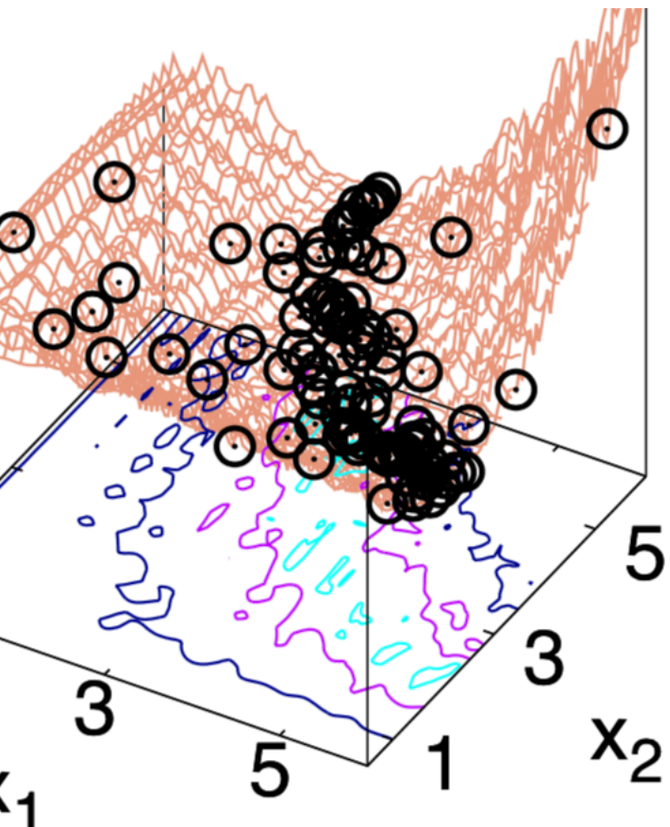
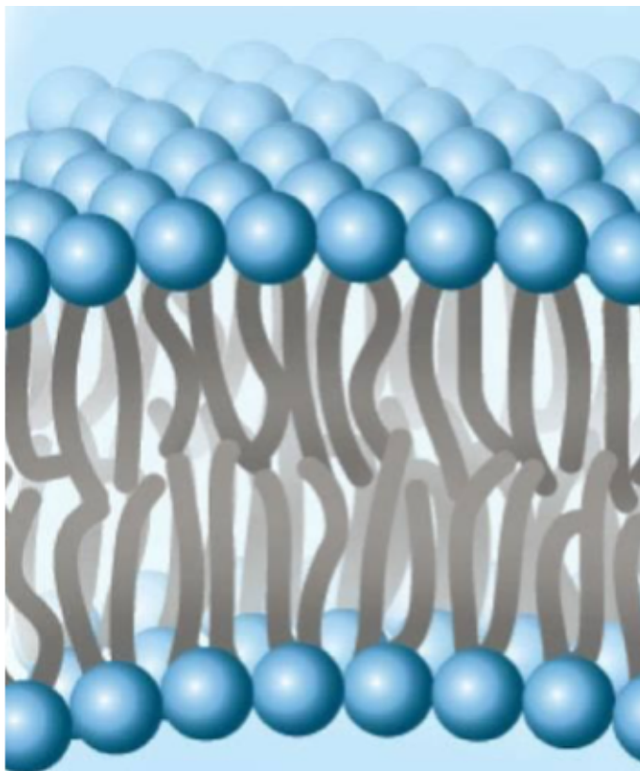
National Heart, Lung, and Blood Institute

Automated Optimization Approaches for the CHARMM36 Lipid Force Field

Andreas Krämer
Laboratory of Computational Biology
National Institutes of Health



OpenFF Webinar
17 December 2019

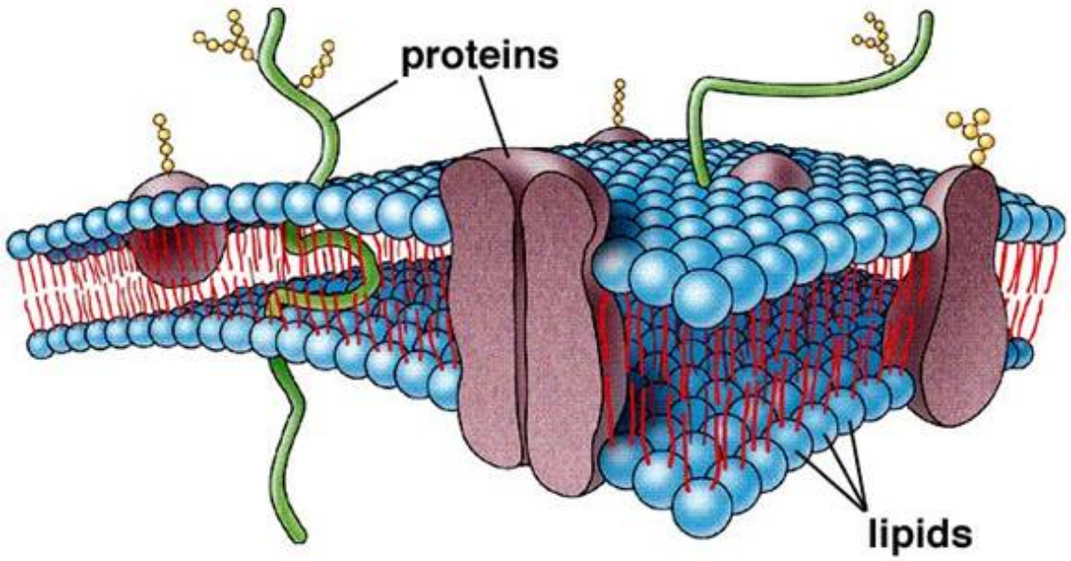


Outline

- Introduction
  - The CHARMM36 Lipid Force Field
  - Force Field Parameterization
- Applications
  - Improving Permeabilities
  - Adapting CHARMM36 for LJ-PME
- Outlook
  - Continuous Integration
  - Polarizable CHARMM Force Field

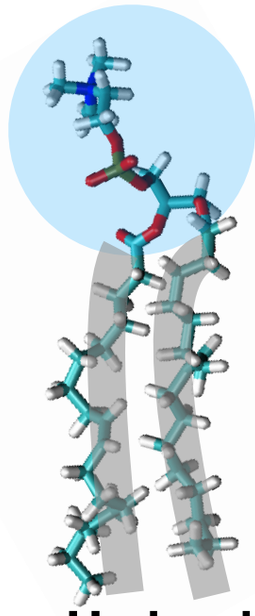
The CHARMM36 (C36) Lipid Force Field

- Lipids = building blocks of biological membranes
- C36: one of the most-used all-atom lipid force fields; >2000 citations over the last decade
- Reasons for its success?



$$V(\hat{R}) = \sum_{bonds} K_b (b - b_0)^2 + \sum_{angles} K_\theta (\theta - \theta_0)^2 + \sum_{dihedrals} \left[\sum_j K_{\varphi,j} \left(1 + \cos(n_j \varphi - \delta_j)\right) \right]$$

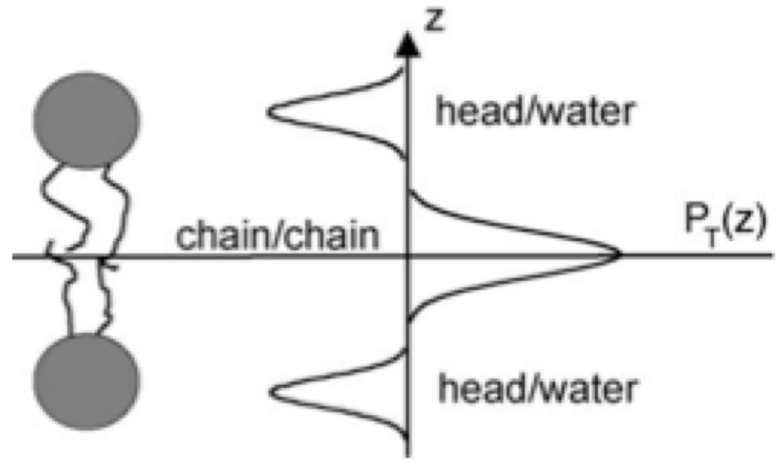
$$+ \sum_{\substack{nonbond \\ pairs}} \varepsilon_{ij} \left[\left(\frac{R_{\min,ij}}{r_{ij}} \right)^{12} - \left(\frac{R_{\min,ij}}{r_{ij}} \right)^{6} \right] + \sum_{\substack{nonbond \\ pairs}} \frac{q_i q_j}{\varepsilon_D r_{ij}}$$



Klauda et al., J. Phys. Chem. B. **114**, 23 (2010).

The CHARMM36 (C36) Lipid Force Field

- **Many structural properties of PC and PE lipids are very good in C36**
  - Surface Area per Lipid
  - Compressibility Moduli
  - Bending Constant
  - Spontaneous Curvature



	Area/lipid (Å²)	
lipid	sim	expt
DPPC	62.9 ± 0.1	63.0 ± 1.0
DMPC	61.5 ± 0.1	60.6 ± 0.5
DLPC	64.4 ± 0.3	63.2 ± 0.5
DOPC (lo)	60.1 ± 0.1	59.3 ± 0.7
DOPC (hi)	68.9 ± 0.1	67.4 to 72
POPC	65.9 ± 0.1	64.3-68.3
POPE	58.7 ± 0.1	59.8-60.8
	K_A (dyn/cm)	
DPPC	230 ± 20	231
DMPC	210 ± 30	234 ± 23
DOPC	285 ± 20	300
POPC	260 ± 20	180-330

Klauda et al., J. Phys. Chem. B. **114**, 23 (2010)

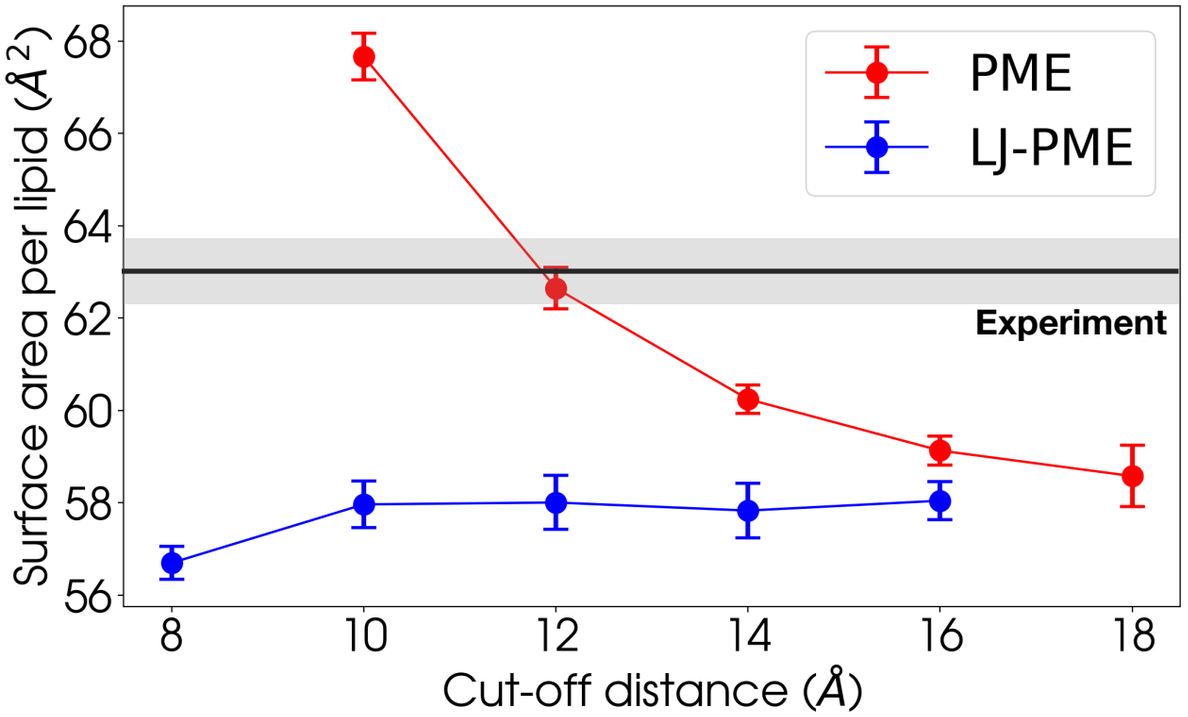
Pastor and MacKerell, J. Phys. Chem. Letters **2**, 1526 (2011)

Levine, Venable, Watson, Lerner, Shea, Pastor, Brown, JACS, **136**, 1358 (2014)

Venable, Brown, Pastor, *Chem. Phys. Lipids,* **192**, 60 (2015)

C36 is great. Why Optimization?

- **Long-range dispersion**
  - Interfacial systems are better represented by incorporating long-range Lennard-Jones (LJ-PME)
  - C36m protein force field has been validated for LJ-PME
  - Monolayer surface tensions better with long-range LJ
  - Inconsistencies regarding LJ treatment
- TIP3P water model?
- **Permeability**
  - Partitioning between water and organic phases is not well modeled by additive FF



Leonard, Simmonett, Pickard, Huang, Venable, Klauda, Brooks, Pastor, J. Chem. Theory Comput., **14**, 2 (2018).

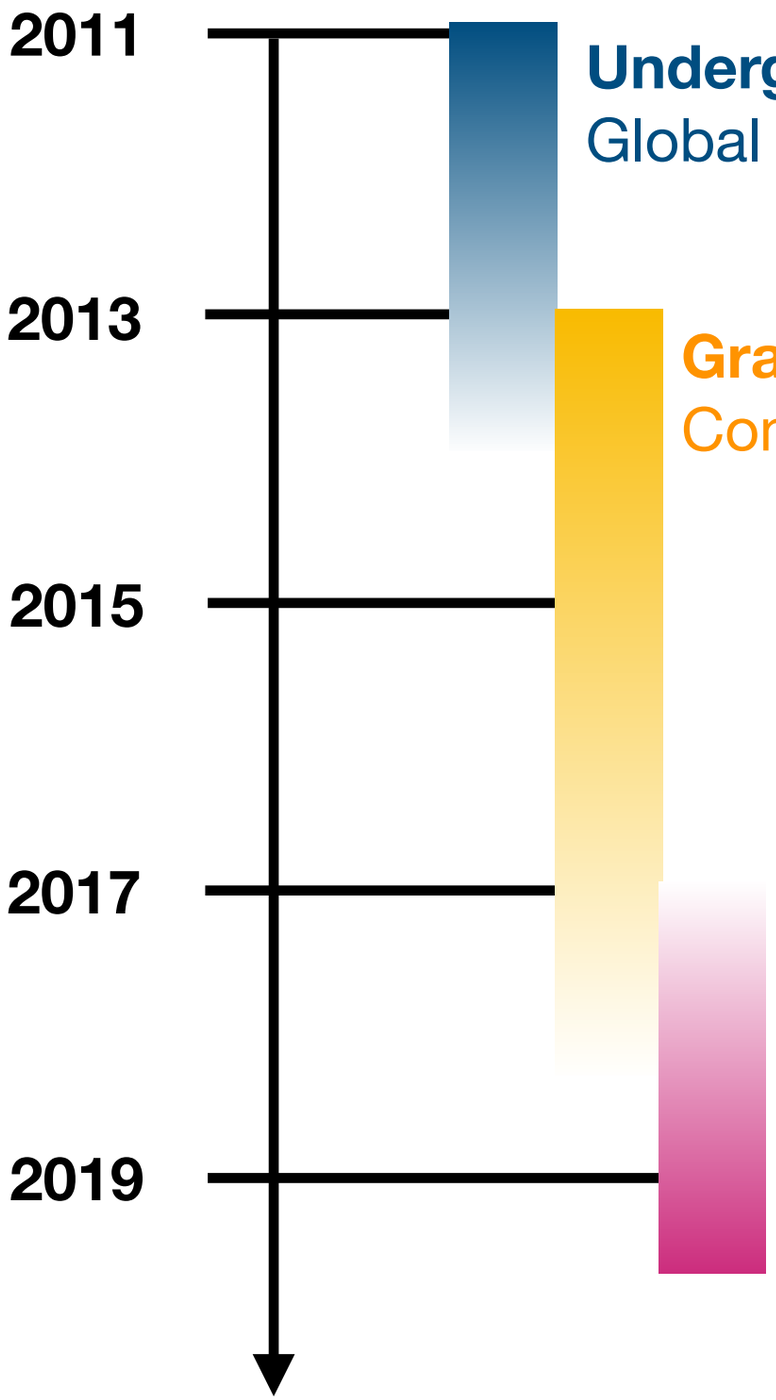
Krämer, Pickard, Huang, Venable, Simmonett, Reith, Kirschner, Pastor, Brooks, J. Chem. Theory Comput., **15**, 6 (2019).

