Advancing predictive modeling through focused development of model systems to drive new modeling innovations

A funding proposal recently submitted to the NIH

David L. Mobley
Associate Professor, Departments of Pharmaceutical Sciences and Chemistry, University of California, Irvine
John D. Chodera
Assistant Faculty Member, Computational Biology Program, Memorial Sloan-Kettering Cancer Center
Lyle Isaacs
Professor, Department of Chemistry and Biochemistry, University of Maryland
Bruce C. Gibb
Professor, Department of Chemistry, Tulane University

ABSTRACT

This work seeks to advance quantitative methods for biomolecular design, especially for predicting biomolecular interactions, via a focused series of community blind prediction challenges. Physical methods for predicting binding free energies, or "free energy methods", are poised to dramatically reshape early stage drug discovery, and are already finding applications in pharmaceutical lead optimization. However, performance is unreliable, the domain of applicability is limited, and failures in pharmaceutical applications are often hard to understand and fix. On the other hand, these methods can now typically predict a variety of simple physical properties such as solvation free energies or relative solubilities, though there is still clear room for improvement in accuracy. In recent years, blind prediction challenges have played a key role in driving innovations in prediction of physical properties and binding, especially in the form of the SAMPL series of challenges. Here, we will continue and extend SAMPL prediction challenges to include new physical properties, more complicated host-guest binding data, and application to biomolecular systems. Carefully selected systems and novel experimental data will provide challenges of gradually increasing complexity spanning between systems which are now tractable to those which are marginally out of reach of today's methods but still slightly simpler than those covered by the Drug Design Data Resource (D3R) series of challenges on existing pharmaceutical data. We will work with D3R to run blind challenges on the data we generate and to ensure it is designed to maximally benefit the field.

In **Aim 1**, we will collect new measurements on partitioning, distribution, and protonation of drug-like compounds, in collaboration with partners in the pharmaceutical industry. In **Aim 2**, we leverage our expertise in host-guest binding to generate new data on host-guest binding in cucubiturils and deep cavity cavitands. And in **Aim 3**, we use high-throughput robotic experiments to generate new protein-ligand binding data of biological relevance. **Aim 4** focuses on using this data to run blind SAMPL challenges, motivating the community to test, understand, and improve these methods. We will also run reference calculations with the latest techniques.

This work will ensure the continued success of SAMPL challenges which have already driven considerable innovation in the field and been the focus of more than 90 different publications (each typically cited 5-50 times) since their inception around 2007, and will play a key role in driving the next several generations of improvements in computational techniques for molecular design. The research proposed here will lead to significant improvements in the predictive power of physical models for drug discovery, molecular design and the prediction of physical properties.

SPECIFIC AIMS

While computational techniques are currently widely used in pharmaceutical drug discovery, current generation technologies (such as docking) are unsuitable for true molecular design. Specifically, these techniques fail to to predict small molecule binding affinities to target and antitarget biomolecules with sufficient accuracy for the variety of applications currently of interest. Computational screening techniques can do better than random selection of compounds, but they lack the accuracy to guide molecular design or optimization. A new generation of physical techniques, alchemical free energy calculations, are poised to fill this void by providing a quantitative, predictive tool that can be used in multiple stages of the drug discovery pipeline, including lead optimization to improve affinity and selectivity or the retention of potency as other physical properties are optimized.

Recent success of alchemical methods in predicting accurate affinities has sparked considerable enthusiasm, but the domain of applicability of these techniques is currently highly limited; broad application and routine use will require further evaluation, refinement, and development. There is a vast gulf between targets within the domain of applicability and those which are outside it. Bridging this gulf to expand the domain will require focused study of carefully selected systems of intermediate complexity. Without such a bridge, these techniques may encounter the same problems faced by docking and related techniques: routine failure without clear insights into why, and years to decades spent making small methodological modifications without dramatic improvements in predictive power.

We propose to collected targeted experimental datasets, use them to conduct blind prediction challenges, and release curated benchmark sets in a manner designed to drive expansion of the domain of applicability and improvements to physical modeling techniques. The data we generate will provide a spectrum of difficulty between systems tractable with current methodologies to the pharmaceutically-relevant drug targets featured in the NIH-funded D3R effort, which fields blind challenges using protein-ligand datasets from pharma. Our systematic set of challenges aims to rapidly advance free energy techniques to the point of standard application in drug design. At the same time, the data we collect will play a long lasting role in the community, going through a life cycle of collection, curation, blind challenges, and then public dissemination to serve as benchmark sets and standard reference data, and to drive construction of new and improved models. While our work here focuses primarily on generating new targeted data for a series of blind SAMPL ("Statistical Assessment of Modeling of Proteins and Ligands") challenges and running those challenges, we also plan for subsequent data dissemination. Here, we will:

<u>Aim 1.</u> Collect new physical property datasets to assess accuracy and spur improvements in force fields and modeling of protonation states and tautomers.
We will develop new solution-phase datasets for druglike small molecules. These data can test critical aspects of small molecule modeling (including accounting for interactions and treatment of protonation/tautomeric state) and improve our ability to predict physical properties relevant to drug discovery in new regions of chemical space. We will initially focus on aqueous/nonpolar distribution coefficients and pK_a measurements, advancing to solubilities and membrane permeabilities, while using these data to drive improvements in the modeling of ligand interactions.

<u>Aim 2.</u> Measure affinities of drug-like compounds in supramolecular hosts to challenge quantitative models of binding in systems not plagued by major receptor sampling issues.
We will measure new host-guest binding free energies (using cucurbiturils and deep-cavity cavitands as hosts) to field binding challenges with varying complexity between physical property prediction and protein-ligand binding. Host guest systems are some of the simplest cases of molecular recognition, and thus these binding data will drive improvements in modeling of simple binding systems with techniques of relevance to drug discovery.

<u>Aim 3.</u> Develop model protein-ligand systems that isolate specific modeling challenges of drug targets.
We will identify suitable biological protein-ligand model systems that isolate individual modeling challenges (selected to push the limits of physical techniques) and develop these for blind challenges based on new protein-ligand affinity measurements. While the initial year will feature fragment binding to human serum albumin, subsequent challenge systems will be selected using a novel informatics platform to focus on timely modeling issues.

<u>Aim 4.</u> Field community blind challenges to advance quantitative biomolecular design.
The data collected in Aims 1–3 will drive annual SAMPL blind challenges, allowing the field to test the latest methods and force fields to assess progress, compare them against one another head-to-head, and perform sensitivity analysis to learn how much different factors (protonation state, tautomer selection, solvent model, force field, sampling method, etc.) affect predictive power. Results will then feed back into improved treatment of these factors for subsequent challenges, driving regular cycles of application, learning, and advancement.

Overall, the data generated here and the cycles of tests in SAMPL challenges will guide new innovations in physical methods for predicting binding and physical properties, providing a foundation for the next several generations of computational methods for pharmaceutical drug discovery.

SIGNIFICANCE

Physical modeling is poised to transform drug discovery and chemical biology by enabling true molecular design. While modeling is used extensively in drug discovery, its main role at present is to aid with idea generation or filter large libraries of compounds for screening. Instead, we imagine using computational techniques extensively to guide the design process. Consider a medicinal chemist in the not-too-distant future who has just finished synthesizing several new derivatives of an existing inhibitor, and has obtained binding affinity or potency data against the desired biomolecular target. Before leaving work, she generates ideas for perhaps 100 new compounds which could be synthesized next, setting her computer to work overnight. By morning, the idea compounds have been prioritized based on reliable predictions of their affinity for the desired target, selectivity against antitargets, solubility, and membrane permeability. The chemist looks through the predicted properties for the top few compounds, selecting some for synthesis. If synthesizing and testing each compound takes several days, this workflow compresses roughly a year's work into a few days.

While this workflow is not yet a reality, significant strides have been made toward accurate binding affinities [97–104], solubilities [105–107], selectivity and drug resistance [108], and membrane permeability [109, 110]. A considerable amount of science and engineering still remains to make this vision a reality. Given recent progress, the question now seems more one of *when* rather than *whether*.

Widespread availability of inexpensive graphics processing units (GPUs) provides a 100-fold increase in price-to-performance ratio over CPUs, while advances in automation [111] and sampling protocols have helped simulation-based techniques reach the point where they now begin to be genuinely useful in guiding drug discovery for a limited *domain of applicability* [100–104, 112, 113]. Specifically, in some situations, free energy calculations appear to be capable of achieving RMS errors of 1–2 kcal/mol with current force fields, even in prospective applications, sufficient to drastically reduce the number of molecules that must be synthesized and assayed [114]. As a consequence, pharmaceutical companies are beginning to use these methods in active discovery projects.

Despite progress, current physical modeling methodologies suffer from severe limitations hindering their widespread use in molecular design. For example, even "small" protein conformational changes not gracefully handled by current methodologies can yield errors up to 5 kcal/mol in calculated binding free energies [115], force field limitations still pose major challenges [116], and the inability to treat important chemical effects like protonation and tautomer equilibria drastically limits the domain of applicability. **For many pharmaceutically relevant systems, the most important sources of error—and modeling challenges—are not yet clear.**

Progress on addressing these challenges has been frustratingly slow, hindered by a lack of high-quality data and community focus. **Neither retrospective tests nor prospective application in discovery projects provides the necessary impetus and data to rapidly overcome remaining barriers to widespread utility.** Large-scale retrospective tests can assess *retrospective* performance, but they do not provide accurate guidance on utility for prospective design, nor do they effectively identify the most important sources of error. Retrospective tests can also easily result in over-fitting, where researchers apply a variety of protocols until apparently significant results are obtained by chance [117]. In retrospective tests, performance may also not be indicative of expected performance in applications because even well-meaning researchers can take advantage of prior knowledge. For example, if the binding mode of a ligand is already known crystallographically, a researcher may use that binding mode in retrospective tests, whereas prospective or design work would require first selecting among candidate binding modes, introducing substantial uncertainty unaccounted for in the retrospective statistics [97, 118, 119]. This also means that in retrospective tests, researchers almost invariably try far fewer methods than in prospective tests, resulting in much less new insight. Prospective tests, in contrast, force researchers to anticipate a multitude of potential situations rather than only those observed in a known benchmark dataset. Prospective application in actual discovery projects, while important, also does not provide the necessary impetus, partly because often, the predicted compounds are in fact never tested [102] or the experimental data necessary to assess the quality of the predictions is absent—for example, because binding affinities are not measured or no crystallography is available.

To accelerate progress in quantitative predictive physical modeling, we need a series of community blind prediction challenges focused on pushing the limits of predictive techniques, providing a bridge between challenging but tractable problems and pharmacologically relevant but currently intractable problems. These challenges should be designed to have the necessary high quality experimental data, but also be prospective, predictive tests. While the Drug Design Data Resource (D3R [120], discussed further below) provides an existing community blind challenge on protein-ligand binding, it focuses on using pre-existing pharmaceutical datasets, rather than on measuring new data carefully selected maximize community learning [120]. D3R serves well to assess where we are now—but we need a carefully-designed effort focused on improving modeling.

Physical modeling accuracy advances most rapidly when progress toward a complex goal can be decomposed into resolving a series of tractable problems, as revealed by carefully collected and curated data. To make rapid progress, our field needs an effort which focuses on specific *component* problems of the overall problem of interest, collects and curates data that highlights these problems, and drives progress via prospective challenges. This process allows the entire community to learn from both methodological success and failure. The model we propose here has been proven to drive dramatic improvements in modeling, as evidenced by our **Statistical Assessment of Modeling of Proteins and Ligands (SAMPL)** series of challenges. SAMPL, born out of frustration with the lack of venues for comparing predictive accuracy on a level playing field, was initiated by Anthony Nicholls of OpenEye software in 2007/2008 [121], and has run challenges approximately every two years since then [122–129]. Governance transitioned to an unfunded academic collaboration during SAMPL3 in 2012; this collaboration ran subsequent challenges as SAMPL4 (2014) and SAMPL5 (2016). The PI of this proposal (Mobley) was a primary organizer of SAMPL4/5 (2014–2016). Much as unit testing is an indispensible tool for discovering where bugs in a program are hiding when complex integration tests fail, exercises like SAMPL are valuable in pinpointing and correcting modeling errors when the overall performance fails to live up to expectations for complex pharmacological targets. To accomplish this, SAMPL has historically focused on both simple challenges that attempt to isolate likely sources of modeling errors, such as physical properties of small molecules (hydration free energies, aqueous tautomer ratios, partition or distribution coefficients between aqueous and nonpolar phases) as well as small molecule binding to targets of reduced complexity (such as host-guest binding, and binding of fragments to trypsin and HIV integrase). **SAMPL has already been a tremendous community resource, resulting in nearly 100 publications which are typically cited 5–50 times or more each** [1–96].

Here, we **design a new series of SAMPL challenges specifically to guide the improvement of models**. Until now, this has been impossible, because SAMPL has been entirely unfunded; its very existence has required "donation" of data and time, giving us no ability to gather datasets tailored to our purpose. Our proposed new challenges bridge the gap between calculations of simple physical properties that isolate forcefield inaccuracies from sampling challenges, like hydration—which can already be calculated fairly accurately [126]—and the D3R Grand Challenges on protein-ligand binding, which are a major source of consternation for the community so far [120, 130–132]. Unless this gap is bridged, there is the very real possibility that modeling may simply continue to fall far short of expectations in pharmaceutical challenges like D3R for reasons which are unclear. Here, we design challenges to highlight major reasons for failure and drive progress towards resolving them.

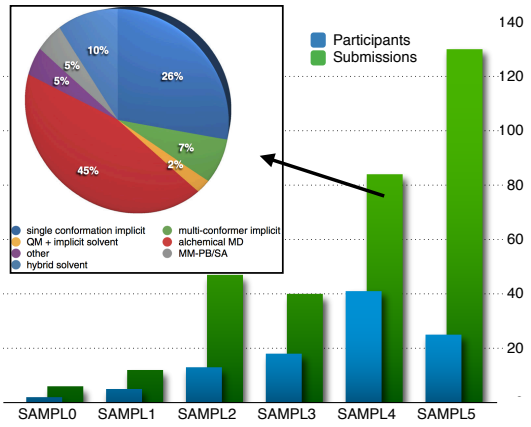

Figure 1. **SAMPL historical participation [126].** Historical participation in SAMPL host-guest + solvation/distribution challenges has climbed rapidly, and we expect this trend to continue. The number of participating groups is shown in blue, and the number of submissions in green. The inset shows the diversity of methods employed for the SAMPL4 hydration challenge, which is typical for SAMPL.

Our major goal is to rapidly advance predictive modeling to where it can guide biomolecular design, and extension of the SAMPL challenges will do exactly that. This work will play a vital role in enhancing the work being done on *existing* data by D3R, helping prepare methods for application to the more challenging systems emerging from pharma in D3R's challenges.

INNOVATION

Blind predictive challenges—and SAMPL in particular—have already led to important new science on method development, evaluation, robustness, and force field improvements. However, they have not yet produced dramatic improvements in predictive molecular design, largely because the challenges posed have not been selected and staged for optimal impact. Indeed, SAMPL challenges to date have had to be opportunistic, making best use of measurements that contributing experimentalists were carrying out to advance other scientific goals. A central goal of this proposal, then, is to maximize the impact of future SAMPL exercises by crafting a series of challenges specifically designed to test and advance computational methods, and gathering the requisite experimental data.

Several historical examples serve to highlight how SAMPL can foster innovation (though far more examples are available in our SAMPL bibliography below; see also Figure 3). The first several SAMPL challenges on hydration free energies had rather hit-and-miss performance, highlighting pitfalls of existing methods and force fields which led to marked improvements in PB models [122, 133, 134], recognition of some limitations of fixed-charge force fields [135, 136], repair of some of these force field deficiencies via additional polarization or introduction of off-site charges [135–137], and helped motivate alternate implicit or hybrid solvent models [138–140]. Shifts in protonation

state and tautomer proved particularly important in the recent SAMPL5 log D challenge [128, 141]. This challenge provided a tractable opportunity to isolate and explore these specific physical effects, which are so important in protein-ligand binding, while avoiding the full complexity of pharmaceutical binding studies. Host-guest binding studies have also been particularly important [142], highlighting the importance of salt effects [129, 142, 143] and in some cases revealing more severe force field limitations than observed in hydration and distribution challenges [144, 145], pointing the way forward for improving predictive models of molecular interactions [142, 146].

