Protein and lipid interactions driving molecular mechanisms of *in meso* crystallization

Niklaus Johner¹, Sayan Mondal¹, Giulia Morra¹,², Martin Caffrey³, Harel Weinstein¹, George Khelashvili¹

Supporting Information

Sign of the curvatures:

On the interfacial surface, all the normal vectors are oriented to point from the surface towards the water. We use the standard convention of positive curvature corresponding to \( \mathbf{r} \cdot \mathbf{n} > 0 \), where \( \mathbf{r} \) is the radius of curvature pointing from the center of curvature to the surface and \( \mathbf{n} \) the normal on the surface.

On the midplane, both sides of the surface are equivalent, so that the two choices for the orientation of the normal, hence the signs of the principal curvatures, are equivalent too. Moreover, when working on one unit cell with periodic boundary conditions, the two sides of the surface become connected, which is not the case when working on 8 unit cells with periodic boundary conditions (an array of 2 by 2 by 2) or any size of LCP without periodic boundary conditions. Special care was taken to orient all the normal in the unit cell studied so that the relative signs of the principal curvatures in different places on the surface are consistent.

Relationships between curvatures on the minimal surface and parallel surfaces:

On the minimal surface we have that

\[-k_1 = k_2 = \frac{1}{R} = \sqrt{-K_0},\]

where \( R \) is the radius of curvature. On a surface parallel to the midplane and at a distance \( l \) from it, the principal curvatures become:

\[
k_1(l) = \frac{1}{-R + l} = \left( \frac{1}{\sqrt{-K_0} + l} \right)^{-1}
\]

\[
k_2(l) = \frac{1}{R + l} = \left( \frac{1}{\sqrt{-K_0} + l} \right)^{-1}
\]
So that the mean and Gaussian curvatures are given by:

\[ H_l = \frac{k_1 + k_2}{2} = \frac{K_0 l}{1 + l^2 K_0} \]  \hspace{1cm} (s1)

\[ K_l = k_1 k_2 = \frac{K_0}{1 + l^2 K_0} \]  \hspace{1cm} (s2)

It is easily shown that the surface element on the midplane and on the parallel surface are linked by:

\[ dA_i = (1 + K_0 l^2) dA_0 \]  \hspace{1cm} (s3)

leading to

\[ A_i = \left(1 + \langle K \rangle_0 l^2\right) A_0. \]  \hspace{1cm} (s4)

It should be noted that the above equation describes a surface at a distance \( l \) on one side of the midplane, whereas the interfacial surfaces obtained from the simulations are on both sides of the midplane and therefore twice the surface given above. Using equations s1, s2, s3 and s4, we find:

\[ \langle K \rangle_l = \frac{1}{A_l} \int_{A_l} K_l dA_l = \frac{1}{A_l} \int_{A_0} K_0 dA_0 = \frac{A_0 \langle K \rangle_0}{A_l} = \frac{\langle K \rangle_0}{1 + \langle K \rangle_0 l^2} \]  \hspace{1cm} (s5)

\[ \langle H \rangle_l = \frac{1}{A_l} \int_{A_l} H_l dA_l = \frac{1}{A_l} \int_{A_0} H_0 dA_0 = l \langle K \rangle_l \]  \hspace{1cm} (s6)
Martini-compatible topology files for 9.9, 7.9, and 7.7 MAG lipids

;;;;;; MONOOLEIN 9.9
[moleculetype]
; molname     nrexcl
   MAG99       1
[atoms]
; id    type    resnr   residu  atom    cgnr    charge
  1     P4      1       MAG99    ETH     1       0
  2     Na      1       MAG99    GL1     2       0
  3     C1      1       MAG99    C1A     3       0
  4     C1      1       MAG99    C2A     4       0
  5     C3      1       MAG99    D3A     5       0
  6     C1      1       MAG99    C4A     6       0
  7     C1      1       MAG99    C5A     7       0
[bonds]
; i j   funct   length  force.c.
  1 2   1       0.47    1250
  2 3   1       0.47    1250
  3 4   1       0.47    1250
  4 5   1       0.47    1250
  5 6   1       0.47    1250
  6 7   1       0.47    1250
[angles]
; i j k         funct   angle   force.c.
  2 3 4         2       180.0   25.0
  3 4 5         2       180.0   25.0
  4 5 6         2       120.0   45.0
  5 6 7         2       180.0   25.0

;;;;;; MONOOLEIN 7.9
[moleculetype]
; molname     nrexcl
   MAG79       1
[atoms]
; id    type    resnr   residu  atom    cgnr    charge
  1     P4      1       MAG79    ETH     1       0
  2     Na      1       MAG79    GL1     2       0
  3     C1      1       MAG79    C1A     3       0
  4     C3      1       MAG79    D2A     4       0
  5     C1      1       MAG79    C3A     5       0
  6     C1      1       MAG79    C4A     6       0
[bonds]
; i j funct length force.c.
  1 2 1  0.47    1250
  2 3 1  0.47    1250
  3 4 1  0.47    1250
  4 5 1  0.47    1250
  5 6 1  0.47    1250

