Best Practices for Alchemical Free Energy Calculations : v0.1

John D. Chodera[1*], Antonia S. J. S. Mey[2*], Julien Michel[2*], David L. Mobley[3*], Conor Parks[4*], Julia E. Rice[5*], Michael Shirts[6*], Bryce K. Allen[7], Levi N. Naden[1], Andrea Rizzi[1,9]

[1]Computational and Systems Biology Program, Memorial Sloan Kettering Cancer Center, New York NY 10065; [2]Departments of Pharmaceutical Sciences and Chemistry, University of California, Irvine; [3]EaStCHEM School of Chemistry, David Brewster Road, Joseph Black Building, The King's Buildings, Edinburgh, EH9 3FJ, UK; [4]I don't know my affiliation; [5]I don't know my affiliation; [7]Silicon Therapeutics, Boston, MA, USA; [8]Tri-Institutional Training Program in Computational Biology and Medicine, New York, NY, USA

This LiveCoMS document is maintained online on GitHub at https://github.com/michellab/alchemical-best-practices; to provide feedback, suggestions, or help improve it, please visit the GitHub repository and participate via the issue tracker.

This version dated April 11, 2018

Abstract

ASJSM: @Volunteer write abstract

This particular document provides a skeleton illustrating key sections for a Best Practices document. Please see the sample `sample-document.tex` in github.com/livecomsjournal/article_templates/templates for additional information on and examples of using the LiveCoMS LaTeX class. Here we also assume familiarity with LaTeX and knowledge of how to include figures, tables, etc.; if you want examples, see the sample just referenced.

In your work, in this particular slot, please provide an abstract of no more than 250 words. Your abstract should explain the main contributions of your article, and should not contain any material that is not included in the main text. Please note that your abstract, plus the authorship material following it, must not extend beyond the title page or modifications to the LaTeX class will likely be needed.

***For correspondence:**

john.chodera@choderalab.org (JDC); dmobley@mobleylab.org (DLM); antonia.mey@ed.ac.uk (ASJSM); info@julienmichel.net (JM); ddd@yyy (CP); fff@ddd (MS)

[†]These authors contributed equally to this work

[‡]These authors also contributed equally to this work

Todo list

Introduction

ASJSM: @Volunteer write Intro

Meeting Members: David Mobley, Conor Parks, Samarjeet, Julia Rice, Toni Mey, John Chodera, Michael Shirts [please add your name if you have contributed to the document!] (Original Members: John Chodera, Conor Parks, Heather Mayes) People to Invite: Clara Christ (drug discovery expertise), Hannes Loeffler (FESetup expertise)

Prerequisites

Before proceeding with this document, you are assumed to have a basic familiarity with the principles of molecular simulations, molecular dynamics simulations, statistical mechanics and thermodynamics, and so on. This document focuses on aspects specific to alchemical free energy calculations, which remain an advanced topic within the molecular simulations area. Thus, if you are a beginner in the area of molecular simulations, not all background you need will be covered within the scope of this document. For example, we assume that you are able to set up and conduct a successful equilibrium molecular dynamics simulation of a particular thermodynamic state of a system interest as a prerequisite for this document.

Scope and Goals

This document focuses on preparation, execution, and analysis of alchemical free energy calculations, especially on aspects of the molecular simulations used here which are unique to alchemical calculations. A particular focus is on calculation of transfer free energies (hydration free energies, partition coefficients, etc.) and binding free energies (absolute and relative).

Not within our scope are advanced ligand binding topics such as:

- Covalent inhibition
- Association which is not 1:1
- Endpoint free energy methods
- PMF binding free energy methods

Additionally, the choice of force field (for protein, ligand, ions, cosolvents, or cofactors) is outside of the scope of this work; here, we focus on attempting to ensure successful alchemical free energy calculations given a particular choice of system and force field.

Goals:

- Checklist and background for new practitioners of relative and absolute alchemical free energy calculations: What should you pay attention to in setting up and running calculations in common codes and why.
- Provide guidance to authors as to what should be reported about their protocols in a Methods section, either conforming to standard or reporting where they did not conform and why. Useful to reviewers as well.