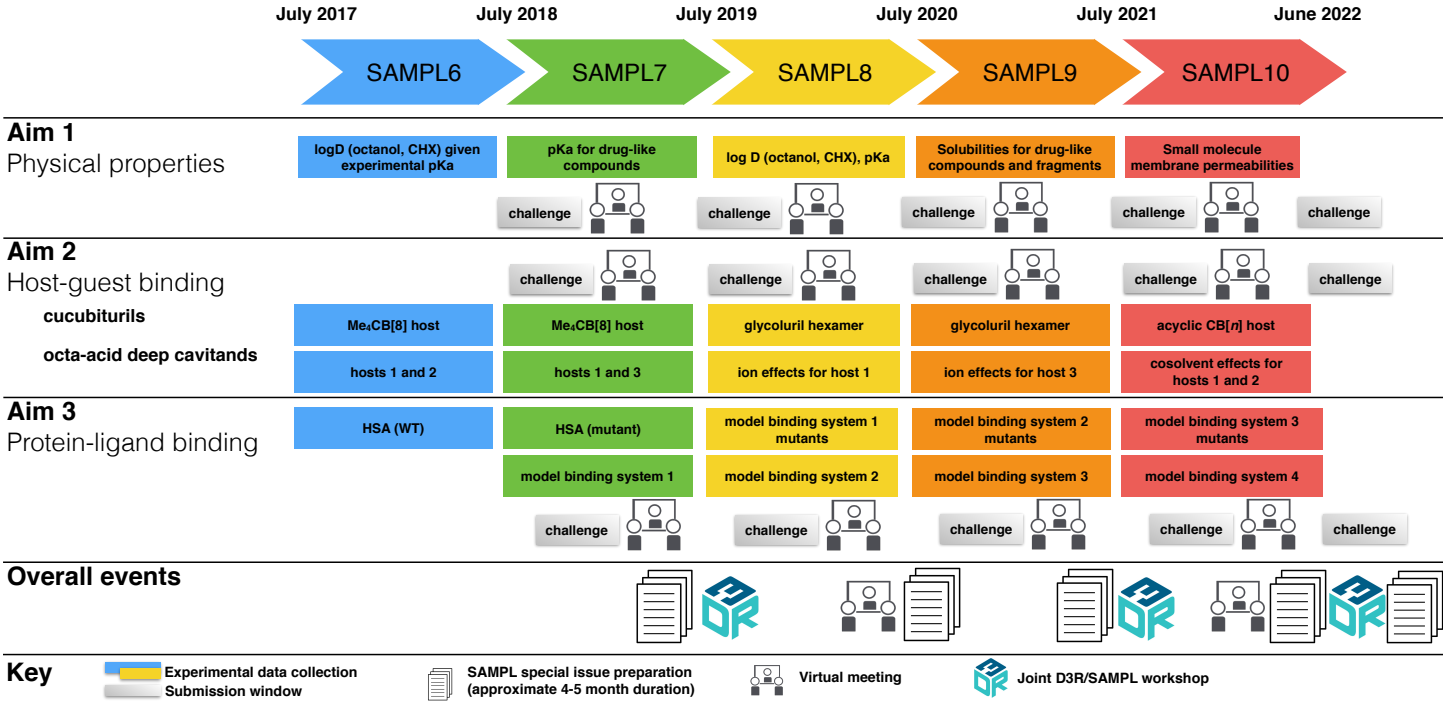

Figure 2. **Timeline for our activities.** Activities covered by this grant include data collection and SAMPL challenges on our three major components (physical properties, host-guest binding, and protein-ligand binding), with each challenge cycle color-coded separately. Data collection within each Aim is shown by a colored bar indicating what is measured and curated. Data collection/curation is followed by a submission window for that challenge component, then all results and analysis are returned to participants and posted on the SAMPL website; this also will nucleate more detailed long-term discussion on the relevant Slack channel. At this point, we will also release the data to the public as a high quality benchmark. Each component will then wrap up with a virtual meeting focused on lessons learned and areas which need further exploration; these will be recorded and posted on our website to assist in rapid dissemination of new insights. Virtual meetings precede the submission window for the next SAMPL challenge, giving the opportunity to incorporate lessons learned for the next challenge. Submission windows and virtual meetings are staggered across categories so that participants can be involved in all three major areas without multiple simultaneous deadlines. In-person meetings are co-hosted with D3R and will occur every two years, supplemented by effort-wide virtual meetings in between. Special issues of JCAMD will have deadlines shortly after the virtual meeting on the protein-ligand challenge for that year, and a 4-5 month timeline (based on historical experience) from the submission deadline until the special issue appears (with the first papers appearing online substantially sooner). Rapid dissemination of insights is critical for rapid progress, so we highly encourage the use of preprints and informal reports to supplement the special issue.

This work is also innovative because of the uniqueness of SAMPL. While there are other predictive challenges in the area of biomolecular modeling, such as D3R [120], the pK_a cooperative [147], CAPRI [148] and CASP [149], **no other blind challenge focuses specifically on data tailored and collected to drive quantitative protein-ligand modeling**. The SAMPL expansion we propose here is unique in its *specific design* to drive improvements in modeling accuracy rather than simply serving an evaluative role. SAMPL benefits the whole modeling community—for example, protein-ligand docking software has improved as a direct result of SAMPL hydration challenges [150], and commercial software vendors have introduced new features or scientific improvements based on participation in SAMPL challenges [141, 151]. In effect, **SAMPL serves as an engine to spur innovation by soliciting novel approaches to complex problems from the community and evaluating their success at predictions.**

This work also focuses on innovative experimental methods. Specifically, in Aim 3, we develop a new informatics platform to facilitate the rapid identification and study of particularly informative protein-ligand systems that are both experimentally tractable for high-throughput biophysical measurements and focus on specific challenges of interest. We employ a fully automated wetlab to screen potential model systems for expression, carry out high-accuracy biophysical measurements, and perform automated error analysis to carefully assess experimental uncertainty. **This work is at the forefront of innovation in high-throughput, automated biophysical experiments to produce high quality data with well-characterized uncertainties.** Not only will the data be of prime importance, but the techniques themselves will help future experiments.

In Aim 4, in addition to running SAMPL challenges, we will also perform reference calculations to test the accuracy of current state-of-the-art techniques. Both the Mobley and Chodera labs are experts in development of free energy methods for application to physical properties (e.g., [126, 152, 153] and binding (e.g., [115, 116, 154]), and **these reference calculations will drive innovation** as well, serving several key roles: (1) Benchmarking the latest method developments against current "best practices" methods (by doing calculations via both approaches); (2) Facilitating learning, allowing others to compare against our results to determine how a change in method or force field impacts results; (3) Focusing the field on key issues by doing sensitivity analysis to whether conditions such as ionic strength, protonation state, tautomer choice, etc., impact computed values.

Careful analysis of challenge predictions and results to identify why models fail and what specific problems need further attention is a critical and powerful aspect of SAMPL that spurs further innovation. Both organizers and participants play key roles in this; organizers identify global patterns and provide a venue for participants to explore these issues, while participants probe failure modes of individual methodologies in greater detail. When methods differ in performance, it is critical to understand whether the differences are statistically significant and important, and to provide an accurate accounting of the uncertainty in performance measures. Thus, careful and innovative analysis of challenge outcomes is particularly important in SAMPL [126, 128, 129], in some cases driving experimentation with new performance metrics [126].

APPROACH

Our approach to systematically advancing modeling for biomolecular design involves collecting carefully targeted experimental datasets for challenges focusing on physical property prediction, host-guest binding, and protein-ligand binding. These datasets, spanning a spectrum of complexity, help isolate individual limitations in quantitative physical modeling to encourage and evaluate multiple solutions from the community. Aims 1–3 focus on tailoring and generating this experimental data, while Aim 4 focuses on fielding annual SAMPL challenges. Each annual challenge includes one or more components from each of the Aims. Aims 1–3 bring together multiple laboratories and both theorists and experimentalists: graduate students from the Mobley and Chodera laboratories are paired with well-equipped experimental groups in industry to collect physical property data (Aim 1); Gibb and Isaacs, leading experimentalists in supramolecular chemistry, work with theorist Mobley to perform host-guest affinity measurements (Aim 2); and the Chodera lab applies new automated approaches to identify suitable protein-ligand systems and measure binding (Aim 3). The annual SAMPL challenges organized by the Mobley and Chodera labs (Aim 4) will leverage the data of Aims 1-3 (Figure 2) and best practices reference calculations to drive progress.

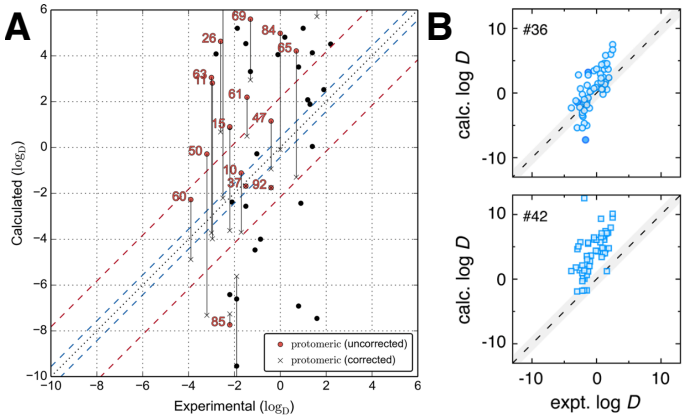

Figure 3. **Lessons learned from SAMPL5 log D predictions.** Predictions of log D values for SAMPL5 provided a number of key lessons. (A) Methods which treated multiple protonation and tautomeric states in their predictions performed dramatically better than those which did not; here, red dots move to x symbols when these effects are treated, improving accuracy in every case [155]. (B) Re-parameterization of a force field to more accurately reproduce pure solvent dielectric constants resulted in dramatically better predictions (top) than the original force field (bottom) [137].

Aim 1: Collect new physical property datasets to assess accuracy and spur improvements in force fields and modeling of protonation states and tautomers.

Simple physical properties such as solvation, partitioning, and protonation equilibria can be calculated quite precisely (but not necessarily accurately) with physical methods, allowing quantitative comparison between calculations and experiment and revealing and isolating deficiencies in our models. These properties allow us to directly probe force field accuracy and chemical effects like protonation and tautomer handling in the absence of slow conformational changes and other effects which complicate assessment in protein-ligand systems.

Rationale: We will generate new solution-phase physical property measurements for drug-like molecules to motivate improvements in force fields and handling of protonation states and tautomers. This builds on our work on water-cyclohexane distribution coefficients for SAMPL5 (in partnership with Genentech), which revealed major issues with handling of protonation states and tautomers [128] as well as serious forcefield limitations [137] (Figure 3). Distribution coefficients give the equilibrium ratio of concentrations of a solute between aqueous and nonpolar phases, and thus relate to transfer free energies from aqueous to more protein- or membrane-like environments. Thus, they capture many of the characteristics of transfer of drugs from water into binding sites but absent challenges with receptor conformational sampling and specific ligand-receptor interactions. In SAMPL5,

distribution coefficients were challenging enough that many methods performed poorly, with even the best methods having accuracies less than would be expected based on hydration free energies in water [128], yet failures were informative and the major sources of error were issues which will also plague prediction of ligand-receptor interactions. In some respects, distribution coefficients posed the ideal SAMPL challenge, hitting the sweet spot in terms of difficulty—difficult enough that clear failures were frequent, with ample room for improvement, but not so difficult that the reasons for failure were unclear in general. Still, many models consistently disagreed with experiment for some compounds [128, 137, 141, 156], revealing the impact that targeted follow-up experiments (such as those we will conduct here) could have on improving models.

This new experimental data is critical to maximize impact on the modeling community. While the community has already benefitted considerably from SAMPL, as indicated by the nearly 100 publications on SAMPL (see attached list), the citation count, and the growing participation (Figure 1), progress will be greatly accelerated by the proposed initiative. A funded, coordinated effort allows the targeted collection of datasets designed to focus on the most important problems. With multiple investigators and collaborators, we are poised to respond and adapt to new challenges and opportunities which emerge in a manner not previously possible. Additionally, we will be able to continue experimental work until the necessary data is collected rather than terminating it at a specific time point dictated by industry internships—allowing us to do things like ensure the full dynamic range of log D values is covered, unlike in SAMPL5 [128, 156]. This will improve our ability to learn from the data—for example, the lack of dynamic range for SAMPL5 meant that, when calculated values often spanned a larger dynamic range than the experimental values, it was unclear if this was an artifact of the data set itself, experimental limitations, or force field problems [128, 137, 141, 156]. With funding to extend the experimental set via follow-up experiments, we will be able to resolve similar issues, providing further impetus to improve models.

Below, we summarize plans for data collection for the SAMPL6-10 challenges of Aim 4 (see also Figure 2). These specific data sets were selected based on a desire to help the field resolve the problems encountered in SAMPL5 then progress to accurate estimation of new properties. Data sets will typically consist of at least 96 compounds for good statistics, though when possible, a much larger amount of data will be collected.

SAMPL6: Cyclohexane/water and octanol/water distribution coefficients. Building on the success of distribution coefficient measurements in SAMPL5 and their surprising ability to motivate rapid advances in physical modeling methodologies [128], we will measure cyclohexane-water distribution coefficients at pH 7.4 for a new batch of commercially-available drug- and fragment-like molecules. Given the routine nature of octanol-water distribution coefficient measurements and indications that their prediction may be computationally tractable [153, 157] despite the heterogeneous structure of the wet octanol phase [158], we will also measure octanol-water distribution coefficients for the same compounds. Because pK_a prediction was difficult but critical for SAMPL5, we will focus SAMPL6 on forcefields and tautomers by measuring pK_a values, revisiting pK_a prediction in SAMPL7 and 8.

drug	features
memantine	adamantane; 1:1
saxagliptin	adamantane; 1:1
premarin	steroid
pancuronium	steroid
varenicline	1:1 vs 1:2
valsartan	pK_a 4.37
omeprazole	pK_a 4.77
ranolazine	pK_a 7.17; epitopes
pradaxa	pK_a 3.87; epitopes
nilotinib	epitopes; pK_a 6.3
sensipar	epitopes; folding
vyvance	diamine; epitopes; folding
minocycline	tetracyclin; amino aniline

Table 1. **Selected drugs whose binding to CB[n] hosts will be assayed for SAMPL6, 8, and 10 challenges (SA 2.1).** These drugs bind to the cucubituril-based host systems considered here, some at high affinity, so measuring their affinities provides a way to test methods for predicting binding interactions absent complexities present in protein-ligand systems.

SAMPL7: pKa measurements for drug-like molecules. While much less complex than protein-ligand affinities, distribution coefficient measurements still conflate the challenging issues of protonation state and tautomer prediction, as well as transfer into different environments which may contain small but important quantities of co-solvents. Thus, we will separate these issues and improve our handling of them one at a time. For SAMPL7, then, we will measure pK_a values for an extensive set of drug-like molecules in water which will serve as the focus of the challenge, paving the way for SAMPL8.

SAMPL8: pKa measurements and distribution coefficients. In the next challenge, we will jointly explore pK_a and transfer issues, measuring distribution coefficients *and* pK_a values for the same set of compounds, with participants predicting (a) log D; (b) pK_a; and (c) log P. Unlike SAMPL6, pK_a values will not be provided.