Venable, Krämer, Pastor, Chem. Rev., **119**, 9 (2019).

Outline

- Introduction
  - The CHARMM36 Lipid Force Field
  - Force Field Parameterization
- Applications
  - Improving Permeabilities
  - Adapting CHARMM36 for LJ-PME
- Outlook
  - Continuous Integration
  - Polarizable CHARMM Force Field

Overview of my Work



Undergraduate Research with Dirk Reith
Global Optimization of Force Field Parameters

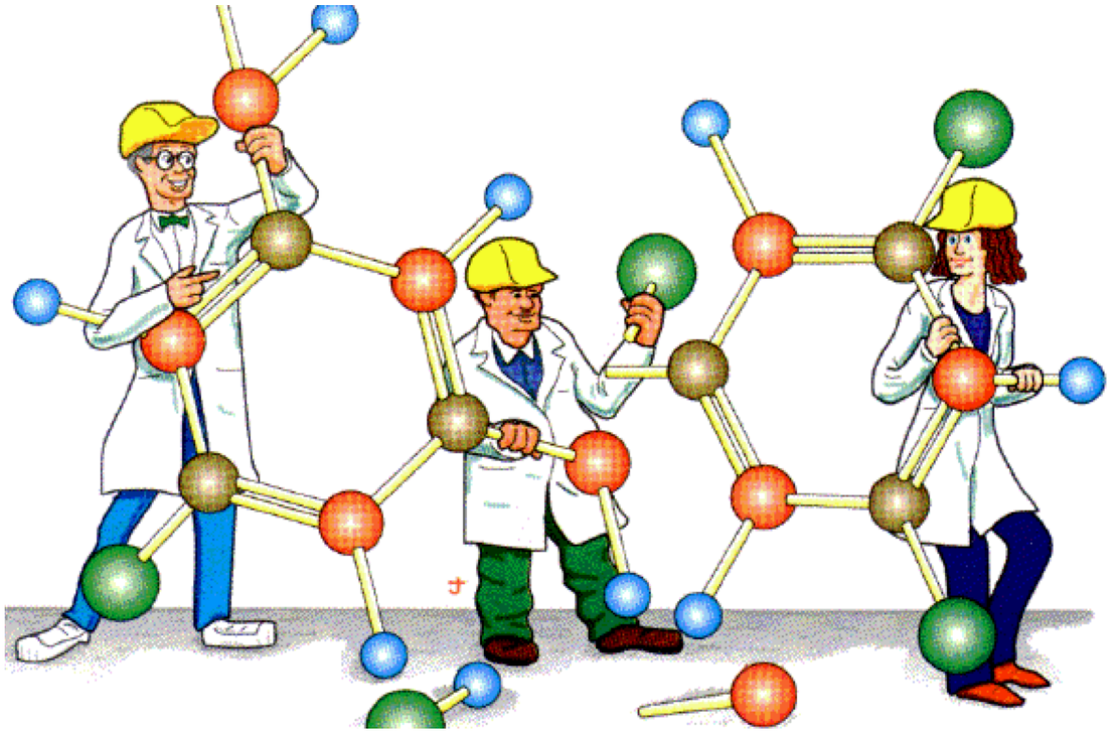
Graduate Research with Holger Foysi
Computational Fluid Dynamics, Lattice Boltzmann

Postdoc with Bernard Brooks and Richard Pastor

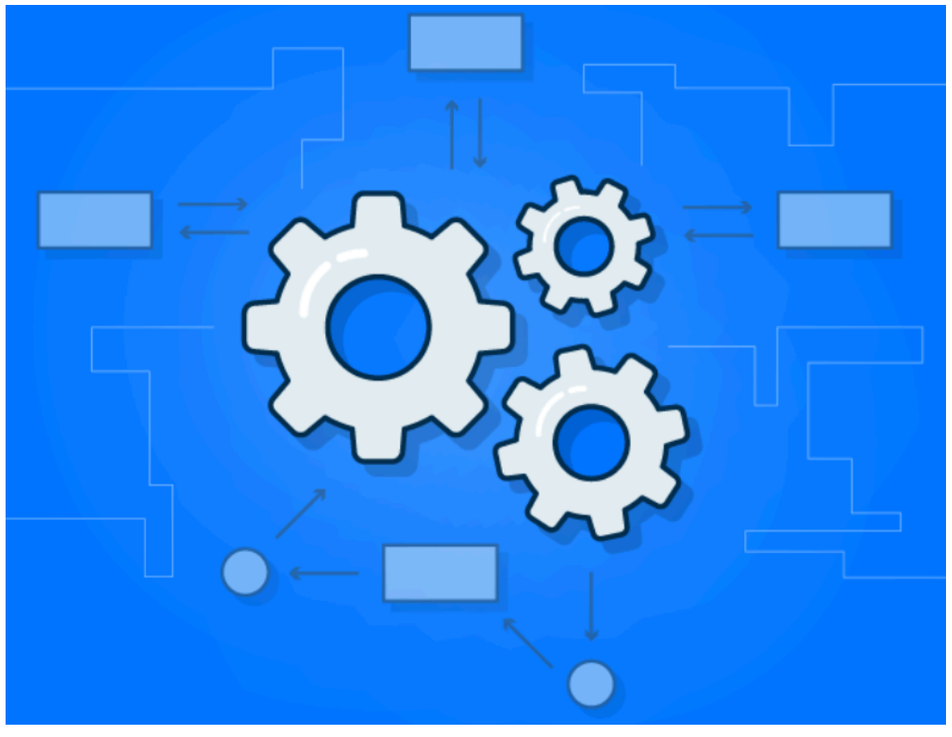
Membrane Permeability
Force Field Parametrization
Alchemical Free Energy Simulations

Force Field Development

Traditional

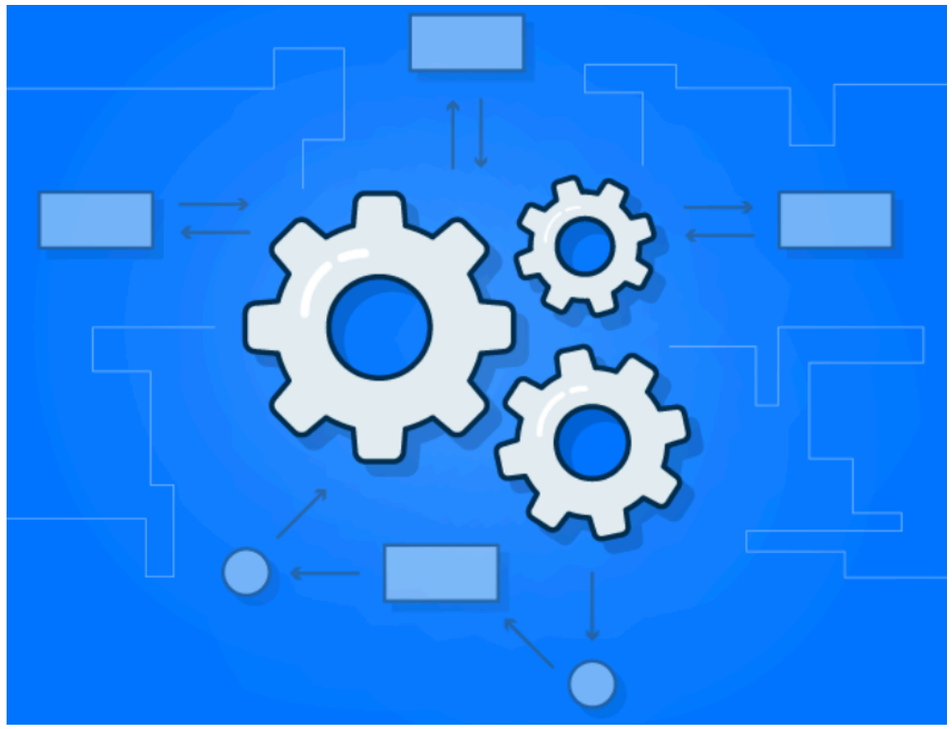


Automated

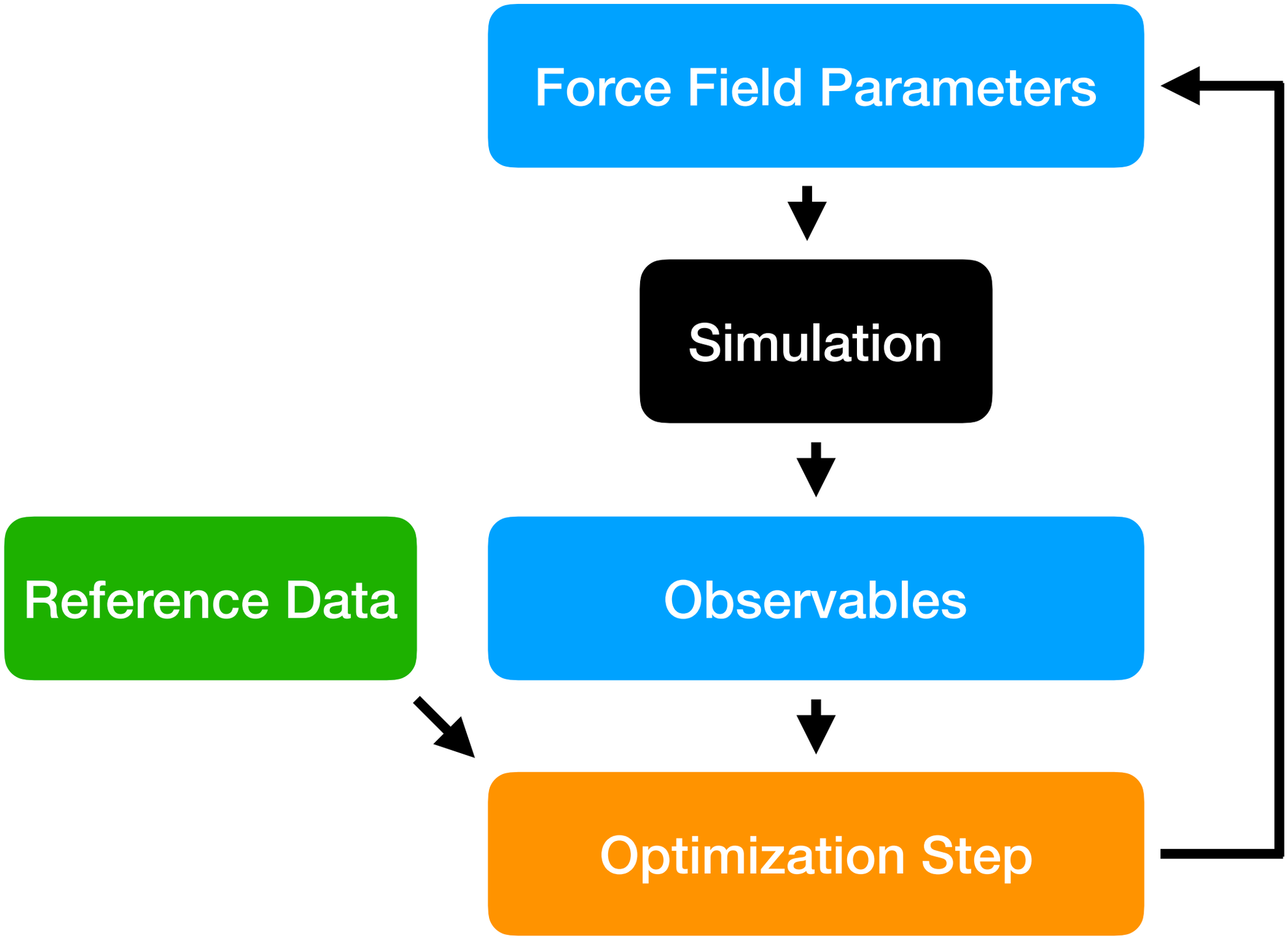


Why is automation important?

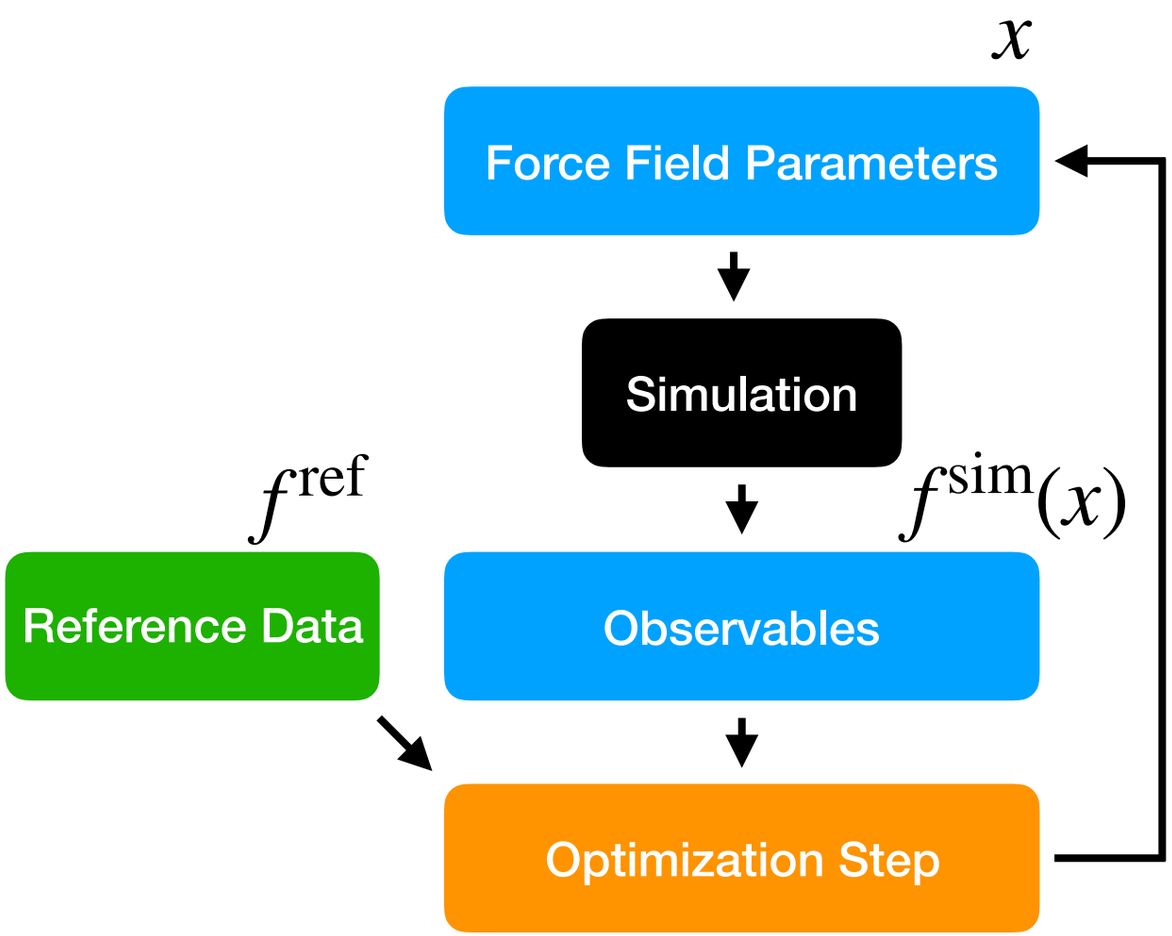
- Development times
- Human error
- Maintainability (Refactoring)
- Transparency



