Dynamic lipid aptamers: non-polymeric chemical path to early life

Amit Kahana, Svetlana Maslov, and Doron Lancet, Dept. of Molecular Genetics, the Weizmann Institute of Science, Rehovot 7610001, Israel

Abstract

A widespread dogma asserts that life could not have emerged without biopolymers - RNA and proteins. However, the widely acknowledged implausibility of a spontaneous appearance and proliferation of these complex molecules in primordial messy chemistry casts doubt on this scenario. A proposed alternative is "Lipid-First", based on the evidence that lipid assemblies may spontaneously emerge in heterogeneous environments, are shown to undergo growth and fission, and to portray autocatalytic self-copying. What seems undecided is whether lipid assemblies have protein-like capacities for stereospecific interactions, a *sine qua non* of life processes. This Viewpoint aims to alleviate such doubts, pointing to growing experimental evidence that lipid aggregates possess dynamic surface configurations capable of stereospecific molecular recognition. Such findings help support a possible key role of lipids in seeding life's origin.

The power of lipid diversity

Non-covalent assemblies of monomeric lipid amphiphiles are held together by hydrophobic interactions mediated by apolar tailgroups. The external surface of such assemblies is typically decorated by variegated hydrophilic headgroups. Such structure bears analogy to that of globular proteins, which when folded, expose mostly polar residues on the outside, ensheathing an apolar core[1]. The capacity of protein surfaces to mediate stereospecific binding and catalysis stems from their stable steric configurations, resting on their covalent polymeric sequence and relatively rigid folded conformations. In contrast, non-covalent supramolecular assemblies of lipids are typically described as fluid agglomerates, with individual monomers dynamically changing their positions and mutual interactions. For this reason, lipid assemblies are seldom thought of as having a role in stereospecific binding and catalysis, an outlook that is extrapolated to the first steps in life's emergence[2]. The present Viewpoint challenges this position, reviewing evidence for subtle and complex recognition roles for lipids, and advocating a Lipid-First scenario for the dawn of life[3-5].

Present day cellular membranes contain a chemically diverse assortment of lipids present in varying compositions. Compositional and structural variation is observed from the level of membrane leaflet to that of a whole organism, consistent with the possibility that membrane lipids fulfill many functions beyond mere compartmentalization[6]. Interdisciplinary approaches have revealed the functional underpinnings of such variegation, stemming from an emerging understanding of lipid-lipid and lipid-target interactions[7].

In contemporary life, an alphabet of 20 amino acid residues, differing in size, electric charge and polarity, build all proteins. Notably, the size of the main repertoire of lipid headgroups is comparable to that of proteinaceous amino acids, further augmented by chemical headgroup modifications and head-tail combinatorial diversity. Considering the fluidic nature of lipid assemblies, how could lipid surfaces express interaction capacities? The answer partly rests on the existence of compositionally-variable lipid nanodomains with capacities for selective molecular recognition, as shown in both biological membrane and model lipid assemblies[8].

Membrane nanodomains

Cellular membranes portray lateral heterogeneity, containing discrete lipid domains with different order parameters[9]. Some are fluid and dynamic, while others are more solid, appearing at certain compositions and temperatures[10]. The first indication for such membrane domains was the differential solubility in non-ionic detergents, defining detergent-resistant membranes (DRMs), often known as rafts, in contrast with detergent-soluble membranes (DSMs)[9]. Subsequent experimentation suggested these domains are generated by liquid-liquid phase separation, generating Liquid-ordered (Lo) domains (rafts), coexisting with more fluid Liquid-disordered (Ld) domains[9].

The key driving force for phase separation is preferential interactions within a specific lipid subset, a collective behavior governed by non-covalent interactions, involving both headgroups and tailgroups of the participant lipids[9, 11]. In an example, cholesterols and saturated lipids such as sphingolipids drive the formation of ordered domains[12], and unsaturated lipids are major components of disordered ones. The preferential interaction between sphingolipids and sterols is governed by hydrogen bonding, strengthened by being localized at a water-poor interfacial region of the bilayer[9]. Another relevant example of specific inter-lipid molecular recognition is calcium ion coordination acting in phosphatidylserine-rich domains[13]. Domain formation is also modulated by the chirality of phospholipids, which influences the order and fluidity of lipid bilayers through lateral lipid-lipid contacts[14]. It thus appears that the formation of lipid nanodomains is readily explained by well-defined molecular interactions.