SAMPL9-10: Solubility prediction and membrane permeability. With solubility predictions now becoming tractable [105–107] (with Schrödinger also working on amorphous solubility prediction), solubility measurements will be a valuable test for SAMPL9, combining the solvation aspects of SAMPL1-8 with a new solid phase component. New computational techniques are targeting membrane permeability [109, 110], and this is experimentally accessible (see support letters from Pfizer and Merck), leading to our interest in permeability for SAMPL10.

Experimental plan: Experimental data will be collected in collaboration with our pharma partners (see support letters), roughly following the model used for SAMPL5, where Chodera lab student Bas Rustenburg went to Genentech to conduct cyclohexane-water log D measurements by adapting a Genentech high-throughput mass spectrometry workflow [156]. To collect this data, the Mobley and Chodera labs will send graduate students on visits or internships to industry collaborators to collect targeted datasets. Working with industry collaborators (see Letters of Collaboration from Genentech, Pfizer, and Merck) gives us substantial access to equipment and high-throughput measurement workflows—such as the Sirius T3 from Sirius Analytical (which can measure partition/distribution coefficients, pK_as, and solubilities for molecules with titratable groups)—automation equipment, and compound libraries—for the purposes of rapidly collecting targeted datasets. Our previous experience demonstrates this model will work [156], and our partners see the value of this data and SAMPL to the modeling community.

Overall, Aim 1 extends prior SAMPL challenges via data focused on quantitative prediction of physical properties of tremendous relevance to accurately predicting biomolecular interactions, paving the way to applications in more complex systems addressed in Aims 2 and 3.

Aim 2: Measure affinities of drug-like compounds in supramolecular hosts to challenge quantitative models of binding in systems not plagued by major receptor sampling issues.

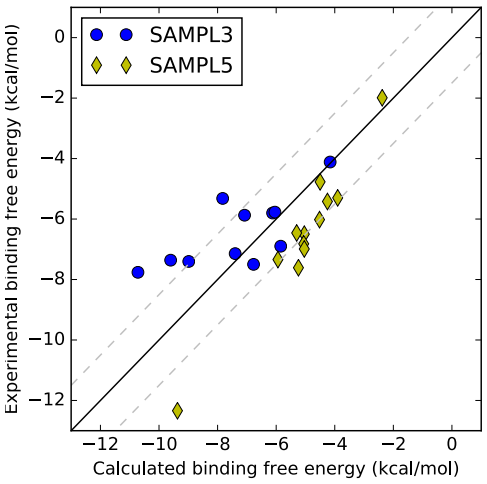

Figure 4. **The best host-guest binding predictions of SAMPL3 [159] and SAMPL5 [144].** Binding free energy predictions have shown clear improvements from SAMPL3 to SAMPL5 as the major challenges become understood and are treated better by models, though a systematic offset remains in the best SAMPL5 predictions (yellow). Dashed lines denote errors of ± 1.5 kcal/mol.

Aim 1 focuses on the behavior of small molecules and its environment-dependence, in the absence of receptors and the associated potential for slow sampling, strong specific interactions, and other challenges such as salt effects. Binding in host-guest systems retains many of the same challenges seen in Aim 1 and introduces strong specific interactions and other challenges like salt effects [142], while still avoiding many of the issues with slow sampling (of protein conformational changes, ions, and ligand binding modes) seen in protein-ligand interactions. That is, binding in host-guest systems introduces a wider variety of challenges relevant to biomolecular interactions, but without the full array of challenges seen in protein-ligand interactions, as reviewed recently [142]. Thus, new data for SAMPL challenges on host-guest binding is critical to provide challenges of intermediate complexity between those of physical properties and those in biomolecular binding. We believe that host-guest binding challenges provide a vital step towards accurately modeling biomolecular interactions, focusing the field on issues not commonly encountered in physical property challenges (such as the importance of accurately modeling ionic conditions) that are highly relevant for protein-ligand interactions. Already, over the past several SAMPL challenges, host-guest systems have provided key tests for modeling of binding interactions, resulting in new attention paid to how co-solvents and ions modulate binding (resulting in errors of up to 5 kcal/mol when these effects are neglected) and the importance of adequately sampling water rearrangements [127, 129, 142, 160].

This new attention has resulted in clear improvements as participants begin to treat the relevant effects more accurately (Figure 4). Host-guest binding proves remarkably difficult to model accurately [161], in part due to force field limitations (resulting in new force field work [146]).

Here, we design a series of SAMPL challenges focused on two classes of host-guest systems—cucurbiturils and analogs (SA 2.1) and Gibb's deep-cavity cavitands (GDCCs, SA 2.2)—both of which build on prior SAMPL challenges. These two sets of systems exhibit different challenges as recently reviewed [142], with the hosts of 2.1 bringing relatively modest co-solvent and ion effects but some receptor sampling problems for the acyclic hosts, and the GDCCs of 2.2 bringing profound ion and co-solvent effects as well as water sampling challenges. Methods which perform well on one class may not perform well on the other [142], since the distinct sets of challenges highlight different limitations. This diversity drives more innovation than would a focus on a single host class.

Subaim 2.1: Cucurbituril-based receptors as model binding systems

Cucurbituril derivatives for host-guest binding. Building on previous success with cucurbit[n]uril (abbreviated CB[n]) experiments for SAMPL challenges [162–164], we will conduct a series of new experiments on these receptors for five new challenges, with experimental work conducted by co-investigator Isaacs, an expert on these systems who provided data for previous SAMPL challenges. CB[n] receptors are particularly well suited to our

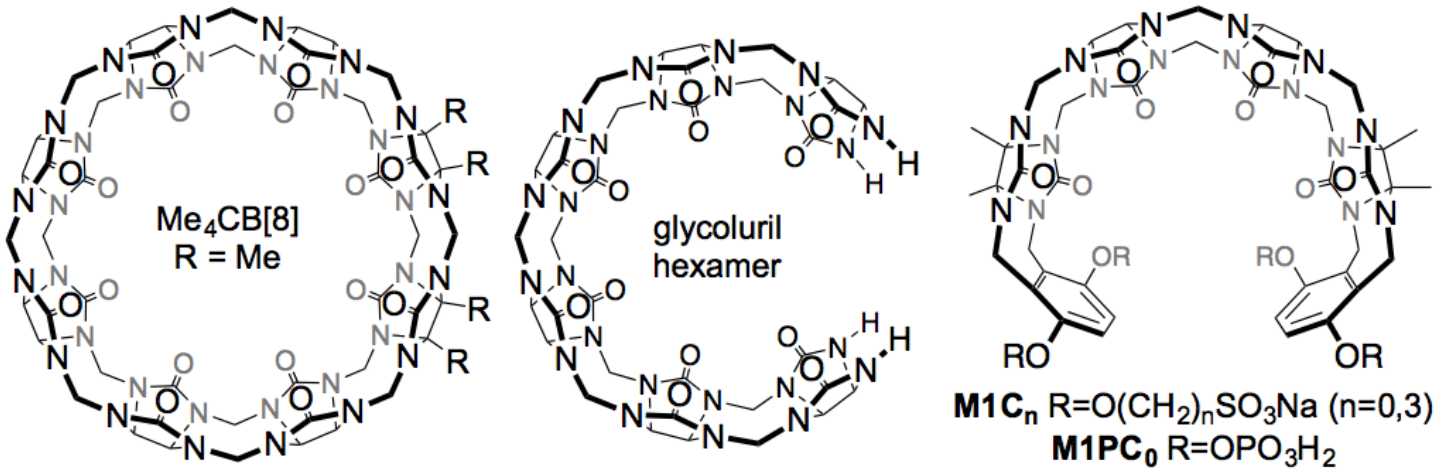

Figure 5. **SAMPL6-10 host-guest challenges will feature cucubituril hosts and analogs, including Me_4CB[8], glycoluril hexamer, and acyclic CB[n]-type receptors (SA 2.1).** These receptors bind a variety of drug-like molecules, some with high affinity.

goals because they exhibit: (1) high binding affinities toward suitable guests in water comparable to protein-ligand affinities (routinely μM to nM; occasionally pM to fM) [165–171], (2) high selectivities between structurally related guests which translate into large $\Delta\Delta G$ values [172], (3) low molecular weights (1–2 kDa) permitting high levels of theory to be used, and (4) highly restricted conformational degrees of freedom, reducing conformational sampling challenges often seen in protein-ligand binding. For SAMPL6-10, we will resynthesize a series of CB[n]-type receptors of increasing complexity, measure K_a values, and determine host-guest stoichiometry and geometry toward pharmaceutically relevant guests (selected drugs) in order to stringently test methods for predicting binding. Figure 5 shows the chemical structures of three hosts—Me_4CB[8] [173], glycoluril hexamer [174], and acyclic CB[n]-type receptors [175–180] which span a range in terms of level of preorganization and formal charge.

SAMPL6-10 cucurbituril challenges. *For SAMPL6*, we will measure K_a and ΔH values, stoichiometry, and geometry for the interaction of Me_4CB[8] (a soluble CB[8] derivative) with 15 guests (selected top drugs, Table 1) by either direct or competition isothermal titration calorimetry (ITC), UV/Vis or fluorescence indicator displacement assay, or NMR competition experiments, as previously [164–166, 181]. Our selection of Me_4CB[8] binding to top drugs allows us to modulate the computational complexity by: 1) changing host flexibility (e.g. Me_4CB[8] can exhibit ellipsoidal deformation) [173], 2) allowing the possibility of binary or ternary (e.g. 1:1 and/or 1:2 host:guest) complexes [182–184], 3) using drugs with several potential binding epitopes or modes to induce sampling issues. Host:guest stoichiometry and geometry (e.g., which binding epitope is complexed) will be addressed by ITC n values, Job plots monitored by UV/Vis or NMR [185], and by ^{1}H NMR complexation induced changes in chemical shifts [186]. All studies will be conducted in phosphate buffered saline (pH 7.4 with physiological salt) which introduces its own complexities due to salt competition for binding [142, 187]. *SAMPL7* will revisit the same host, but use 15 different guests be selected from commercial sources on the basis of reference calculations (on a larger set of guests) to ensure that they cover substantial dynamic range and/or exhibit affinities that depend substantially on the force field or water model, thus effectively testing our force fields and methods. For *SAMPL8*, we will focus on binding of the same 15 drugs (Table 1), but to glycoluril hexamer. This host introduces the complication of increased conformational dynamics, and influences the number and energy of solvating (and unusually coordinated) water molecules implicated in the high binding constants for CB[n]-guest complexes [171, 188]. The selected drugs include several with pK_a values in the 3.8 to 7.4 range; given that CB[n]-type receptors (like biomolecular receptors) can induce pK_a shifts in their guests of up to 4 pK_a units [189–191], this will test how well models can predict these effects. Additionally, it will couple nicely with the focus on pK_a values in Aim 1. *SAMPL9* will revisit glycouril hexamer with the same 15 guests from *SAMPL7*. *SAMPL10* will shift to acyclic CB[n]-type receptors (e.g. M1C_3, M1C_0, and M1PC_0 that contain anionic solubilizing groups attached via different linker lengths. As in SAMPL3 [159], these acyclic CB[n]-type receptors introduce conformational complexity, and water interactions play a key role. Moreover, the presence of 4 anionic groups near the cavity will likely impact the balance between ion-dipole interactions and solvation of the free host.

Subaim 2.2. Gibb deep cavity cavitands for host-guest studies

History of GDCC SAMPL challenges. During SAMPL4 [192] and SAMPL5 [193] we focused on two specific GDCC hosts: the octa-acid 1 (R = H) and another octa-acid variant with four methyl groups at the portal of the binding pocket (1, R = Me). These studies used isothermal titration calorimetry (ITC) to measure the thermodynamics of (1) host **1** (R = H) binding a range of 9 carboxylate guests, and (2) the binding of 6 carboxylate

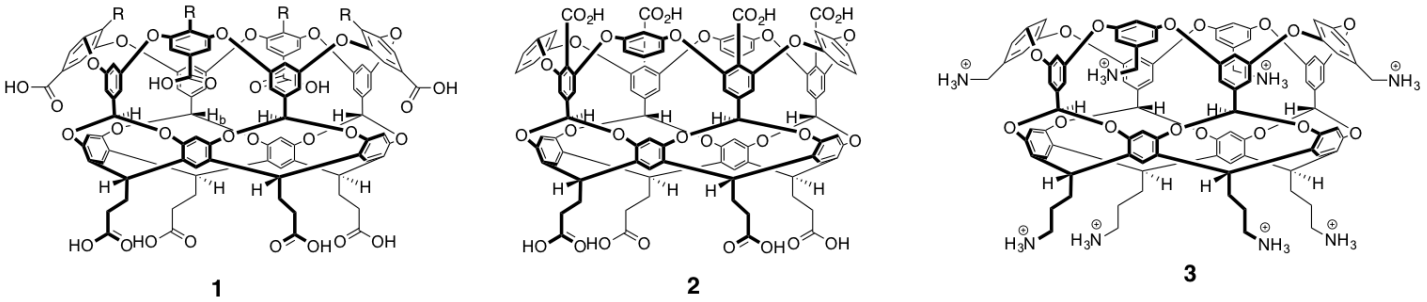

Figure 6. **Gibb deep cavity cavitands for the SAMPL6-10 datasets (SA 2.2).** These hosts bind a variety of carboxylate and trimethylammonium guests in a strongly salt-dependent manner, providing a stringent test of our ability to model salt-dependent binding.

and trimethylammonium guests to both hosts (**1**, R = H and Me; Figure 6). In both cases ^{1}H-NMR titration was also used to confirm ITC-derived free energies of binding. As noted above, co-solvent effects and water rearrangements posed particular challenges for predicting binding in these hosts. SAMPL5 emphasized how differences in the shape of the hydrophobic pocket of the host can have a profound effect on affinity for some guests [129].

Novel deep cavity hosts probe the effects of binding site charge constellations. For future GDCC datasets, we will expand on the range of hosts by including **2** and **3** in our ITC studies (Figure 6). Like cavitand **1**, host **2** is an octa-acid derivative. However, the four benzoate groups are relocated from the extreme exterior in the case of **1**, to the rim of the binding pocket in **2**. We surmise that this will have a direct effect on the binding of charged guests as well as an indirect effect on guest complexation via changes to the solvation of the empty host. Octa-trimethylammonuim cavitand ("positand" **3**) has the same overall architecture as host **1**, but inverts the charges on the water solubilizing exterior coat. While it is not yet clear if this switch in groups relatively remote from the pocket will directly affect guest complexation, results from related systems suggest it can (unpublished).

SAMPL6-10 deep cavity cavitand datasets. Data for SAMPL6 will focus on how well the effect of host carboxylate substituent location can be predicted, and will involve hosts **1** and **2** with a set of five previously uninvestigated guests. Guests will be selected from commercial sources on the basis of reference calculations in a similar manner to SAMPL7 in Subaim 2.1, specifically picking guests which have broad dynamic range and, here, have marked differences in affinities between hosts. SAMPL7 will provide a second iteration of this experiment to test algorithmic improvements in predictive modeling following SAMPL6 by comparing hosts **1** and **3** with a different set of guests. We anticipate that because of the relative remoteness of the charged groups in these two hosts, the effects of switching charges will be subtler than the differences between **1** and **2**. SAMPL8 will consider the effect of common biologically-relevant counterions/salts on guest binding, comparing the effects of NaCl and NaI on the complexation of five guests to **1**. We have previously shown that iodide has a weak affinity for the binding pocket of **1**, while sodium ions have an affinity for the outer carboxylates [194], requiring modeling to capture the differential affinities of these ions in addition to guest affinities to successfully model the observed affinities. SAMPL9 will follow up on this by examining the effects of these same two salts on the complexation of five guests to **3**, again giving the modeling community time to incorporate algorithmic improvements following SAMPL8. While we have not yet quantified salt affinities to host **3**, we expect the iodide to have affinity for both the pocket and the positively charged solubilizing groups. For SAMPL10 we will consider the effects of co-solvents on the binding of five guests to **1** and **2** to probe the effect of co-solvent competition for the binding site, as well as effects co-solvents may have in weakening the hydrophobic effect. While the number of guests considered in each challenge is relatively small, the total number of binding affinities measured is significant across the full family of hosts, meaning that the full data set will be of considerable value as a benchmark set [142].