Automated Force Field Optimization

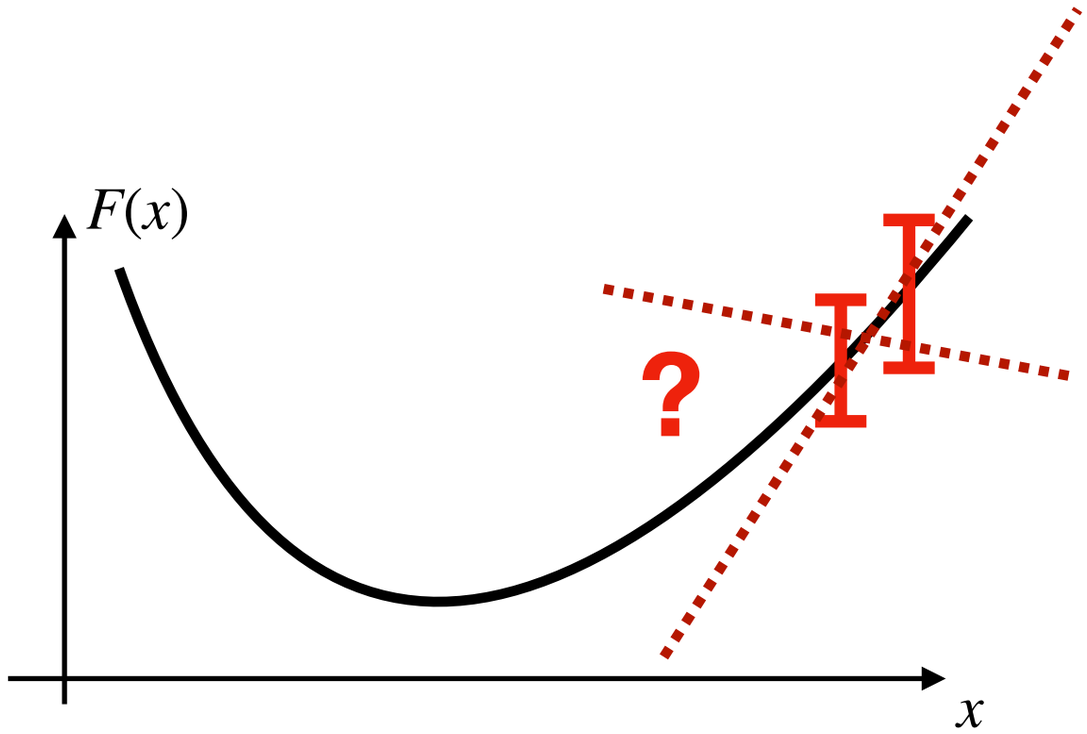


Automated Force Field Optimization



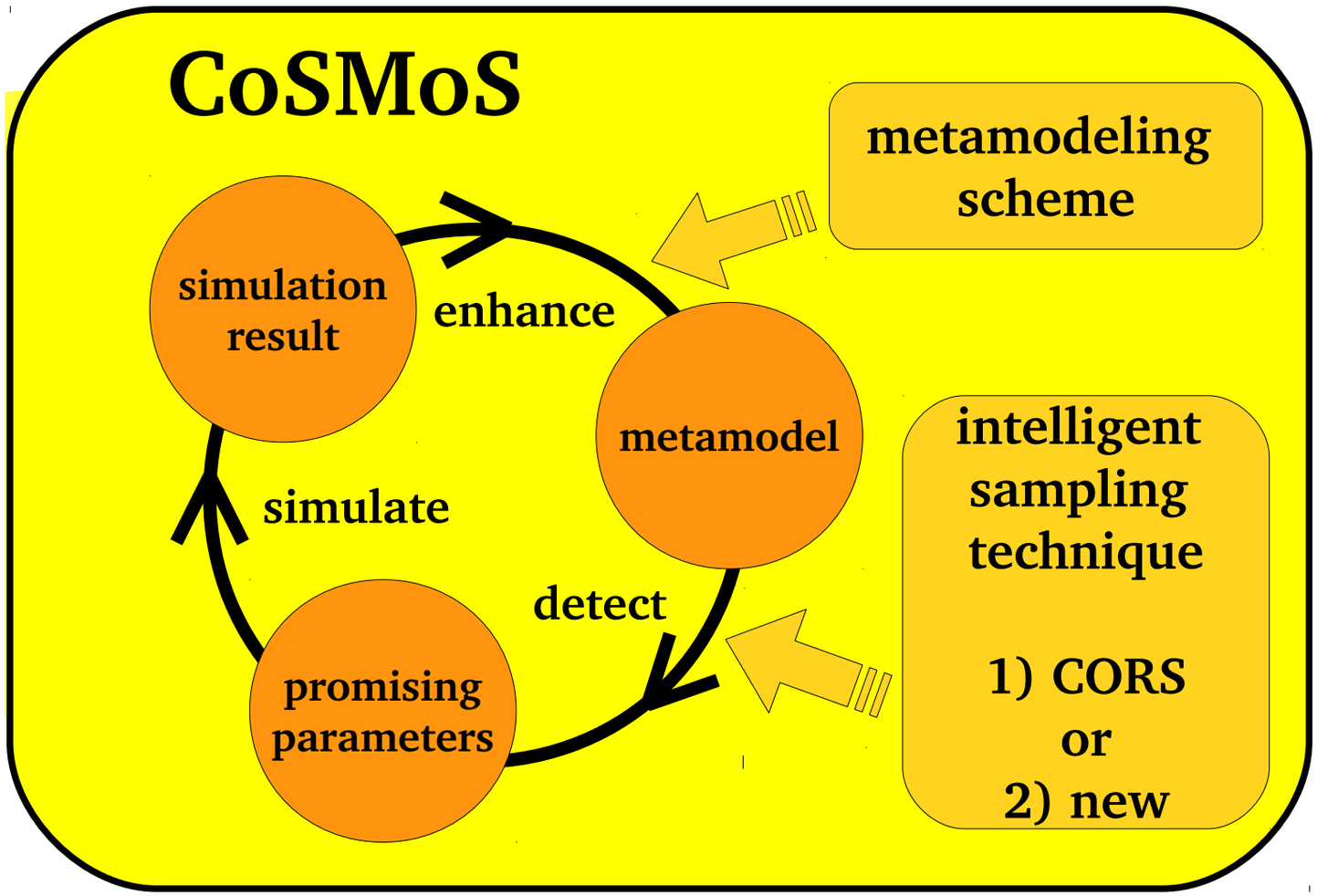
- Mostly gradient-based optimization algorithms
- Simulation as a black box
- Gradients from finite differences (expensive and noisy!)

$$\text{min!} \quad F(x) = \left\| W \left(f^{\mathrm{sim}}(x) - f^{\mathrm{ref}} \right) \right\|$$



Hülsmann, Köddermann, Vrabec, Reith, Comp. Phys. Comm. **181** (2010).

Metamodel-based Automated Force Field Optimization



- Metamodel based
- Global approximation to f^{sim}
- Noise filtered by regression
- Intelligent sampling balances exploration vs. exploitation
- Works well for few parameters

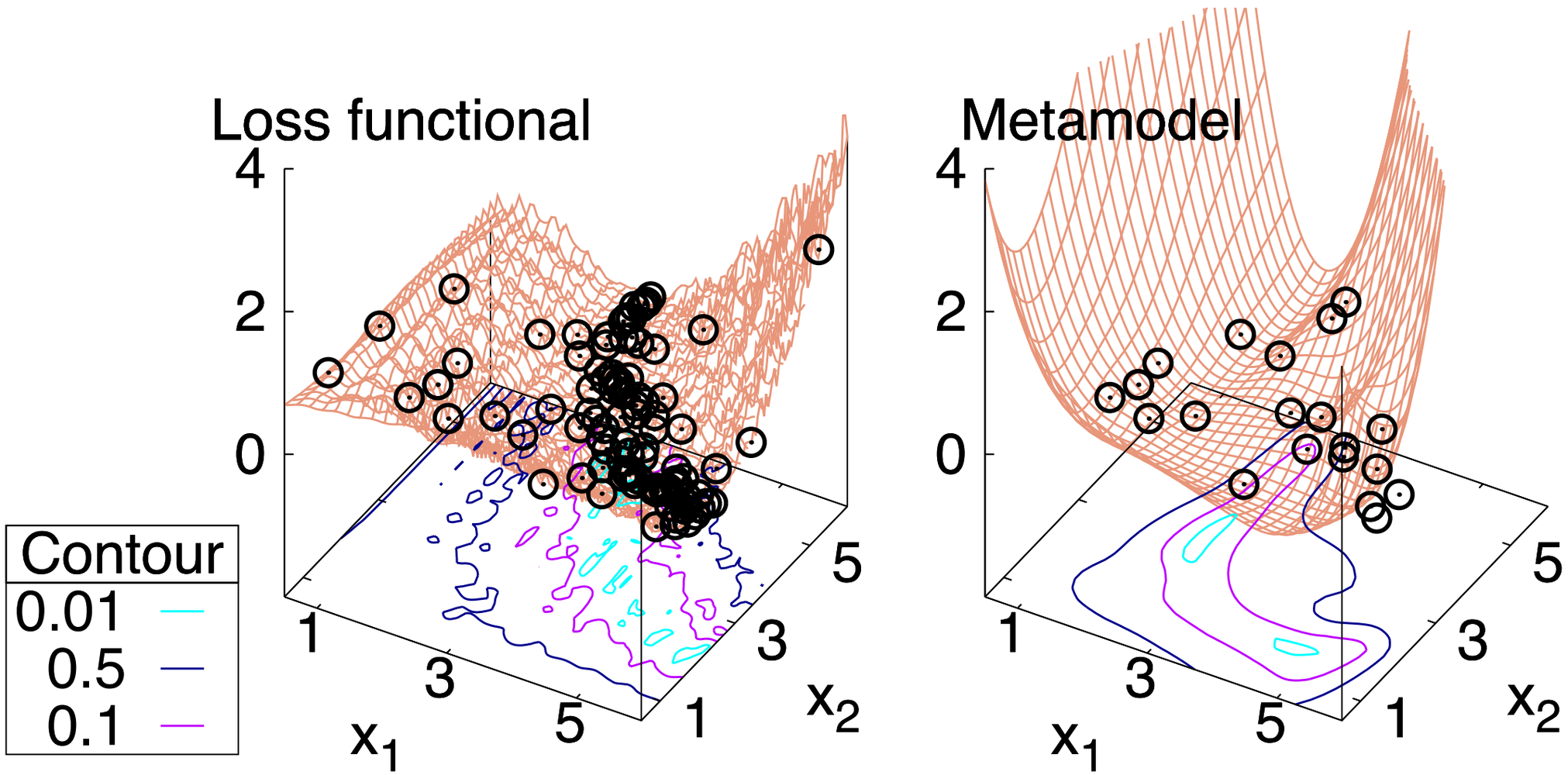
$$\min! \quad F(x) = \left\| W \left(f^{\mathrm{sim}}(x) - f^{\mathrm{ref}} \right) \right\|$$

Krämer, Hülsmann, Köddermann, Reith, Comput. Phys. Commun. **185**, 12 (2014)

Elfgen, Hülsmann, Krämer, Köddermann, Kirschner, Reith, Eur. Phys. J. Spec. Top. **225** (2016)

Hülsmann, Kirschner, Krämer, Heinrich, Krämer-Fuhrmann, Reith, Select Pap. from FOMMS 2015 (2016)

Metamodel-based Automated Force Field Optimization



Philosophy: Simulation is an expensive black box. Remember everything.

Krämer, Hülsmann, Köddermann, Reith, Comput. Phys. Commun. **185**, 12 (2014)

Physics-Informed Prediction of Simulation Results

Evaluation

Simulation Trajectory

$$f^{\text{sim}}(x) =: \langle f^{\text{sim}} \rangle_A$$

Partial Derivative Calculations

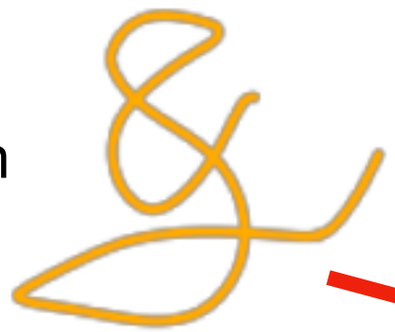
$$f^{\text{sim}}(x + he_i) =: \langle f^{\text{sim}} \rangle_B$$

$$\left\langle f^{\text{sim}} \right\rangle_B$$

Physics-Informed Prediction of Simulation Results

Evaluation

Simulation Trajectory



$$f^{\text{sim}}(x) =: \langle f^{\text{sim}} \rangle_A$$

Partial Derivative Calculations

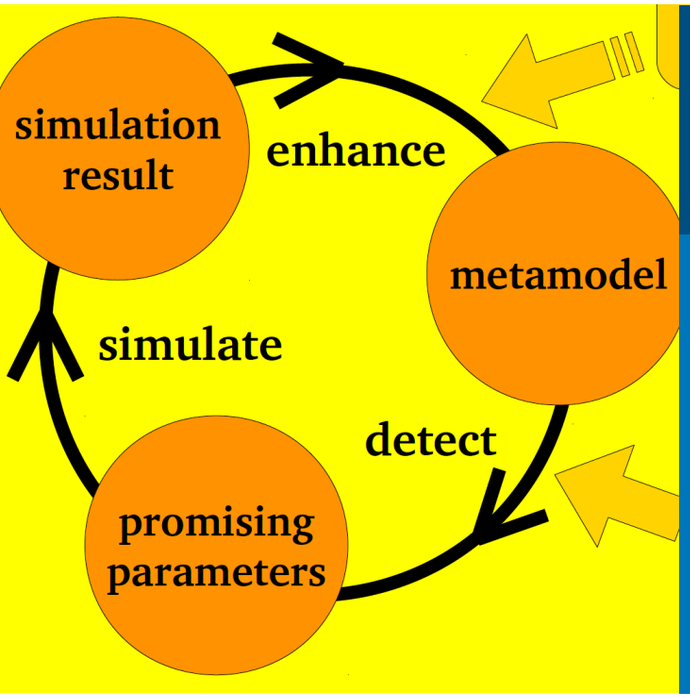
$$f^{\text{sim}}(x + he_i) =: \langle f^{\text{sim}} \rangle_B$$

$$\langle f^{\text{sim}} \rangle_B = \frac{\int dr^{3N} f^{\text{sim}} e^{-\beta U_B}}{\int dr^{3N} e^{-\beta U_B}} = \frac{\int dr^{3N} f^{\text{sim}} e^{-\beta(U_B - U_A)} e^{-\beta U_A}}{\int dr^{3N} e^{-\beta(U_B - U_A)} e^{-\beta U_A}} = \frac{\langle f^{\text{sim}} e^{-\beta(U_B - U_A)} \rangle_A}{\langle e^{-\beta(U_B - U_A)} \rangle_A}$$

Cheap + accurate gradient calculations, no additional simulations.

