Identification of EM-Resolved Lipid Fragments Using Streamlined Alchemical Free Energy Perturbation (SAFEP)

Ezry Santiago-McRae[1], Thomas T. Joseph[2], Wayland Cheng[3], Jérôme Hénin[4], Grace Brannigan[1,5]

1) Center for Computational and Integrative Biology, Rutgers University, Camden 2) Dept. of Anesth. Crit. Care, University of Pennsylvania Perelman School of Medicine 3) School of Medicine in St. Louis, Washington University 4) Laboratoire de Biochimie Théorique, CNRS 5) Camden, Dept. Physics, Rutgers University

Partially Resolved Lipid Densities

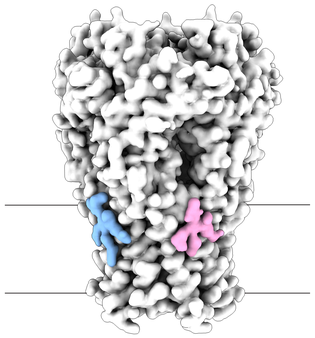


Fig 1: EM Density of ELIC. This density of Erwinia ligand gated ion channel (gray) is an example of a membrane protein with partially resolved lipid-like densities (blue and pink). Density obtained in PC:PG:PE 2:1:1 nanodiscs. The outer-most helix (M4) was not well-resolved and is hidden in this view. Density published in [2].

- The function of membrane proteins often depends on the local lipid environment. [1]
- Functionally relevant membranes may be untenable in nanodiscs or for cryo-EM generally
- EM densities of membrane proteins increasingly reveal bound lipid fragments of unknown identity. (e.g. [2])

Ligand-gated Ion Channels

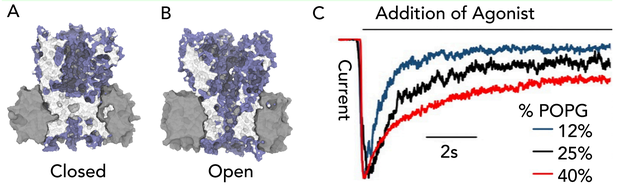


Fig 2: *Erwinia* Ligand-gated Ion Channel (ELIC). Cross sections of closed (A) and open (B) structures of ELIC (blue, solvent accessible, and white, interior) embeded in a POPC membrane (gray). Structures published in [2]. C) Patch clamp recordings of ELIC in a POPC:POPG model membrane. Peak currents are normalized. As POPG concentration is increased, desensitization is delayed. Peak currents also increase (Data not shown). Adapted from citation 4.

- Pentameric Ligand-gated Ion Channels (pLGICs) [3]:
  - Gated by small molecules
  - Many neurotransmitter receptors
  - Desensitize over time after initial opening
- *Erwinia* Ligand-gated Ion Channel (ELIC):
  - A bacterial model pLGIC
  - Function depends on POPG (Fig 2.C) and other lipids

Molecular Dynamics & FEP

- FEP: Free Energy Perturbation
  - Free energy method for physical simulations
  - Non-bonded interactions are weakened or strengthened in order to obtain the free energy difference
- Classical FEP is not well suited to superficial sites.

References

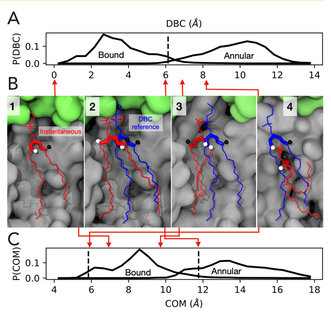
SAFEP Tutorial[6]

1. Yeagle. The Membranes of Cells. 2016
2. Petroff, [...], Santiago-McRae, [...] Hénin, Brannigan, Cheng. Nat Comms. 2022.
3. Jaiteh, Taly, Hénin. PLOS ONE. 2014
4. Tong, [...], Brannigan, Cheng. eLife. 2019
5. Salari, Joseph, Lohia, Hénin, Brannigan. JCTC. 2018.
6. Santiago-McRae, Ebrahimi, Sandberg, Brannigan, Hénin. LiveCoMS. 2023

What Counts as a Bound Lipid in a Superficial Site? - SAFEP

- Binding to a superficial site introduces an ambiguity:
  - When is the lipid "bound" and when is it coincidentally in the site?
  - Or, equivalently, when is the site "occupied" and when is it simply filled with solvent?
- The Distance from Bound Configuration (DBC) (Fig 2):
  - RMSD of the most stably bound lipid atoms
  - In the protein frame of reference
  - Captures the fluctuations apparent in the cryo density
- Restraining the DBC:
  - Doesn't affect the bound ensembles
  - Can be corrected for in the gas phase