Checklist

ASJSM: @Volunteer This needs to be revised and expanded, some initial thoughts were just thrown in.

An attempt at identifying most important checklist items.

STEP 0 – KNOW WHAT YOU WANT TO SIMULATE

What are the first questions that need addressing before setting up a molecular dynamics simulation Extensive explanation for the checklist questions can be found in section 5.

- ☐ Can I get the required accuracy with the simulation I want to carry out
- ☐
- ☐ And finally

SIMULATION PREPARATION

How do I get started setting up an alchemical free energy calculation Extensive explanation for the checklist questions can be found in section 6.

- ☐ Have I followed the Best practices for biomolecular simulation set up?
- ☐ In a relative simulation, will I run into problems with clashing geometries in the ligand transformation or crystal waters?

ABSOLUTE SIMULATIONS

What are the main things I need to consider for an absolute alchemical free energy calculation? Extensive explanation for the checklist questions can be found in section 7.

- ☐ Topology
- ☐ Restraints
- ☐ Standard state handling

RELATIVE SIMULATIONS

What are the main things I need to consider for an relative alchemical free energy calculation? Extensive explanation for the checklist questions can be found in section 7.

- ☐ First thing
- ☐ Also remember
- ☐ And finally

ANALYSIS

This is all about analysis of the simulation Extensive explanation for the checklist questions can be found in section **??**.

- ☐ Are my simulations converged enough?
- ☐ Am I using the right analysis techniques?

Step 0 – What can be expected from alchemical simulations?

- What level of accuracy you can expect?
- What timescales and how many transformations can you address given available computational resources?
- Can you even hope to tackle the problem you are attempting?

Step 1 – Simulation prerequesites

1. Generate geometry of initial state: Reference biomolecular simulation preparation best practices
2. Relative: Generate geometry of final state (Mey)
    - e.g. ligand ideally should not clash with receptor, etc.; satisfy constraints that might be imposed by protocol such as overlapping atoms, etc.
    - reference biomolecular preparation setup practices for placing ligand into binding site, but elaborate on constraints that must be considered for relative free energy calculations

Step 2 – Simulation protocol selection

Absolute and relative free energy calculations have some differences

Alchemical free energy calculations can be grouped into two main categories, "absolute" and "relative" [1], which differ in whether they compute properties for a single molecule (absolute) or compare properties of different, usually closely related, molecules (relative). To use binding as a concrete example, in absolute binding free energy calculations, one computes the binding free energy of a ligand to an individual receptor relative to a standard reference concentration. In contrast, in relative binding free energy calculations, one compares the binding free energy of two related inhibitors to determine the potency difference.

Many protocol issues for alchemical calculations are common, but some are different between absolute and relative calculations, so before treating the common elements we treat the protocol differences.

Relative free energy calculations must select a topology and produce an atom mapping

Topologies and atom mappings.

A critical first step in relative calculations is to select an approach to these calculations, determining whether to use a dual topology, single topology, or hybrid topology approach to relative calculations. The distinction between these can be illustrated by considering a hypothetical transformation from molecule A to molecule B, where both atoms share a common substructure but differ in which functional groups are present; e.g. consider a transformation of ethane (A) to methanol (B). In this case the common substructure is at least CH3, though perhaps may be larger depending on how it is defined, as we discuss below. In single topology calculations, the overall transformation is set up to involve as few additional atoms as possible, so ethane would be typically changed into methanol by changing two of the protons into non-interacting atoms called "dummy atoms" (retaining their bonded interactions but not interacting with the rest of the system) and the connected carbon mutated into an oxygen (with an associated change in the C-H bond parameters as the atoms change to an O-H). Thus in a single topology calculation, atoms may change their type so relatively few dummy atoms are created. In contrast, in a dual topology free energy calculation, no atoms are allowed to change type [ref Shirts book chapter in Computational Drug Discovery and Design] so the ethane to methanol transformation involves starting with ethane plus two non-interacting dummy atoms, then passing through an intermediate state where atoms which are becoming dummy atoms or ceasing being dummy atoms are partially interacting (this state may or may not be well defined [ref Mobley perspective]), and culminating in a state where methane is present along with three additional dummy atoms which were previously a corresponding methyl group of ethane. Hybrid topology calculations have not seen much use [ref] but essentially consist of two absolute free energy calculations in opposite directions at the same time (turning one molecule off while turning the other on), and are best considered in that light. At present, the most widely used approaches, such as in Schrodinger's FEP+[ref] and in FESetup[ref] (for which calculations may be planned with Lead Optimization Mapper (LOMAP) [refs]) seem to use single topology approaches, though some codes only support dual topology. To our knowledge efficiency differences have not been thoroughly explored, though conventional wisdom would suggest that fewer dummy atoms are better [ref LOMAP paper/Mobley perspective].

