong binding.

Delayed unfolding of aLA upon OA interaction

Even though the concentrations in the MD simulations were high (5 mM aLA and 45 mM OA) it took up to 1.5 μs for aLA to bind to micelles. Using the number of contacts between individual aLA and OA molecules as a measure of binding and R_g as a measure of unfolding, we looked into the binding and unfolding pathway of the protein (Figure S4). In all cases, binding preceded unfolding. However, some proteins started unfolding almost immediately after interaction with the micelle whereas

Figure 6. Buried surface area between α -lactalbumin and oleate upon binding. A) Residues of α -lactalbumin that are buried by the OA surface averaged over two times four simulations with (upper panel) and without (lower panel) EIneDyn. Green-shaded area indicates α -helix and grey-shaded area indicates β -sheet. The y-axis is positive in both directions to allow direct comparison of data with and without EIneDyn. B) Structure of the surface of α -lactalbumin showing the residues that are buried most by OA in blue and less buried in red from the simulations of the upper panel in (A). C) Structure of the surface of α -lactalbumin showing the residues that are buried most by OA in blue and less buried in red from the simulations of the lower panel in (A).

other proteins took >500 ns before unfolding occurred (Figure S4). The long time from binding to unfolding might explain the absence of unfolding in all-atom simulations in which the simulation time is much shorter; furthermore, dynamics in an all-atom simulation can be somewhat slower because of the friction caused by the smaller bead sizes. Moreover, the tertiary structure of the protein is less stable in our CG simulations because ElnDyn is not used. Improvement of simulations will require maintenance of timescales accessible with CG while at the same time improving the individual interactions between OA and aLA. The simulations have, however, still been able to give an idea of the overall structure of the liprotide structure as well as the mechanism of the binding between the protein and lipid.

Conclusions

In this study we investigated liprotides of OA with aLA and determined the enthalpically optimal ratio to be 12.0 ± 1.4 OA units per aLA molecule. CG MARTINI simulations showed the formation of liprotide complexes at these experimentally constrained ratios between aLA and OA, leading to a range of liprotide sizes. Our study shows that initial binding of folded aLA to negatively charged OA micelles takes place in a positively charged area of aLA. Through the use of multiple proteins and lipids we have seen binding patterns that are somewhat different from those seen earlier and have been able to see unfolding of the protein. By comparing our results from MD simulations with experimental SAXS data we have been able to validate the obtained MD structures, thus providing us with assurance that the MD results are meaningful and that unfolding of the protein is necessary to find agreement with the experiments.

When unfolding of aLA occurred in the presence of OA, many parts of the protein were eventually involved in interactions with the OA micelle. The interaction seems rather unspecific and it is likely that the observed formation of liprotides with different proteins is possible simply because it mainly requires accessible hydrophobic patches that become available upon protein unfolding. This unfolding might also be important for the liprotides' interaction with membranes in the context of their tumoricidal activity.

Note that whereas the MD simulations provide a way to sample many different conformations, the SAXS spectra obtained from them still relate only to a few structures sampled over relatively short times. In contrast, samples used for SAXS experiments most likely consist of different amounts of aLA and OA and involve aLA bound in many different conformations. This limits the approach to a certain extent, but we expect that increasing the sampling to more structures should improve the agreement with the experiments further.

The R_g values of the liprotides and single aLA molecules determined with simulations and experiments suggest that binding in the simulations between aLA and OA might be overestimated and caution against too literal interpretation of different steps in the process. Nevertheless, the simulation still gives a good idea of the actual structures of the liprotides as well as

giving insight into the binding of aLA to OA. The rather slow unfolding of aLA upon OA interaction also helps explain why heating during formation of liprotides can facilitate liprotide formation. Generally, formation of liprotides requires unfolding of the protein to initiate hydrophobic contacts, as well as interaction between several proteins and OA for covering most of the otherwise unstable OA micelle.

Experimental Section

Materials: Ca^{2+} -depleted α -lactalbumin (apo-aLA, $\geq 85\%$ pure) from bovine milk and oleate (OA, $\geq 95\%$ pure) were from Sigma-Aldrich.

Liprotide formation: Liprotides (lip4) were prepared by mixing protein (2 mg mL^{-1}) with OA (0.6 mg mL^{-1} , 1:15 aLA:OA molar ratio) in NaHPO_4 (50 mM)/NaCl (150 mM) at pH 7.4 and heating at 80°C for 15 min. Samples were spun at $15000g$ for 8 min and kept at room temperature before use. The samples named lip7 and lip8 were formed as described in ref. [14], with aLA (6 mg mL^{-1}) being mixed with OA (6.3 mM) in glycine (pH 8.5, 10 mM) and heating at 80°C for 15 min. After cooling, the pH was changed to pH 7.4 (Lip7) or pH 8.0 (Lip8) by using NaHPO_4 (50 mM) together with appropriate amounts of NaOH or HCl from 1 M stocks.