Dynamic aptamers

The dimensions of lipid domains can extend to several hundred nanometers, significantly exceeding the typical arena of stereospecific binding sites. Thus, lateral heterogeneity and potential functionality of lipid patches not much larger than 10 nanometers is worthy of special attention. In this realm, there are experiments proving the existence of ultrananodomains with sizes <5nm, generated by molecular mechanisms that underlie the formation of larger membrane domains[15]. An example is the favorable effect of sphingomyelin in admixture with cholesterol and certain variants of phosphatidylcholine on the preferred formation of ultrananodomains[15].

We explore here the notion that such nanoscopic non-covalent lipid clusters with defined compositions could interact selectively with intramembranous and environmental target molecules. We draw attention to a possible functional analogy of such lipid molecular configurations to aptamers, combinatorial segments of covalently threaded sequences of folded proteins or RNAs that can perform different molecular recognition tasks[16]. The lipid analogs may behave as non-covalent dynamic aptamers[17], capable of specific recognition (Fig. 1).

Experimental support for these proposals comes, for example, from a study on combinatorial micelles[18]. The authors use lipids with 10 different natural amino acid headgroups to generate mixed micelles by "combinatorial synthesis". They envisage their surface as a "dynamic fluid mosaic" array, and point out that the different amino acid headgroups are brought together as closely as they would be if linked covalently. The ensuing micelle library with different surface proto-epitopes (synonymous to dynamic aptamers) can be screened for binding to molecular targets of interest. The authors point out that due to the dynamic fluid nature of the surface building blocks, each micelle presents a multitude of different proto-epitopes, enhancing the probability of successful binding. Indeed, binding affinities 10-100 times higher than background

are observed for specific lipid combinations with target polypeptides such as immunoglobulin G, lysozyme and the bacterial antibiotic bacitracin.

Another study supporting similar concepts[19] measured the lipid-binding affinities of 91 different Pleckstrin Homology (PH) domains from various proteins, using a liposome microarray assay. The results disclose that combinatorial compositions of eight different phosphoinositides along with auxiliary lipids with various headgroup charges, lead to specific affinity values for different PH domains. Further, the study shows that the affinity contributions of different lipid types add non-linearly, provoking cooperativity between lipids in the binding of the PH domain. Finally, small changes in protein sequence among phylogenetically different PH domains can lead to significant lipid recognition changes, suggesting that dynamic aptamers may portray a great measure of fine tuning[6].

Lipid-protein recognition

It appears that lipid-lipid interactions may lead to the formation of relatively stable non-random molecular lipid configurations[20]. It is therefore no wonder that a plethora of studies show that such lipid configurations constitute functional hubs for the initiation and regulation of protein-mediated cellular processes[21]. It appears that ordered lipid domains with specific monomer compositions may recognize proteins with selectivity reminiscent of protein-protein associations. This could be regarded as a conceptual reversal, whereby lipid aptamers appear to play a receptor role, while a target protein is reminiscent of a ligand.

Some of the clearest examples for such recognition are the lipid-binding domains of peripheral membrane proteins, through which the proteins attach to combinations of individual lipids at the bilayer surface. This is demonstrated by multiple methodologies, including crystallography, binding assays and molecular dynamics[20]. That these lipid-binding domain types appear in numerous functionally different proteins attests to a central role of lipids in complex cellular pathways[9, 11, 21]. Specifically, each of these protein domain classes binds to a different composition of lipid headgroups (Fig. 2).

Another case of lipid-protein recognition is observed in lipidated membrane proteins, which manifest domain-specific interactions related to the type of their acyl chain modification. Proteins with saturated lipid modification (e.g. with a palmitoyl group) favorably insert to the ordered raft domains, whereas proteins with branched or unsaturated modifiers such as prenyl anchors prefer disordered (non-raft) regions[9].