Once a particular approach to the topology is selected, a crucial next step is to identify the common atoms which will not be perturbed. Rigorously, this process essentially comprises a maximal common substructure (MCSS) search of the molecules involved to identify the common substructure – though the parameters of the MCSS search will differ depending on whether single or dual topology calculations are planned. Specifically, with a single topology approach in mind, atom types are allowed to change, so a permissive MCSS search can be done, whereas with dual topology a more strict search is required. Some tools automate this process;

[1] The distinction is a bit of a misnomer, since both compute ratios of partition functions relative to another state, and neither computes an absolute free energy.

for example, LOMAP can take a set of ligands and generate proposed pairings of molecules which are scored by their MCSS similarity and other properties [refs]. Schrödinger's FEP+ planning tool is based on a version of LOMAP [ref].

MCSS searches can be relatively time consuming, so if scoring a library of ligands to identify promising pairs for relative calculations is the goal, it can be helpful to use faster approaches such as shape similarity to perform an initial scoring and then use MCSS only to identify final mappings for relative calculations.

The MCSS approach, though relatively standard, takes into account only topological similarity. It is possible that changes in binding mode could actually require a different choice of mapping, so in some cases mappings may need to be planned differently depending on 3D positioning of atoms in space [ref; does Cournia paper address this?].

Single topology relative calculations, and calculations based on substructure searches, only work if in fact the ligands share a common substructure. If no common substructure is shared, then essentially one ends up needing sophisticated dual or hybrid topology free energy calculations, where one would co-localize a pair of compounds in a binding site, exclude their interactions with one another, and compute the relative binding free energy by turning one molecule on from being dummy atoms while turning the other off. To our knowledge no general pipeline for such calculations yet exists and this would likely remain a research problem.

Ring breaking and forming.

Relative free energy calculations for ring breaking and forming are particularly challenging/problematic, in part because relative calculations rely on the free energy contributions of dummy atoms canceling between different legs of the thermodynamic cycle [refs], which may not be true whenever dummy atoms are involved in rings. Some approaches have attempted to address this [ref Schrodinger] but a general solution is not yet in mainstream use.

Constraints and relative free energy calculations.

One issue which requires particular care is the use of constraints. Commonly, bonds involving hydrogen are constrained to a fixed length to allow the use of longer timesteps. However, in single topology relative free energy calculations, the atoms involved might be mutated to other atom types – for example, in a mutation of methane to methanol, one hydrogen might become an oxygen atom. Typical molecular dynamics engines are not set up to recognize this change, or at least not to correctly include contributions to the free energy from changing constraints/constraint length, so results for a transformation would usually be erroneous. At present the most general solution to this problem is simply to avoid the use of constraints (and thus use a smaller timestep if necessary) in any relative free energy calculation involving a transformation of a constrained bond.

Absolute free energy calculations must handle the standard state and use restraints

AR, LNN: complete absolute free energy simulation protocol

Absolute free energy calculations involve completely removing the interactions between the ligand or solute and its environment, taking it to a non-interacting state that may or may not retain intramolecular nonbonded interactions. This non-interacting state can then be shifted between environments (from the protein to water, or from one solution to another) without changing its free energy, and then interactions can be restored.