SAXS: Liprotides were measured at a concentration of 2 mg mL^{-1} with an optimised NanoSTAR SAXS instrument (Bruker, at Aarhus University), which uses a liquid Ga metal jet X-ray source (Excillum) and homebuilt scatterless pinholes. The wavelength was $\lambda = 1.34 \text{ \AA}$ and the acquisition time 10 min. Water at 20°C was used for calibration. Background subtraction and conversion to absolute scale was achieved with the SUPERSAXS program package (C. L. P. Oliveira, and J.S.P., unpublished), according to standard data treatment methods.^[26] Background subtraction was carried out by use of the measured intensity levels behind the semitransparent beam-stop. The data are displayed as a function of q , the modulus of the scattering vector. By using the indirect Fourier transformation procedure we were able to obtain the pair distance distribution function $p(r)$,^[27] a model-independent procedure that gives real-space distance information.

The liprotides were modelled by use of a core-shell model as described previously.^[4a] In this model the core consists of the hydrocarbon part of OA whereas the shell consists of partially disordered protein. We used this model to estimate the amounts of protein and lipid in a sample and to provide constraints for aLA:OA stoichiometry in simulations.

Isothermal titration calorimetry (ITC): Binding of OA to aLA was monitored with a VP isothermal titration calorimeter (MicroCal, Northampton, USA). OA (16 mM) in glycine (pH 10, 50 mM)/NaCl (150 mM) was titrated into aLA ($0\text{--}1.5 \text{ mg mL}^{-1}$, 1.4 mL) in glycine (pH 10, 50 mM)/NaCl (150 mM) at 37°C . The solutions were degassed (vacuum) before measurements. An initial $4 \mu\text{L}$ injection was followed by 26 injections of $10 \mu\text{L}$ each. Three different transition points were determined at different aLA:OA ratios. Raw heat signals obtained from the experiments were integrated by use of the Origin software provided by MicroCal. The first and second transition points were determined by locally fitting all data series to a polynomial around the transition and determining maxima and minima, respectively. The third transition was determined for each series by determining the point at which signal intensity (enthalpy of injection) was halfway between point 2 and the end level (set to 0 kcal mol^{-1}). The relationship $[\text{OA}]_{\text{tot}} = [\text{OA}]_{\text{free}} + n_{\text{OA, bound}} \times$

[aLA] for the dependence of the transitions for the different aLA concentrations was used to determine free and bound [OA] at a given transition.

MD simulations: All simulations were run with use of GROMACS (v. 4.6.5) software package.^[28] Simulations used the MARTINI force field (v. 2.2P)^[29] with small modifications. Because the MARTINI force field was not originally developed for proteins free in solution but instead for proteins imbedded in a lipid membrane, it was necessary to modify the protein beads to avoid aggregation. This was done by changing the backbone beads of the protein to smaller S-type beads (except Ala and charged termini) and also changing the side chain beads of Lys and Arg to S-type (Djurre de Jong, unpublished). Note that this modification only affects backbone–backbone aggregation, whereas backbone interactions with OA or side chains remain unchanged. An alternative solution to prevent protein aggregation by specifically increasing backbone–water interaction by 0.35 kJ mol^{-1} was attempted, but this changes the water–lipid partitioning, causing loss of aLA–OA binding. Whenever relevant, the EINEDyn approach was used to stabilise the protein's tertiary structure.^[30] The standard values of the force constant $F_c = 500 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ and cutoff = 0.9 nm were used, defining a network of elastic bonds between backbone beads. Simulations of proteins using the MARTINI force field often require EINEDyn to control the overall shape of the molecule. EINEDyn does this by forcing an elastic network on the structure to prevent unfolding of the protein. In our main simulations, we did not include EINEDyn and we thereby obtained a protein that was more flexible than in the all-atom simulation. Subsequently, we performed simulations employing the EINEDyn force field to find binding spots on the outside of the protein.

We used the calcium-depleted aLA crystal structure from PDB ID 1F6R (using chain A with reconstruction of the missing atoms CG, CD, CE, and NZ in Lys122) as our starting structure. The Martinize script^[29] was used to convert the PDB structure into the CG MARTINI representation, with DSSP^[31] being used to determine the secondary structure. In MARTINI, the secondary structure of the backbone is encoded in the CG bead types and was therefore fixed during all simulations. OA was built by combining the hydrocarbon tail from the DOPC^[32] molecule with the polarisable sidechain group of Asp^[29] from the standard MARTINI force field parameters.