We note that the association of proteins to the membrane is not static, and often involves non-covalent rearrangements in both binding partners. This is exemplified by Annexin A5[22], which binds to phosphatidylserine headgroups, causing long acyl chain enrichment, hence changes in lipid orientation and overall fluidity. In parallel, lipid binding may lead to conformational changes in the affected protein[23], as in classical ligand-receptor interactions. This is exemplified by lipid-induced conformational changes in Annexin A5[22] and the effect of lysophosphatidic acid on the conformation of structurally disordered peptides[24]. Taken together, these transformations may be regarded as "mutual induced fit", whereby the two binding partners modify each other, potentially resulting in better binding.

Small molecule recognition by lipid aptamers

Since a focus of this Viewpoint is on the relevance of lipid aptamers to prebiotic processes prior to the advent of macromolecules, it is essential to seek evidence for cases in which lipids can selectively recognize and bind small molecules, as reviewed[23]. In one study[25], the authors question whether membranes composed of chiral lipids, such as glycerophospholipids and cholesterols, could portray enantioselective binding of small molecules. The results show that despite the membrane's fluid disposition, enantiomers of target drug-like compounds bind differentially to the tested membranes, suggesting enantiospecific interactions. In another paper[26], a single amino acid (L-histidine) is found to reveal enantiospecific binding to cholesterol-containing lipid nanodomains. The authors predict that different lipid compositions could allow effective binding and chiral recognition of other amino acids.

Additional experimental evidence for small-molecule recognition at lipid surfaces comes from a study[17] in a water/octanol two-phase model system. The authors examine the differential efficacy of transfer of two dyes from the aqueous phase to the organic phase. When different amphiphile mixtures were present in the octanol phase, different transfer selectivities were observed. A possible explanation is that amphiphiles form an interphase monolayer and reverse micelles within the organic phase, selectively gating the dye transfer to the oil phase. The authors point out that supramolecular amphiphile combinations are sufficient for dye molecular recognition in analogy to nucleic acid aptamers.

Finally, there are published cases of lipid cooperativity in catalyzing covalent modifications of small molecules, indicating stereospecific molecular recognition. One study describes vesicles comprised of two lipid types that catalyze a transphosphorylation reaction[27]. This case exhibits synergism, as the reaction is carried more efficiently in mixtures than in homogeneous vesicles. The authors explain the cooperativity through dynamic reorganization and optimization of lipid interactions. Another study depicts micelles composed of different binary combinations of lipids that catalyze a plethora of synthetic reactions to various degrees[28]. A wide range of such lipid-mediated reactions as reviewed[29] suggests that lipids may reveal a spectrum of enzyme-like catalytic capacities, acting as reaction promoters and regulators[23].

Consequences for prebiotic evolution

There are two contrasting points of view in the scientific community regarding the chemical roots of life. One is the "RNA-First" hypothesis, whereby life began with protein-encoding polynucleotides that must have appeared prebiotically. While in this scenario the significance of lipids is acknowledged, lipids are assumed to largely play a role of passive containers[2]. The antithetical view claims that protocellular roots of life could have emerged *prior* to the advent of biopolymers. This is envisioned as a consequence of established networks of mutual interactions among relatively small molecules, in collections capable of ensemble reproduction, hence selection and evolution[5, 30-32].

Obviously, the institution of such mutual interaction networks requires the fulfillment of two criteria: intimate molecular proximity and capability for mutual stereospecific recognition. Supporters of a Lipid-First scenario point out that residing at or within a lipid phase (bilayers or micelles) is an excellent way to attain molecular proximity[1, 3, 5, 29], thus fulfilling the first criterion. Crucially, proximity is not equally promoted in the more dilute aqueous lumen of lipid vesicles[1, 23].

The portrayals in this Viewpoint provide strong support to the second criterion, showing that lipid assemblies with diverse compositions can carry out non-trivial molecular recognition feats, including stereospecific binding and non-enzymatic catalysis. That such lipid recognition takes place towards both large and small molecular targets attests to the generality of the findings. The joint fulfillment of the proximity and recognition criteria leads to a credible inference that prebiotic assemblies of lipid monomers are capable of forming interaction networks with quantifiable chemical parameters[5], in the spirit of a Systems Chemistry outlook for life's origin[1, 5, 33].