Numerical Optimization = Metamodeling

- MBAR: Statistical optimal predictions + error estimates
- Requires descent configurational overlap between states

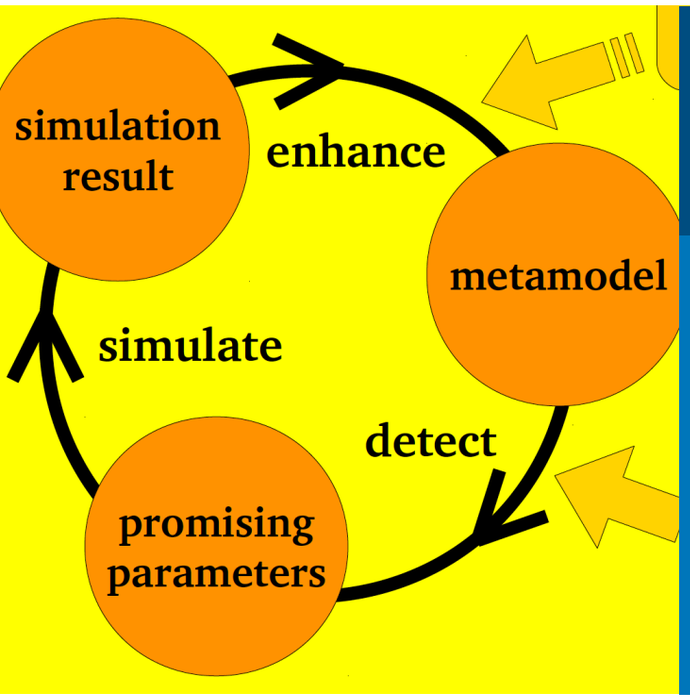


Metamodel	Black-Box	Physics-Informed
Local (Linear, Quadratic)	GROW (Finite Differences)	Reweighting ForceBalance MBAR
Global	CoSMoS (Response Surface)	MBAR? Pair Correlation Function Reweighting (PCFR)

Shirts and Chodera, J. Chem. Phys. **129**, 12 (2008)
Wang, Martinez, Pande, J. Phys. Chem. Lett. **5**, 11 (2014)
Messerly, Razavi, Shirts, J. Chem. Theory Comput. **14**, 6 (2018)

Numerical Optimization = Metamodeling

- MBAR: Statistical optimal predictions + error estimates
- Requires descent configurational overlap between states



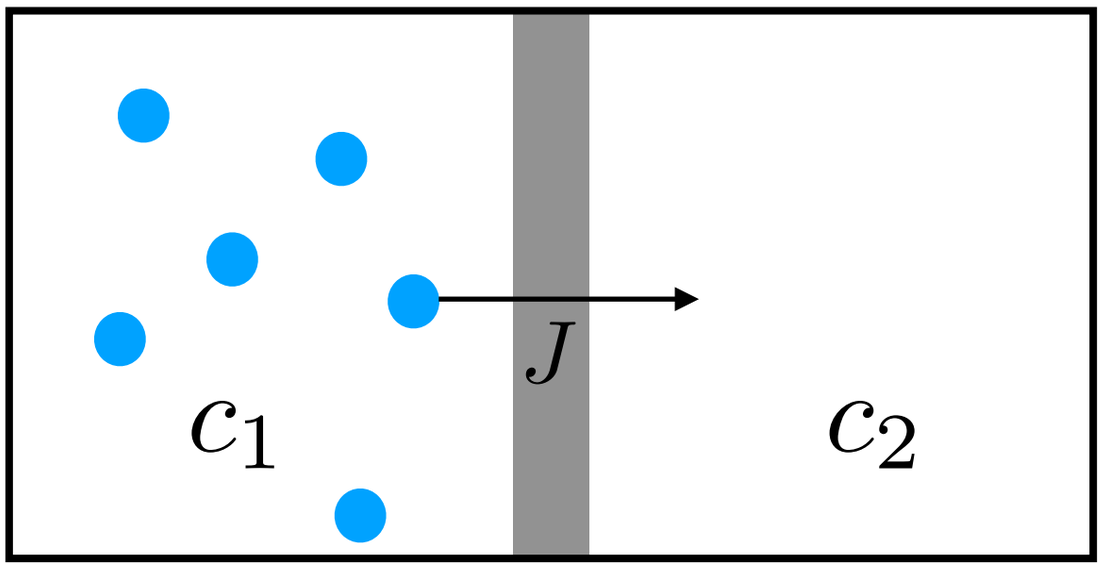
	Metamodel	Black-Box	Physics-Informed
Local (Linear, Quadratic)		GROW (Finite Differences)	Reweighting ForceBalance MBAR
Global		CoSMoS (Response Surface)	MBAR? Pair Correlation Function Reweighting (PCFR)

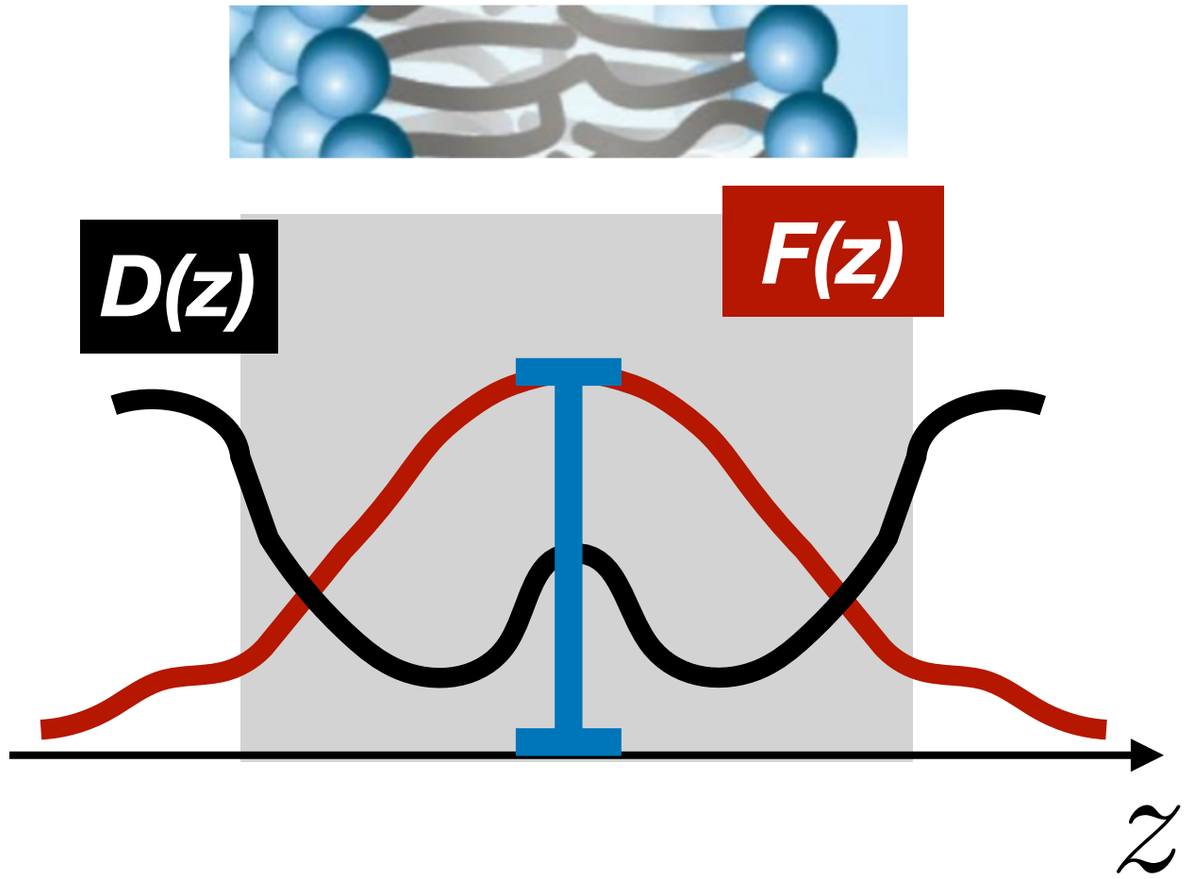
Shirts and Chodera, J. Chem. Phys. **129**, 12 (2008)
Wang, Martinez, Pande, J. Phys. Chem. Lett. **5**, 11 (2014)
Messerly, Razavi, Shirts, J. Chem. Theory Comput. **14**, 6 (2018)

Outline

- Introduction
  - The CHARMM36 Lipid Force Field
  - Force Field Parameterization
- Applications
  - Improving Permeabilities
  - Adapting CHARMM36 for LJ-PME
- Outlook
  - Continuous Integration
  - Polarizable CHARMM Force Field

Membrane Permeability





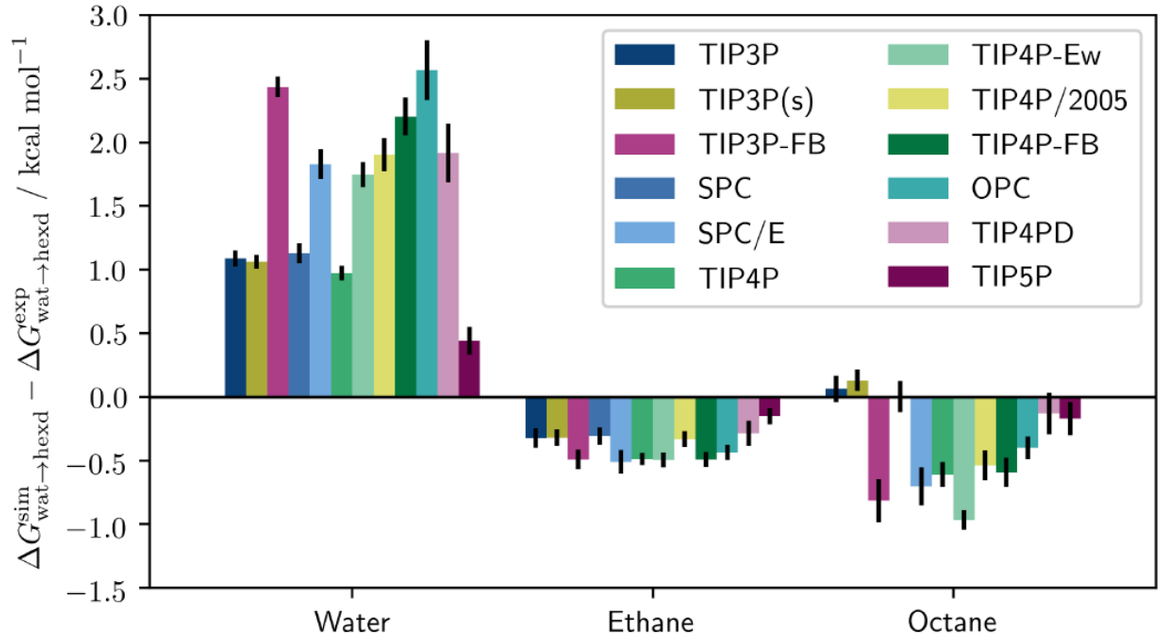
Permeability depends on the free energy and diffusion profiles.

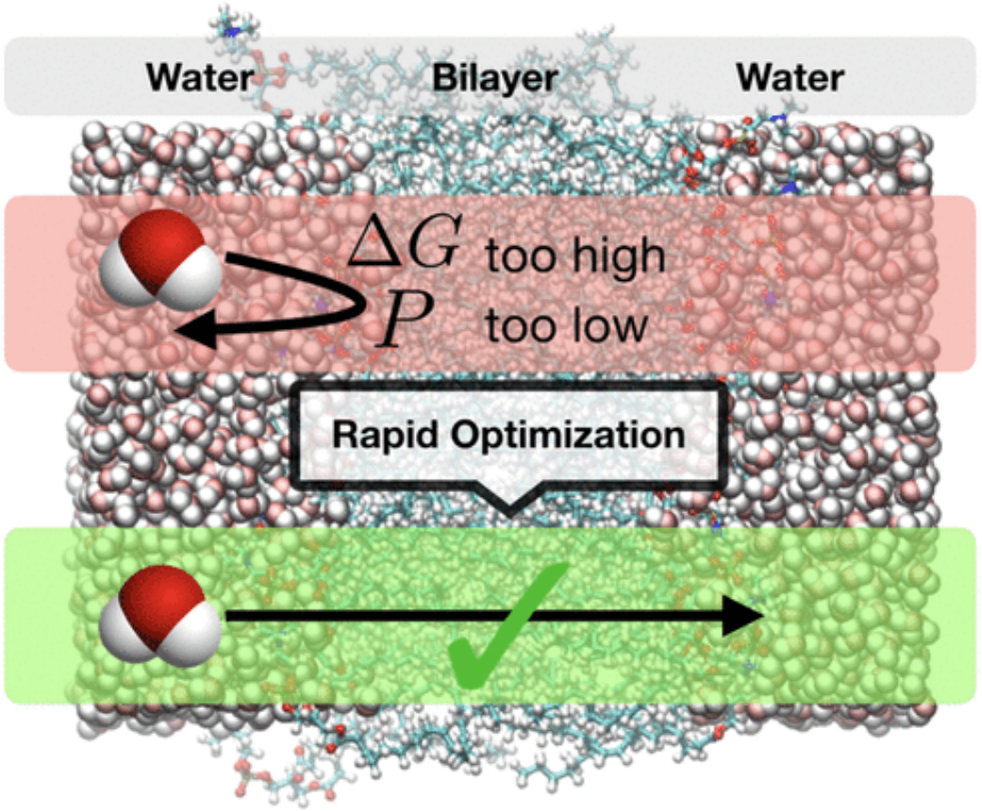
Krämer, Pickard, Huang, Venable, Simmonett, Reith, Kirschner, Pastor, Brooks, J. Chem. Theory Comput., **15**, 6 (2019).

Venable, Krämer, Pastor, Chem. Rev., **119**, 9 (2019).

Water/Alkane Interactions

- Additive water models underestimate water partitioning into alkanes.
- Most recent water additive models perform worst (explicitly optimized for bulk, not transferrable to low-dielectric environments)
- Water permeation through bilayers is up to an order of magnitude too low in the CHARMM36 force field.

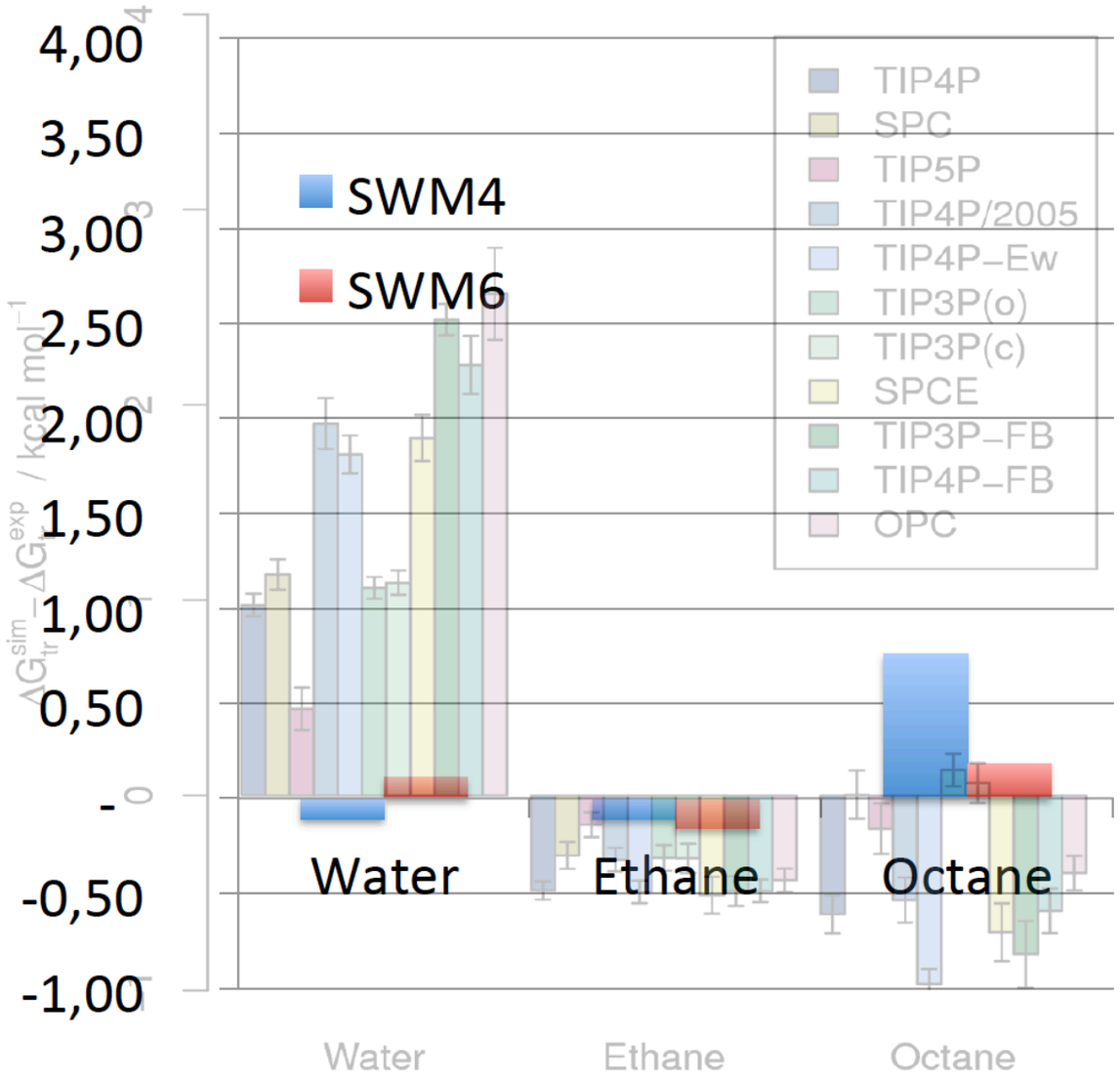




Krämer, Pickard, Huang, Venable, Simmonett, Reith, Kirschner, Pastor, Brooks, J. Chem. Theory Comput., **15**, 6 (2019).