Absolute free energy calculations are typically done with respect to a specific reference state or reference concentration. For solvation free energy calculations this is typically straightforward to handle and treating it correctly simply means ensuring that the non-interacting solute is taken to the same (or equivalent) final reference state in both environments, e.g. that the transformation involves a 1 M to 1M equivalent transfer free energy (where the non-interacting solute still occupies essentially the same volume as the solute in the interacting system). So typically in such cases no special care is required to ensure the correct standard state, as long as the *experimental* data being analyzed uses the same standard state and if it does not, a simple entropic correction is needed.

However, for binding the situation is much more complex and requires special care. Experimental absolute binding free energies are reported relative to a specific reference state – a 1 M standard state – which must also be used in calculations. In practice this has implications for how the calculations are done, as the reference concentration must enter the thermodynamic cycle employed.

Typically, to deal with both practical sampling issues and the standard state issue, restraints are employed in absolute binding free energy calculations to keep the ligand in a well defined volume as its interactions with the system are removed [ref Gilson 1997 BPJ]. This solves two problems. First, if the ligand were not kept in a well-defined region, as its interactions were removed it might wander the system, perhaps quite slowly, and only inadequately sample the noninteracting or weakly interacting state – yet adequate sampling of these states might be required for convergence. So for practical purposes, the use of restraints can dramatically improve sampling as interactions are weakened and removed. Second, if the ligand is not kept in a well-defined region then it is hard

to determine how to link a computed binding free energy to the correct 1M standard state. In contrast, with restraints, the free energy of releasing the restrained ligand to a 1M standard state can be computed analytically or numerically by solving the relevant integral [refs], allowing the standard state to enter the thermodynamic cycle [refs].

Several choices of restraints are possible.

In practice, a variety of types of restraints are common, from simple harmonic distance restraints between the ligand and the protein [refs], to flat-bottom restraints which work similarly but only exert a force if the ligand leaves a specific region [refs].

DLM: Enlist Naden to discuss problems with analytical approximation to standard-state correction for Boresch restraints

Alternatively, a set of restraints proposed by Boresch have also commonly been employed, where all six rigid-body degrees of freedom governing the orientation of the ligand relative to the receptor are restrained [refs]. Further restraints, such as on the overall ligand RMSD have also been used [ref Roux].

In principle, all of these forms will yield correct binding free energies in the limit of adequate sampling (if their effects and connection to the standard state are correctly handled) but they have different strengths and weaknesses. For example, with more involved restraints, sampling at intermediate lambda values will not likely need to be as extensive but more computational effort must go to computing the restraining free energy. Additionally, such restraints would typically keep the ligand from exploring alternative binding modes, which may be undesirable with Hamiltonian lambda exchange or expanded ensemble techniques where allowing the ligand to exchange binding modes when it is non-interacting could provide sampling benefits [refs, including Yank docs]. Concretely, flat bottom restraints might allow a ligand to explore multiple binding sites, harmonic restraints multiple binding modes within a site, and Boresch restraints a single binding mode within a single site [ref Yank docs?].

Many choices of restraints involve selecting reference atoms. Again, in principle this choice is unimportant given adequate simulation time but practical considerations may be important. The choice is likely especially important with Boresch-style restraints, where some relative placements of reference atoms are likely to be numerically unstable; additionally, ligand reference atoms should likely be in a part of the molecule which defines the binding orientation well, rather than in a floppy solvent-exposed tail, for example.

DLM: Get input from JDC on what they've learned about these.

DLM: Clarify terminology: Double decoupling, etc. See Feature Box below.

Absolute and relative calculations must deal with some of the same issues

Changes in net charge can be challenging/problematic.

If the net charge of the system will change as the alchemical calculation progresses, this can pose major challenges. Specifically, finite-size effects can introduce profound artifacts into computed binding free energies [refs], in part because typical schemes for long-range electrostatics (including PME and reaction field) do not handle free energy contributions from such changes effectively or as they would be handled in a hypothetical macroscopic bulk solution [refs].

There are two main potential solutions to avoid artifacts due to changes in net charge: Correcting for the introduced artifacts, or avoiding changing the net charge.

Many relative free energy planning tools have been set up to avoid changing the net charge of the systems considered, including LOMAP [ref] and early implementation of Schrödinger's FEP+, though later implementations allow changes in net charge by including charge corrections.

