Supporting Information

The lipid-chaperone hypothesis in mechanisms of membrane damage by intrinsically disordered proteins

Michele F. Sciacca[1,9], Fabio Lolicato[2,3,9], Carmelo Tempra[4,5,9], Federica Scollo[5,6,9], Bikash R. Sahoo[7], Matthew D. Watson[8], Sara García-Viñuales,[1] Danilo Milardi[1], Antonio Raudino[5], Jennifer C. Lee[8], Ayyalusamy Ramamoorthy[7] and Carmelo La Rosa[5,*]

[1]Istituto di Cristallografia, CNR, Catania, Italy

[2]Heidelberg University Biochemistry Center, Heidelberg, Germany

[3]Department of Physics, University of Helsinki, P.O. Box 64, FI-00014 Helsinki, Finland

[4]Institute of Organic Chemistry and Biochemistry, Prague, Czech Republic

[5]Department of Chemical Sciences, University of Catania, Catania, Italy

[6]J. Heyrovský Institute of Physical Chemistry, Academy of Sciences of the Czech Republic, Prague, Czech Republic

[7]Biophysics and Department of Chemistry, University of Michigan, Ann Arbor, Michigan

[8]National Institutes of Health, Bethesda, Maryland

[9]These authors contributed equally: Michele F. Sciacca, Fabio Lolicato, Carmelo Tempra, Federica Scollo

A narrative overview of this manuscript

The mechanism of amyloid formation *in vivo* can be quite different from that observed in dilute aqueous solution. Thus, studies of amyloidogenic IDPs in heterogeneous water/membrane environments may provide more significant advances in our understanding of the pathogenic mechanisms of these proteins. It has been often observed that the rate of amyloid fibril formation upon the interaction with membrane is enhanced. Furthermore, IDP oligomers may also induce membrane damage, which is believed to play a crucial role in pathogenesis during the early stages of protein misfolding diseases.

Many laboratories investigate the lipids-IDPs-water ternary systems by using lipid bilayers as a customary model system for real membranes. Our recent studies suggest investigate, besides the lipid-peptides interactions within the membrane (where the local lipid concentration is high), also the interactions in water phase where the concentration of free lipids is extremely small. It is well-known that lipids in water solutions have a high tendency to self-aggregate, forming bilayer structures. Although most lipids reside in the bilayer forming a quasi-2D thin oily phase, a sizeable amount of lipids (the concentration of which ranges from nM to μM and depends on lipid hydrophobic length and headgroups charge) remains in the aqueous phase. These ubiquitous *free lipids* are in dynamic equilibrium with those belonging to the self-assembled bilayers and their concentration is termed as the Critical Micelle Concentration (CMC). CMCs of several lipid systems are reported in the literature[1] . An important property of the CMC is that it self-regulates according to the usual laws of chemical equilibrium.

The main ideas reported in this article stem from a failure when trying to simulate by Molecular Dynamics (MD) the kinetics of insertion of IDPs into lipid bilayers. Simulations were performed using different force-fields, long simulation times (more than 1 μs for each simulation) and different lipid composition. Although several experimental evidences support the notion of a rapid protein insertion in the lipid matrix, repeated simulation runs were unable to reproduce the experimentally observed protein transfer from water to the lipid phase. Notably, once proteins had been inserted in the lipid core, they remain indefinitely stable. These apparent failures raised our concern that something was missing. Indeed, we repeated the MD simulations using *pre-formed* lipid-protein complexes in water instead of the bare proteins. Two exciting outcomes were observed: (a) IDPs (hIAPP in those simulations) and lipids form stable complexes in water, and (b) the lipid-protein complexes ready penetrate the lipid bilayers. These preliminary MD results suggested a possible chaperone-like role of the free lipids, as reported in a recent note[2]. Shortly later, an experimental model based on the use of fluorescent dyes to probe the IDPs-induced damage of lipid bilayers was

published. What is the main message stemming from our unsuccessful simulations? In our opinion, MD results highlight the key role of different time scales in this kind of complex kinetics which is controlled by a "rare event" In the present case the rare event is the encounter probability among IDPs and free lipids in the aqueous phase because both IDPs and free lipids concentrations are low in water. On the other hand, once the lipid-protein complex is formed, its insertion in the lipid matrix is relatively fast. Thus, even if the initial events (complexations) are rare, the following events are *far reaching*, a behavior typical of kinetics controlled by "rare events" (for a comprehensive analysis see, e.g., Introduction to Rare Event Simulation, J. Bucklew, 2004, Spinger-Verlag New York **ISBN:** 9780387200781. Although the above preliminary data were promising, our hypothesis deserves to be validated by a lot of experimental and theoretical checks. In particular, we have to ascertain the following facts:

A) Do lipid-protein complexes exist in water phase?