The dynamic nature of lipid aptamers constitutes a great advantage in the early steps towards life: while a randomly formed protein sequence can only explore a small region of combinatorial space, lipid dynamic aptamers can span a much larger space, ending up with a much higher probability of molecular recognition towards a specific target (Fig. 3a). The probability may be further enhanced via mutual induced fit, augmenting the capacity to establish significant and specific molecular interactions[34]. Such dynamic flexibility of recognition events should likely enhance the chance that randomly assorted molecular ensembles would reveal catalytic closure, a necessary criterion for self-reproduction[4, 5, 30].

The mutually catalytic network model for life's origin without biopolymers gets much of its support from modelling[4, 30], but there are accumulating experimental confirmations, as exemplified[35] and further reviewed[29]. A relevant chemically rigorous computer simulatable model is the Graded Autocatalysis Replication Domain (GARD)[3, 5, 32]. It posits that lipid entry into or endogenous synthesis within an assembly is catalyzed by lipids already present within the assembly. The observed catalytic power is related to the summated contributions of the different resident molecules. The existence of dynamic aptamers and the observed cooperative recognition among lipids within ultrananodomains lend support to such crucial facets of the model (Fig. 3b).

Finally, cellular information is traditionally attributed solely to sequence-based biological polymers such as RNA and proteins, while lipids, with a few exceptions[5, 6, 36], are rarely thought of as information carriers. Yet, according to combinatorial analyses, lipid compositional information is comparable to sequence-based information of biopolymers[5, 17]. Complemented by the reported evidence in this Viewpoint, lipid compositional information seems increasingly realistic, and appears directly related to molecular recognition functions. This nontrivial connection should be given more attention in protobiological research. Ultimately, the aptameric nature of lipid nanodomains heightens the credibility of a proposed scenario in which lipids are central to self-reproducing multi-molecular systems, probably preceding and facilitating the evolutionary emergence of biopolymers[5, 29].

Conflicts of Interest
There are no conflicts to declare.
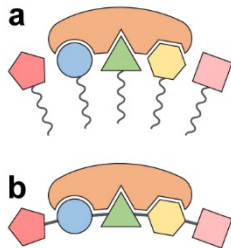
Acknowledgements

Figures

Fig. 1 (a) Dynamic lipid aptamers with target molecules recognizing a group of 3 adjacent lipid headgroups. (b) The same target recognized by the same groups being covalently threaded within a polymer.
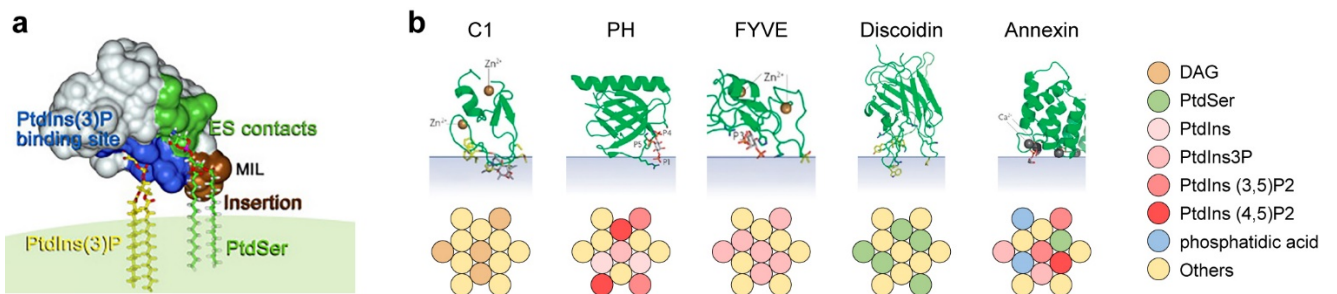