Aim 3: Develop model protein-ligand systems that isolate specific modeling challenges of drug targets.

We seek to drive advances in quantitative modeling of protein-ligand interactions. While D3R [120] benchmarks accuracy for targets of pharmaceutical interest, it does not provide a clear route to improving poor performance because the large number complexities exhibited by these targets make it difficult to identify clear points of failure [120, 130–132]. For example, while kinases are targets of great interest to drug discovery, blind challenges involving kinase targets conflate issues of slow protein conformational dynamics [195], protonation state effects of both protein [196] and ligand [197, 198], charged ligands, and the modeling of complex divalent salt environments and phosphorylation state effects along with the standard challenges of conformational sampling and forcefield accuracy. Thus D3R exercises serve the community well to understand current accuracy, but **blind challenges on complex pharmaceutical targets have limited ability to rapidly advance quantitative predictive modeling.**

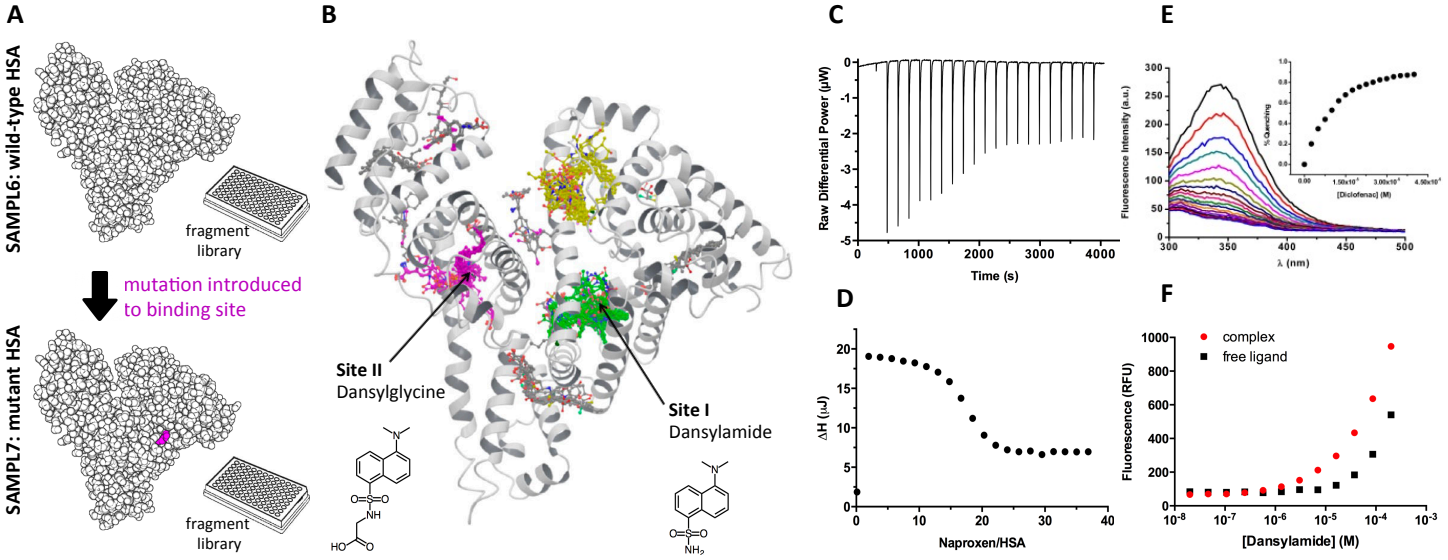

Figure 7. **The SAMPL6/7 protein-ligand challenge focuses on soluble drug fragment binding to human serum albumin (HSA) (Aim 3).** *(A)* SAMPL6 will study binding of a library of 96 small soluble druglike fragments to recombinant HSA, with an engineered HSA mutant used for SAMPL7. *(B)* HSA has at least eight known binding sites, with two major well-characterized sites (green, Sudlow's Site I; purple, Site II) that bind a variety of drugs (figure from [199]). Two fluorescent probes—dansylamide and dansylglycine—bind with $\sim\mu$M affinity and high selectivity to Site I and Site II, respectively; both exhibit binding-enhanced fluorescence at 480 nm, and can be used to site-specifically probe ligand affinities by competition. *(C, D)* Binding affinities of soluble molecules can be measured by isothermal titration calorimetry (ITC); here, we show the *(C)* differential power and *(D)* integrated injection heats for the ITC titration of HSA by naproxen sodium collected using the Chodera lab automation pipeline; *(E)* HSA tryptophan fluorescence quenching can also be used to measure ligand binding affinity; here, HSA titration by diclofenac is shown, with the inset plot showing percent quenching at 346 nm [200, 201]. *(F)* Direct fluorescence binding assay of Dansylamide (fluorescent ligand) and HSA collected on the Chodera lab automation system. The binding curve can be constructed based on binding-induced fluorescence emission at 480 nm.

We take the alternative approach of identifying and developing specific protein-ligand systems which *isolate* individual accuracy-limiting effects in a series of prospective challenges. By developing *model binding systems*—real protein-ligand systems that may be of pharmacological interest, but comprised of single-domain proteins binding to a simple ligand series free of complex phenomena—we can study systems of complexity intermediate between completely artificial systems (such as the T4 lysozyme L99A system developed by Shoichet and Matthews [118, 142, 202]) and complex pharmaceutical targets where multiple modeling challenges make it difficult to learn from failure (Figure 8A). This process focuses the field on identifying and evaluating multiple solutions to selected accuracy-limiting effects (such as how to deal with ligand and protein protonation-state issues [203], slow protein conformational dynamics, etc.) while avoiding other complicating factors.

While model systems have had ample success in driving progress in individual research laboratories, community participation in blind challenges amplifies their power. For example, SAMPL3 featured the binding of small, rigid charged molecules to bovine trypsin [204], and rapidly focused the field on the deficiencies of current alchemical free energy methodologies in treating the binding of charged ligands. Within two years, multiple laboratories had developed practical solutions to effectively handle charged ligand binding [205–207].

SAMPL6-10 model protein-ligand challenges. We will introduce a new model protein-ligand system each year (revisiting the prior year's system if this becomes too difficult), with multiple challenges on each system (Figure 2) to allow iterative improvement and assessment. Our SAMPL6 data will focus on binding of small soluble drug fragments to one particular protein (below). However, maximizing gains in this area requires adapting subsequent challenges based on deficiencies identified by previous D3R/SAMPL challenges. Therefore, subsequent model systems will be rapidly identified and developed using our new informatics platform (below).

SAMPL6: Assessing predictive modeling of binding to multiple weak sites via measuring fragment binding to human serum albumin (HSA). HSA, the most abundant blood plasma protein, has a remarkable ability to bind a great variety of small molecule drugs in multiple binding sites (Figure 7B) [208]. As a result, HSA not only helps isolate the challenge of multiple weak ligands binding to a stable rigid protein, but it is also a pharmacologically relevant because of how it drastically modulates drug pharmacokinetics [199]. HSA has at least *eight* known binding sites, with numerous crystal structures available for drugs binding to two predominant sites (Site I and II) [199]. Small soluble molecules resembling drug fragments are highly likely to bind to HSA ($\geq$90% of such fragments, as detected by SPR [209]), providing an experimentally-tractable diverse set of ligands spanning several orders of magnitude in affinity [209]. As current advanced methodologies such as alchemical free energy calculations currently assume a single well-defined binding site with high affinity [210], this dataset will allow the

isolation of the effect of weak multiple binding from the majority of other confounding factors in protein-ligand binding. As HSA is relatively rigid, and computational methods already show some promise in computing binding affinities to HSA [199, 211, 212], this is an optimal model system for SAMPL6.

Recombinant HSA will be expressed in *E. coli* and purified via refolding from inclusion bodies [215], then defatted at low pH [216] to ensure the resulting protein is free of the glycosylation and bound fatty acids found in plasma-isolated HSA [216]. Recombinant expression will also allow a mutant form of HSA (engineered via single-primer mutagenesis) to be fielded for SAMPL7 (Figure 7). We will obtain a diverse library of 96 soluble drug-fragment-like molecules in pre-plated format as dry compound, and use our automated isothermal titration calorimetry (ITC) pipeline (Figure 7C,D) to characterize overall binding affinities to HSA. The same ligands pre-plated in DMSO format will be used to conduct a separate set of fluorescence titration assays (monitoring tryptophan fluorescence quenching, Figure 7D) and competition assays with site-specific fluorescent probes (Figure 7B) to resolve site-specific affinities to Sites I and II. We will field several levels of challenges, including challenges focused on affinities to Sites I and II, as well as challenges focused on predicting overall affinity and stoichiometry.

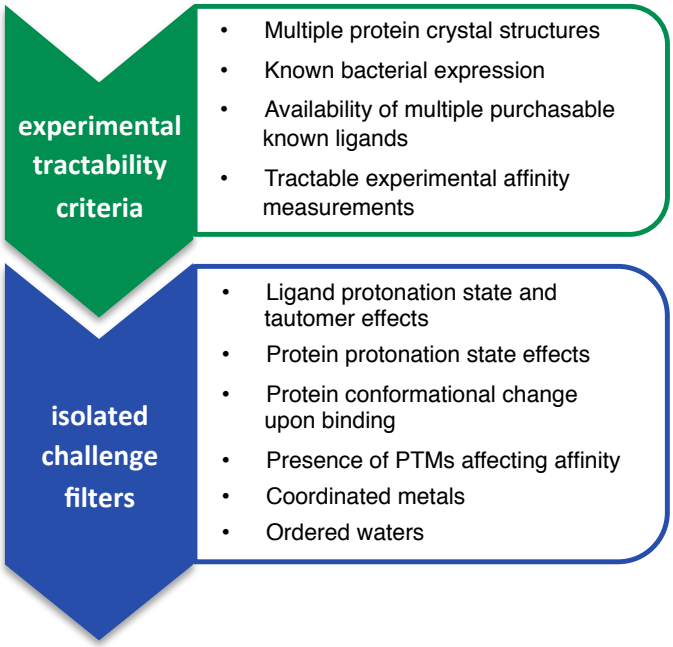

Figure 8. Mining model protein-ligand systems to focus on individual modeling challenges via a structural and chemical informatics platform (Aim 3). We are developing a structural and chemical informatics system called TargetExplorer [`https://github.com/choderalab/targetexplorer`] that applies successive filters to all potential protein-ligand systems for which structural data is available. Suitable model systems should meet all experimental tractability criteria (green box) and possess only a few challenging properties, ideally only one (blue box). Tractability of experimental affinity measurements includes properties like known ligands with potentially fluorescent scaffolds (for fluorescence competition assays), highly soluble ligands (for ITC), or ligands above a minimal mass (for SPR or MST). Additional filters annotate experimentally tractable systems with their potential computational challenges, including charged ligands or potential ligand protonation state or tautomer effects [213] (deduced from predicted aqueous protonation/tautomer energies); potential protein protonation state effects (deduced from MCCE2 calculations [214]); protein conformational changes (deduced from variation in protein conformation or the presence of unresolved loops in protein-ligand crystal structures); the presence of post-translational modifications that may affect affinity (deduced from Uniprot annotations); coordinated metals (identified in crystal structures); and ordered waters (present in multiple crystal structures).

SAMPL7-10: Rapid development of new tailored model systems using a novel informatics platform. We are developing a novel informatics platform aimed at identifying protein targets that can be rapidly developed into experimentally- and computationally-tractable model systems focusing on individual challenges (Figure 8). This tool filters all known protein targets with structural data available in the PDB, first selecting for experimental tractability, then annotating experimentally tractable targets to determine which targets possess (or are likely to be free of) specific challenges for physical modeling. This will allow us to select systems which introduce only specific modeling challenges.

The Chodera lab has developed an automated wetlab to facilitate the development of such model protein-ligand systems using bacterial expression (see Equipment and Facilities). Potential targets matching desired challenge criteria will be screened for bacterial expression using high-throughput cloning, transformation, and expression testing, with purity and yield assessed by capillary electrophoresis on a Caliper GXII. Targets will be screened for stability in various buffers using Thermofluor thermal shift assays [217]. Ligands identified via TargetExplorer as spanning a wide dynamic range of binding affinities will be purchased as dry powder stocks and prepared for assay by highly accurate gravimetric solution preparation techniques using a Quantos automated balance. Our lab has access to a wide variety of biophysical techniques for quantitative measurement of protein-ligand binding affinities, including fluorescence (if fluorescent probe ligands are available), absorption (e.g. Soret band shifts), automated isothermal titration calorimetry (provided ligands are sufficiently soluble), surface plasmon resonance, microscale thermophoresis (MST), luminescence, and alphascreen; all except MST are fully automated.

We take a twofold approach to developing challenge datasets: First, we will purchase and assay small molecules similar to known ligands, presuming that these molecules are likely to have measurable affinities. Second, using single-primer quick-change mutagenesis, we will introduce site-directed mutants to modulate the binding affinities of known ligands. This can be performed and screened for expression in 96-well format. Thus, datasets will consist of a matrix of protein mutants and ligands, providing opportunity to deeply explore the effects of interest.

Aim 4: Field community blind challenges to advance quantitative biomolecular design. The value of the targeted datasets generated in Aims 1–3 will be amplified enormously by the strategic release of this data through iterative, coordinated SAMPL blind challenges (Figure 2). These blind challenges are designed to test the state of the art, provoke new methodological and force field innovations, allow comparative evaluation of methods, and drive downstream improvements. The new, progressive, targeted nature of the data generated means that SAMPL challenges will now build on one another, and for success in later challenges, participants must build on lessons learned from prior challenges. SAMPL challenges and subsequent data release activities will therefore facilitate rapid cycles of application, learning, and improvement. Each iteration will likely yield its own incremental benefits (e.g.. as in Figure 3) for molecular design, in addition to contributing to progress towards our larger goals.

SAMPL blind challenges. Challenges will have yearly submission deadlines and involve roughly the same size data sets as prior challenges. The full timeline for SAMPL challenges (Figure 2) will be made available on the website [`https://drugdesigndata.org/about/sampl`] at the outset, allowing participants to plan their work and select what challenges to be involved in. As experimental data for each component becomes available and is curated, input files and challenge details will be made available at least six months prior to the challenge deadline; data not yet available at that time will be held for a subsequent challenge (with the exception of three months for year 1 due to startup timescales). As in prior SAMPLs, submissions will be handled by a web upload service on the SAMPL website (which will be migrated to separate hosting if the D3R effort is not renewed) which validates submissions to ensure that they meet format standards we specify along with the challenge details. As in SAMPL4 and SAMPL5, analysis will also be conducted by our automated Python framework, and results returned automatically online. All participant submissions and methodology descriptions will (as before) be made available publicly on the website, along with participant information (except for participants who specifically request to remain anonymous prior to submission). Aggregate statistics and historical performance will also be made available on our website, along with a record linking publications to historical submissions.

Our goal is not just to run blind challenges, but to advance modeling by helping participants identify both modeling failures and potential solutions. To achieve this, we provide guidance to participants as to what known modeling issues we expect may be relevant when providing details on each SAMPL component. For a host-guest system, for example, we might highlight known buffer/salt effects, protonation state challenges, and point out previous work on sampling challenges, with pointers to the relevant experimental work and to modeling work from past SAMPL challenges and elsewhere [142]. This helps participants design their approach. **Additionally, we will run reference calculations using current best practices.** This serves several purposes: It provides a test of the current methods and force fields we select; it helps facilitate learning—we announce what calculations we plan to perform, make input files available in a wide variety of formats [128, 129, 218], and others can repeat our calculations with a different method but same system and force field to compare methods, or swap force field but keep the method and system fixed to compare force fields, etc.; and it allows us to conduct sensitivity analysis, as by varying the conditions of our simulations (protonation state, tautomer, etc., [128]) we can see how much this impacts calculated values and thereby how important it is, even if participants don't do these tests. Reference calculations have, for example, helped us highlight the importance of a small amount of water in cyclohexane for accurately calculating log D values, show how an incorrect tautomer could affect calculated values by many log units [128], and discover that small forcefield modifications could significantly improve results on hydration free energies [126]. To further aid follow-up studies, we will make the input files, results, and simulation workflows used for our reference calculations—along with the data—available via GitHub and Docker Hub.