B) How stable are the complexes? (because of the low water concentration of free lipids (ranging from nM to μM) and proteins (of order of μM for IDPs) in water, the concentration of complexes is crucial for an efficient insertion kinetics).

C) Does the complex(es) exhibit enhance hydrophobicity in respect to the bare protein? (otherwise their formation is useless for protein insertion in the lipid matrix).

D) Can the lipid-assisted flux of proteins determine the kind and extent of membrane damage?

E) Is this a general mechanism or is it restricted to a limited number of proteins?

F) May a chemical change in the lipid structure (e.g., by lipid peroxidation, a ubiquitous phenomenon associated to inflammatory processes) modify the lipid CMC and therefore their ability to transfer proteins in the membrane core?

The answer to the above questions is mandatory prior to validate our hypotheses. In this paper we addressed the problem by different experimental techniques (Circular Dichroism (CD) spectroscopy, Isothermal Calorimetry (ITC), 2D NMR and fluorescent labels measurements) on several IDPs involved in different misfolding pathologies. The interpretation of the experimental data was supported by MD simulations. Last but not least, we compared our results with a vast literature on membrane-IDPs interactions.

Table S1. Simulation details of free protein and lipid in solution systems.

System	Molecule	# Water	# Na	# Cl	# Repeat	Tot. Time (ns)
P1	hIAPP	9602	27	29	1	500
P2	rIAPP	9602	27	29	1	500
P3	Aβ 1-40	8734	19	16	1	500
P4	Aβ 1-42	17155	35	32	1	500
L1	PC14	10329	0	0	1	500
L2	DEPC22	10329	0	0	1	500
L3	DMPC14	10329	0	0	1	500

Table S2. Simulation details of protein-lipid complex systems.

System	# Water	# Na	# Cl	# Repeat	Tot. Time (μs)
P1-L1 / P1-L2	9334	17	19	3	3 (3x1 μs)
P2-L1 / P1-L2	9364	16	19	3	3 (3x1 μs)
P3-L1 / P1-L2 / P3-L3	8457	18	15	3	3 (3x1 μs)
P4-L1 / P1-L2	16637	33	30	3	3 (3x1 μs)

Evaluation of lipid-protein complex hydrophobicity index

Table S3. Amino acid Eisenberg's hydrophobic index taken from supporting ref. 2.

Ala: 0.620
Arg: -2.530
Asn: -0.780
Asp: -0.900
Cys: 0.290
Gln: -0.850
Glu: -0.740
Gly: 0.480
His: -0.400
Ile: 1.380
Leu: 1.060
Lys: -1.500
Met: 0.640
Phe: 1.190
Pro: 0.120
Ser: -0.180
Thr: -0.050
Trp: 0.810
Tyr: 0.260
Val: 1.080

Evaluation of CH_2 hydrophobic index

Mytilene hydrophobic index was evaluated as follow: since in the literature there is no hydrophobicity index for the methylenes of the hydrocarbon chains of lipids and considering that the values of the amino acids used in this work (Eisenberg) are relative values, we considered the principle of additivity of the hydrophobicity indexes of Eisenberg. In fact, the additive is satisfied for neutral or slightly charged amino acids. ALA and GLY as well as GLU and ASP differ for a CH2, therefore from subtraction we can obtain the contribution of CH2:

ALA-GLY=0.62-0.48=0.14

GLU-ASP=-0.74-(-0.90)=0.16

In our calculations we used the average value of 0.15.

Table S4. Example of how calculate the hydrophobicity index of bare protein (Aβ40) and protein-lipid complex (C14 and C22). Total protein was calculated as the sum of Eisenberg hydrophobicity index and protein-lipid complex was calculated using equation 1. First column: Protein amino acid sequence; Second column; Eisenberg hydrophobic index; Third Probability contact between i-amino acid and C14 lipid; Fourth column: Probability contact between i-amino acid and C22 lipid; Fifth and sixth column: hydrophobic index of lipid-protein complex of i-amino acid.