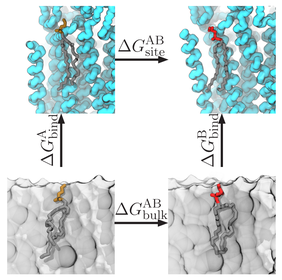
Fig 3: Comparison of collective variables for a bound lipid. Distributions and poses taken from a simulation of POPE bound to ELIC. A) the distribution of the DBC of POPE in both a bound and annular state. Red arrows indicate the location of each pose along the collective variable. B.1) The instantaneous lipid pose (red) with glycerol oxygens colored black and white to indicate orientation. B.2-4) The instantaneous lipid pose compared with the reference pose (blue). C) As in A, showing the distribution of the COM.



What Lipid is Responsible for those Densities? - Relative Affinities

Fig 4: Schematic of the Relative Binding Free Energy (RBFE) via SAFEP. Over the course of a simulation, one headgroup (orange) is decoupled while the other (red) is coupled, effectively converting one lipid into another. This is carried out in both the site (to obtain the ΔG^{AB}_{site}) and the bulk (ΔG^{AB}_{bulk}). The relative free energy of binding is

$$\Delta\Delta G^{AB}_{bind} = \Delta G^{AB}_{site} - \Delta G^{AB}_{bulk} = \Delta G^{B}_{bind} - \Delta G^{A}_{bind}$$



- Each candidate ligand is converted into another yielding a free energy of conversion (Fig 4)
- The free energies of conversion can be combined into a relative binding free energy (ΔΔG) by Eqn 1 (Table 1)
- Candidate ligands can be ranked by their ΔΔGs
- Relative afffinities can be used to compute binding probabilities (Fig 5)

Fig 5: Site occupancy in a 2:1:X PC:PE:PG membrane. The probability of site occupancy by each lipid as a function of PG mol fraction. Site occupancy by a lipid (^{A}) is given by

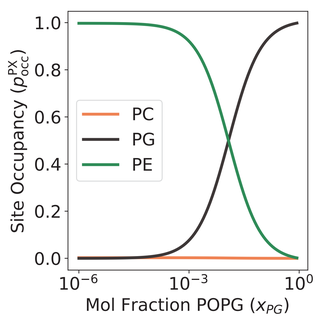
$$p^{A}_{occ} = \frac{p_A}{p_A + p_B + p_C}$$

where

$$p_A = \frac{1}{1 + p_{BA} + p_{CA}}$$

and

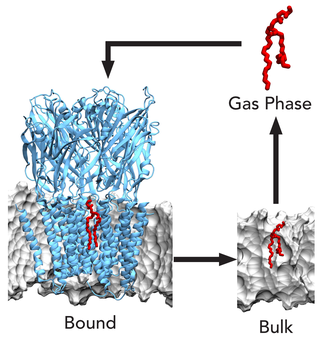
$$p_{AB} = \frac{x_A}{x_B} e^{-\beta\Delta\Delta G^{AB}_{bind}}$$



- For the closed state (Fig 1A):
  - $\Delta\Delta G^{PG\to PE}_{bind}$ = 2 kcal/mol
  - $\Delta\Delta G^{PE\to PC}_{bind}$ = 6 kcal/mol
  - PG out-competes PE even at low mol fractions. (Fig 5)
- POPC has very weak relative affinity for this binding mode suggesting that it is a non-binder.

Are those Lipids Functionally Relevant? - Absolute Affinities

Fig 6: Schematic of Absolute Alchemical Binding Free Energy (ABFE) via SAFEP. The ligand (purple circle) starts the simulation either bound to the protein (A, green) or unbound (B) in the bulk (blue). To obtain the ΔG^{0}_{bind}, the free energy difference between the site-occupied (A) and unoccupied (B) states, a non-physical path is taken through the gas phase (C).