Absolute free energy calculations can potentially avoid changing the charge of the system by making a charge perturbation of equal and opposite sign elsewhere in the system; for example, as a charged ligand is removed, a charged counterion of opposite sign could also be removed, or one of the same sign could be inserted. This is the approach employed by the Yank free energy package [ref].

Charge corrections have also been explored, and are potentially a viable solution to this problem [refs] where artifacts introduced by finite-size effects are corrected numerically. However, application of such corrections typically remains a research problem (except in the FEP+ protocol [ref]).

When free energy calculations *do* need to change the charge of a ligand or solute, the literature does not yet seem to indicate what approach should be preferable, so considerable care should be taken. We are not yet aware of a careful comparison of charge corrections versus other approaches such as decoupling an ion at the same time, so in our view the issue of proper handling of charge mutations in the context of alchemical calculations remains a research problem.

The alchemical pathway is quite important

LNN: write common principles alchemical path choice

AR: write absolute-specific section on alchemical path choice

BA: write relative-specific section on alchemical path choice

Both absolute and relative calculations must choose an alchemical pathway connecting initial and final states, which is in principle arbitrary but in practice affects the efficiency of the calculations considerably. Some choices are particularly crucial – for example, transformations involving insertions or deletions of atoms should employ soft-core potentials for Lennard-Jones or other hard-core interactions [refs]. Other issues, such as whether absolute calculations retain intramolecular nonbonded interactions or remove these interactions, may be less critical and differ among studies in the literature [refs].

Relative calculations introduce additional choices, including whether to define explicit intermediate states [ref] or leave these implicitly defined by the code [ref]. Typically in single topology relative calculations it proves most efficient to first remove electrostatic interactions of any atoms which will be deleted, then modify other nonbonded interactions, then restore electrostatic interactions of any atoms which are being inserted. Other schemes, such as simultaneously changing electrostatic and Lennard-Jones interactions, even with electrostatic "soft core" potentials, in our experience typically introduce errors and/or instabilities or are at least unreliable. We have less experience with dual topology calculations but expect that similar considerations will apply, and the principle of first removing electrostatics and then removing steric interactions will likely serve well.

A key additional consideration in choosing the alchemical pathway is the choice of spacing of intermediate states. The spacing depends to some extent on the choice of analysis method, though states should essentially be spaced equidistant in the relevant thermodynamic length [ref]. For BAR/MBAR techniques this means that spacings should typically be equal in [what, variance? ref]. Some schemes to adaptively optimize the spacing of intermediate states based on initial exploratory simulations have been proposed [refs].

1. Choice of discrete alchemical protocol (Shirts, Mey, Chodera)
2. Relative Three-stage protocol (discharge unique initial atoms, transform LJ, charge unique final atoms) vs soft-core electrostatics/LJ
3. Absolute
    - Select a **common alchemically-eliminated end state**
    - Decoupled vs annihilated for electrostatics and LJ
    - Sequential electrostatics and LJ versus simultaneous (recommend sequential)
4. Concerns: Part of AMBER still can't run at endpoints (lambda = 0 or 1); SANDER cannot but PMEMD can.

1. Determine whether you need to **handle multiple binding modes**: (Mobley)
    - In absolute: Confine-and-release, BLUES
    - In relative (ugh!!)

LNN, BA: write stopping condition section

2. Determine **stopping conditions** Uncertainty-directed stopping criteria can ensure target uncertainty is achieved
3. Select which **data should be saved and with which frequency**
    - What data to save: dU/dlambda, Delta E's between neighbor for BAR, between further for MBAR, ...
    - BAR captures most of info with well-optimized lambda protocol, but MBAR when perhaps not, except when there are way too many lambda values.
    - Recommend against solely relying on TI when possible
    - Recommend cross-comparing methods (TI (spline, trapezoid, etc.), BAR, MBAR) as diagnosis of trouble