		C14	C22	Cpx(C14)	Cpx(C22)
Sequence	**Heisenberg**	**Probability**	**Probability**		
ASP	-0,90	0,20	0,30	-0,87	-0,86
ALA	0,62	0,24	0,38	0,66	0,68
GLU	-0,74	0,32	0,41	-0,69	-0,68
PHE	1,19	0,19	0,60	1,22	1,28
ARG	-2,53	0,18	0,42	-2,50	-2,47
HIS	-0,40	0,26	0,40	-0,36	-0,34
ASP	-0,90	0,38	0,27	-0,84	-0,86
SER	-0,18	0,53	0,13	-0,10	-0,16
GLY	0,48	0,67	0,97	0,58	0,63
TYR	0,26	0,65	0,38	0,36	0,32
GLU	-0,74	0,37	0,17	-0,68	-0,71
VAL	1,08	0,37	0,17	1,14	1,11
HIS	-0,40	0,27	0,30	-0,36	-0,36
HIS	-0,40	0,23	0,23	-0,37	-0,37
GLN	-0,85	0,23	0,31	-0,82	-0,80
LYS	-1,50	0,22	0,51	-1,47	-1,42
LEU	1,06	0,51	0,70	1,14	1,17
VAL	1,08	0,62	0,82	1,17	1,20
PHE	1,19	0,48	0,84	1,26	1,32
PHE	1,19	0,31	0,76	1,24	1,30
ALA	0,62	0,33	0,50	0,67	0,70

GLU	-0,74	0,24	0,40	-0,70	-0,68
ASP	-0,90	0,42	0,31	-0,84	-0,85
VAL	1,08	0,20	0,33	1,11	1,13
GLY	0,48	0,34	0,22	0,53	0,51
SER	-0,18	0,28	0,22	-0,14	-0,15
ASN	-0,78	0,31	0,29	-0,73	-0,74
LYS	-1,50	0,40	0,37	-1,44	-1,44
GLY	0,48	0,52	0,38	0,56	0,54
ALA	0,62	0,57	0,52	0,71	0,70
ILE	1,38	0,64	0,78	1,476	1,497
ILE	1,38	0,33	0,81	1,4295	1,5015
GLY	0,48	0,25	0,66	0,5175	0,579
LEU	1,06	0,24	0,72	1,096	1,168
MET	0,64	0,28	0,7	0,682	0,745
VAL	1,08	0,3	0,28	1,125	1,122
GLY	0,48	0,69	0,23	0,5835	0,5145
GLY	0,48	0,47	0,28	0,5505	0,522
VAL	1,08	0,39	0,34	1,1385	1,131
VAL	1,08	0,28	0,28	1,122	1,122
Total	**6,93**			**9,13605**	**9,5835**