Fig. 2 (a) Model of the FYVE domain in the early endosome antigen 1 (EEA1) gene product, anchoring to three membrane entities, left to right: Basic binding pocket (blue) interacts with PtdIns(3)P, a membrane interaction loop (MIL, brown) that inserts into the bilayer, and residues (green) involved in electrostatic contacts with PtdSer. PtdIns(3)P and PtdSer are shown as stick models. Reproduced from Kutateladze[37]. (b) Lipid specificity of five types of peripheral membrane protein domains: C1, a domain of the regulatory region of protein kinase C, among others; PH, Pleckstrin Homology domain; FYVE, shown in Fig. 2a; Discoidin, a domain found in cell adhesion-related proteins; Annexin, a domain found *inter-alia* in proteins that govern membrane trafficking. Color coding indicates the optimal admixture of the lipids (belonging to a lipid subset in the key on right) leading to protein-membrane multi-point recognition, based on table 3 of Lemmon[20] and on Kutateladze[37]. The green folded protein structures are reproduced from Lemmon[20].
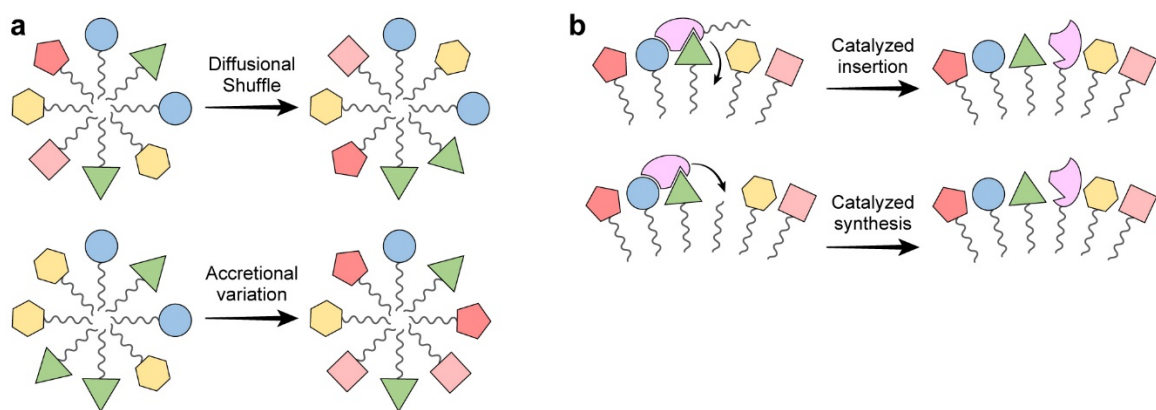
Fig. 3 (a) Lipid aptamer diversity-generating processes. Top, Diffusional shuffling in fluid lipid assemblies (both membrane bilayers and micelles). Bottom, Lipid micelles, because of their nanoscopic dimensions, when accreting in highly heterogeneous environments will show high compositional variability leading to a large combinatorial number of aptamers. This variation-generating process, when occurring in many instances of mutually-catalytic assemblies, could jump-start selection[5]. (b) The basic mechanism of the GARD model, proposed to be operating in the realm of dynamic lipid aptamers. Top, Two lipids cooperatively enhance the joining of a new lipid to a micelle, as modelled[32]. Bottom, enhanced synthesis of a new lipid by joining an externally supplied headgroup to an apolar tailgroup residing within the micellar core, as experimentally described[29, 35], proposed to happen via cooperative interactions with several lipids, as alluded based on experiments[26-28].

References

1. S. Serrano-Luginbühl, K. Ruiz-Mirazo, R. Ostaszewski, F. Gallou and P. Walde, *Nature reviews chemistry*, 2018, 1.
2. D. Deamer, *Life*, 2017, **7**, 5.
3. D. Segrè, D. Ben-Eli, D. W. Deamer and D. Lancet, *Origins of Life and Evolution of the Biosphere*, 2001, **31**, 119-145.
4. D. Segrè and D. Lancet, *EMBO reports*, 2000, **1**, 217-222.
5. D. Lancet, R. Zidovetzki and O. Markovitch, *Journal of The Royal Society Interface*, 2018, **15**, 20180159.
6. T. Dingjan and A. H. Futerman, *BioEssays*, 2021, **43**, 2100021.
7. T. Harayama and H. Riezman, *Nature reviews Molecular cell biology*, 2018, **19**, 281.
8. M. Cebecauer, M. Amaro, P. Jurkiewicz, M. J. o. Sarmento, R. Šachl, L. Cwiklik and M. Hof, *Chemical reviews*, 2018, **118**, 11259-11297.
9. E. Sezgin, I. Levental, S. Mayor and C. Eggeling, *Nature reviews Molecular cell biology*, 2017, **18**, 361.
10. M. Yang, K. Wang, J. Lin, L. Wang, F. Wei, J. Zhu, W. Zheng and L. Shen, *Langmuir*, 2018, **34**, 8408-8414.
11. K. Simons and M. J. Gerl, *Nature reviews Molecular cell biology*, 2010, **11**, 688-699.
12. Ü. Coskun, M. Grzybek, D. Drechsel and K. Simons, *Proceedings of the National Academy of Sciences*, 2011, **108**, 9044-9048.