Results for Polarizable Water Models

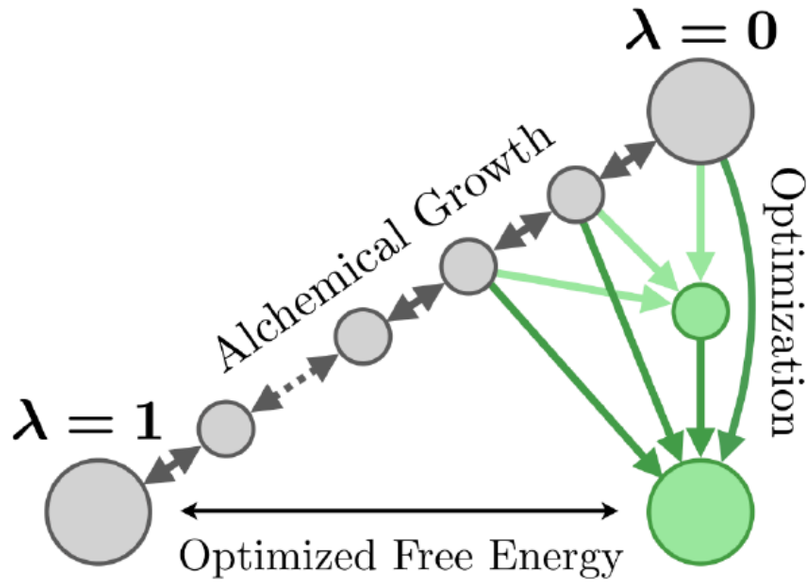
- Much better description of water transfer into alkanes
- Can we get similar results in an additive force field?

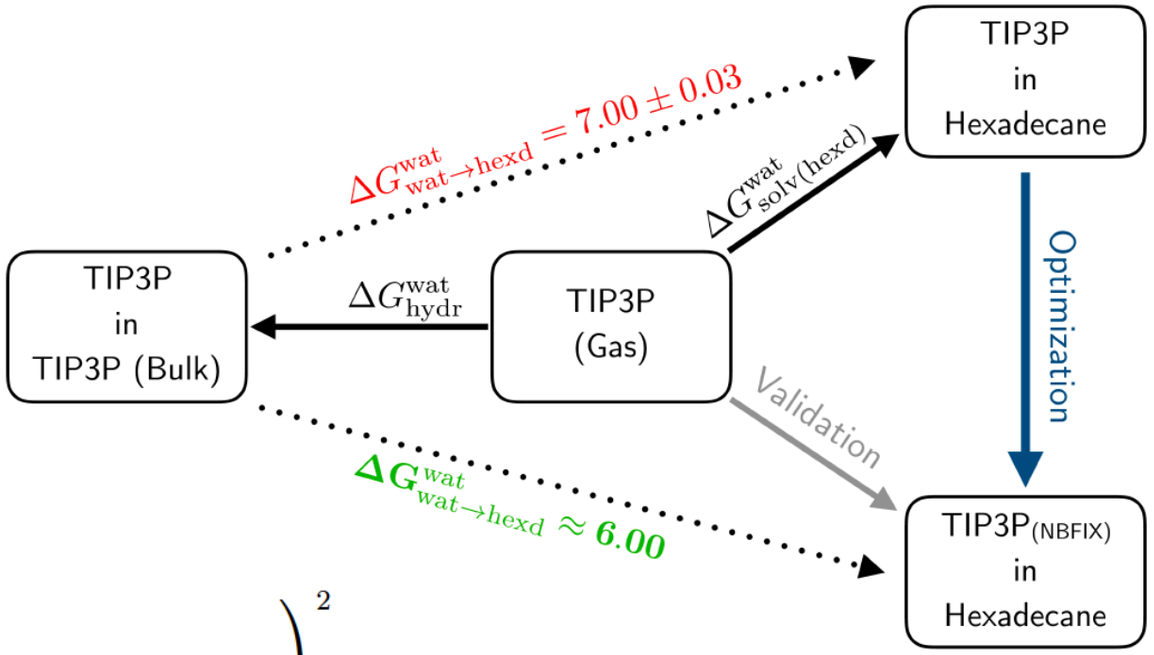


Krämer, Pickard, Huang, Venable, Simmonett, Reith, Kirschner, Pastor, Brooks, J. Chem. Theory Comput., **15**, 6 (2019).

Optimization Procedure

- Automated optimization of the force field parameters (pairwise water-alkane LJ parameters) using multistate reweighting.
- Reuse the last lambda states from an alchemical simulation to predict free energies for new parameters
- Optimization converged within less than a day on a single GPU node.



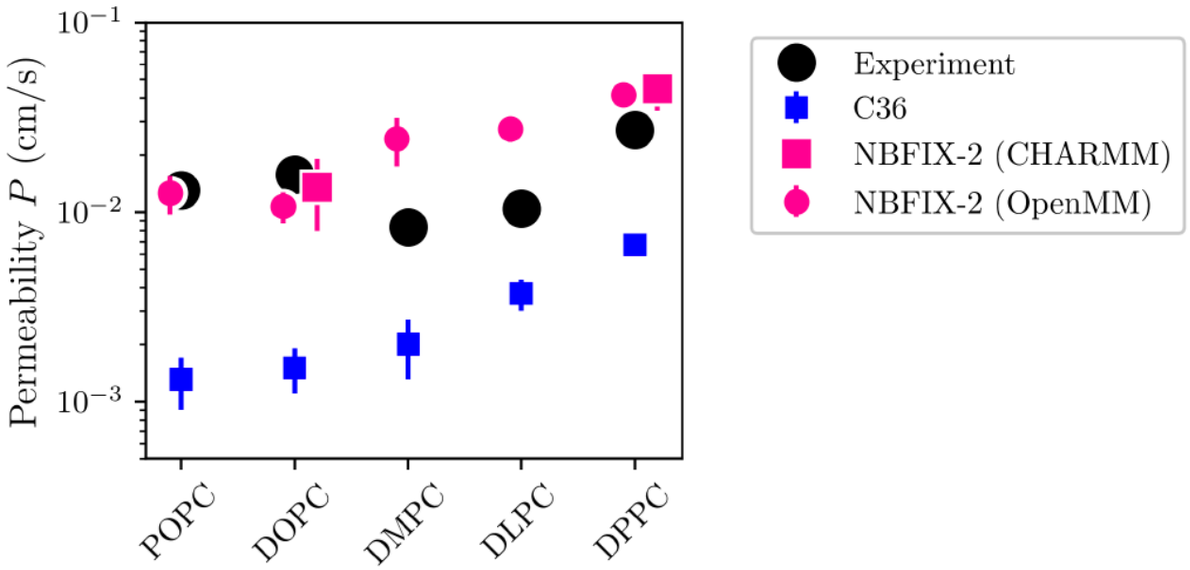


$$\text{min!} \quad f(\lambda_\sigma, \lambda_\varepsilon) \overset{\text{Eq.(1)}}{=} \left(- \underbrace{\Delta G^{\text{wat,exp}}_{\text{wat}\to\text{hexd}}}_{\text{experimental reference}} - \underbrace{\Delta G^{\text{wat}}_{\text{hydr}} + \Delta G^{\text{wat}}_{\text{solv(hexd)}}}_{\text{from alchemical growth}} + \Delta G^{\text{MR}}_{\text{nbfix}}(\lambda_\sigma, \lambda_\varepsilon) \right)^2$$

Krämer, Pickard, Huang, Venable, Simmonett, Reith, Kirschner, Pastor, Brooks, J. Chem. Theory Comput., **15**, 6 (2019).

Water/Alkane Interactions (ctd.)

- Optimized parameters ("NBFIX-2"):
  - Epsilon increased by 7%
  - Sigma decreased by 15%
- Results:
  - Perfect recovery of water partitioning into alkanes
  - Better match of interfacial tensions (4% overestimate instead of 10% underestimate in C36).
  - Better permeability of water through lipid bilayers.



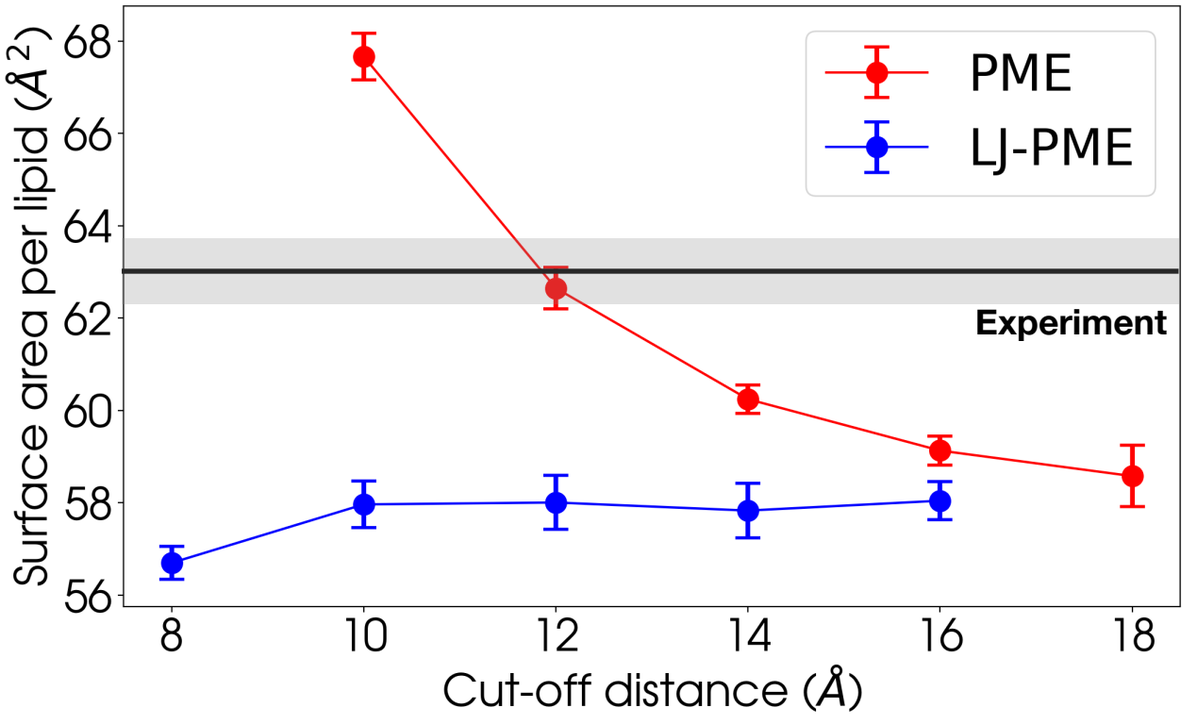
Krämer, Pickard, Huang, Venable, Simmonett, Reith, Kirschner, Pastor, Brooks, J. Chem. Theory Comput., **15**, 6 (2019).

Outline

- Introduction
  - The CHARMM36 Lipid Force Field
  - Force Field Parameterization
- Applications
  - Improving Permeabilities
  - Adapting CHARMM36 for LJ-PME
- Outlook
  - Continuous Integration
  - Polarizable CHARMM Force Field

C36 is great. Why Optimization?

- **Long-range dispersion**
  - Interfacial systems are better represented by incorporating long-range Lennard-Jones (LJ-PME)
  - C36m protein force field has been validated for LJ-PME
  - Monolayer surface tensions better with long-range LJ
  - Inconsistencies regarding LJ treatment
- TIP3P water model?
- **Permeability**
  - Partitioning between water and organic phases is not well modeled by additive FF

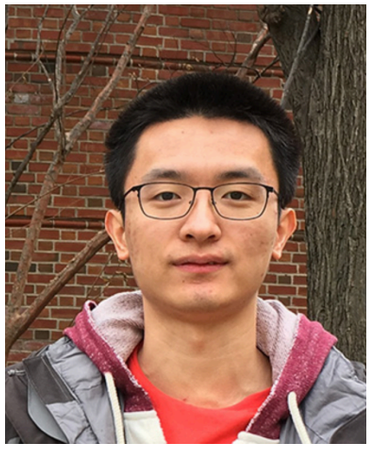


Leonard, Simmonett, Pickard, Huang, Venable, Klauda, Brooks, Pastor, J. Chem. Theory Comput., **14**, 2 (2018).

Krämer, Pickard, Huang, Venable, Simmonett, Reith, Kirschner, Pastor, Brooks, J. Chem. Theory Comput., **15**, 6 (2019).

Venable, Krämer, Pastor, Chem. Rev., **119**, 9 (2019).