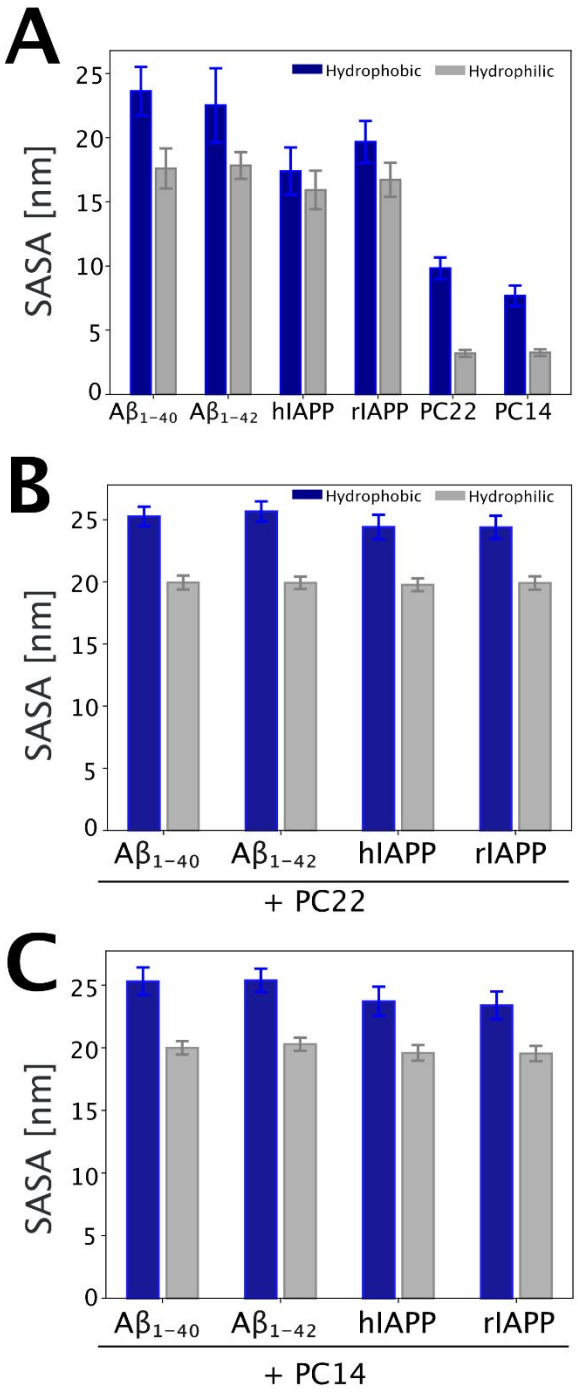

Figure S1. Solvent accessible surface area (SASA) analysis for **(A)** proteins and lipids alone; **(B)** proteins in the presence of a single molecule of PC22 lipid type and **(C)** proteins in the presence of a PC14 lipid type. The first 100 ns of each simulation was discarded from the analysis.

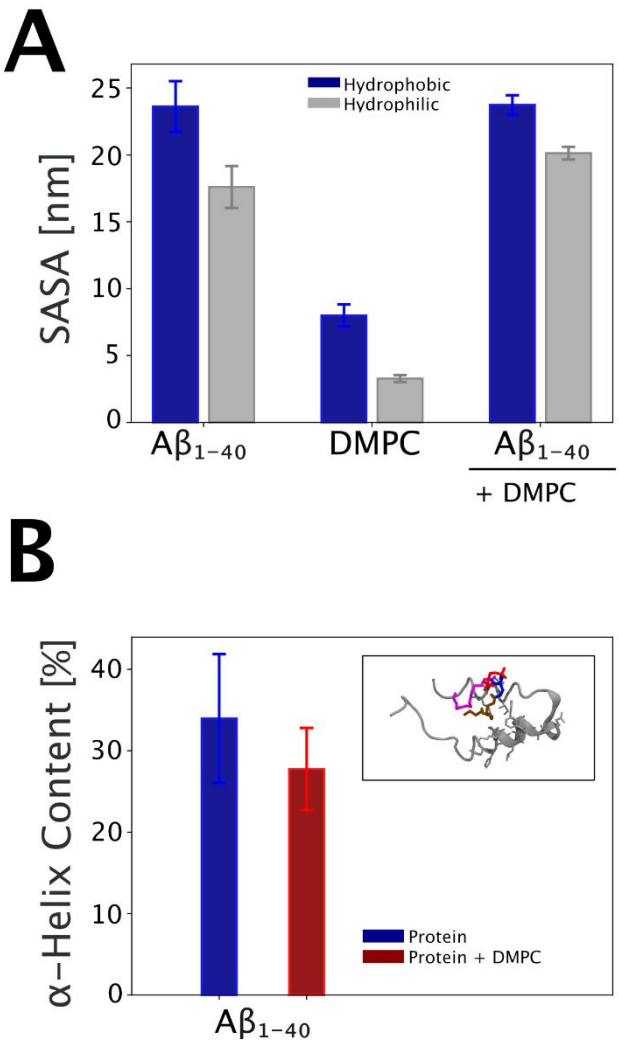

Figure S2. Structural analysis of protein's α-helix content before and after the interaction of $A\beta_{1\text{-}40}$ with free DMPC lipid. The average is calculated by removing the first 100 ns of all simulations from the calculation. The error bars represent the relative standard errors. Alongside is shown the corresponding 3D central structures of the $A\beta_{1\text{-}40}$-DMPC complex. The central structure is the configuration with the smallest average RMSD from all other structures within the most populated cluster of the $A\beta_{1\text{-}40}$-DMPC interfaces. It is calculated by concatenating the last 100 ns of the three repeats.