13. M. J. Hallock, A. I. Greenwood, Y. Wang, J. H. Morrissey, E. Tajkhorshid, C. M. Rienstra and T. V. Pogorelov, *Biochemistry*, 2018, **57**, 6897-6905.
14. M. Inagaki, M. Shibakami and S. L. Regen, *Journal of the American Chemical Society*, 1997, **119**, 7161-7162.
15. P. Pathak and E. London, *Biophysical journal*, 2015, **109**, 1630-1638.
16. H.-M. Meng, H. Liu, H. Kuai, R. Peng, L. Mo and X.-B. Zhang, *Chemical Society Reviews*, 2016, **45**, 2583-2602.
17. T. N. Bell, K. Feng, G. Calvin, D. H. Van Winkle and S. Lenhert, *ACS Omega*, 2020.
18. R. New, G. S. Bansal, M. Bogus, K. Zajkowska, S. Rickelt and I. Toth, *Molecules*, 2013, **18**, 3427-3441.
19. I. Vonkova, A.-E. Saliba, S. Deghou, K. Anand, S. Ceschia, T. Doerks, A. Galih, K. G. Kugler, K. Maeda and V. Rybin, *Cell reports*, 2015, **12**, 1519-1530.
20. M. A. Lemmon, *Nature reviews Molecular cell biology*, 2008, **9**, 99-111.
21. D. Lingwood and K. Simons, *science*, 2010, **327**, 46-50.
22. M. A. Lizarbe, J. I. Barrasa, N. Olmo, F. Gavilanes and J. Turnay, *International journal of molecular sciences*, 2013, **14**, 2652-2683.
23. P. Walde, H. Umakoshi, P. Stano and F. Mavelli, *Chemical Communications*, 2014, **50**, 10177-10197.
24. T. Juhász, J. Mihály, G. Kohut, C. Németh, K. Liliom and T. Beke-Somfai, *Scientific reports*, 2018, **8**, 1-15.
25. H. Tsuchiya and M. Mizogami, *Molecules*, 2018, **23**, 49.
26. T. Ishigami, A. Tauchi, K. Suga and H. Umakoshi, *Langmuir*, 2016, **32**, 6011-6019.
27. C. Z.-J. Ren, P. Solís-Muñana, G. G. Warr and J. L.-Y. Chen, *ACS Catalysis*, 2020, **10**, 8395-8401.
28. K. V. Bukhryakov, V. G. Desyatkin, J.-P. O'Shea, S. R. Almahdali, V. Solovyeva and V. O. Rodionov, *ACS combinatorial science*, 2015, **17**, 76-80.
29. A. Kahana and D. Lancet, *Nature Reviews Chemistry*, 2021, **(under revision)**.
30. S. A. Kauffman, *Journal of theoretical biology*, 1986, **119**, 1-24.
31. R. Shapiro, *IUBMB life*, 2000, **49**, 173-176.
32. D. Segrè, D. Ben-Eli and D. Lancet, *Proceedings of the National Academy of Sciences*, 2000, **97**, 4112-4117.
33. Y. Bai, A. Chotera, O. Taran, C. Liang, G. Ashkenasy and D. G. Lynn, *Chemical Society Reviews*, 2018, **47**, 5444-5456.
34. D. Marsh, *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 2008, **1778**, 1545-1575.
35. I. Colomer, A. Borissov and S. P. Fletcher, *Nature Communications*, 2020, **11**, 1-9.
36. R. Gross and X. Han, *Future Lipidology*, 2006, **1**, 539-547.
37. T. G. Kutateladze, *Progress in lipid research*, 2007, **46**, 315-327.