Physical methods are only valuable if they can reliably outperform alternate methods, so **a new focus of SAMPL6-10 will be selecting quality null models and running them to provide a point of comparison for participants**, going far beyond previous SAMPL nulls [127, 128, 137]).

Following submission and analysis of each SAMPL challenge, challenge results will be released and discussed, with SAMPL workshops allowing more formal presentations on and discussion of results in years 1, 3, and 5. Workshops will run every two years at the request of past participants, and will be co-run and co-hosted with D3R Grand Challenge workshops (see support letter). During the off years, SAMPL challenges will still run, but discussion of and dissemination of results will be via asynchronous means (as discussed below) and a "virtual workshop" consisting of talks and interaction over Google Hangouts or YouTube Live. While coordination with D3R will mostly be at the level of workshops, we will also ensure that SAMPL challenge submission deadlines are offset from D3R deadlines to allow maximum community participation in both efforts. If the D3R effort is not renewed beyond its current funding cycle, we will run SAMPL workshops independently, controlling costs via the model we use for our Workshops on Free Energy Methods in Drug Design—specifically, most participants will pay their own way to the workshop, and we will seek pharmaceutical and software industry sponsorships to defray costs.

Dissemination of results and data. Rapid dissemination of results is critical so that new insights can be used in subsequent challenges. We will continue to publish special issues of the *Journal of Computer Aided Molecular Design* (JCAMD) collecting publications related to each year's SAMPL challenges (see Letter of Support). To ensure immediate availability of reports, we will strongly encourage prepublication sharing of results and analysis, including both slides and posters from SAMPL meetings (via F1000 Research) and paper preprints (via bioRxiv). We also want to ensure that participants learn from one another by rapid exchange of ideas outside of formal workshops and meetings. While this has happened in the past—for example, when participants using similar methods work together after the SAMPL meeting to identify the origin of these discrepancies [129, 142, 160, 219, 220]—we hope to accelerate this kind of collaboration. To facilitate more open communication between the community, we will use collaboration software—such as Slack, which facilitates scientific communication for the NASA/JPL Mars Rover teams and NSF antarctic scientific research teams—to build a community discussion platform, facilitating a process of learning from one another more rapidly than normal publication channels.

Each dataset will have a life cycle of collection, curation, blind challenges, and public dissemination. In the past, the unfunded nature of SAMPL has forced us to primarily emphasize the *blind challenges* and pre-challenge *curation* aspects, with isolated forays into collection [156, 204]. This work now considers the full life cycle, with Aims 1–3 dealing primarily with collection and pre-challenge curation. Post-challenge, datasets will receive additional curation, then be released as standard test or benchmark sets that allow retrospective evaluation of methodologies on high-quality data [142]. The FreeSolv dataset, for example, includes a large number of calculated and experimental hydration free energies from SAMPL0–4, and provides a standard benchmark dataset for hydration free energy calculations [221]. Post-challenge curation will receive new attention here; in the past, lack of resources has always prevented follow-up experimental work, even when the data clearly indicated it was warranted (such as the puzzling issues with dynamic range for log D values in SAMPL5 [128, 156]). The requested budget will allow follow up experiments motivated by computation when warranted. Dissemination is the final stage in the data life cycle (see Resource Sharing Plan); we will make the data (including primary data, processed data, and our analysis of challenge submissions) available freely and publicly with permanent, citeable DOIs; ensure relevant data is deposited in standard community repositories (e.g. BindingDB [222]); and guarantee data longevity via backup hosting on library archival facilities (such as the UC's DASH (`https://dash.cdlib.org/`)).

We will also push for containerization of tools and methods in conjunction with other efforts such as AutoDesk's Molecular Design Toolkit and the NSF Molecular Sciences (MolSci) initiative. Our vision is that long-term, instead of participants submitting a set of predictions, they would also submit the entire workflow they applied via Docker containers allowing reproducibility and repurposing, ensuring dissemination workflows, not just results.

COLLABORATION MANAGEMENT PLAN

We have a strong previous history of successful collaboration, with Mobley and Chodera having co-authored roughly a dozen publications and organized several workshops and other initiatives. Mobley, Isaacs, and Gibb have also worked together to coordinate past SAMPL challenges, and Mobley and Gibb a previous NSF workshop. PI Mobley will oversee the project, with teams for the other aims (Aim 1: Mobley & Chodera; Aim 2.1: Isaacs; Aim 2.2: Gibb; Aim 3: Chodera; Aim 4: Mobley & Chodera) involving the other co-investigators as needed. Meetings will consist of a monthly Google Hangout and a yearly in-person planning meeting. Chodera and Mobley will communicate more frequently due to the interlinked nature of their work. Publications are expected to be largely dictated by the overall Timeline, with an experimental publication associated with each challenge component being prepared for distribution to participants along with their results. Conflict resolution is expected to be straightforward, but if any serious difficulties arise, Michael Gilson (UCSD) will arbitrate given our close connections with D3R.

OUTLOOK

Physical methods have been slow to achieve their promise in binding prediction, in part because truly significant innovations are so hard to recognize due to a lack of standard tests and benchmarks, and in part because of an "applications first" approach which seems to plague our community where we rush to apply our methods to problems of pharmaceutical relevance without ensuring they can tackle simpler, better-understood problems first. Here, we propose an innovative extension of the successful series of SAMPL blind challenges, generating novel experimental data to drive improvement of the methods in our field and help them become pharmaceutically relevant – beginning with relatively simple physical property prediction and progressing to challenging problems in biomolecular recognition via a series of carefully designed intermediate steps. SAMPL already has a strong track record of success, and funding will ensure dramatically increased impact and continued success. The proposed series of carefully tailored challenges will focus our community on a variety of problems which we *can* realistically resolve in the near term, resulting in dramatic improvements in computational molecular design.

Full List of SAMPL References

[1] Monroe, J. I. and Shirts, M. R.: Converging free energies of binding in cucurbit[7]uril and octa-acid host–guest systems from SAMPL4 using expanded ensemble simulations. J Comput Aided Mol Des. 28(4): 401–415, March 2014.

[2] Muddana, H. S., Yin, J., Sapra, N. V., Fenley, A. T., and Gilson, M. K.: Blind prediction of SAMPL4 cucurbit[7]uril binding affinities with the mining minima method. J Comput Aided Mol Des. 28(4): 463–474, February 2014.

[3] Gallicchio, E., Chen, H., Chen, H., Fitzgerald, M., Gao, Y., He, P., Kalyanikar, M., Kao, C., Lu, B., Niu, Y., Pethe, M., Zhu, J., and Levy, R. M.: BEDAM binding free energy predictions for the SAMPL4 octa-acid host challenge. J Comput Aided Mol Des. 29(4): 315–325, March 2015.

[4] Mikulskis, P., Cioloboc, D., Andrejić, M., Khare, S., Brorsson, J., Genheden, S., Mata, R. A., Söderhjelm, P., and Ryde, U.: Free-energy perturbation and quantum mechanical study of SAMPL4 octa-acid host–guest binding energies. J Comput Aided Mol Des. 28(4): 375–400, April 2014.

[5] Hsiao, Y.-W. and Söderhjelm, P.: Prediction of SAMPL4 host–guest binding affinities using funnel metadynamics. J Comput Aided Mol Des. 28(4): 443–454, February 2014.

[6] Bhakat, S. and Söderhjelm, P.: Resolving the problem of trapped water in binding cavities: Prediction of host-guest binding free energies in the SAMPL5 challenge by funnel metadynamics. J Comput Aided Mol Des. 2016.

[7] Pal, R. K., Haider, K., Kaur, D., Flynn, W., Xia, J., Levy, R. M., Taran, T., Wickstrom, L., Kurtzman, T., and Gallicchio, E.: A combined treatment of hydration and dynamical effects for the modeling of host-guest binding thermodynamics: The SAMPL5 blinded challenge. Journal of Computer-Aided Molecular Design. 2016.

[8] Yin, J., Henriksen, N. M., Slochower, D. R., and Gilson, M. K.: The SAMPL5 Host-Guest Challenge: Binding Free Energies and Enthalpies from Explicit Solvent Simulations. J Comput Aided Mol Des. 2016.

[9] Bosisio, S., Mey, A. S. J. S., and Michel, J.: Blinded predictions of host-guest standard free energies of binding in the SAMPL5 challenge. J Comput Aided Mol Des. 2016.

[10] Tofoleanu, F., Lee, J., Pickard IV., F. C., König, G., Huang, J., Baek, M., Seok, C., and Brooks, B. R.: Absolute binding free energy calculations for octa-acids and guests. J Comput Aided Mol Des. 2016.

[11] Mobley, D. L., Wymer, K. L., Lim, N. M., and Guthrie, J. P.: Blind prediction of solvation free energies from the SAMPL4 challenge. J Comput Aided Mol Des. 28(3): 135–150, March 2014.

[12] Muddana, H. S., Fenley, A. T., Mobley, D. L., and Gilson, M. K.: The SAMPL4 host–guest blind prediction challenge: An overview. J Comput Aided Mol Des. 28(4): 305–317, March 2014.

[13] Sullivan, M. R., Sokkalingam, P., Nguyen, T., Donahue, J. P., and Gibb, B. C.: Binding of carboxylate and trimethylammonium salts to octa-acid and TEMOA deep-cavity cavitands. J Comput Aided Mol Des. pp 1–8, July 2016.

[14] Deng, N., Forli, S., He, P., Perryman, A., Wickstrom, L., Vijayan, R. S. K., Tiefenbrunn, T., Stout, D., Gallicchio, E., Olson, A. J., and Levy, R. M.: Distinguishing Binders from False Positives by Free Energy Calculations: Fragment Screening Against the Flap Site of HIV Protease. J. Phys. Chem. B. 119(3): 976–988, January 2015.

[15] Li, L., Dill, K. A., and Fennell, C. J.: Testing the semi-explicit assembly model of aqueous solvation in the SAMPL4 challenge. J Comput Aided Mol Des. 28(3): 259–264, January 2014.

[16] Paranahewage, S. S., Gierhart, C. S., and Fennell, C. J.: Predicting water-to-cyclohexane partitioning of the SAMPL5 molecules using dielectric balancing of force fields. J Comput Aided Mol Des. pp 1–7, August 2016.

[17] Klamt, A., Eckert, F., Reinisch, J., and Wichmann, K.: Prediction of cyclohexane-water distribution coefficients with COSMO-RS on the SAMPL5 data set. J Comput Aided Mol Des. pp 1–9, July 2016.

[18] Tielker, N., Tomazic, D., Heil, J., Kloss, T., Ehrhart, S., Güssregen, S., Schmidt, K. F., and Kast, S. M.: The SAMPL5 challenge for embedded-cluster integral equation theory: Solvation free energies, aqueous pKa, and cyclohexane–water log D. J Comput Aided Mol Des. pp 1–10, August 2016.

[19] König, G., Pickard, F. C., Huang, J., Simmonett, A. C., Tofoleanu, F., Lee, J., Dral, P. O., Prasad, S., Jones, M., Shao, Y., Thiel, W., and Brooks, B. R.: Calculating distribution coefficients based on multi-scale free energy simulations: An evaluation of MM and QM/MM explicit solvent simulations of water-cyclohexane transfer in the SAMPL5 challenge. J Comput Aided Mol Des. pp 1–18, August 2016.

[20] Luchko, T., Blinov, N., Limon, G. C., Joyce, K. P., and Kovalenko, A.: SAMPL5: 3D-RISM partition coefficient calculations with partial molar volume corrections and solute conformational sampling. J Comput Aided Mol Des. pp 1–13, September 2016.

[21] Santos-Martins, D., Fernandes, P. A., and Ramos, M. J.: Calculation of distribution coefficients in the SAMPL5 challenge from atomic solvation parameters and surface areas. J Comput Aided Mol Des. pp 1–8, September 2016.

[22] Perryman, A. L., Santiago, D. N., Forli, S., Santos-Martins, D., and Olson, A. J.: Virtual screening with AutoDock Vina and the common pharmacophore engine of a low diversity library of fragments and hits against the three allosteric sites of HIV integrase: Participation in the SAMPL4 protein–ligand binding challenge. J Comput Aided Mol Des. 28(4): 429–441, February 2014.

[23] König, G., Pickard, F. C., Mei, Y., and Brooks, B. R.: Predicting hydration free energies with a hybrid QM/MM approach: An evaluation of implicit and explicit solvation models in SAMPL4. J Comput Aided Mol Des. 28(3): 245–257, February 2014.

[24] Voet, A. R. D., Kumar, A., Berenger, F., and Zhang, K. Y. J.: Combining in silico and in cerebro approaches for virtual screening and pose prediction in SAMPL4. J Comput Aided Mol Des. 28(4): 363–373, January 2014.

[25] Park, H.: Extended solvent-contact model approach to SAMPL4 blind prediction challenge for hydration free energies. J Comput Aided Mol Des. 28(3): 175–186, February 2014.

[26] Rustenburg, A. S., Dancer, J., Lin, B., Feng, J. A., Ortwine, D. F., Mobley, D. L., and Chodera, J. D.: Measuring experimental cyclohexane-water distribution coefficients for the SAMPL5 challenge. bioRxiv. 063081 pp, July 2016.

[27] Reinisch, J. and Klamt, A.: Prediction of free energies of hydration with COSMO-RS on the SAMPL4 data set. J Comput Aided Mol Des. 28(3): 169–173, January 2014.

[28] Muddana, H. S., Sapra, N. V., Fenley, A. T., and Gilson, M. K.: The SAMPL4 hydration challenge: Evaluation of partial charge sets with explicit-water molecular dynamics simulations. J Comput Aided Mol Des. 28(3): 277–287, January 2014.

[29] Manzoni, F. and Söderhjelm, P.: Prediction of hydration free energies for the SAMPL4 data set with the AMOEBA polarizable force field. J Comput Aided Mol Des. 28(3): 235–244, March 2014.

[30] Sandberg, L.: Predicting hydration free energies with chemical accuracy: The SAMPL4 challenge. J Comput Aided Mol Des. 28(3): 211–219, February 2014.

[31] Brini, E., Paranahewage, S. S., Fennell, C. J., and Dill, K. A.: Adapting the semi-explicit assembly solvation model for estimating water-cyclohexane partitioning with the SAMPL5 molecules. J Comput Aided Mol Des. pp 1–11, September 2016.

[32] Kamath, G., Kurnikov, I., Fain, B., Leontyev, I., Illarionov, A., Butin, O., Olevanov, M., and Pereyaslavets, L.: Prediction of cyclohexane-water distribution coefficient for SAMPL5 drug-like compounds with the QMPFF3 and ARROW polarizable force fields. J Comput Aided Mol Des. pp 1–12, September 2016.

[33] Diaz-Rodriguez, S., Bozada, S. M., Phifer, J. R., and Paluch, A. S.: Predicting cyclohexane/water distribution coefficients for the SAMPL5 challenge using MOSCED and the SMD solvation model. J Comput Aided Mol Des. pp 1–11, August 2016.

[34] Kenney, I. M., Beckstein, O., and Iorga, B. I.: Prediction of cyclohexane-water distribution coefficients for the SAMPL5 data set using molecular dynamics simulations with the OPLS-AA force field. J Comput Aided Mol Des. pp 1–14, August 2016.

[35] Caldararu, O., Olsson, M. A., Riplinger, C., Neese, F., and Ryde, U.: Binding free energies in the SAMPL5 octa-acid host–guest challenge calculated with DFT-D3 and CCSD(T). J Comput Aided Mol Des. pp 1–20, September 2016.