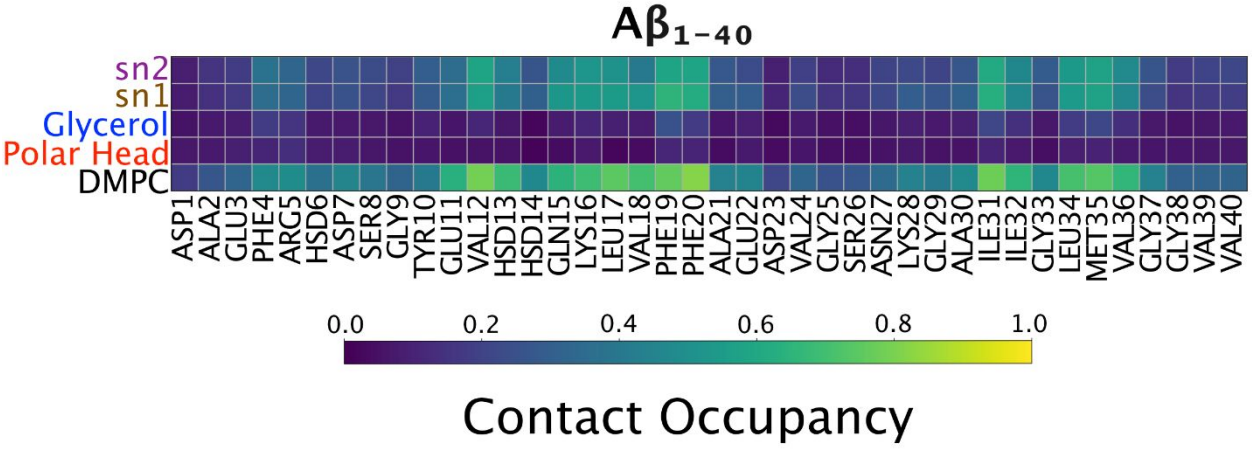

Figure S3. Pairwise contact map with all residues for the of $A\beta_{1\text{-}40}$-DMPC. Contact occupancy equal to 1.0 corresponds to the situation where a given interaction has taken place for all the calculated time. The average was calculated by concatenating the last 900 ns from all the three repeats. A contact is defined if any of the atoms between two groups were closer than 0.6 nm.

ΔSASA lipid-protein calculation.

From molecular dynamics, ΔSASA was calculated as follow:

$\Delta SASA_{Complex}=SASA_{Complex}-(SASA_{Lipid}+SASA_{Protein})$

Results are reported below:

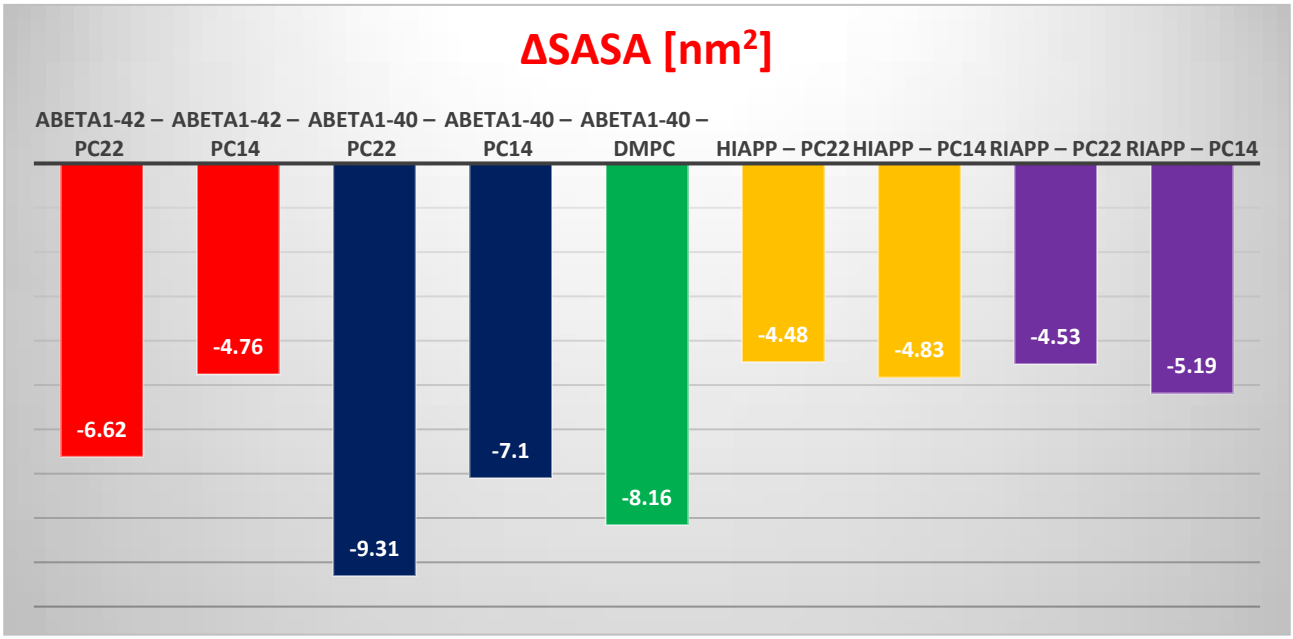

Figure S4. Histogram of ΔSASA of lipid-protein obtained from molecular dynamics simulations.

Phospholipid-Aβ_{1-42} binding constant determination: Isothermal Titration Calorimetry (ITC)