Optimization Approach for C36 with LJ-PME



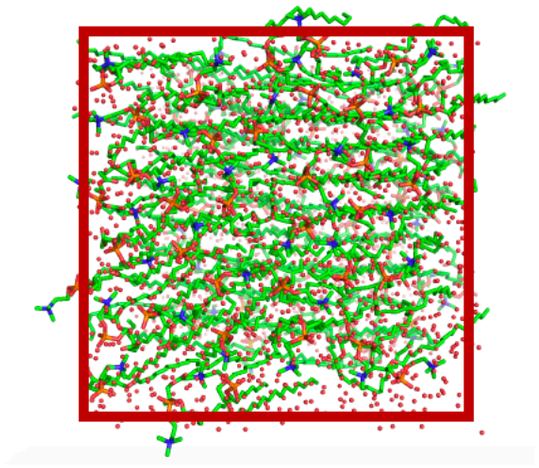
- Objectives:
  - Adapt parameters to LJ-PME
  - Retain C36-like quality for bilayers
  - Achieve consistency for monolayers/bilayers
- Parameters: Charges, LJ, Torsion Parameters in PC Headgroup
- (Alkane parameters well-validated for LJ-PME, Leonard et al. JCTC)

Surface tensions for C36 monolayers are way off

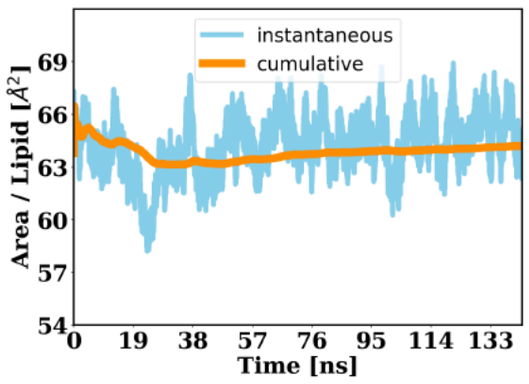
Area [Å^2/lipid]	IPS DFFT	PME r_c=12 Å	Experiment
54	17.8±1.3	−7.0±1.2	17.9
64	43.9±0.8	26.4±1.0	40.9
80	59.5±0.5	43.8±0.4	54.9

Optimization Targets

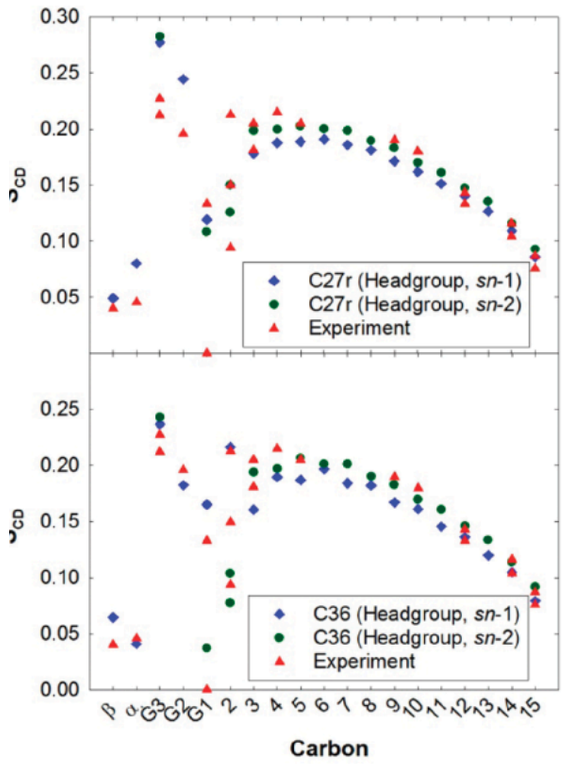
Area per Lipid



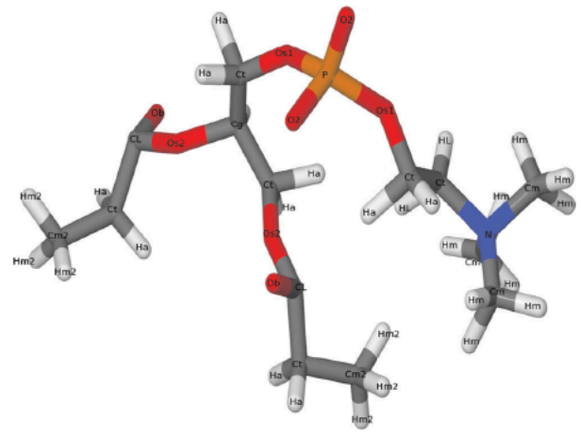
Compressibility

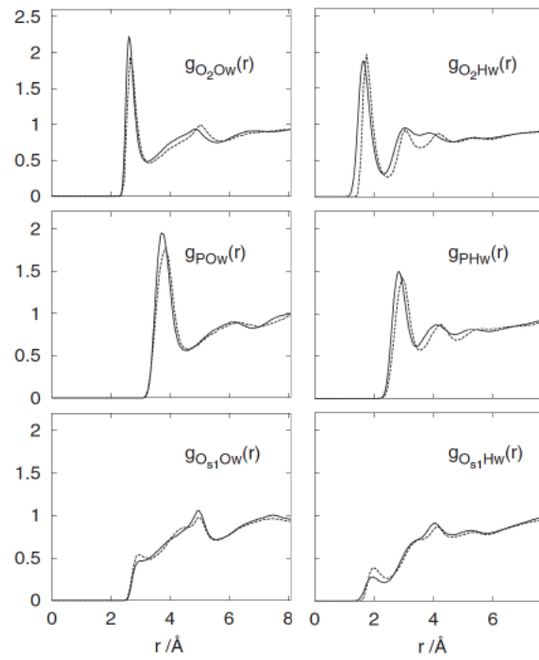


Order Parameter



Hydration of the headgroup

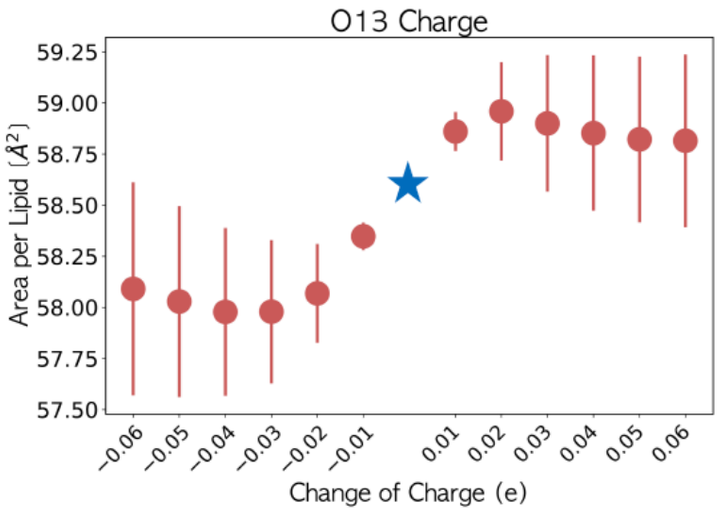


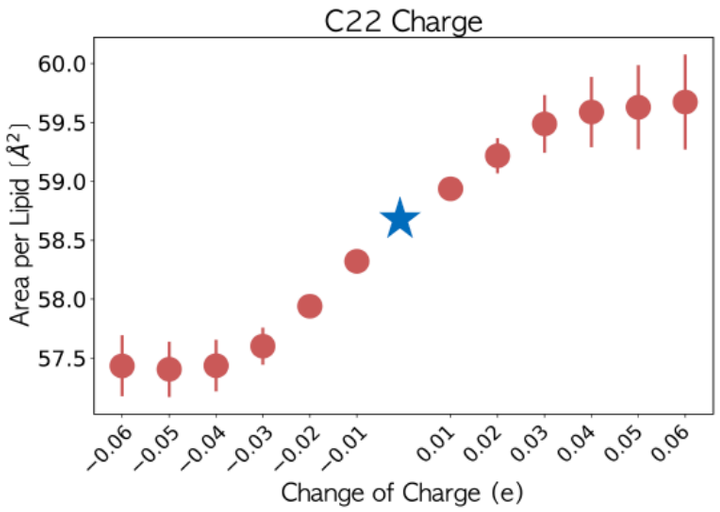


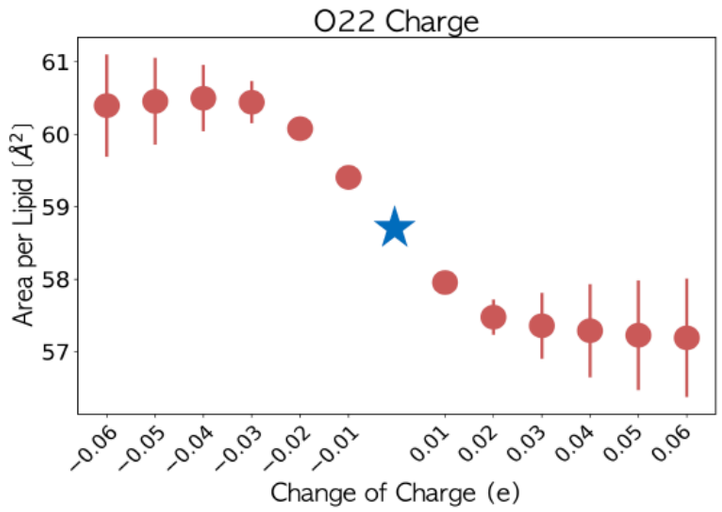
Mclain et. al. The Journal of Chemical Physics 133, 145103 (2010)

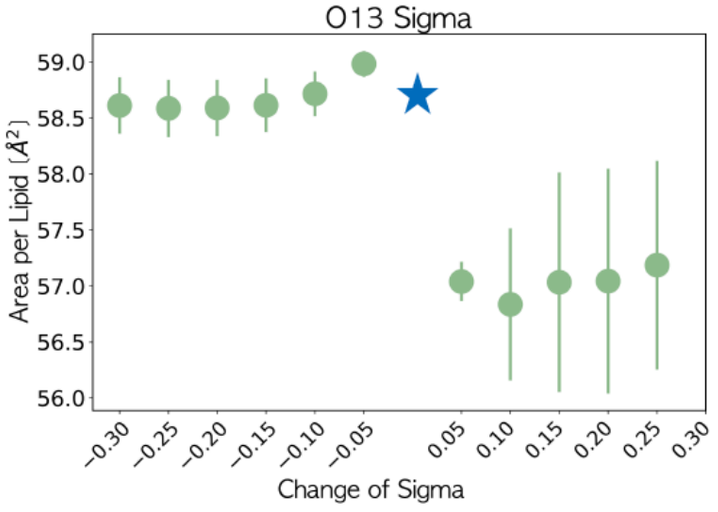
Sensitivities

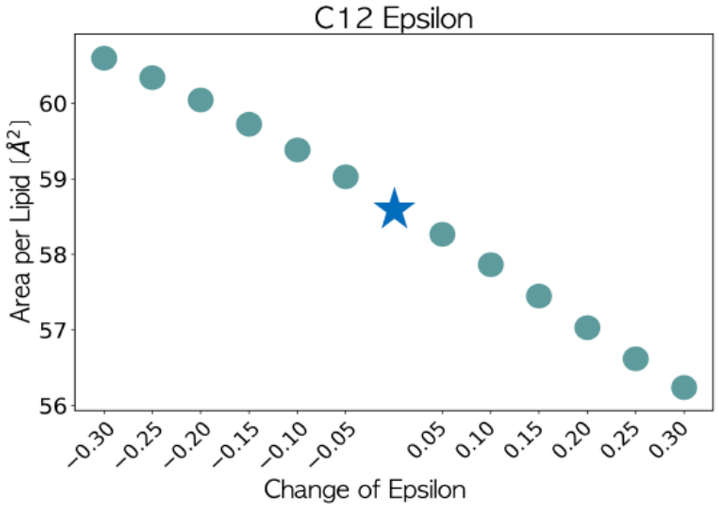
Reweighting

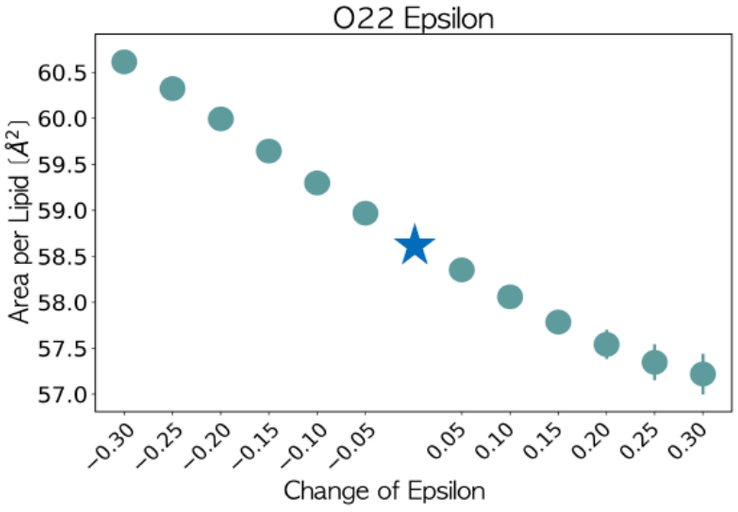






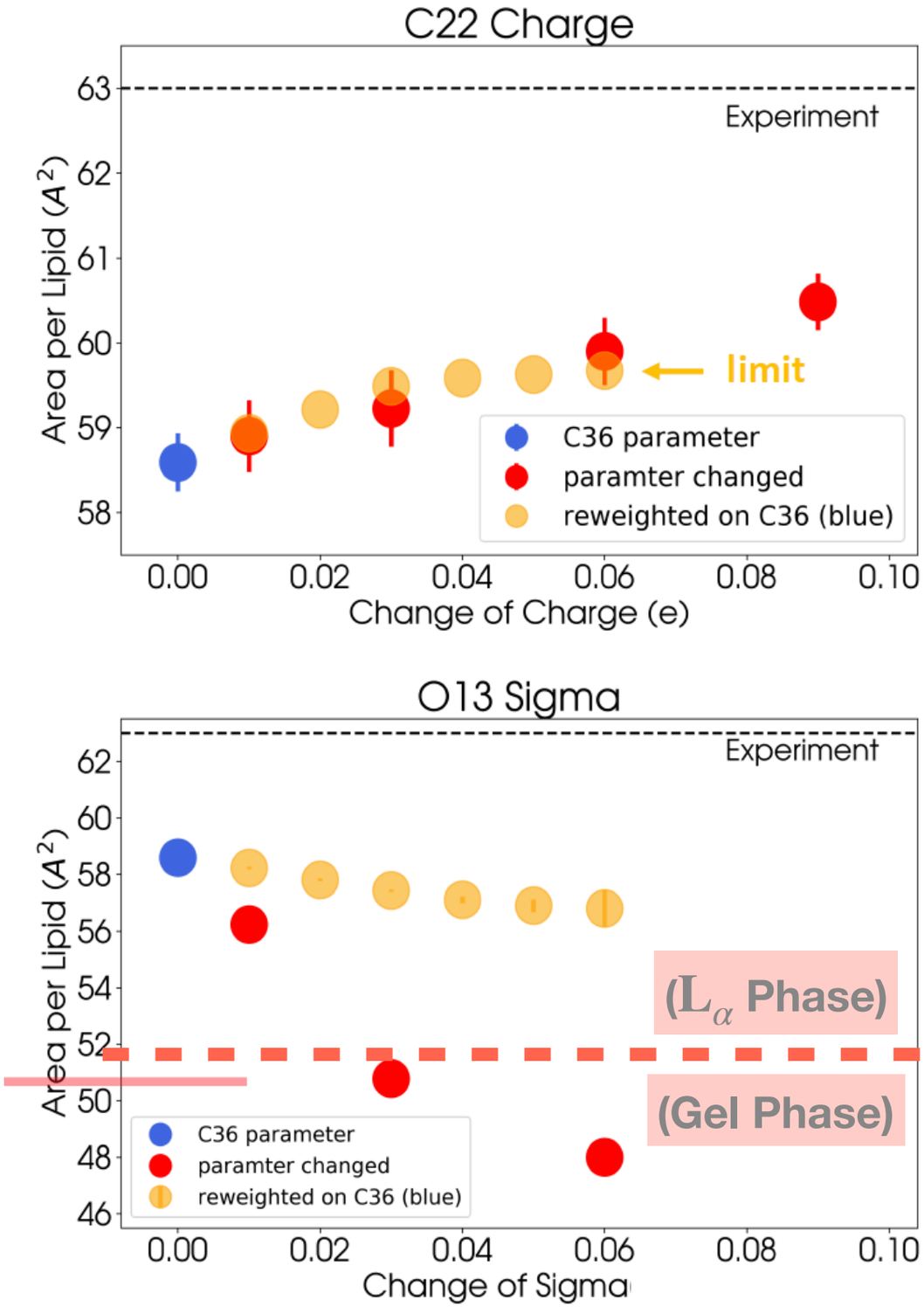






★ : simulation with the original C36 parameters

Simulation vs. Reweighting



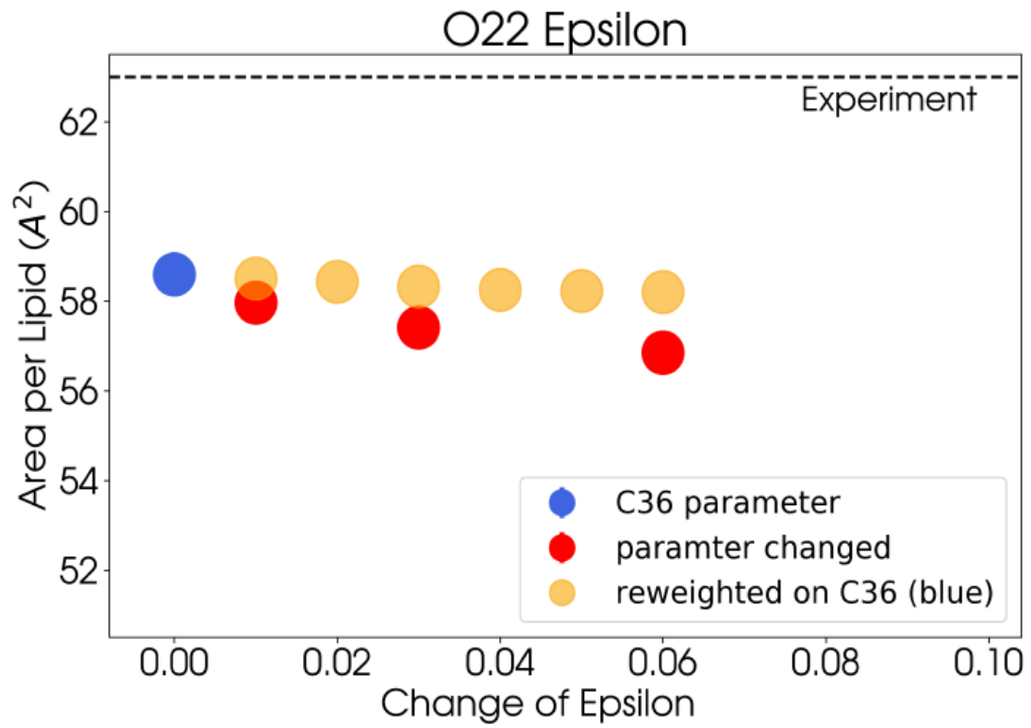
- Estimates are only meaningful for small parameter changes, correct trend