[36] Genheden, S. and Essex, J. W.: All-atom/coarse-grained hybrid predictions of distribution coefficients in SAMPL5. J Comput Aided Mol Des. pp 1–8, July 2016.

[37] Chung, K.-C. and Park, H.: Extended solvent-contact model approach to blind SAMPL5 prediction challenge for the distribution coefficients of drug-like molecules. J Comput Aided Mol Des. pp 1–15, July 2016.

[38] Koziara, K. B., Stroet, M., Malde, A. K., and Mark, A. E.: Testing and validation of the Automated Topology Builder (ATB) version 2.0: Prediction of hydration free enthalpies. J Comput Aided Mol Des. 28(3): 221–233, January 2014.

[39] Yin, J., Henriksen, N. M., Slochower, D. R., Shirts, M. R., Chiu, M. W., Mobley, D. L., and Gilson, M. K.: Overview of the SAMPL5 host–guest challenge: Are we doing better? J Comput Aided Mol Des. pp 1–19, September 2016.

[40] Bannan, C. C., Burley, K. H., Chiu, M., Shirts, M. R., Gilson, M. K., and Mobley, D. L.: Blind prediction of cyclohexane–water distribution coefficients from the SAMPL5 challenge. J Comput Aided Mol Des. pp 1–18, September 2016.

[41] Lee, J., Tofoleanu, F., Pickard, F. C., König, G., Huang, J., Damjanović, A., Baek, M., Seok, C., and Brooks, B. R.: Absolute binding free energy calculations of CBClip host–guest systems in the SAMPL5 blind challenge. J Comput Aided Mol Des. pp 1–15, September 2016.

[42] Jones, M. R., Brooks, B. R., and Wilson, A. K.: Partition coefficients for the SAMPL5 challenge using transfer free energies. J Comput Aided Mol Des. pp 1–10, September 2016.

[43] Pickard, F. C., König, G., Tofoleanu, F., Lee, J., Simmonett, A. C., Shao, Y., Ponder, J. W., and Brooks, B. R.: Blind prediction of distribution in the SAMPL5 challenge with QM based protomer and pKa corrections. J Comput Aided Mol Des. pp 1–14, September 2016.

[44] Cao, L. and Isaacs, L.: Absolute and relative binding affinity of cucurbit[7]uril towards a series of cationic guests. Supramolecular Chemistry. 26(3-4): 251–258, March 2014.

[45] Muddana, H. S. and Gilson, M. K.: Prediction of SAMPL3 host–guest binding affinities: Evaluating the accuracy of generalized force-fields. J Comput Aided Mol Des. 26(5): 517–525, January 2012.

[46] Gibb, C. L. D. and Gibb, B. C.: Binding of cyclic carboxylates to octa-acid deep-cavity cavitand. J Comput Aided Mol Des. 28(4): 319–325, November 2013.

[47] Klimovich, P. V. and Mobley, D. L.: Predicting hydration free energies using all-atom molecular dynamics simulations and multiple starting conformations. Journal of Computer-Aided Molecular Design. 24(4): 307–316, April 2010.

[48] Mobley, D. L., Liu, S., Cerutti, D. S., Swope, W. C., and Rice, J. E.: Alchemical prediction of hydration free energies for SAMPL. Journal of Computer-Aided Molecular Design. 26(5): 551–562, PMC3583515, May 2012.

[49] Muddana, H. S., Varnado, C. D., Bielawski, C. W., Urbach, A. R., Isaacs, L., Geballe, M. T., and Gilson, M. K.: Blind prediction of host–guest binding affinities: A new SAMPL3 challenge. J Comput Aided Mol Des. 26(5): 475–487, February 2012.

[50] Skillman, A. G.: SAMPL3: Blinded prediction of host–guest binding affinities, hydration free energies, and trypsin inhibitors. J Comput Aided Mol Des. 26(5): 473–474, May 2012.

[51] Newman, J., Dolezal, O., Fazio, V., Caradoc-Davies, T., and Peat, T. S.: The DINGO dataset: A comprehensive set of data for the SAMPL challenge. J Comput Aided Mol Des. 26(5): 497–503, December 2011.

[52] Gallicchio, E., Deng, N., He, P., Wickstrom, L., Perryman, A. L., Santiago, D. N., Forli, S., Olson, A. J., and Levy, R. M.: Virtual screening of integrase inhibitors by large scale binding free energy calculations: The SAMPL4 challenge. J Comput Aided Mol Des. 28(4): 475–490, February 2014.

[53] Klamt, A. and Diedenhofen, M.: Blind prediction test of free energies of hydration with COSMO-RS. J Comput Aided Mol Des. 24(4): 357–360, April 2010.

[54] Fennell, C. J., Kehoe, C. W., and Dill, K. A.: Modeling aqueous solvation with semi-explicit assembly. PNAS. 108(8): 3234–3239, February 2011.

[55] Ellingson, B. A., Skillman, A. G., and Nicholls, A.: Analysis of SM8 and Zap TK calculations and their geometric sensitivity. J Comput Aided Mol Des. 24(4): 335–342, April 2010.

[56] Surpateanu, G. and Iorga, B. I.: Evaluation of docking performance in a blinded virtual screening of fragment-like trypsin inhibitors. J Comput Aided Mol Des. 26(5): 595–601, December 2011.

[57] Purisima, E. O., Corbeil, C. R., and Sulea, T.: Rapid prediction of solvation free energy. 3. Application to the SAMPL2 challenge. J Comput Aided Mol Des. 24(4): 373–383, April 2010.

[58] König, G. and Brooks, B. R.: Predicting binding affinities of host-guest systems in the SAMPL3 blind challenge: The performance of relative free energy calculations. J Comput Aided Mol Des. 26(5): 543–550, December 2011.

[59] Kehoe, C. W., Fennell, C. J., and Dill, K. A.: Testing the semi-explicit assembly solvation model in the SAMPL3 community blind test. J Comput Aided Mol Des. 26(5): 563–568, December 2011.

[60] Kumar, A. and Zhang, K. Y. J.: Computational fragment-based screening using RosettaLigand: The SAMPL3 challenge. J Comput Aided Mol Des. 26(5): 603–616, January 2012.

[61] Meunier, A. and Truchon, J.-F.: Predictions of hydration free energies from continuum solvent with solute polarizable models: The SAMPL2 blind challenge. J Comput Aided Mol Des. 24(4): 361–372, March 2010.

[62] Genheden, S., Martinez, A. I. C., Criddle, M. P., and Essex, J. W.: Extensive all-atom Monte Carlo sampling and QM/MM corrections in the SAMPL4 hydration free energy challenge. J Comput Aided Mol Des. 28(3): 187–200, February 2014.

[63] Beckstein, O., Fourrier, A., and Iorga, B. I.: Prediction of hydration free energies for the SAMPL4 diverse set of compounds using molecular dynamics simulations with the OPLS-AA force field. J Comput Aided Mol Des. 28(3): 265–276, February 2014.

[64] Coleman, R. G., Sterling, T., and Weiss, D. R.: SAMPL4 & DOCK3.7: Lessons for automated docking procedures. J Comput Aided Mol Des. 28(3): 201–209, February 2014.

[65] Hogues, H., Sulea, T., and Purisima, E. O.: Exhaustive docking and solvated interaction energy scoring: Lessons learned from the SAMPL4 challenge. J Comput Aided Mol Des. 28(4): 417–427, January 2014.

[66] Reinisch, J., Klamt, A., and Diedenhofen, M.: Prediction of free energies of hydration with COSMO-RS on the SAMPL3 data set. J Comput Aided Mol Des. 26(5): 669–673, May 2012.

[67] Kulp, J. L., Blumenthal, S. N., Wang, Q., Bryan, R. L., and Guarnieri, F.: A fragment-based approach to the SAMPL3 Challenge. J Comput Aided Mol Des. 26(5): 583–594, January 2012.

[68] Klamt, A. and Diedenhofen, M.: Some conclusions regarding the predictions of tautomeric equilibria in solution based on the SAMPL2 challenge. J Comput Aided Mol Des. 24(6-7): 621–625, April 2010.

[69] Fu, J., Liu, Y., and Wu, J.: Fast prediction of hydration free energies for SAMPL4 blind test from a classical density functional theory. J Comput Aided Mol Des. 28(3): 299–304, March 2014.

[70] Hamaguchi, N., Fusti-Molnar, L., and Wlodek, S.: Force-field and quantum-mechanical binding study of selected SAMPL3 host-guest complexes. J Comput Aided Mol Des. 26(5): 577–582, February 2012.

[71] Colas, C. and Iorga, B. I.: Virtual screening of the SAMPL4 blinded HIV integrase inhibitors dataset. J Comput Aided Mol Des. 28(4): 455–462, January 2014.

[72] Ellingson, B. A., Geballe, M. T., Wlodek, S., Bayly, C. I., Skillman, A. G., and Nicholls, A.: Efficient calculation of SAMPL4 hydration free energies using OMEGA, SZYBKI, QUACPAC, and Zap TK. J Comput Aided Mol Des. 28(3): 289–298, March 2014.

[73] Sulea, T. and Purisima, E. O.: Predicting hydration free energies of polychlorinated aromatic compounds from the SAMPL-3 data set with FiSH and LIE models. J Comput Aided Mol Des. 26(5): 661–667, December 2011.

[74] Geballe, M. T., Skillman, A. G., Nicholls, A., Guthrie, J. P., and Taylor, P. J.: The SAMPL2 blind prediction challenge: Introduction and overview. J Comput Aided Mol Des. 24(4): 259–279, May 2010.

[75] Ribeiro, R. F., Marenich, A. V., Cramer, C. J., and Truhlar, D. G.: Prediction of SAMPL2 aqueous solvation free energies and tautomeric ratios using the SM8, SM8AD, and SMD solvation models. J Comput Aided Mol Des. 24(4): 317–333, April 2010.

[76] Skillman, A. G., Geballe, M. T., and Nicholls, A.: SAMPL2 challenge: Prediction of solvation energies and tautomer ratios. J Comput Aided Mol Des. 24(4): 257–258, April 2010.

[77] Gallicchio, E. and Levy, R. M.: Prediction of SAMPL3 host-guest affinities with the binding energy distribution analysis method (BEDAM). J Comput Aided Mol Des. 26(5): 505–516, February 2012.

[78] Mikulskis, P., Genheden, S., Rydberg, P., Sandberg, L., Olsen, L., and Ryde, U.: Binding affinities in the SAMPL3 trypsin and host–guest blind tests estimated with the MM/PBSA and LIE methods. J Comput Aided Mol Des. 26(5): 527–541, December 2011.

[79] Geballe, M. T. and Guthrie, J. P.: The SAMPL3 blind prediction challenge: Transfer energy overview. J Comput Aided Mol Des. 26(5): 489–496, April 2012.

[80] Guthrie, J. P.: SAMPL4, a blind challenge for computational solvation free energies: The compounds considered. J Comput Aided Mol Des. 28(3): 151–168, April 2014.

[81] Nicholls, A., Wlodek, S., and Grant, J. A.: SAMPL2 and continuum modeling. J Comput Aided Mol Des. 24(4): 293–306, April 2010.

[82] Soteras, I., Orozco, M., and Luque, F. J.: Performance of the IEF-MST solvation continuum model in the SAMPL2 blind test prediction of hydration and tautomerization free energies. J Comput Aided Mol Des. 24(4): 281–291, March 2010.

[83] Lawrenz, M., Wereszczynski, J., Ortiz-Sánchez, J. M., Nichols, S. E., and McCammon, J. A.: Thermodynamic integration to predict host-guest binding affinities. J Comput Aided Mol Des. 26(5): 569–576, February 2012.

[84] Sulea, T., Hogues, H., and Purisima, E. O.: Exhaustive search and solvated interaction energy (SIE) for virtual screening and affinity prediction. J Comput Aided Mol Des. 26(5): 617–633, December 2011.

[85] Beckstein, O. and Iorga, B. I.: Prediction of hydration free energies for aliphatic and aromatic chloro derivatives using molecular dynamics simulations with the OPLS-AA force field. J Comput Aided Mol Des. 26(5): 635–645, December 2011.

[86] Benson, M. L., Faver, J. C., Ucisik, M. N., Dashti, D. S., Zheng, Z., and Merz, K. M.: Prediction of trypsin/molecular fragment binding affinities by free energy decomposition and empirical scores. J Comput Aided Mol Des. 26(5): 647–659, April 2012.

[87] Kast, S. M., Heil, J., Güssregen, S., and Schmidt, K. F.: Prediction of tautomer ratios by embedded-cluster integral equation theory. J Comput Aided Mol Des. 24(4): 343–353, March 2010.

[88] Mobley, D. L., Bayly, C. I., Cooper, M. D., and Dill, K. A.: Predictions of Hydration Free Energies from All-Atom Molecular Dynamics Simulations. J Phys Chem B. 113: 4533–4537, January 2009.

[89] Newman, J., Fazio, V., Caradoc-Davies, T., Branson, K., and Peat, T. S.: Practical Aspects of the SAMPL Challenge: Providing an Extensive Experimental Data Set for the Modeling Community. Journal of Biomolecular Screening. 14(10): 1245, January 2009.

[90] Klamt, A., Eckert, F., and Diedenhofen, M.: Prediction of the Free Energy of Hydration of a Challenging Set of Pesticide-Like Compounds†. J Phys Chem B. January 2009.

[91] Guthrie, J. P.: A Blind Challenge for Computational Solvation Free Energies: Introduction and Overview. J Phys Chem B. 113(14): 4501–4507, January 2009.

[92] Marenich, A. V., Cramer, C. J., and Truhlar, D. G.: Performance of SM6, SM8, and SMD on the SAMPL1 Test Set for the Prediction of Small-Molecule Solvation Free Energies. J. Phys. Chem. B. 113(14): 4538–4543, April 2009.

[93] Sulea, T., Wanapun, D., Dennis, S., and Purisima, E. O.: Prediction of SAMPL-1 Hydration Free Energies Using a Continuum Electrostatics-Dispersion Model. J. Phys. Chem. B. 113(14): 4511–4520, April 2009.

[94] Nicholls, A., Wlodek, S., and Grant, J. A.: The SAMP1 Solvation Challenge: Further Lessons Regarding the Pitfalls of Parametrization. J. Phys. Chem. B. 113(14): 4521–4532, April 2009.

[95] Nicholls, A., Mobley, D. L., Guthrie, J. P., Chodera, J. D., Bayly, C. I., Cooper, M. D., and Pande, V. S.: Predicting Small-Molecule Solvation Free Energies: An Informal Blind Test for Computational Chemistry. J. Med. Chem. 51(4): 769–779, February 2008.

[96] Chamberlin, A. C., Cramer, C. J., and Truhlar, D. G.: Performance of SM8 on a Test To Predict Small-Molecule Solvation Free Energies. J. Phys. Chem. B. 112(29): 8651–8655, July 2008.

Bibliography and References Cited

[97] Mobley, D. L. and Klimovich, P. V.: Perspective: Alchemical free energy calculations for drug discovery. J. Chem. Phys. 137(23): 230901, January 2012.

[98] Christ, C. D. and Fox, T.: Accuracy Assessment and Automation of Free Energy Calculations for Drug Design. J. Chem. Inf. Model. 54(1): 108–120, January 2014.

[99] Deng, N., Forli, S., He, P., Perryman, A., Wickstrom, L., Vijayan, R. S. K., Tiefenbrunn, T., Stout, D., Gallicchio, E., Olson, A. J., and Levy, R. M.: Distinguishing Binders from False Positives by Free Energy Calculations: Fragment Screening Against the Flap Site of HIV Protease. J. Phys. Chem. B. 119(3): 976–988, January 2015.

[100] Sherborne, Bradley, : Opening the lid on FEP. J Comput Aided Mol Des. 2016.