[angles]
; i j k funct angle force.c.
  2 3 4  2  180.0   25.0
  3 4 5  2  120.0   45.0
  4 5 6  2  180.0   25.0

;;;;; MONOOLEIN 7.7
[moleculetype]
; molname nrexcl
  MAG77  1
[atoms]
; id type resnr residu atom cgnr charge
  1  P4  1 MAG77 ETH  1  0
  2  Na  1 MAG77 GL1  2  0
  3  C1  1 MAG77 C1A  3  0
  4  C3  1 MAG77 D2A  5  0
  5  C1  1 MAG77 C3A  7  0

[bonds]
; i j funct length force.c.
  1 2 1  0.47    1250
  2 3 1  0.47    1250
  3 4 1  0.47    1250
  4 5 1  0.47    1250

[angles]
; i j k funct angle force.c.
  2 3 4  2  180.0   25.0
  3 4 5  2  120.0   45.0
## Supporting Tables

### Table S1: Conditions for self-assembly simulations of 9.9 MAG.

$N_{MAG}$ is the number of monoolein lipids and $N_W$ of waters in the simulation box. The third column shows the water concentration in the solution. $T$ is the temperature of the simulation and $a$ the unit cell size of the self-assembled LCP. Shaded are the conditions in which protein constructs were studied.

| $N_{MAG}$ | $N_W$ | % w/w | $T$ [°C] | $a$ [Å] |
|-----------|-------|--------|----------|--------|
| 1200$^a$  | 4285  | 42     | 20       | 113    |
| 950       | 3200  | 41     | 25       | 104    |
| 950$^{a,b}$ | 2700  | 36     | 25       | 102    |
| 820       | 2900  | 42     | 30       | 100    |
| 740       | 2000  | 35     | 35       | 94     |
| 600       | 1780  | 38     | 40       | 88     |
| 600$^b$   | 1300  | 30     | 55       | 86     |
| 560       | 1300  | 32     | 50       | 85     |
| 550$^b$   | 900   | 25     | 60       | 82     |

$^a$ – $\text{A}_2\text{AR-BRIL}$ construct studied in this LCP.

$^b$ – $\text{A}_2\text{AR}$ construct studied in this LCP.

### Table S2: Conditions for self-assembly simulations of 7.9 MAG.

For column headers, see Table 1.

| $N_{MAG}$ | $N_W$ | % w/w | $T$ [°C] | $a$ [Å] |
|-----------|-------|--------|----------|--------|
| 1400      | 6500  | 50     | 20       | 119    |
| 1300      | 5600  | 47     | 30       | 116    |
| 1070      | 4440  | 48     | 35       | 108    |
| 850       | 3300  | 46     | 40       | 99     |
| 820       | 2900  | 44     | 45       | 97     |
| 750       | 2400  | 41     | 50       | 93     |
| 650       | 2100  | 41     | 55       | 89     |
| 650$^a$   | 1700  | 36     | 55       | 87     |

### Table S3: Conditions for self-assembly simulations of 7.7 MAG.

For column headers, see Table 1.

| $N_{MAG}$ | $N_W$ | % w/w | $T$ [°C] | $a$ [Å] |
|-----------|-------|--------|----------|--------|
| 1100      | 6000  | 57     | 40       | 111    |
| 1000      | 5200  | 55     | 45       | 106    |
| 902       | 4400  | 54     | 50       | 102    |
| 840       | 3700  | 51     | 55       | 98     |
| 800       | 3200  | 49     | 60       | 95     |
Supporting Figures

Figure S1: Definition of the CG beads for the 9.9 MAG, 7.9 MAG and 7.7 MAG lipids (panels a, b and c respectively). Oval shapes show approximate groupings of the heavy atoms into the CG beads (see also the Supporting Information for the CG bead force-field types and interaction parameters). The carbon atoms forming the double bond in the three lipids are labeled (C9=C10 in 9.9 MAG, and C7=C8 in 7.9 MAG and 7.7 MAG).