- Phosphate oxygen: 3% increase in sigma causes a phase transition

- Not captured by reweighting

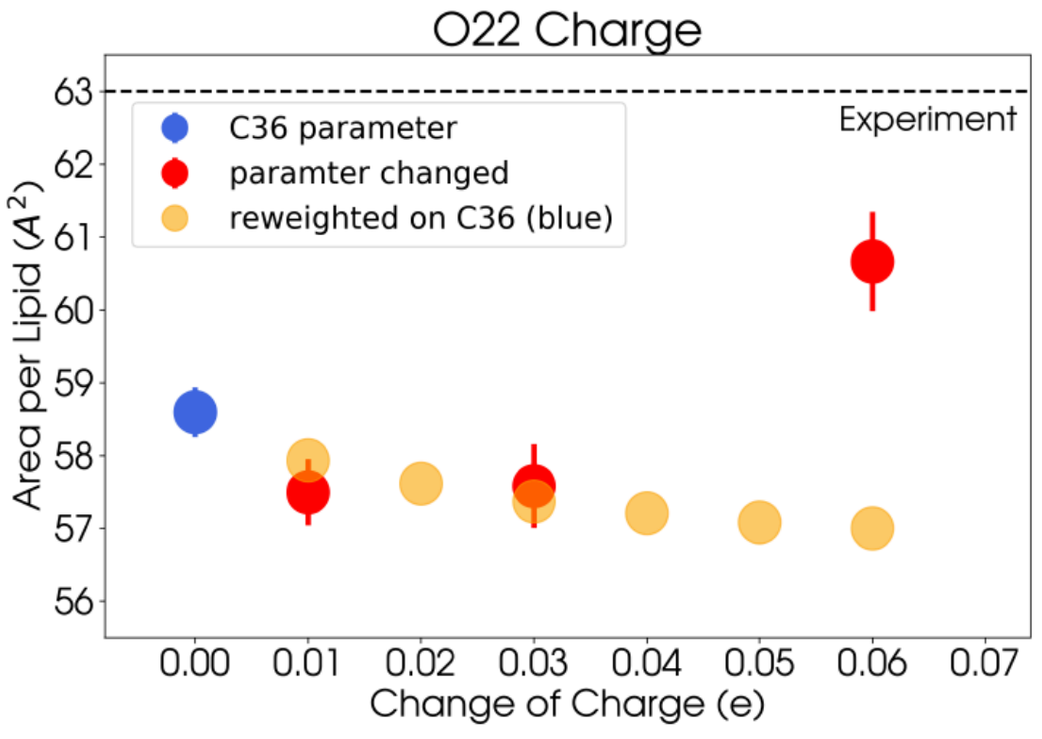
Simulation vs. Reweighting

- Most sensitivities are underestimated by reweighting



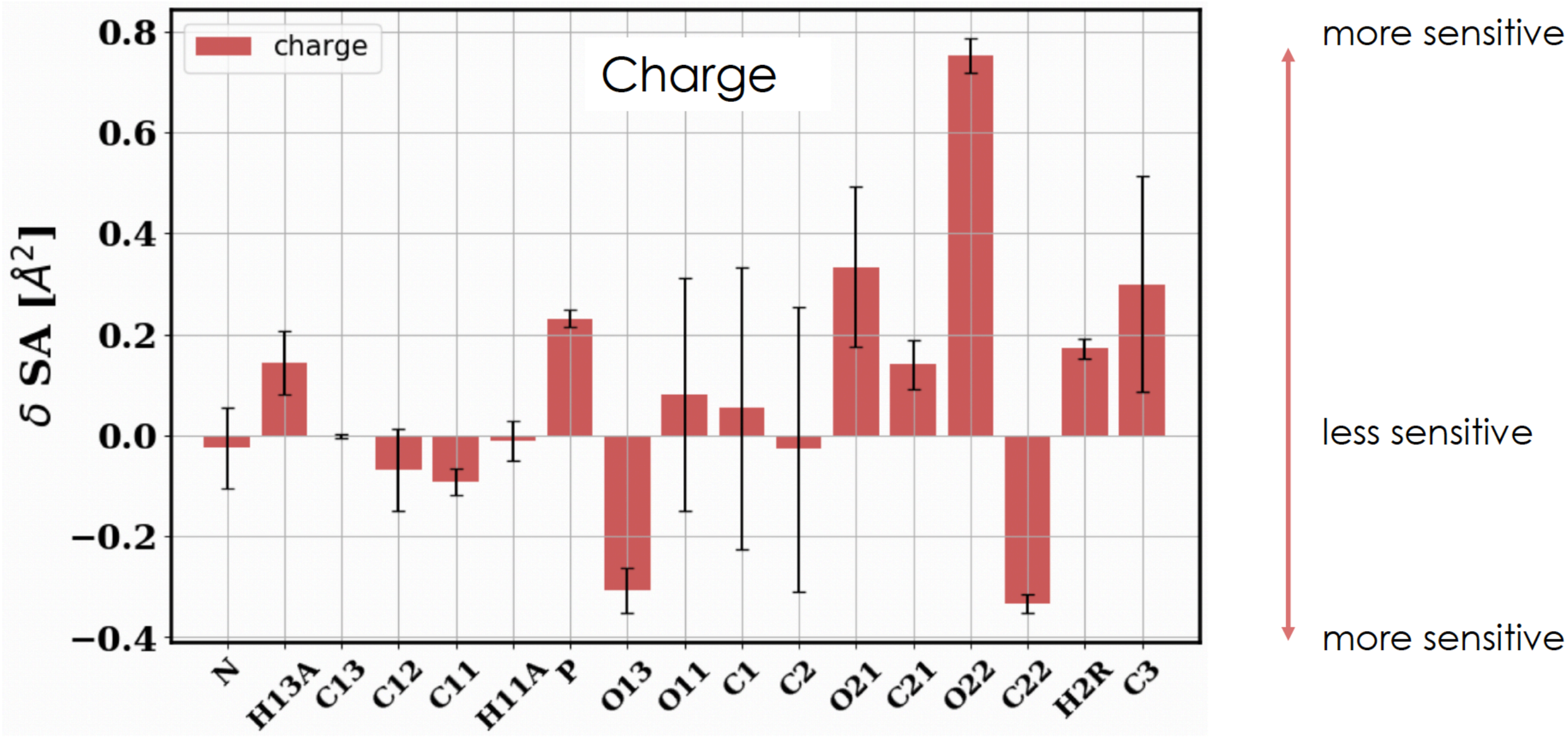
- Yield a feasible descent direction for an optimizer

- Area dependence on charges is highly nontrivial



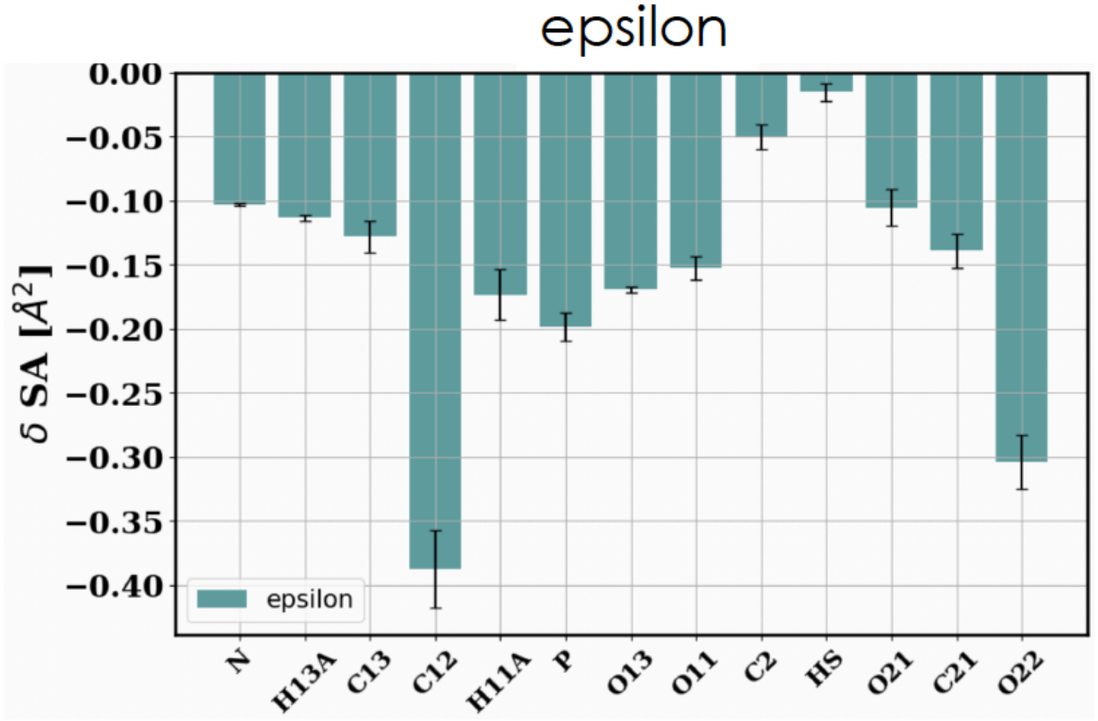
Sensitivities of Surface Area

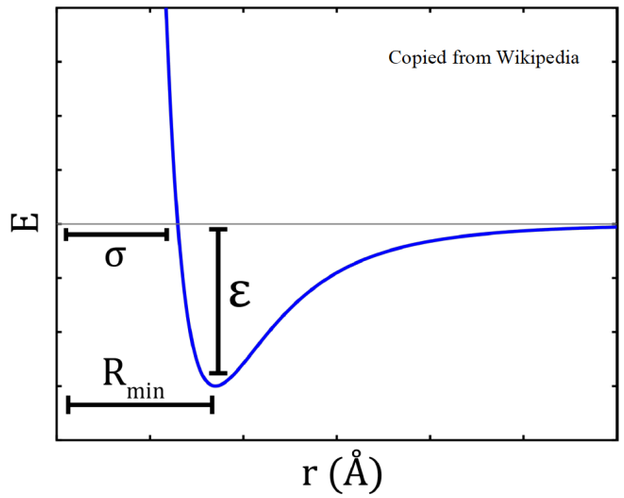
When changing the charges by only 0.05 e …

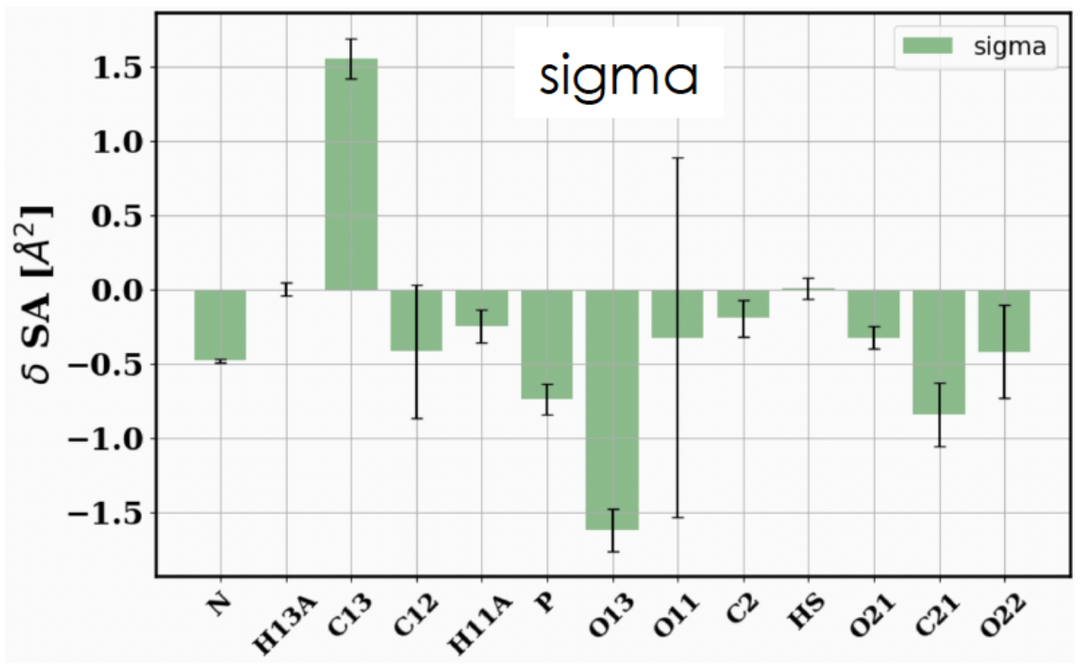


Sensitivities of Surface Area

When increasing the LJ parameters by only 5% …







Sigma sensitivities are unexpected.

Hydration effects!

Optimization Protocol

In each iteration *i* , adapt the torsion parameters to match C36/QM torsion profiles.