Step 3 – Overview of available analysis techniques

JDC, BA: uncertainty estimation section

1. Detecting boundary between equilibrated and production regions (Chodera: http://dx.doi.org/10.1021/acs.jctc.5b00784)
2. Decorrelating samples for analysis
    - Subsample different lambdas based on correlation times
    - Ensure all simulations at least 50x correlation time
3. Examining output data for common problems with discussions of what exactly to plot or look at; examples of typical curves for dV/dlambda and free energy versus lambda, for example
    - Make sure ligand doesn't tumble out of binding site (Mey has observed this)
    - Significant discrepancies between different free energy estimators (TI, BAR, MBAR)
    - Poor replica mixing (for replica-exchange)
    - Correlation time as a function of lambda as it would be expected to be a smooth
    - Dependence on initial conformation
    - Torsional analysis: Is it stuck in specific states? Only very rarely transitions?
    - More "usual suspects"
4. Estimators for free energies
    - MBAR recommended if all energy differences are available
    - BAR just as good for highly optimized lambda values

- TI should be roughly concordant, but quadrature error hard to quantify
- Other variants useful in special circumstances (e.g. Z. Tan stochastic version)

5. Computing and reporting uncertainties on free energies correlated bootstrap v. timeseries analysis
    - Quantifying standard error in dG estimate
    - Varience in dG estimate using multiple methods (TI, DEXP, IEXP, BAR, MBAR, GDEL, etc.)
    - Agreement in dG estimate when repeating calculations with different parameters (random seeds, initial configurations, forcefield, etc.)
    - Calculating differences in free energy change as a function of time
        - Starting from beginning or end of simulation
        - Significance of differences in midpoint estimate (High middle error: high uncertainty)
    - Ensemble method to combine uncertainties into interpretable weighted metric
        - Simple normalized metric to determine confidence in calculation
        - Easily interpreted by chemist/biologist when prioritizing new chemistries
6. Other considerations for many transformations Cycle closure error

Terminology and abbreviations

- Feature Box covering major technical terms and abbreviations
- Examples:
    - EXP, BAR, MBAR
    - Double decoupling, single-topology, dual-topology, hybrid-topology, coupled-topology
    - FEP (free energy peturbation), alchemical, AFE (alchemical free energy)

Available software – a summary

- Commercial:
    - FEP+
- Free or low-cost for academics / commercial for industry:
    - CHARMM / DOMDEC / CHARMM-OPENMM
    - TIES and AMBER FEW? (Peter Coveney)
    - AMBER / PMEMD
- Free (libre) open source:
    - SIRE
    - YANK
    - gromacs
    - pmx for mutations
- Setup tools
    - FESetup: AMBER, gromacs, Sire
    - Lomap/Lomap2 : Relative alchemical transformation graph planning
- Analysis tools:
    - Free Energy Workflows: Sire-specific free energy map analysis using weighted path averages https://github.com/michellab/freenrgworkflows
    - Alchemlyb: Multipackage free energy analysis https://github.com/alchemistry/alchemlyb
    - pymbar: MBAR implementation, but have to roll your own analysis wrapper https://github.com/choderalab/pymbar

Online resources

- http://www.ks.uiuc.edu/Training/Workshop/Urbana_2010A/lectures/TCBG-2010.pdf
- Basic Ingredients of Free Energy Calculations: A Review (DOI:10.1002/jcc.21450)
- Good Practices in Free-Energy Calculations (DOI:10.1021/jp102971x)
- Alchemical Free Energy Methods for Drug Discovery: Progress and Challenges (doi:10.1016/j.sbi.2011.01.011)
- Alchemistry wiki: http://www.alchemistry.org/wiki/Best_Practices

Author Contributions

(Explain the contributions of the different authors here)

For a more detailed description of author contributions, see the GitHub issue tracking and changelog at https://github.com/michellab/alchemical-best-practices.

Other Contributions

(Explain the contributions of any non-author contributors here) For a more detailed description of contributions from the community and others, see the GitHub issue tracking and changelog at https://github.com/michellab/alchemical-best-practices.

Potentially Conflicting Interests

Declare any potentially conflicting interests here, whether or not they pose an actual conflict in your view.

Funding Information

FMS acknowledges the support of NSF grant CHE-1111111.

References