To estimate the aggregation number of OA, we set up a system consisting of 500 OA molecules and 30913 water beads in a $16.4 \times 16.4 \times 16.4 \text{ nm}^3$ box that was simulated for $1 \mu\text{s}$. The system was equilibrated with NaCl (150 mM) with use of the standard MARTINI Na^+ and Cl^- ions. On the basis of the result, simulations used a micelle consisting of 36 OA molecules, which was created by simulating 36 OA molecules in a $7 \times 7 \times 7 \text{ nm}^3$ box for 100 ns.

Five independent simulations, each involving four aLA molecules and a micelle consisting of 36 OA molecules, were carried out. The proteins and micelle were randomly distributed in an $11 \times 11 \times 11 \text{ nm}^3$ truncated octahedron box with periodic boundaries. The box was filled with water by use of polarisable MARTINI water, in which four water molecules are represented by one CG bead and two dummy atoms carrying partial charges.^[33] The system was equilibrated with NaCl (150 mM) by use of the standard MARTINI Na^+ and Cl^- ions. The temperature was maintained at 20°C by using the velocity rescaling thermostat with $\tau_T = 1.0 \text{ ps}$, and the pressure was kept constant at 1.0 bar by using the Berendsen barostat.^[34] The total run time for the simulation was $3 \mu\text{s}$. Common parameter settings as defined in de Jong et al.^[35] were used with an integration step of 20 fs.

Additionally, four similar systems of OA and aLA were made with the only difference being that the tertiary structure of aLA was stabilised by EINEDyn to follow interactions of native aLA with the micelle. All the systems containing OA and aLA consisted of 11131–11354 CG water beads.

The accessible surface area of aLA, the distance between OA and aLA, the root mean square fluctuation (RMSF), the radius of gyration (R_g) and the cluster sizes of OA were analysed with standard GROMACS tools (g_sas, g_mindist, g_rmsf, g_gyrate and g_clustsize). RMSF and R_g of aLA with the CG force field were determined as an average of the first 10 ns or the last 500 ns of simulation. Analysis of binding of OA to aLA was based on simulations acquired in the 2–3 μs time window.

Relevant structures were translated into an all-atom structure by using the Backward script.^[20] The OA mapping for the Backward script was constructed from CHARMM27 mapping coordinates. The back-mapped structures were simulated for a further 20 ns in a $10 \times 10 \times 10 \text{ nm}^3$ box to check their stability in an all-atom simulation using the CHARMM27 force field.^[36] Visualisations of obtained structures were created with VMD.^[37]

Conversion of MD simulation into SAXS spectra: Snapshots from the trajectory of the CG simulation were taken at intervals of 20 ns during the 500 ns directly after binding of the proteins to the micelle, to sample over a somewhat larger range of slightly different structures. The CG structures were converted into all-atom structures and equilibrated, and the water and ions were removed. To convert the structures into SAXS spectra, we used the excess electron density of aLA and OA. The standard contrast of $2.00 \times 10^{10} \text{ cm g}^{-1}$ was used for the protein whereas the volume of OA was set to $520 \text{ \AA}^3$ with the tail volume being $480 \text{ \AA}^3$ as described.^[4a] The total excess scattering length densities from the three components were distributed evenly on non-hydrogen atoms. An overall scale factor and a constant background were used when optimising the agreement between simulation and experimental scattering curves. An average of the obtained SAXS spectra from the 26 structures was used to represent the structural ensemble of the complex. The shown spectra are based on reduced χ^2 values.

Acknowledgements

This work is funded by the Danish Innovation Foundation (DFORT; to J.N.P. and D.E.O.). We are grateful for support from the ERC, Netherlands Organisation for Scientific Research (P.W.J.M.F., Veni 722.015.005) and the Dutch Ministry of Education, Culture and Science (Gravitation Program 024.001.035).

Conflict of Interest

The authors declare no conflict of interest.

Keywords: alpha-lactalbumin • fatty acids • molecular dynamics • oleate • proteins

- [1] E. D. Chrysina, K. Brew, K. R. Acharya, *J. Biol. Chem.* **2000**, 275, 37021–37029.
- [2] D. A. Dolgikh, R. I. Gilmanshin, E. V. Brazhnikov, V. E. Bychkova, G. V. Semisotnov, S. Y. Venyaminov, O. B. Ptitsyn, *FEBS Lett.* **1981**, 136, 311–315.
- [3] A. Håkansson, B. Zhivotovsky, S. Orrenius, H. Sabharwal, C. Svanborg, *Proc. Natl. Acad. Sci. USA* **1995**, 92, 8064–8068.