Figure S2: (a) Illustration of low-pass filtering on the density map of waters for a reference frame in a trajectory. The original map is shown in blue, whereas the filtered map is shown in cyan. (b) Illustration of the determination of the interface surface. The lipid density is shown in purple, the water density in cyan and the interface surface as blue spheres for the 88Å 9.9 MAG LCP.
Figure S3: Midplane surface estimate as a function of the sphere radius used in the determination of the local infinitesimal surfaces (see Methods) for three different 9.9 MAG LCP with $a=82\text{Å}$ (blue circles), $a=86\text{Å}$ (red triangles) and $a=102\text{Å}$ (green crosses) respectively. The dashed lines show the theoretical surface estimate $A_0 = 1.919a^2$. 
Figure S4: Mean curvature (full symbols) and principal curvatures $k_1$ (positive curvature, empty symbols) and $k_2$ (negative curvature, empty symbols) for different lipids (blue circles 7.7 MAG, green right triangles 7.9 MAG and red diamonds 9.9 MAG) and compositions shown as a function of the lattice constant $a$. The error bars represent the width of the distributions of curvatures on the surface. It is noteworthy that the average mean curvatures are between 50 and 1000 times smaller than the average principal curvatures.
Figure S5: Thickness of the LCP membrane monolayer \( (l) \) as a function of lattice constant \( (a) \) for the different types of lipids (blue circles 7.7 MAG, green right triangles 7.9 MAG and red diamonds 9.9 MAG) and compositions.
Figure S6: Water/lipid interface for the 9.9 MAG LCP with lattice constant \( a = 82 \text{Å} \). In panel (a), the surface is colored according to the Gaussian curvature and in panel (b) the color shows the membrane monolayer thickness. Regions of low (in absolute value) Gaussian curvature correspond to regions of thin membrane.
Figure S7: Membrane midplane and water/lipid interface for the 9.9 MAG LCP with lattice constant $a=82$ Å. Both surfaces are colored according to the Gaussian curvature (intensity of red for the midplane and of blue for the interface). Regions of the membrane with low Gaussian curvature on the midplane also have low Gaussian curvature on the interface.
Figure S8: Organization of lipids around the A$_2A$ receptor in a 9.9 MAG 82Å LCP. Lipid headgroups are shown as red spheres, tails as purple lines and the membrane midplane (see Methods) is shown in white. The protein is shown as a cartoon and is an all-atom model superimposed on the CG protein from this snapshot for clarity of the representation. The loops are shown in white, TM1 in green, TM2 in cyan, TM3 in orange, TM4 in yellow, TM5 in blue, TM6 in light purple and TM7 in salmon. The ligand (white) and several key residues (E3.50, W4.50, P5.50, P6.50 and P7.50, same color as the TM they’re in) are shown in space filling. Panel a shows the extracellular end of the protein and the following panels (b-d) show slices deeper in the TM region of the protein. Panel d shows the intracellular end of the protein, with notably HX8 and E3.50 (orange) visible.
Figure S9: Illustration of the orientation of the A2AR-BRIL in the 103Å (panel a) and 113Å LCPs (panel b). Interfacial surfaces are shown in blue and an all-atom model superimposed on the average structures is shown as a cartoon (see Methods), colored as: loops in white, TM1 in green, TM2 in cyan, TM3 in orange, TM4 in yellow, TM5 in blue, TM6 in light purple and TM7 in salmon and the BRIL in red. The red line shows the axis of the protein (best fit line to the protein excluding the BRIL extension) and the green lines represent the normal of the best fit plane to the interfacial surface around TM5 (within 15Å of Arg206 \textsuperscript{5.67}). The region of the interface used for the fitting of the plane is also shown in green. The difference in tilt of the receptor relative to the surrounding LCP orients the BRIL extension towards the neighboring Monkey saddle point in the 103Å LCP (panel a), but not in the larger 113Å LCP. This would allow the formation of contacts with a neighboring receptor in the first case but not in the larger LCP.
Supporting Movies

Supporting movie 1: The movie illustrates the procedures used for the analysis of the CG-MD simulations. It starts by with a isosurface of the density of waters in blue. This density is replaced by its filtered counterpart shown in cyan. Then the filtered lipid density is added in purple and we zoom in onto a tetrahedral water channel. The water lipid interface obtained from the filtered densities is added as blue spheres, followed by a swipe in the level of the isocontours, illustrating how the interface is determined as the surface where the two densities are equal. We then illustrate how the normals are determined on the interfacial surface, showing in green patches of membranes from which the local normal is determined as the best fit plane and shown as a black line. Finally we add the membrane midplane to illustrate the overall arrangement of the Pn3M LCP.

Supporting movie 2: This movie shows the organizations of the lipids around the A2AR as we move from the extracellular end of the protein towards its intracellular side. The representation and coloring scheme is the same as in Figure S6. The movie starts from the node of water channel tetrahedron above the protein and then we move down through the water channel that is now occupied by the receptor. It is clear that the membrane midplane is wrapped around the protein and the lipids organize as a bilayer around TM1 and TM4.

Supporting PDB Files

Several files allowing the visualization of the LCP surfaces are provided as Supporting Web Enhanced Objects. All files are for the 88Å LCP of 9.9MAG and the surfaces represent one unit cell of the LCP. 9.9_MAG_88A_boundary.pdb is the water/lipid interface surface. 9.9_MAG_88A_midplane_raw.pdb visualizes the membrane midplane obtained from the density of the last bead of the lipid tails. As explained in the Methods, the identification of the midplane from local maxima of this density results in a noisy surface, that is then cleaned by keeping only points that are at a maximum of the density in the direction normal to the midplane surface. 9.9_MAG_88A_midplane_cleaned.pdb file illustrates the midplane surface after this cleaning step.