Solubilization of M2 Transmembrane Peptide of Influenza A in Pure Water: Implications for Emergence of Proteins and Protein-embedded Primeval Membranes in Unsalted Oceans

Jianxing Song[1, 2]* and Linlin Miao[1]

[1]Department of Biological Sciences, Faculty of Science, [2]Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore; 10 Kent Ridge Crescent, Singapore 119260

* Corresponding author: Song, J. (dbssjx@nus.edu.sg)

Abstract

We demonstrated that M2 transmembrane peptide, one of the most hydrophobic sequences in nature, can be solublized to at least ~100 μM in unsalted water without any lipid molecules. Strikingly, the M2 peptide also forms a highly-helical conformation in water which remains almost unchanged even at 95 ºC, as characterized by CD spectroscopy. Our result has critical implications in understanding emergence of proteins and protein-embedded primeval membranes in unsalted oceans.

It remains a great mystery how hydrophobic membrane proteins which are commonly thought to be insoluble in solution found and merged with the lipid molecules to form protein-embedded primitive membrane in evolution. In a TIBS article [1], Mulkidjanian and colleagues hypothesized elegant scenarios for this, but however, they found it enigmatic how the integral membrane proteins could reach the primeval membranes, since these proteins having extremely high hydrophobicity are water-insoluble and consequently "even if occasionally synthesized, would remain stuck in the ribosome" [1]. As a consequence, emergence of protein-embedded primeval membranes is called a "chicken and egg paradox".

On the other hand, we recently discovered that in contrast to the well-established dogma in which protein solubility will be first enhanced at low salt concentration (<500 mM, salting-in), followed by reduction of solubility at high salt concentration (salting-out), all previously-thought "insoluble proteins" could in fact solubilized in unsalted water [2]. So far, our discovery holds

for >50 insoluble proteins in my lab and has now been used by other groups to study "insoluble proteins" [3, 4]. To explore the limit for this discovery, we chemically-synthesized and HPLC-purified the M2 transmembrane peptide of influenza A, which is one of the most hydrophobic sequences in nature, with a sequence of SSDPL VVAAS IIGIL HLILW ILDRL. Previously, the structure of this peptide has been determined with lipid molecules by crystallography [5] and liquid and solid NMR spectroscopy respectively [6, 7]. Each M2 fragment is a well-formed helix and further assembles into a tetramer (Figure 1a). Surprisingly, we found that the M2 peptide could be solublized to at least ~100 μM in unsalted water without any lipid molecules. Strikingly, the M2 peptide also forms a highly-helical conformation in water which remains almost unchanged even at 95 ºC, as characterized by CD spectroscopy (Figure 1b).

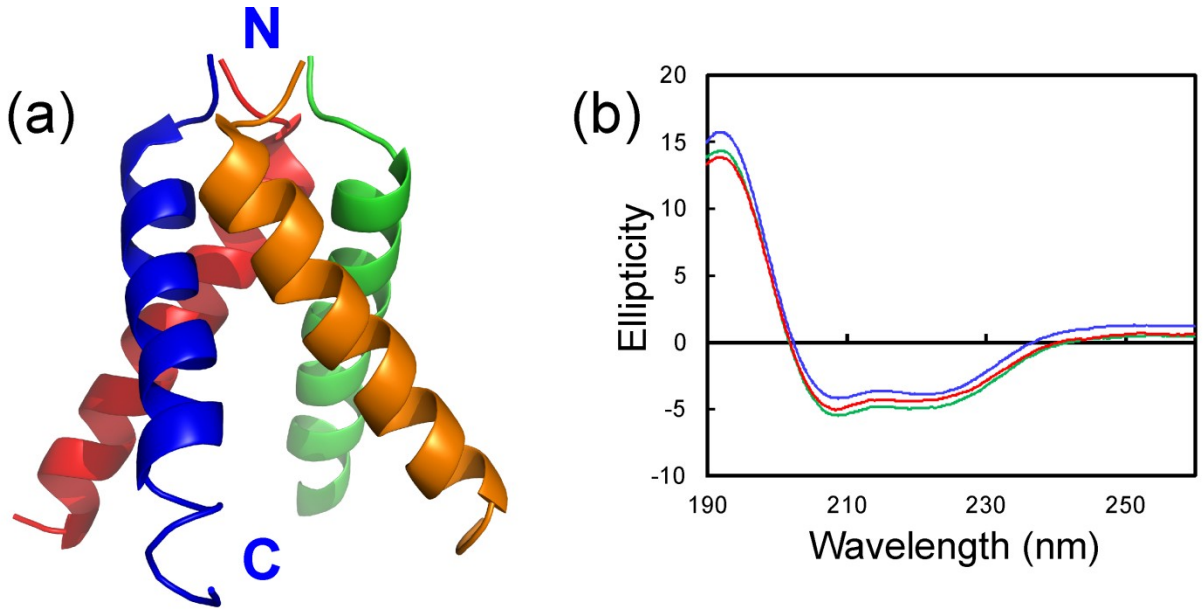

Figure 1. (a) Tetrameric structure of M2 fragment; (b) Far-UV CD spectra of the 25-residue M2 fragment at 20 (blue); 95 (red) and 20 degree after the thermal unfolding.