- [4] a) J. D. Kaspersen, J. N. Pedersen, J. G. Hansted, S. B. Nielsen, S. Sakthivel, K. Wilhelm, E. L. Nemashkalova, S. E. Permyakov, E. A. Permyakov, C. L. Pinto Oliveira, L. A. Morozova-Roche, D. E. Otzen, J. S. Pedersen, *ChemBioChem* **2014**, *15*, 2693–2702; b) M. Svensson, A. K. Mossberg, J. Pettersson, S. Linse, C. Svanborg, *Protein Sci.* **2003**, *12*, 2805–2814; c) S. E. Permyakov, E. L. Knyazeva, L. M. Khasanova, R. S. Fadeev, A. P. Zhadan, H. Roche-Hakansson, A. P. Hakansson, V. S. Akatov, E. A. Permyakov, *Biol. Chem.* **2012**, *393*, 85–92.
- [5] M. Svensson, H. Sabharwal, A. Hakansson, A. K. Mossberg, P. Lipniunas, H. Leffler, C. Svanborg, S. Linse, *J. Biol. Chem.* **1999**, *274*, 6388–6396.
- [6] a) B. Spolaore, O. Pinato, M. Canton, M. Zamboni, P. Polverino de Laureto, A. Fontana, *Biochemistry* **2010**, *49*, 8658–8667; b) T. Kamijima, A. Ohmura, T. Sato, K. Akimoto, M. Itabashi, M. Mizuguchi, M. Kamiya, T. Kukawa, T. Aizawa, M. Takahashi, K. Kawano, M. Demura, in *Biochem. Biophys. Res. Commun.*, Vol. 376, United States **2008**, pp. 211–214; c) M. Zhang, F. Yang, Jr., F. Yang, J. Chen, C. Y. Zheng, Y. Liang, in *Chem. Biol. Interact.*, Vol. 180, Ireland **2009**, pp. 131–142; d) S. E. Permyakov, E. L. Knyazeva, M. V. Leonteva, R. S. Fadeev, A. V. Chekanov, A. P. Zhadan, A. P. Hakansson, V. S. Akatov, E. A. Permyakov in *Biochimie*, Vol. 93, 2011 Elsevier Masson SAS, France, **2011**, pp. 1495–1501.
- [7] a) J. N. Pedersen, H. S. Frislev, J. S. Pedersen, D. E. Otzen, *J. Dairy Sci.* **2016**, *99*, 7755–7767; b) J. N. Pedersen, J. S. Pedersen, D. E. Otzen, *Biochemistry* **2015**, *54*, 4815–4823.
- [8] A. Rendón, D. G. Carton, J. Sot, M. Garcia-Pacios, L. R. Montes, M. Valle, J. L. Arrondo, F. M. Goni, K. Ruiz-Mirazo, *Biophys. J.* **2012**, *102*, 278–286.
- [9] J. C. Ho, P. Storm, A. Rydstrom, B. Bowen, F. Alsin, L. Sullivan, I. Ambite, K. H. Mok, T. Northen, C. Svanborg, *J. Biol. Chem.* **2013**, *288*, 17460–17471.
- [10] a) D. N. Lu, Z. Liu, J. Z. Wu, *J. Chem. Phys.* **2007**, *126*, 064906; b) A. A. Walker, A. C. Warden, H. E. Trueman, S. Weisman, T. D. Sutherland, *J. Mater. Chem. B* **2013**, *1*, 3644–3651; c) V. Krishnamani, J. K. Lanyi, *Biochemistry* **2012**, *51*, 1061–1069.
- [11] a) A. Chaudhuri, X. Prasanna, P. Agiru, H. Chakraborty, A. Rydstrom, J. C. Ho, C. Svanborg, D. Sengupta, A. Chattopadhyay, *Sci. Rep.* **2016**, *6*, 35015; b) Y. Delgado, M. Morales-Cruz, C. M. Figueroa, J. Hernandez-Roman, G. Hernandez, K. Griebenow, *FEBS Open Bio* **2015**, *5*, 397–404; c) J. Fast, A. K. Mossberg, H. Nilsson, C. Svanborg, M. Akke, S. Linse, *FEBS Lett.* **2005**, *579*, 6095–6100.
- [12] a) H. I. Ingólfsson, C. A. Lopez, J. J. Uusitalo, D. H. de Jong, S. M. Gopal, X. Periole, S. J. Marrink, *Wires Comput. Mol. Sci.* **2014**, *4*, 225–248; b) S. J. Marrink, D. P. Tieleman, *Chem. Soc. Rev.* **2013**, *42*, 6801–6822.
- [13] A. Fontana, B. Spolaore, P. Polverino de Laureto, *Biochim. Biophys. Acta Proteins Proteomics* **2013**, *1834*, 1125–1143.