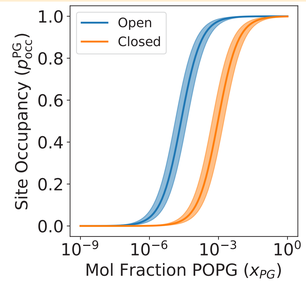
- The whole lipid is decoupled to the gas phase
- This yields the absolute free energy difference of moving the lipid from the bulk to the site
- More expensive than RBFE, but more precise
- Applied to the highest affinity lipid, POPG (Fig 7)

Fig 7: Site occupancy by POPG in a primarily POPC membrane. Occupancy probability of open (blue) and closed (orange). Shaded regions indicate ±1SEM. Occupancy probability can be expressed as a function of POPG mole fraction by

$$p_{occ}(x_{PG}) = \frac{x_{PG}}{x_{PG} + x_{50}}$$

where

$$x_{50} = e^{-\beta\Delta G^{*}_{bind}} \cdot x_{sim}$$



- x_{50}: The mol fraction at which the site is 50% occupied
  - Open (Fig 1A): 10^{-5} % PG
  - Closed (Fig 1B): 10^{-3} % PG
- POPG occupies the site even at low mol fractions (Fig 7)
- Site occupancy is close to 100% under experimentally relevant conditions

Abstract

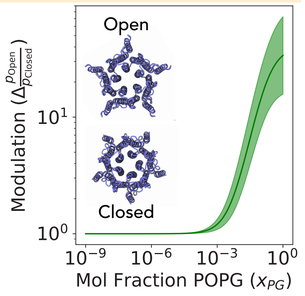
Over the past ten years, the number of cryo-EM structures with partially resolved lipids has increased dramatically. Lipid binding sites may be important both for native protein function and as potentially drugable sites. Although cryo-EM data can indicate the presence of ordered lipids, those lipids are rarely identifiable from the density alone. We use our Streamlined Alchemical Free Energy Perturbation protocol (SAFEP) for computationally estimating free energies of phospholipid binding (Gbind) to identify lipids and compare affinities across sites and protein conformations. There are several challenges posed by the binding of phospholipids to membrane proteins: 1) lipid flexibility makes traditional restraint schemes less effective, 2) slow relaxation of lipids slows convergence of the free energy estimates, and 3) lipid-water phase separation necessitates careful interpretation of any lipid Gbind. Addressing the first two issues required simple, but highly designed restraints combined with thorough sampling. The phase separation of the system was addressed theoretically by framing the results in terms of binding probabilities - leaving any "standard" binding free energy as an intermediate result. Convergence of all calculations was monitored by several metrics as any one metric was been found to be fallible. We use lipid binding by Erwinia Ligand-Gated Ion Channel (ELIC) as a model system. ELIC is a GABA-activated, prokaryotic member of the pentameric ligand-gated ion channel (pLGIC) protein family. Like other pLGICs, ELIC is known to be modulated by its lipid environment through unknown mechanism(s). By a combination of relative and absolute binding free energy calculations we were able to identify a partially resolved lipid as POPG, determine that POPC is a non-binder to the site, POPE can compete for the site at higher mole fractions, and POPG binds with greater affinity to the open conformation of ELIC.

How Strong is the Modulation?

Fig 8: Log modulation of ELIC versus mole fraction of POPG Calculated by eqn 2. Greater values correspond to gain of function. Shaded region indicates ±1SEM. The relative open probability is a function of binding free energies

$$\frac{p^{open}_{occ}}{p^{closed}_{occ}} = \frac{\Delta G^{open}_{bind} \cdot x_{PG} + 1}{\Delta G^{closed}_{bind} \cdot x_{PG} + 1}$$

Insets show the open (Fig 1B) and closed (Fig 1A) transmembrane domains.



- The differential affinity of POPG (Fig 7) for the open and closed states suggests allosteric modulation (Fig 8).
- Prediction: ELIC in 10% POPG will have approximately 15x higher open probability than in pure POPC

Conclusions

- Application:
  - A partially resolved lipid in two ELIC densities was identified as POPG (relative SAFEP)
  - POPG affinity is state dependent (absolute SAFEP) with an estimated 10 to 30 fold increase in open probability under experimentally relevant conditions.
- Methodology:
  - The DBC is an effective metric for quantifying the bound state of a lipid informed by the density
  - Restraining the DBC enabled convergence of both relative and absolute FEP
  - Using SAFEP, we are now able to make quantitative predictions of phospholipid binding to superficial sites

Acknowledgments

• Computation resources provided by the Rutgers Office of Advanced Research Computing and ACCESS (allocation BIO220103)
• Funding provided by NSF DGE2152059 & NIH K08 GM139031