Together with our previous results, our current discovery may offer a solution to the above chicken and egg paradox [1]. As we demonstrated, protein aggregation triggered by hydrophobic interactions can be successfully suppressed with repulsive electrostatistic interactions generated by having the solution pH several units away from the protein pI, for example, slightly acidic [2]. Amazingly, the solution condition we used to solubilize "insoluble proteins" is very similar to the primitive oceans, which were largely unsalted and slightly acidic [1, 8]. If so, the chicken and egg paradox might be easily solved because in the unsalted oceans,

proteins even with very high hydrophobicity are highly soluble. In particular, the protein concentration might also be extremely diluted in the primitive oceans and consequently almost all proteins were able to freely diffuse around to reach the primeval membranes. Indeed, we have previously demonstrated that the 297-residue Nogo-B receptor with a 20-residue transmembrane fragment was totally insoluble in buffer but could be dissolved to a concentration of at least 500 μM in unsalted water [9].

Our discovery might also shed light on another mysterious issue associated with the diversification of proteins. The space of realized protein folds appears to just account for one-tenth of the space of possible folds [11]. This implies that a large portion of the sequence space remains unexplored in the life forms on Earth. Here we propose that this might also be associated with the ocean salinity. The machinery generating proteins is believed to exist before the emergence of the membrane-enveloped primitive cells [1, 12]. Without membranes, in unsalted oceans proteins created with their sequences highly randomized were all soluble and thus could diffuse freely to encounter other inorganic and organic molecules for forming various functional complexes/machineries. The increase of salt concentrations in oceans might be the major driving force for the assembly of proteins with other macro- and small molecules. The ubiquitous but specific interactions of various anions to proteins have just been revealed with NMR spectroscopy in pure water by my group [13, 14]. Once the membrane-contained cells were formed and the oceans became highly salted, the proteins with high hydrophobicity would be aggregated and might start to cause damages to cells, as exemplified by the modern aggregation-causing neurodegenerative diseases [10]. This eventually led to the emergence of mechanisms to halt the highly-randomized sampling of the sequence and conformational spaces. In this regard, all modern proteins including well-folded; intrinsically-unstructured and membrane-associated might be all diverged from the constrained numbers of the primordial which were randomly created in unsalted oceans. This scenario perfectly rationalizes our previous discovery that "[modern] proteins appear so designed that in pure water their intrinsic repulsive interactions are sufficient to suppress the attractive forces, thus preventing them from severe precipitation/aggregation" [2]. Simply but with beauty, the 20 natural amino acids were selected to make proteins probably because their polymerized forms, proteins, were all soluble in the primitive unsalted oceans.

Acknowledgements

The study is supported by Ministry of Education (MOE) of Singapore Tier 2 Grant MOE2011-T2-1-096 and Singapore National Medical Research Council Grants R- 154-000-382-213 to J. Song.

References

1. Mulkidjanian, A.Y. *et al.* (2009) Co-evolution of primordial membranes and membrane proteins. ***Trends Biochem Sci.*** 34, 206-15

2. Song, J. (2009) Insight into "insoluble proteins" with pure water. ***FEBS Lett.*** 583, 953-9

3. Delak, K. *et al.* (2009) The Tooth Enamel Protein, Porcine Amelogenin, Is an Intrinsically Disordered Protein with an Extended Molecular Configuration in the Monomeric Form. ***Biochemistry.*** 48, 2272-81

4. Aguado-Llera, D. *et al.* (2010) The Basic Helix-Loop-Helix Region of Human Neurogenin 1 Is a Monomeric Natively Unfolded Protein Which Forms a “Fuzzy” Complex upon DNA Binding. ***Biochemistry.*** 49, 1577-89.

5. Schnell, J. R. & Chou, J. J. (2008). Structure and mechanism of the M2 proton channel of influenza A virus. ***Nature,*** 451, 591–595.

6. Stouffer, A. L., Acharya, R., Salom, D., Levine, A. S., Di Costanzo, L., Soto, C. S. et al. (2008). Structural basis for the function and inhibition of an influenza virus proton channel. ***Nature***, 451, 596–599.

7. Cady SD, Mishanina TV, Hong M. (2009) Structure of amantadine-bound M2 transmembrane peptide of influenza A in lipid bilayers from magic-angle-spinning solid-state NMR: the role of Ser31 in amantadine binding. ***J Mol Biol.*** 385, 1127-41.

8. Pinti, D.L. (2005) The origin and evolution of the oceans. In Lectures in Astrobiology (Gargaud, M. et al., eds), 83–111, Springer-Verlag

9. Li, M. and Song, J. (2007) Nogo-B receptor possesses an intrinsically unstructured ectodomain and a partially folded cytoplasmic domain. ***Biochem Biophys Res Commun.*** 360, 128-34

10. Shi, J. Lua S, Tong JS and Song, J (2010) Elimination of the native structure and solubility of the hVAPB MSP domain by the Pro56Ser mutation that causes amyotrophic lateral sclerosis. ***Biochemistry.*** 49, 3887-97

11. Taylor, W.R. *et al.* (2009) Probing the "dark matter" of protein fold space. ***Structure.*** 17, 1244-52

12. Koonin, E.V. and Martin, W. (2005) On the origin of genomes and cells within inorganic compartments. ***Trends Genet.*** 21, 647–654.

13. Miao L, Qin H, Koehl P and Song J. (2011) Selective and specific ion binding on proteins at physiologically-relevant concentrations. *FEBS Lett.* 585, 3126-32.

14. Miao L, Qin H and Song J. (20112) “Dark Mediators” of Proteins as Revealed by NMR in Water: Residue-selective Anion Bindings that are Masked by Pre-existing Buffer" Submitted to ***Nat Chem Biol***, (Nature Precedings: npre.2012.6769.1).