[101] Wang, L., Wu, Y., Deng, Y., Kim, B., Pierce, L., Krilov, G., Lupyan, D., Robinson, S., Dahlgren, M. K., Greenwood, J., Romero, D. L., Masse, C., Knight, J. L., Steinbrecher, T., Beuming, T., Damm, W., Harder, E., Sherman, W., Brewer, M., Wester, R., Murcko, M., Frye, L., Farid, R., Lin, T., Mobley, D. L., Jorgensen, W. L., Berne, B. J., Friesner, R. A., and Abel, R.: Accurate and Reliable Prediction of Relative Ligand Binding Potency in Prospective Drug Discovery by Way of a Modern Free-Energy Calculation Protocol and Force Field. J Am Chem Soc. 137(7): 2695–2703, February 2015.

[102] Christ, C. D. Binding affinity prediction from molecular simulations: A new standard method in structure-based drug design?, May 2016.

[103] Cui, G. Affinity Predictions with FEP+: A Different Perspective on Performance and Utility, May 2016.

[104] Verras, A. Free Energy Perturbation at Merck: Benchmarking against Faster Methods, May 2016.

[105] Schnieders, M. J., Baltrusaitis, J., Shi, Y., Chattree, G., Zheng, L., Yang, W., and Ren, P.: The Structure, Thermodynamics, and Solubility of Organic Crystals from Simulation with a Polarizable Force Field. J. Chem. Theory Comput. 8(5): 1721–1736, May 2012.

[106] Park, J., Nessler, I., McClain, B., Macikenas, D., Baltrusaitis, J., and Schnieders, M. J.: Absolute Organic Crystal Thermodynamics: Growth of the Asymmetric Unit into a Crystal via Alchemy. J. Chem. Theory Comput. 10(7): 2781–2791, July 2014.

[107] Liu, S., Cao, S., Hoang, K., Young, K. L., Paluch, A. S., and Mobley, D. L.: Using MD Simulations To Calculate How Solvents Modulate Solubility. Journal of Chemical Theory and Computation. 12(4): 1930–1941, February 2016.

[108] Leonis, G., Steinbrecher, T., and Papadopoulos, M. G.: A Contribution to the Drug Resistance Mechanism of Darunavir, Amprenavir, Indinavir, and Saquinavir Complexes with HIV-1 Protease Due to Flap Mutation I50V: A Systematic MM–PBSA and Thermodynamic Integration Study. J. Chem. Inf. Model. 53(8): 2141–2153, August 2013.

[109] Lee, C. T., Comer, J., Herndon, C., Leung, N., Pavlova, A., Swift, R. V., Tung, C., Rowley, C. N., Amaro, R. E., Chipot, C., Wang, Y., and Gumbart, J. C.: Simulation-Based Approaches for Determining Membrane Permeability of Small Compounds. J. Chem. Inf. Model. 56(4): 721–733, April 2016.

[110] Comer, J., Schulten, K., and Chipot, C.: Calculation of Lipid-Bilayer Permeabilities Using an Average Force. J Chem. Theory Comput. 10(2): 554–564, February 2014.

[111] Liu, S., Wu, Y., Lin, T., Abel, R., Redmann, J. P., Summa, C. M., Jaber, V. R., Lim, N. M., and Mobley, D. L.: Lead optimization mapper: Automating free energy calculations for lead optimization. J Comput Aided Mol Des. 27(9): 755–770, September 2013.

[112] Mikulskis, P., Genheden, S., and Ryde, U.: A Large-Scale Test of Free-Energy Simulation Estimates of Protein–Ligand Binding Affinities. J. Chem. Inf. Model. 54(10): 2794–2806, October 2014.

[113] Homeyer, N., Stoll, F., Hillisch, A., and Gohlke, H.: Binding Free Energy Calculations for Lead Optimization: Assessment of Their Accuracy in an Industrial Drug Design Context. J. Chem. Theory Comput. 10(8): 3331–3344, August 2014.

[114] Shirts, M. R., Mobley, D. L., and Brown, S. P. Free-energy calculations in structure-based drug design. In Merz, , Kenneth MJ., Ringe, D., and Reynolds, C. H. (Eds.): Drug Design: Structure and Ligand-Based Approaches. Cambridge University Press, January 2010.

[115] Lim, N. M., Wang, L., Abel, R., and Mobley, D. L.: Sensitivity in binding free energies due to protein reorganization. Journal of Chemical Theory and Computation. July 2016.

[116] Rocklin, G. J., Boyce, S. E., Fischer, M., Fish, I., Mobley, D. L., Shoichet, B. K., and Dill, K. A.: Blind Prediction of Charged Ligand Binding Affinities in a Model Binding Site. J. Mol. Biol. 425(22): 4569–4583, November 2013.

[117] Nuzzo, R.: How scientists fool themselves –and how they can stop. Nature. 526(7572): 182–185, October 2015.

[118] Mobley, D. L., Graves, A. P., Chodera, J. D., McReynolds, A. C., Shoichet, B. K., and Dill, K. A.: Predicting absolute ligand binding free energies to a simple model site. J. Mol. Biol. 371(4): 1118–1134, August 2007.

[119] Boyce, S. E., Mobley, D. L., Rocklin, G. J., Graves, A. P., Dill, K. A., and Shoichet, B. K.: Predicting ligand binding affinity with alchemical free energy methods in a polar model binding site. J. Mol. Biol. 394(4): 747–763, December 2009.

[120] Gathiaka, S., Liu, S., Chiu, M., Yang, H., Stuckey, J. A., Kang, Y. N., Delproposto, J., Dunbar, J. B., Carlson, H. A., Burley, S., Walters, W., Amaro, R. E., Feher, V., and Gilson, M. K.: D3R Grand Challenge 2015: Evaluation of Protein-Ligand Pose and Affinity Prediction. J Comput Aided Mol Des. 2016.

[121] Nicholls, A., Mobley, D. L., Guthrie, J. P., Chodera, J. D., Bayly, C. I., Cooper, M. D., and Pande, V. S.: Predicting Small-Molecule Solvation Free Energies: An Informal Blind Test for Computational Chemistry. J. Med. Chem. 51(4): 769–779, February 2008.

[122] Nicholls, A., Wlodek, S., and Grant, J. A.: The SAMP1 Solvation Challenge: Further Lessons Regarding the Pitfalls of Parametrization. J. Phys. Chem. B. 113(14): 4521–4532, April 2009.

[123] Mobley, D. L., Bayly, C. I., Cooper, M. D., and Dill, K. A.: Predictions of Hydration Free Energies from All-Atom Molecular Dynamics Simulations. J Phys Chem B. 113: 4533–4537, January 2009.

[124] Geballe, M. T., Skillman, A. G., Nicholls, A., Guthrie, J. P., and Taylor, P. J.: The SAMPL2 blind prediction challenge: Introduction and overview. J Comput Aided Mol Des. 24(4): 259–279, May 2010.

[125] Geballe, M. T. and Guthrie, J. P.: The SAMPL3 blind prediction challenge: Transfer energy overview. J Comput Aided Mol Des. 26(5): 489–496, April 2012.

[126] Mobley, D. L., Wymer, K. L., Lim, N. M., and Guthrie, J. P.: Blind prediction of solvation free energies from the SAMPL4 challenge. J Comput Aided Mol Des. 28(3): 135–150, March 2014.

[127] Muddana, H. S., Fenley, A. T., Mobley, D. L., and Gilson, M. K.: The SAMPL4 host–guest blind prediction challenge: An overview. J Comput Aided Mol Des. 28(4): 305–317, March 2014.

[128] Bannan, C. C., Burley, K. H., Chiu, M., Shirts, M. R., Gilson, M. K., and Mobley, D. L.: Blind prediction of cyclohexane–water distribution coefficients from the SAMPL5 challenge. J Comput Aided Mol Des. pp 1–18, September 2016.

[129] Yin, J., Henriksen, N. M., Slochower, D. R., Shirts, M. R., Chiu, M. W., Mobley, D. L., and Gilson, M. K.: Overview of the SAMPL5 host–guest challenge: Are we doing better? J Comput Aided Mol Des. pp 1–19, September 2016.

[130] Ignjatović, M. M., Caldararu, O., Dong, G., Muñoz-Gutierrez, C., Adasme-Carreño, F., and Ryde, U.: Binding-affinity predictions of HSP90 in the D3R Grand Challenge 2015 with docking, MM/GBSA, QM/MM, and free-energy simulations. J Comput Aided Mol Des. pp 1–24, August 2016.

[131] Deng, N., Flynn, W. F., Xia, J., Vijayan, R. S. K., Zhang, B., He, P., Mentes, A., Gallicchio, E., and Levy, R. M.: Large scale free energy calculations for blind predictions of protein–ligand binding: The D3R Grand Challenge 2015. J Comput Aided Mol Des. pp 1–9, August 2016.

[132] Sunseri, J., Ragoza, M., Collins, J., and Koes, D. R.: A D3R prospective evaluation of machine learning for protein-ligand scoring. J Comput Aided Mol Des. pp 1–11, September 2016.

[133] Ellingson, B. A., Skillman, A. G., and Nicholls, A.: Analysis of SM8 and Zap TK calculations and their geometric sensitivity. J Comput Aided Mol Des. 24(4): 335–342, April 2010.

[134] Ellingson, B. A., Geballe, M. T., Wlodek, S., Bayly, C. I., Skillman, A. G., and Nicholls, A.: Efficient calculation of SAMPL4 hydration free energies using OMEGA, SZYBKI, QUACPAC, and Zap TK. J Comput Aided Mol Des. 28(3): 289–298, March 2014.

[135] Mobley, D. L., Liu, S., Cerutti, D. S., Swope, W. C., and Rice, J. E.: Alchemical prediction of hydration free energies for SAMPL. Journal of Computer-Aided Molecular Design. 26(5): 551–562, PMC3583515, May 2012.

[136] Fennell, C. J., Wymer, K. L., and Mobley, D. L.: A Fixed-Charge Model for Alcohol Polarization in the Condensed Phase, and Its Role in Small Molecule Hydration. J. Phys. Chem. B. 118(24): 6438–6446, June 2014.

[137] Paranahewage, S. S., Gierhart, C. S., and Fennell, C. J.: Predicting water-to-cyclohexane partitioning of the SAMPL5 molecules using dielectric balancing of force fields. J Comput Aided Mol Des. pp 1–7, August 2016.

[138] Sulea, T. and Purisima, E. O.: Predicting hydration free energies of polychlorinated aromatic compounds from the SAMPL-3 data set with FiSH and LIE models. J Comput Aided Mol Des. 26(5): 661–667, December 2011.

[139] Li, L., Dill, K. A., and Fennell, C. J.: Testing the semi-explicit assembly model of aqueous solvation in the SAMPL4 challenge. J Comput Aided Mol Des. 28(3): 259–264, January 2014.

[140] Brini, E., Paranahewage, S. S., Fennell, C. J., and Dill, K. A.: Adapting the semi-explicit assembly solvation model for estimating water-cyclohexane partitioning with the SAMPL5 molecules. J Comput Aided Mol Des. pp 1–11, September 2016.

[141] Klamt, A., Eckert, F., Reinisch, J., and Wichmann, K.: Prediction of cyclohexane-water distribution coefficients with COSMO-RS on the SAMPL5 data set. J Comput Aided Mol Des. pp 1–9, July 2016.

[142] Mobley, D. L. and Gilson, M. K.: Predicting binding free energies: Frontiers and benchmarks. bioRxiv. 074625 pp, September 2016.

[143] Muddana, H. S., Yin, J., Sapra, N. V., Fenley, A. T., and Gilson, M. K.: Blind prediction of SAMPL4 cucurbit[7]uril binding affinities with the mining minima method. J Comput Aided Mol Des. 28(4): 463–474, February 2014.

[144] Yin, J., Henriksen, N. M., Slochower, D. R., and Gilson, M. K.: The SAMPL5 Host-Guest Challenge: Binding Free Energies and Enthalpies from Explicit Solvent Simulations. J Comput Aided Mol Des. 2016.

[145] Muddana, H. S., Sapra, N. V., Fenley, A. T., and Gilson, M. K.: The SAMPL4 hydration challenge: Evaluation of partial charge sets with explicit-water molecular dynamics simulations. J Comput Aided Mol Des. 28(3): 277–287, January 2014.

[146] Yin, J., Fenley, A. T., Henriksen, N. M., and Gilson, M. K.: Toward Improved Force-Field Accuracy through Sensitivity Analysis of Host-Guest Binding Thermodynamics. The Journal of Physical Chemistry B. 119(32): 10145–10155, August 2015.

[147] Nielsen, J. E., Gunner, M. R., and García-Moreno E., B.: The pKa Cooperative: A collaborative effort to advance structure-based calculations of pKa values and electrostatic effects in proteins. Proteins. 79(12): 3249–3259, December 2011.

[148] Janin, J.: Assessing predictions of protein–protein interaction: The CAPRI experiment. Protein Science. 14(2): 278–283, February 2005.

[149] Moult, J., Fidelis, K., Kryshtafovych, A., Schwede, T., and Tramontano, A.: Critical assessment of methods of protein structure prediction (CASP) —round x. Proteins. 82: 1–6, February 2014.

[150] Coleman, R. G., Sterling, T., and Weiss, D. R.: SAMPL4 & DOCK3.7: Lessons for automated docking procedures. J Comput Aided Mol Des. 28(3): 201–209, February 2014.

[151] Reinisch, J., Klamt, A., and Diedenhofen, M.: Prediction of free energies of hydration with COSMO-RS on the SAMPL3 data set. J Comput Aided Mol Des. 26(5): 669–673, May 2012.

[152] Beauchamp, K. A., Behr, J. M., Rustenburg, A. S., Bayly, C. I., Kroenlein, K., and Chodera, J. D.: Towards Automated Benchmarking of Atomistic Forcefields: Neat Liquid Densities and Static Dielectric Constants from the ThermoML Data Archive. Journal of Physical Chemistry B. 119(40): 12912–12920, September 2015.

[153] Bannan, C. C., Calabró, G., Kyu, D. Y., and Mobley, D. L.: Calculating Partition Coefficients of Small Molecules in Octanol/Water and Cyclohexane/Water. Journal of Chemical Theory and Computation. 12(8): 4015–4024, August 2016.

[154] Wang, K., Chodera, J. D., Yang, Y., and Shirts, M. R.: Identifying ligand binding sites and poses using GPU-accelerated Hamiltonian replica exchange molecular dynamics. J Comput Aided Mol Des. 27(12): 989–1007, December 2013.

[155] Pickard, F. C., König, G., Tofoleanu, F., Lee, J., Simmonett, A. C., Shao, Y., Ponder, J. W., and Brooks, B. R.: Blind prediction of distribution in the SAMPL5 challenge with QM based protomer and pKa corrections. J Comput Aided Mol Des. pp 1–14, September 2016.

[156] Rustenburg, A. S., Dancer, J., Lin, B., Feng, J. A., Ortwine, D. F., Mobley, D. L., and Chodera, J. D.: Measuring experimental cyclohexane-water distribution coefficients for the SAMPL5 challenge. bioRxiv. 063081 pp, July 2016.

[157] Bhatnagar, N., Kamath, G., and Potoff, J. J.: Prediction of 1-octanol–water and air–water partition coefficients for nitro-aromatic compounds from molecular dynamics simulations. Physical Chemistry Chemical Physics. 15(17): 6467, 2013.

[158] Kollman, P. A.: Advances and continuing challenges in achieving realistic and predictive simulations of the properties of organic and biological molecules. Accounts of Chemical Research. 29(10): 461–469, 1996.

[159] Muddana, H. S., Varnado, C. D., Bielawski, C. W., Urbach, A. R., Isaacs, L., Geballe, M. T., and Gilson, M. K.: Blind prediction of host–guest binding affinities: A new SAMPL3 challenge. J Comput Aided Mol Des. 26(5): 475–487, February 2012.

[160] Bhakat, S. and Söderhjelm, P.: Resolving the problem of trapped water in binding cavities: Prediction of host-guest binding free energies in the SAMPL5 challenge by funnel metadynamics. J Comput Aided Mol Des. 2016.