- [14] J. Nedergaard Pedersen, J. Skov Pedersen, D. E. Otzen, *Protein Sci.* **2017**, DOI: 10.1002/pro.3337.
- [15] D. E. Otzen, P. Sehgal, P. Westh, *J. Colloid Interface Sci.* **2009**, *329*, 273–283.
- [16] B. M. Davis, J. L. Richens, P. O'Shea, *Biophys. J.* **2011**, *101*, 245–254.
- [17] J. J. Janke, W. F. D. Bennett, D. P. Tieleman, *Langmuir* **2014**, *30*, 10661–10667.
- [18] D. T. King, D. B. Warren, C. W. Pouton, D. K. Chalmers, *Langmuir* **2011**, *27*, 11381–11393.
- [19] E. M. Rath, A. P. Duff, A. P. Hakansson, R. B. Knott, W. B. Church, *Proteins Struct. Funct. Bioinf.* **2014**, *82*, 1400–1408.
- [20] T. A. Wassenaar, K. Pluhackova, R. A. Bockmann, S. J. Marrink, D. P. Tieleman, *J. Chem. Theory Comput.* **2014**, *10*, 676–690.
- [21] A. C. Stark, C. T. Andrews, A. H. Elcock, *J. Chem. Theory Comput.* **2013**, *9*, 4176–4185.
- [22] K. Kuwajima, *FASEB J.* **1996**, *10*, 102–109.
- [23] M. Kataoka, K. Kuwajima, F. Tokunaga, Y. Goto, *Protein Sci.* **1997**, *6*, 422–430.
- [24] D. Svergun, C. Barberato, M. H. J. Koch, *J. Appl. Crystallogr.* **1995**, *28*, 768–773.
- [25] O. Halskau, N. A. Froystein, A. Muga, A. Martinez, *J. Mol. Biol.* **2002**, *321*, 99–110.
- [26] P. Lindner in *Neutrons, X-rays and Light: Scattering Methods Applied to Soft Condensed Matter* (Eds.: P. Lindner, T. Zemb), Elsevier, Amsterdam, **2002**, p. 23–48.
- [27] O. Glatter, *J. Appl. Crystallogr.* **1977**, *10*, 415–421.
- [28] H. J. C. Berendsen, D. Vanderspoel, R. Vandrunen, *Comput. Phys. Commun.* **1995**, *91*, 43–56.
- [29] D. H. de Jong, G. Singh, W. F. D. Bennett, C. Arnarez, T. A. Wassenaar, L. V. Schafer, X. Periole, D. P. Tieleman, S. J. Marrink, *J. Chem. Theory Comput.* **2013**, *9*, 687–697.
- [30] X. Periole, M. Cavalli, S. J. Marrink, M. A. Ceruso, *J. Chem. Theory Comput.* **2009**, *5*, 2531–2543.
- [31] W. Kabsch, C. Sander, *Biopolymers* **1983**, *22*, 2577–2637.
- [32] T. A. Wassenaar, H. I. Ingólfsson, R. A. Bockmann, D. P. Tieleman, S. J. Marrink, *J. Chem. Theory Comput.* **2015**, *11*, 2144–2155.
- [33] S. O. Yesylevsky, L. V. Schafer, D. Sengupta, S. J. Marrink, *PLoS Comput. Biol.* **2010**, *6*, e1000810.
- [34] H. J. C. Berendsen, J. P. M. Postma, W. F. Vangunsteren, A. Dinola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684–3690.
- [35] D. H. de Jong, S. Baoukina, H. I. Ingólfsson, S. J. Marrink, *Comput. Phys. Commun.* **2016**, *199*, 1–7.
- [36] a) S. E. Feller, A. D. MacKerell, *J. Phys. Chem. B* **2000**, *104*, 7510–7515; b) A. D. MacKerell, M. Feig, C. L. Brooks, *J. Comput. Chem.* **2004**, *25*, 1400–1415; c) J. B. Klauda, R. M. Venable, J. A. Freites, J. W. O'Connor, D. J. Tobias, C. Mondragon-Ramirez, I. Vorobyov, A. D. MacKerell, R. W. Pastor, *J. Phys. Chem. B* **2010**, *114*, 7830–7843.
- [37] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph.* **1996**, *14*, 33–38.

Manuscript received: September 12, 2017

Accepted manuscript online: November 20, 2017

Version of record online: January 9, 2018