Then, change charges and LJ to minimize

$$F^{(i)}(x) = \left\| W \cdot \left(\mathcal{M}^{(i)}(x) - f^{\mathrm{ref}} \right) \right\|_2^2$$

Weights chosen by user;
Adapted to uncertainty of the sensitivity

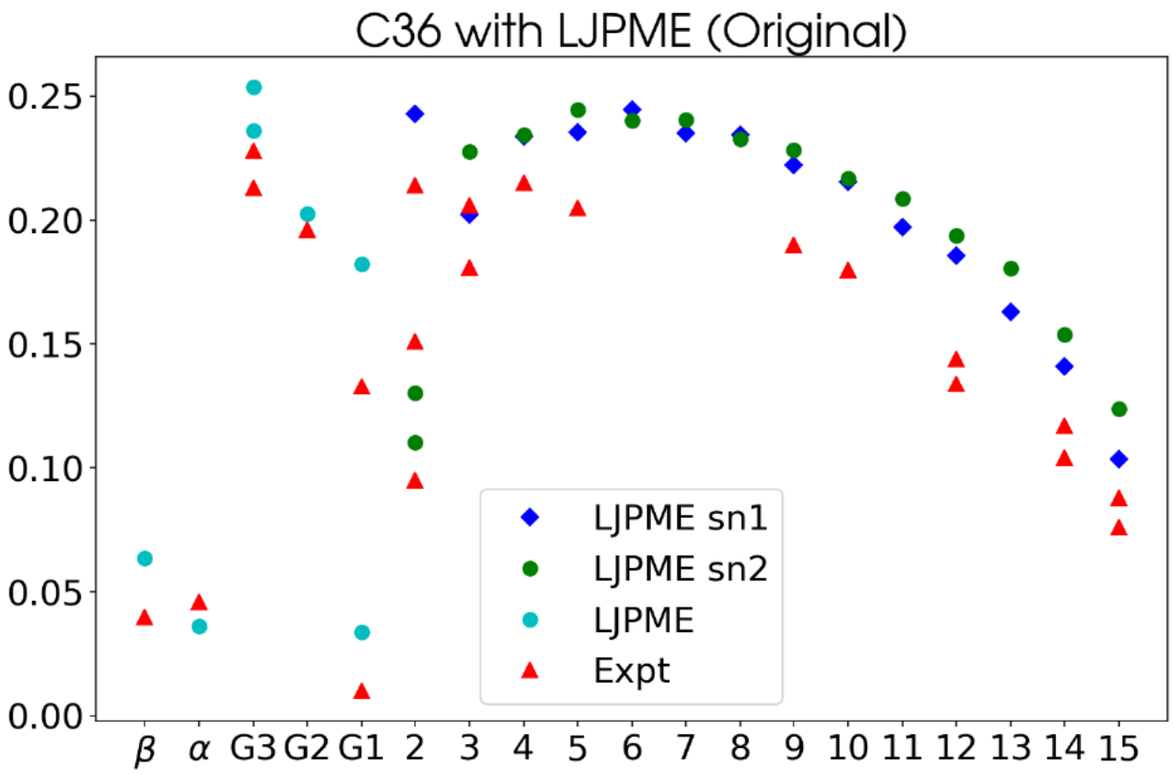
Linear metamodel
based on sensitivities
from reweighting

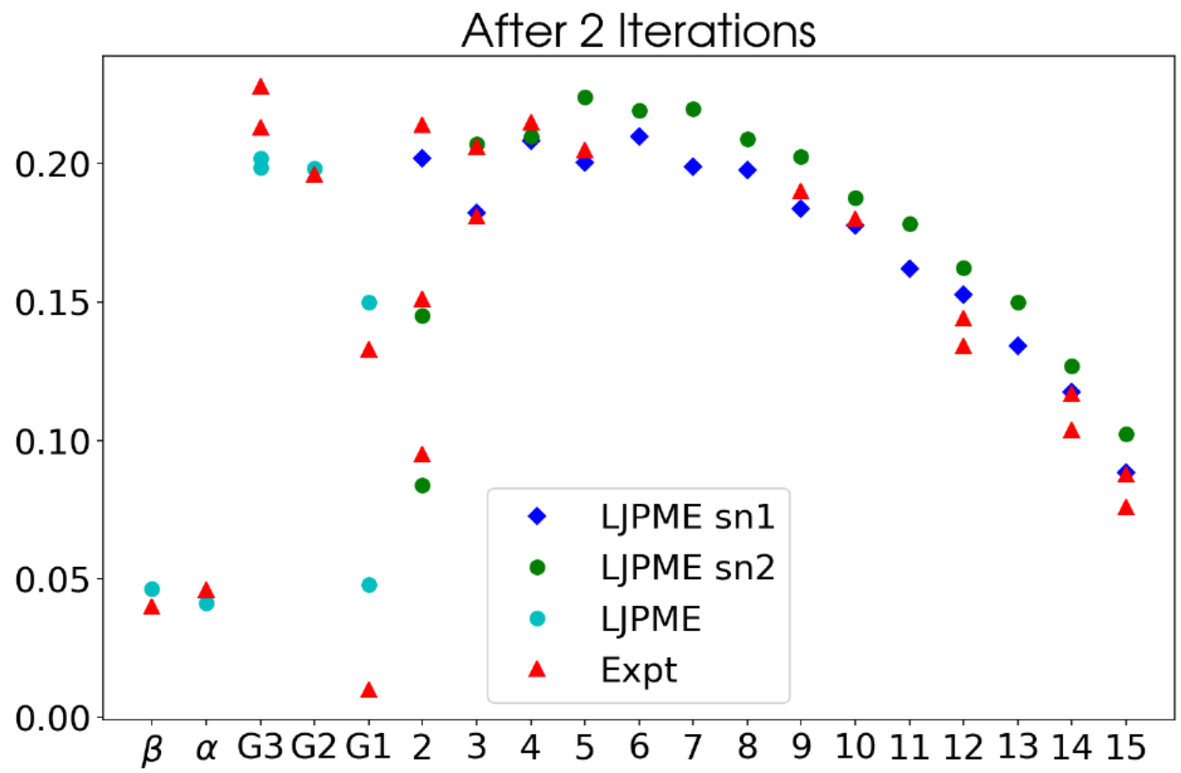
Original C36 parameters included
as a target to avoid overfitting
(Regularization)

Optimization Results: Surface Areas

	C36 PME	iteration 0	Iteration 1	Iteration 2	Experiment
DPPC bilayer NPT (323K)	**62.9**	**58.7**	**61.8**	**62.9 (0.3)**	**63**
DPPC monolayer at 18 dyne/cm	BAD	47.8	53.4	56.8 (0.3)	54
DPPC monolayer at 40 dyne/cm	BAD	61.0	64.3	64.8 (0.4)	64
DPPC monolayer at 55 dyne/cm	BAD	71.3	76.3	79.1(0.3)	80
Bilayer Compressibility (dyne/m)	**0.21**	**0.25**	**0.19**	**0.18**	**0.23**
DMPC bilayer NPT (303K)	**61.5**	**-**	**-**	**61.6 (0.4)**	**60.6**
POPC bilayer NPT (303K)	**68.7**	**-**	**-**	**68.4 (0.5)**	**67.4**
DOPC bilayer NPT (303K)	**66.0**	**-**	**-**	**66.0 (0.4)**	**64.3, 68.3**

Optimization Results: Order Parameters





- Original C36 parameters: LJ-PME worsens the SCD considerably
- Optimized SCD are even slightly better than the C36 FF (headgroups)

Minimal Modifications:

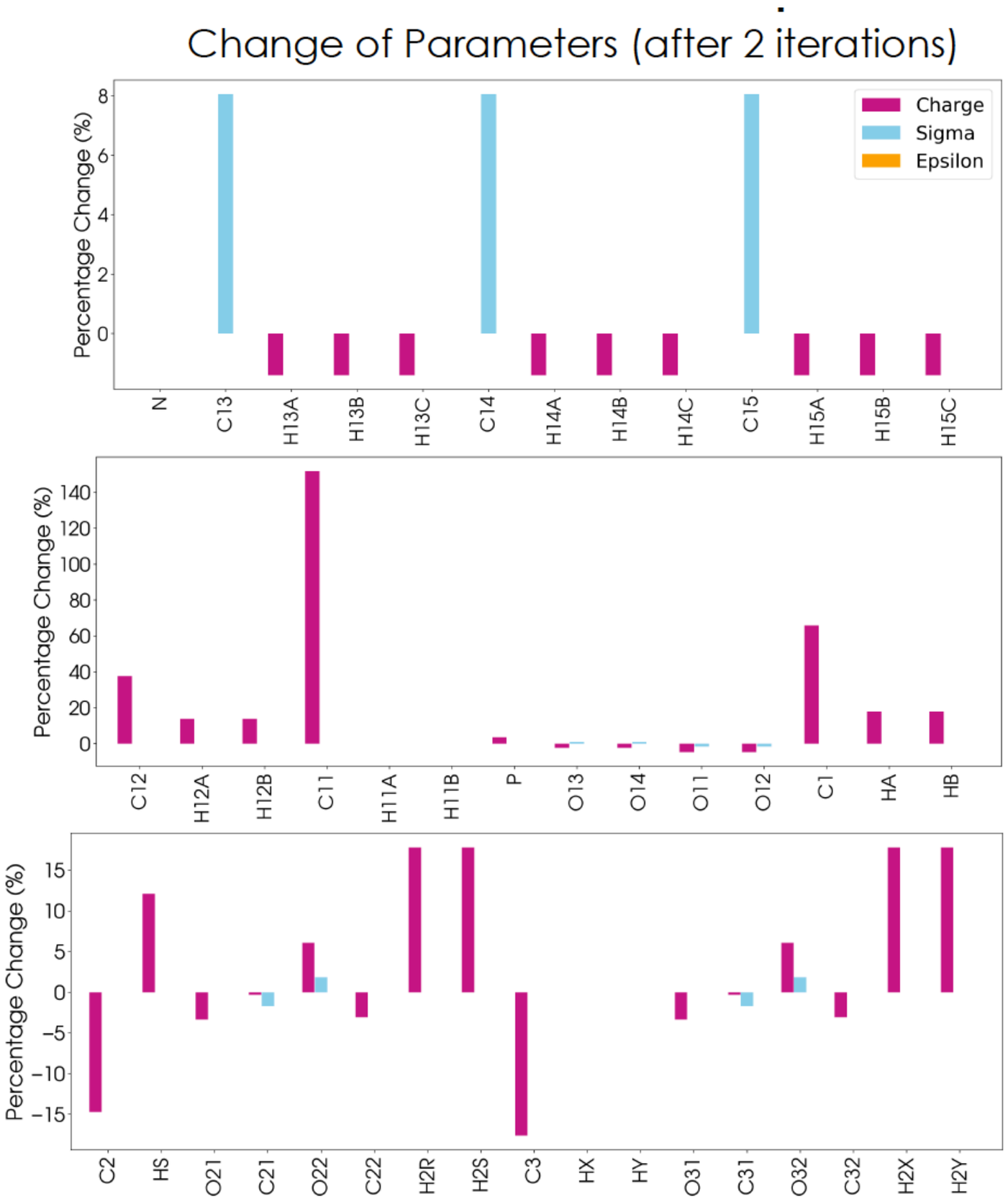
charges < 0.1 e

LJ sigmas < 0.01 nm

Successful incorporation of LJ-PME into C36:

So far PC head group

PE, PG, (PS) to follow



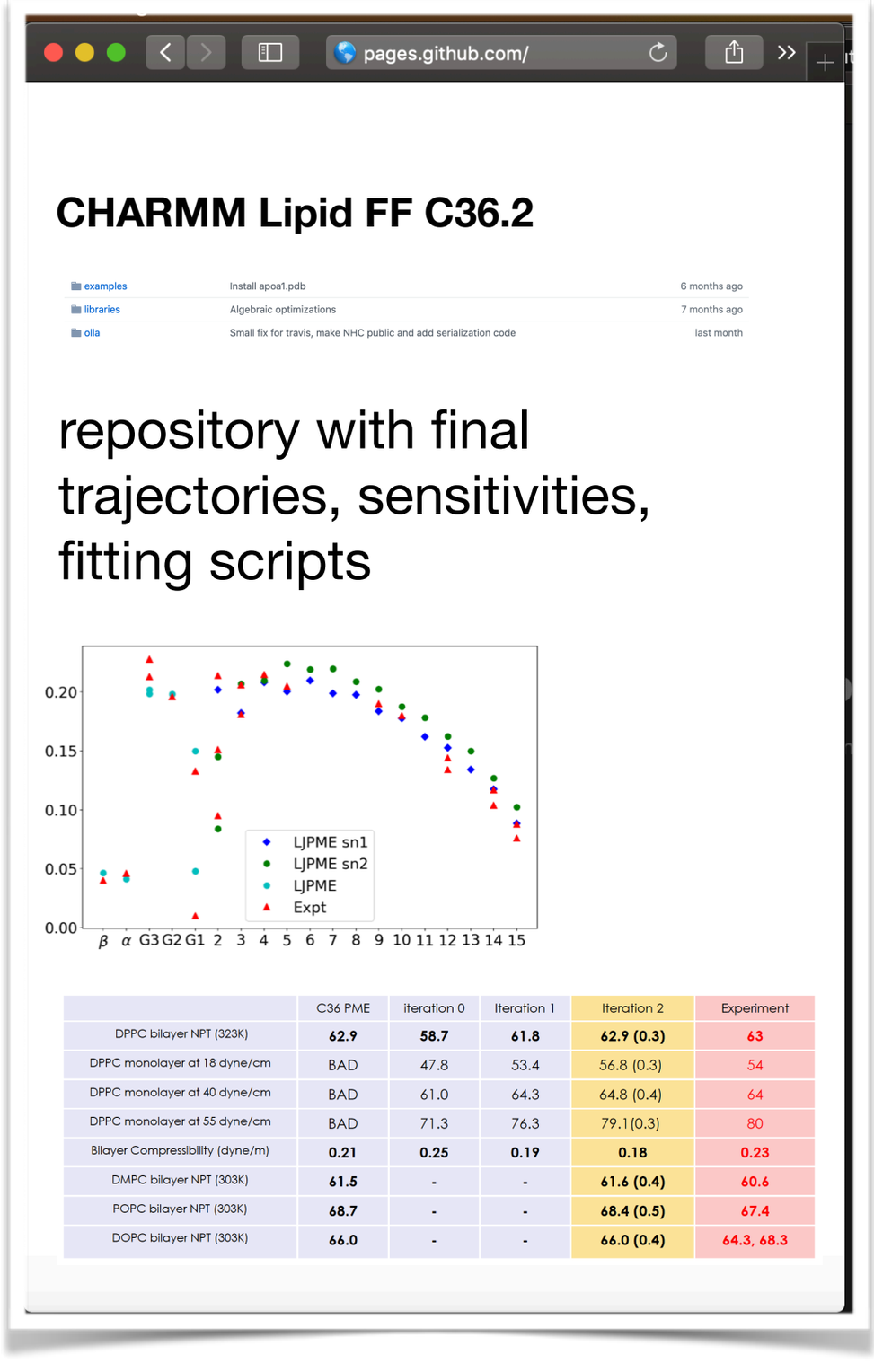
Outline

- Introduction
  - The CHARMM36 Lipid Force Field
  - Force Field Parameterization
- Applications
  - Improving Permeabilities
  - Adapting CHARMM36 for LJ-PME
- Outlook
  - Continuous Integration
  - Polarizable CHARMM Force Field

Continuous Integration of Force Fields

Reliable **Editable**

- Transparency (Stability)
- Maintainability



	C36 PME	iteration 0	Iteration 1	Iteration 2	Experiment
DPPC bilayer NPT (323K)	**62.9**	**58.7**	**61.8**	**62.9 (0.3)**	**63**
DPPC monolayer at 18 dyne/cm	BAD	47.8	53.4	56.8 (0.3)	54
DPPC monolayer at 40 dyne/cm	BAD	61.0	64.3	64.8 (0.4)	64
DPPC monolayer at 55 dyne/cm	BAD	71.3	76.3	79.1(0.3)	80
Bilayer Compressibility (dyne/m)	**0.21**	**0.25**	**0.19**	**0.18**	**0.23**
DMPC bilayer NPT (303K)	**61.5**	-	-	**61.6 (0.4)**	**60.6**
POPC bilayer NPT (303K)	**68.7**	-	-	**68.4 (0.5)**	**67.4**
DOPC bilayer NPT (303K)	**66.0**	-	-	**66.0 (0.4)**	**64.3, 68.3**

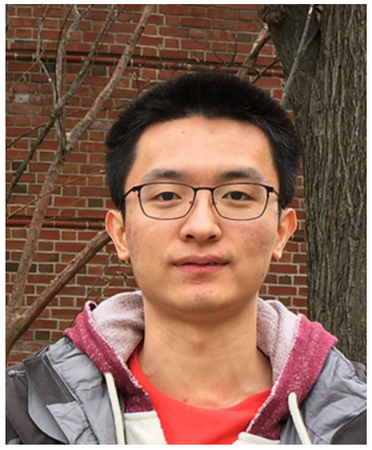
Outlook: Optimization of the CHARMM Polarizable Force Field (Drude)

- Longer-term solution to improve permeability?
- Current FF: Compressibilities way off
- MPID = equivalent formulation, more flexible than Drude

Conclusions

- Automated approach to improve the description of permeabilities
    - Multistate reweighting as a global metamodel
- C36 adapted to LJ-PME (PC headgroup)
    - Reweighting as a local metamodel
    - Only two iterations required
    - Consistent parameters/settings for bilayer and monolayer

Thank you for your attention.



- **NHLBI**
  - Bernard Brooks
  - Richard Pastor
  - Richard Venable
  - Andrew Simmonett
  - Frank Pickard (now @Pfizer)
- **Ghent University**
  - An Ghysels
- **Westlake Institute, Hangzhou**
  - Jing Huang
- **University of Maryland College Park**
  - Yalun Yu
  - Eric Wang
  - Jeffrey Klauda
- **Bonn-Rhein-Sieg UAS**
  - Dirk Reith
  - Karl Kirschner