[161] Henriksen, N. M., Fenley, A. T., and Gilson, M. K.: Computational Calorimetry: High-Precision Calculation of Host–Guest Binding Thermodynamics. Journal of Chemical Theory and Computation. 11(9): 4377–4394, September 2015.

[162] Ma, D., Glassenberg, R., Ghosh, S., Zavalij, P. Y., and Isaacs, L.: Acyclic cucurbituril congener binds to local anaesthetics. Supramolecular Chemistry. 24(5): 325–332, May 2012.

[163] Cao, L. and Isaacs, L.: Absolute and relative binding affinity of cucurbit[7]uril towards a series of cationic guests. Supramolecular Chemistry. 26(3-4): 251–258, March 2014.

[164] She, N., Moncelet, D., Gilberg, L., Lu, X., Sindelar, V., Briken, V., and Isaacs, L.: Glycoluril-Derived Molecular Clips are Potent and Selective Receptors for Cationic Dyes in Water. Chem. Eur. J. pp n/a–n/a, August 2016.

[165] Cao, L., Šekutor, M., Zavalij, P. Y., Mlinarić-Majerski, K., Glaser, R., and Isaacs, L.: Cucurbit[7]uril·Guest Pair with an Attomolar Dissociation Constant. Angew. Chem. Int. Ed. 53(4): 988–993, January 2014.

[166] Liu, S., Ruspic, C., Mukhopadhyay, P., Chakrabarti, S., Zavalij, P. Y., and Isaacs, L.: The Cucurbit[n]uril Family: Prime Components for Self-Sorting Systems. Journal of the American Chemical Society. 127(45): 15959–15967, November 2005.

[167] Mock, W. L. and Shih, N. Y.: Structure and selectivity in host-guest complexes of cucurbituril. The Journal of Organic Chemistry. 51(23): 4440–4446, November 1986.

[168] Assaf, K. I. and Nau, W. M.: Cucurbiturils: From synthesis to high-affinity binding and catalysis. Chem Soc Rev. 44(2): 394–418, January 2015.

[169] Moghaddam, S., Yang, C., Rekharsky, M., Ko, Y. H., Kim, K., Inoue, Y., and Gilson, M. K.: New Ultrahigh Affinity Host-Guest Complexes of Cucurbit[7]uril with Bicyclo[2.2.2]octane and Adamantane Guests: Thermodynamic Analysis and Evaluation of M2 Affinity Calculations. Journal of the American Chemical Society. 133(10): 3570–3581, March 2011.

[170] Shetty, D., Khedkar, J. K., Park, K. M., and Kim, K.: Can we beat the biotin–avidin pair?: cucurbit[7]uril-based ultrahigh affinity host–guest complexes and their applications. Chem. Soc. Rev. 44(23): 8747–8761, 2015.

[171] Biedermann, F., Uzunova, V. D., Scherman, O. A., Nau, W. M., and De Simone, A.: Release of High-Energy Water as an Essential Driving Force for the High-Affinity Binding of Cucurbit[n]urils. J. Am. Chem. Soc. 134(37): 15318–15323, September 2012.

[172] Isaacs, L.: Stimuli Responsive Systems Constructed Using Cucurbit[n]uril-Type Molecular Containers. Acc. Chem. Res. 47(7): 2052–2062, July 2014.

[173] Vinciguerra, B., Zavalij, P. Y., and Isaacs, L.: Synthesis and Recognition Properties of Cucurbit[8]uril Derivatives. Org. Lett. 17(20): 5068–5071, October 2015.

[174] Lucas, D., Minami, T., Iannuzzi, G., Cao, L., Wittenberg, J. B., Anzenbacher, P., and Isaacs, L.: Templated Synthesis of Glycoluril Hexamer and Monofunctionalized Cucurbit[6]uril Derivatives. J. Am. Chem. Soc. 133(44): 17966–17976, November 2011.

[175] Ma, D., Zhang, B., Hoffmann, U., Sundrup, M. G., Eikermann, M., and Isaacs, L.: Acyclic Cucurbit[n]uril-Type Molecular Containers Bind Neuromuscular Blocking Agents In Vitro and Reverse Neuromuscular Block In Vivo. Angew. Chem. Int. Ed. 51(45): 11358–11362, November 2012.

[176] Ma, D., Hettiarachchi, G., Nguyen, D., Zhang, B., Wittenberg, J. B., Zavalij, P. Y., Briken, V., and Isaacs, L.: Acyclic cucurbit[n]uril molecular containers enhance the solubility and bioactivity of poorly soluble pharmaceuticals. Nat Chem. 4(6): 503–510, June 2012.

[177] Zhang, B. and Isaacs, L.: Acyclic Cucurbit[n]uril-type Molecular Containers: Influence of Aromatic Walls on their Function as Solubilizing Excipients for Insoluble Drugs. J. Med. Chem. 57(22): 9554–9563, November 2014.

[178] Gilberg, L., Zhang, B., Zavalij, P. Y., Sindelar, V., and Isaacs, L.: Acyclic cucurbit[n]uril-type molecular containers: Influence of glycoluril oligomer length on their function as solubilizing agents. Org. Biomol. Chem. 13(13): 4041–4050, 2015.

[179] Sigwalt, D., Moncelet, D., Falcinelli, S., Mandadapu, V., Zavalij, P. Y., Day, A., Briken, V., and Isaacs, L.: Acyclic Cucurbit[n]uril-Type Molecular Containers: Influence of Linker Length on Their Function as Solubilizing Agents. ChemMedChem. 11(9): 980–989, May 2016.

[180] Zhang, B., Zavalij, P. Y., and Isaacs, L.: Acyclic CB[n]-type molecular containers: Effect of solubilizing group on their function as solubilizing excipients. Org. Biomol. Chem. 12(15): 2413–2422, 2014.

[181] Ma, D., Zavalij, P. Y., and Isaacs, L.: Acyclic Cucurbit[n]uril Congeners Are High Affinity Hosts. J. Org. Chem. 75(14): 4786–4795, July 2010.

[182] Ko, Y. H., Kim, E., Hwang, I., and Kim, K.: Supramolecular assemblies built with host-stabilized charge-transfer interactions. Chem. Commun. (13): 1305–1315, 2007.

[183] Barrow, S. J., Kasera, S., Rowland, M. J., del Barrio, J., and Scherman, O. A.: Cucurbituril-Based Molecular Recognition. Chem. Rev. 115(22): 12320–12406, November 2015.

[184] Urbach, A. R. and Ramalingam, V.: Molecular Recognition of Amino Acids, Peptides, and Proteins by Cucurbit[n]uril Receptors. Isr. J. Chem. 51(5-6): 664–678, May 2011.

[185] Connors, K. A.: Binding Constants. New York, NY, John Wiley & Sons, 1987.

[186] Masson, E., Ling, X., Joseph, R., Kyeremeh-Mensah, L., and Lu, X.: Cucurbituril chemistry: A tale of supramolecular success. RSC Adv. 2(4): 1213–1247, 2012.

[187] Márquez, C., Hudgins, R. R., and Nau, W. M.: Mechanism of Host-Guest Complexation by Cucurbituril. J. Am. Chem. Soc. 126(18): 5806–5816, May 2004.

[188] Biedermann, F., Nau, W. M., and Schneider, H.-J.: The Hydrophobic Effect Revisited—Studies with Supramolecular Complexes Imply High-Energy Water as a Noncovalent Driving Force. Angew. Chem. Int. Ed. 53(42): 11158–11171, October 2014.

[189] 'il Saleh, N., Koner, A., and Nau, W.: Activation and Stabilization of Drugs by Supramolecular pKa Shifts: Drug-Delivery Applications Tailored for Cucurbiturils. Angewandte Chemie. 120(29): 5478–5481, July 2008.

[190] Nau, W. M., Florea, M., and Assaf, K. I.: Deep Inside Cucurbiturils: Physical Properties and Volumes of their Inner Cavity Determine the Hydrophobic Driving Force for Host–Guest Complexation. Isr. J. Chem. 51(5-6): 559–577, May 2011.

[191] Ghosh, I. and Nau, W. M.: The strategic use of supramolecular pKa shifts to enhance the bioavailability of drugs. Advanced Drug Delivery Reviews. 64(9): 764–783, June 2012.

[192] Gibb, C. L. D. and Gibb, B. C.: Binding of cyclic carboxylates to octa-acid deep-cavity cavitand. J Comput Aided Mol Des. 28(4): 319–325, November 2013.

[193] Sullivan, M. R., Sokkalingam, P., Nguyen, T., Donahue, J. P., and Gibb, B. C.: Binding of carboxylate and trimethylammonium salts to octa-acid and TEMOA deep-cavity cavitands. J Comput Aided Mol Des. pp 1–8, July 2016.

[194] Carnegie, R. S., Gibb, C. L. D., and Gibb, B. C.: Anion Complexation and The Hofmeister Effect. Angew. Chem. 126(43): 11682–11684, October 2014.

[195] Lin, Y.-L., Meng, Y., Jiang, W., and Roux, B.: Explaining why Gleevec is a specific and potent inhibitor of Abl kinase. Proc. Natl. Acad. Sci. 110(5): 1664–1669, January 2013.

[196] Shan, Y., Seeliger, M. A., Eastwood, M. P., Frank, F., Xu, H., Jensen, M. O., Dror, R. O., Kuriyan, J., and Shaw, D. E.: A conserved protonation-dependent switch controls drug binding in the Abl kinase. PNAS. 106(1): 139–144, June 2009.

[197] Szakács, Z., Béni, S., Varga, Z., Örfi, L., Kéri, G., and Noszál, B.: Acid-Base Profiling of Imatinib (Gleevec) and Its Fragments. Journal of Medicinal Chemistry. 48(1): 249–255, January 2005.

[198] Grante, I., Actins, A., and Orola, L.: Protonation effects on the UV/Vis absorption spectra of imatinib: A theoretical and experimental study. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy. 129: 326–332, August 2014.

[199] Hall, M. L., Jorgensen, W. L., and Whitehead, L.: Automated Ligand- and Structure-Based Protocol for *in Silico* Prediction of Human Serum Albumin Binding. Journal of Chemical Information and Modeling. 53(4): 907–922, April 2013.

[200] Epps, D. E., Raub, T. J., Caiolfa, V., Chiari, A., and Zamai, M.: Determination of the Affinity of Drugs toward Serum Albumin by Measurement of the Quenching of the Intrinsic Tryptophan Fluorescence of the Protein. Journal of Pharmacy and Pharmacology. 51(1): 41–48, January 1999.

[201] Bou-Abdallah, F., Sprague, S. E., Smith, B. M., and Giffune, T. R.: Binding thermodynamics of Diclofenac and Naproxen with human and bovine serum albumins: A calorimetric and spectroscopic study. The Journal of Chemical Thermodynamics. 103: 299–309, December 2016.

[202] Merski, M., Fischer, M., Balius, T. E., Eidam, O., and Shoichet, B. K.: Homologous ligands accommodated by discrete conformations of a buried cavity. PNAS. 112(16): 5039–5044, April 2015.

[203] Onufriev, A. V. and Alexov, E.: Protonation and pK changes in protein–ligand binding. Quarterly Reviews of Biophysics. 46(02): 181–209, May 2013.

[204] Newman, J., Dolezal, O., Fazio, V., Caradoc-Davies, T., and Peat, T. S.: The DINGO dataset: A comprehensive set of data for the SAMPL challenge. J Comput Aided Mol Des. 26(5): 497–503, December 2011.

[205] Rocklin, G. J., Mobley, D. L., Dill, K. A., and Hünenberger, P. H.: Calculating the binding free energies of charged species based on explicit-solvent simulations employing lattice-sum methods: An accurate correction scheme for electrostatic finite-size effects. The Journal of Chemical Physics. 139(18): 184103, 2013.

[206] Lin, Y.-L., Aleksandrov, A., Simonson, T., and Roux, B.: An Overview of Electrostatic Free Energy Computations for Solutions and Proteins. J. Chem. Theory Comput. 10(7): 2690–2709, July 2014.

[207] Reif, M. M. and Oostenbrink, C.: Net charge changes in the calculation of relative ligand-binding free energies via classical atomistic molecular dynamics simulation. Journal of Computational Chemistry. 35(3): 227–243, January 2014.

[208] Fasano, M., Curry, S., Terreno, E., Galliano, M., Fanali, G., Narciso, P., Notari, S., and Ascenzi, P.: The extraordinary ligand binding properties of human serum albumin. IUBMB Life (International Union of Biochemistry and Molecular Biology: Life). 57(12): 787–796, December 2005.

[209] Elinder, M., Geitmann, M., Gossas, T., Kallblad, P., Winquist, J., Nordstrom, H., Hamalainen, M., and Danielson, U. H.: Experimental Validation of a Fragment Library for Lead Discovery Using SPR Biosensor Technology. Journal of Biomolecular Screening. 16(1): 15–25, January 2011.

[210] Gilson, M., Given, J., Bush, B., and McCammon, J.: The statistical-thermodynamic basis for computation of binding affinities: A critical review. Biophysical Journal. 72(3): 1047–1069, March 1997.

[211] Lexa, K. W., Dolghih, E., and Jacobson, M. P.: A Structure-Based Model for Predicting Serum Albumin Binding. PLoS ONE. 9(4): e93323, April 2014.

[212] Evoli, S., Mobley, D., Guzzi, R., and Rizzuti, B.: Multiple binding modes of ibuprofen in human serum albumin identified by absolute binding free energy calculations. bioRxiv. 068502 pp, August 2016.

[213] Martin, Y. C.: Let's not forget tautomers. Journal of Computer-Aided Molecular Design. 23(10): 693–704, October 2009.

[214] Song, Y., Mao, J., and Gunner, M. R.: MCCE2: Improving protein pKa calculations with extensive side chain rotamer sampling. Journal of Computational Chemistry. pp NA–NA, 2009.

[215] Latta, M., Knapp, M., Sarmientos, P., Bréfort, G., Becquart, J., Guerrier, L., Jung, G., and Mayaux, J.-F.: Synthesis and Purification of Mature Human Serum Albumin from E. Coli. Bio/Technology. 5(12): 1309–1314, December 1987.

[216] Lang, B. E. and Cole, K. D.: Unfolding properties of recombinant human serum albumin products are due to bioprocessing steps. Biotechnology Progress. 31(1): 62–69, January 2015.

[217] Reinhard, L., Mayerhofer, H., Geerlof, A., Mueller-Dieckmann, J., and Weiss, M. S.: Optimization of protein buffer cocktails using Thermofluor. Acta Crystallographica Section F Structural Biology and Crystallization Communications. 69(2): 209–214, February 2013.

[218] Shirts, M. R., Klein, C., Swails, J. M., Yin, J., Gilson, M. K., Mobley, D. L., Case, D. A., and Shirts, M. R.: Lessons learned from comparing molecular dynamics englines on the SAMPL5 dataset. J Comput Aided Mol Des. 2016.

[219] Monroe, J. I. and Shirts, M. R.: Converging free energies of binding in cucurbit[7]uril and octa-acid host–guest systems from SAMPL4 using expanded ensemble simulations. J Comput Aided Mol Des. 28(4): 401–415, March 2014.

[220] Bosisio, S., Mey, A. S. J. S., and Michel, J.: Blinded predictions of host-guest standard free energies of binding in the SAMPL5 challenge. J Comput Aided Mol Des. 2016.

[221] Mobley, D. L. and Guthrie, J. P.: FreeSolv: A database of experimental and calculated hydration free energies, with input files. J Comput Aided Mol Des. 28(7): 711–720, PMC4113415, June 2014.

[222] Liu, T., Lin, Y., Wen, X., Jorissen, R. N., and Gilson, M. K.: BindingDB: A web-accessible database of experimentally determined protein–ligand binding affinities. Nucl. Acids Res. 35(suppl 1): D198–D201, January 2007.