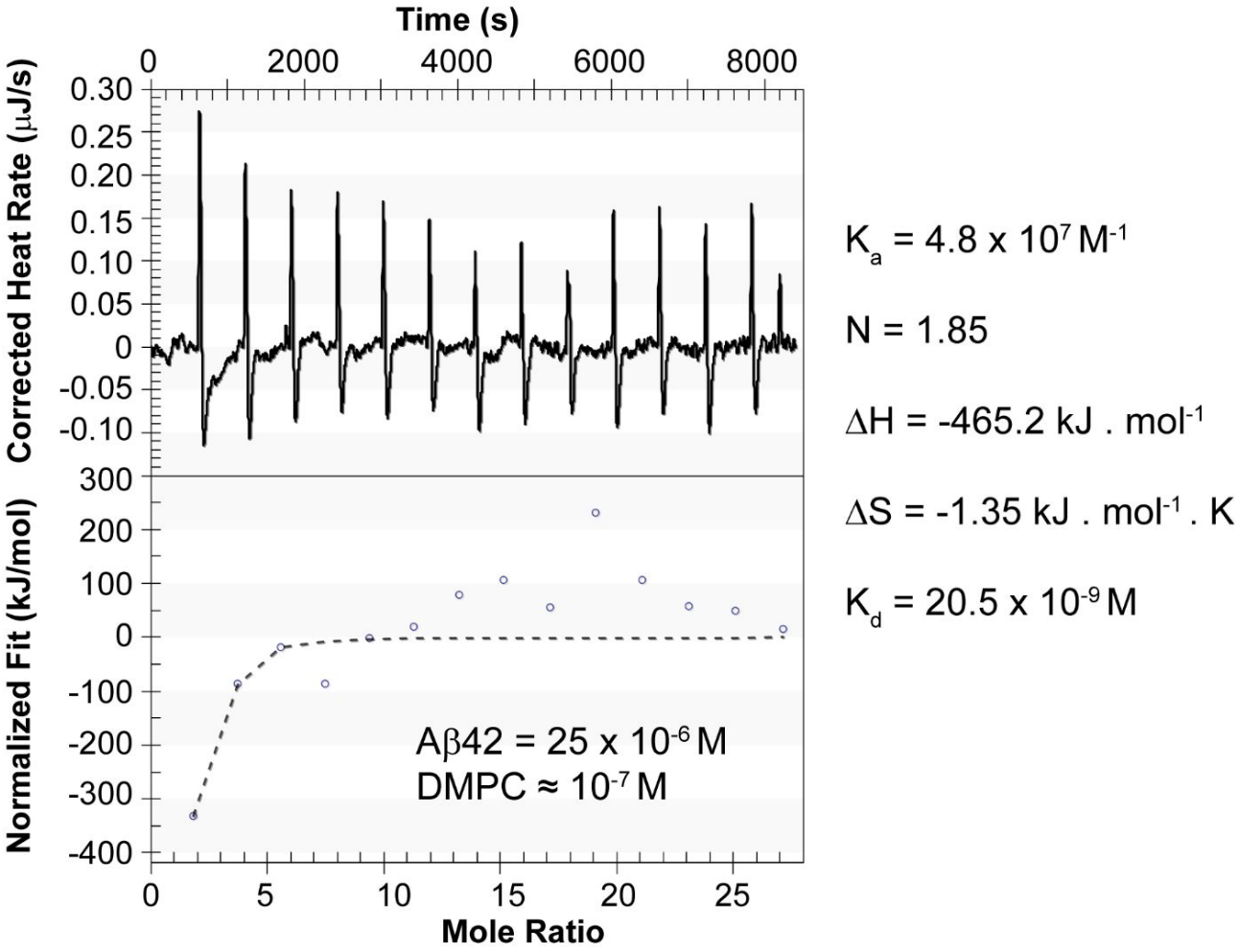

Figure S6. ITC titration of Aβ1-42 and PC 12:1 at its cmc.

Hydrophobic index calculation

The previously developed theory[1] predicts that the formation of the protein-lipid complex is favoured only if the total hydrophobicity of the complex is greater than that of the individual molecules. Since in the lipid-protein complex each methylene group of the lipid acyl chain masks a protein residue, a reasonable estimate of the variation of the protein hydrophobicity upon lipid binding can be calculated by means of the following equation:

$$\Im = \sum_i^{n_{aa}} (\mathcal{H}_i^{aa} + CH2P_i^{aa}) \qquad (1)$$

where $\Im$ is the hydrophobicity index of the protein-lipid complex, $\mathcal{H}_i^{aa}$ is the Eisenberg's hydrophobicity index[2] of the i^{th} amino acid (aa), $CH2$ is the hydrophobicity index of the methylene group of a lipid tail (evaluated as below) and P_i^{aa} is the contact probability between the i^{th} amino acid and lipid acyl chains calculated by MD simulations.

To prove this hypothesis we have calculated the hydrophobicity of the complex and of the hIAPP, rIAPP, $A\beta_{1\text{-}40}$, $A\beta_{1\text{-}42}$, DMPC and DEPC. Results are reported in **Fig. S5**, showing in all cases a consistent increment of hydrophobicity index in the the lipid-protein 1:1 complex. Moreover, our simulation data are in good agreement with NMR experimental results (contact lipid-proteins).

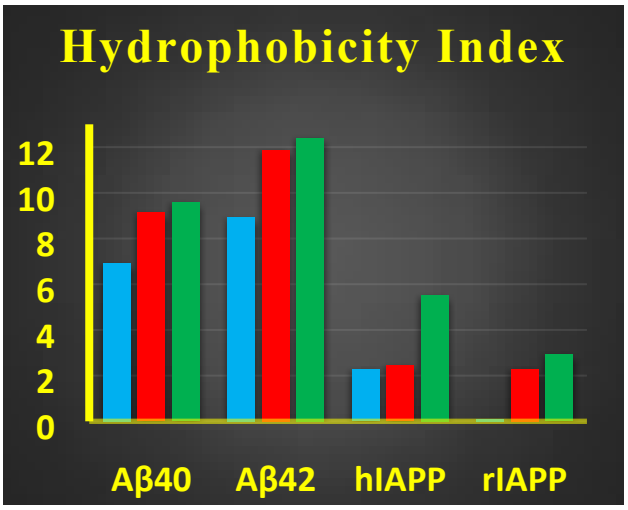

Figure S5. Hydrophobicity index of bare proteins (light blue), protein-DMPC (red, containing 14 carbon atoms per chain) and protein-DEPC (green containing 22 carbon atoms per chains).

Supporting References

(1) La Rosa, C.; Scalisi, S.; Lolicato, F.; Pannuzzo, M.; Raudino, A. Lipid-Assisted Protein Transport: A Diffusion-Reaction Model Supported by Kinetic Experiments and Molecular Dynamics Simulations. *J. Chem. Phys.* **2016**, *144* (18), 184901. https://doi.org/10.1063/1.4948323.

(2) Eisenberg, D.; Schwarz, E.; Komaromy, M.; Wall, R. Analysis of Membrane and Surface Protein Sequences with the Hydrophobic Moment Plot. *J. Mol. Biol.* **1984**, *179* (1), 125–142. https://doi.org/10.1016/0022-2836(84)90309-7.