Trading-off Accuracy with Computational Cost: Adaptive Algorithms to Reduce Time to Clinical Insight

Jumana Dakka*,[1], Kristof Farkas-Pall*,[2], Vivek Balasubramanian[1] , Matteo Turilli[1],
Shunzhou Wan[2], David W Wright[2], Stefan Zasada[2],
Peter V Coveney[2], Shantenu Jha[1,3]

[1] *Rutgers, the State University of New Jersey, Piscataway, NJ 08854, USA*
[2] *University College London, London, UK, WC1H 0AJ*
[3] *Brookhaven National Laboratory, Upton, New York, 11973*
[*] *Contributed Equally*

***Abstract*—The efficacy of drug treatments depends on how tightly small molecules bind to their target proteins. Quantifying the strength of these interactions (the so called 'binding affinity') is a grand challenge of computational chemistry, surmounting which could revolutionize drug design and provide the platform for patient specific medicine. Recently, evidence from blind challenge predictions and retrospective validation studies has suggested that molecular dynamics (MD) can now achieve useful predictive accuracy ($\leq$ 1 kcal/mol) This accuracy is sufficient to greatly accelerate hit to lead and lead optimization.**

To translate these advances in predictive accuracy so as to impact clinical and/or industrial decision making requires that binding free energy results must be turned around in timescales of hours without loss of accuracy. This demands advances in algorithms, scalable software systems, and intelligent and efficient utilization of supercomputing resources. Specifically, it necessitates refining algorithms and developing technologies to marshal huge simulation campaigns.

This work is motivated by the real world problem of providing insight from drug candidate data on a time scale that is as short as possible. Specifically, we reproduce results from a collaborative project between UCL and GlaxoSmithKline to study a congeneric series of drug candidates binding to the BRD4 protein – inhibitors of which have shown promising preclinical efficacy in pathologies ranging from cancer to inflammation. We demonstrate the use of a framework called HTBAC, designed to support the aforementioned requirements of accurate and rapid drug binding affinity calculations. HTBAC facilitates the execution of the numbers of simulations while supporting the adaptive execution of algorithms. Furthermore, HTBAC enables the selection of simulation parameters during runtime which can, in principle, optimize the use of computational resources whilst producing results within a target uncertainty.

I. Scientific Motivation

Bromodomain-containing proteins, and in particular the four members of the BET (bromodomain and extra terminal domain) family, are currently a major focus of research in the pharmaceutical industry. Small molecule inhibitors of these proteins have shown promising preclinical efficacy in pathologies ranging from cancer to inflammation. Indeed, several compounds are progressing through early stage clinical trials and are showing exciting early results [1]. One of the most extensively studied targets in this family is the first bromodomain of bromodomain-containing protein 4 (BRD4-BD1) for which extensive crystallographic and ligand binding data are available [2].

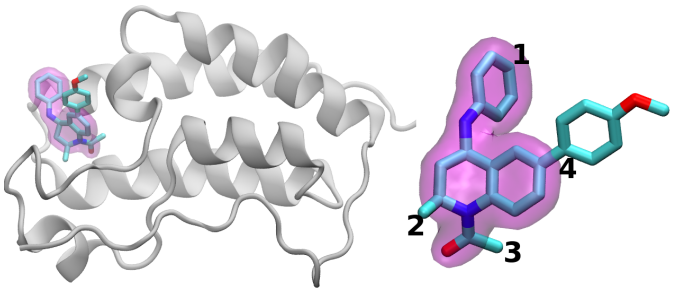

Fig. 1. (L) Cartoon representation of the BRD4 bound to an inhibitor shown in chemical representation (based on PDB:4BJX). (R) Ligand in cartoon representation with the tetrahydroquinoline scaffold highlighted in magenta. The regions which are modified between ligands investigated are labelled 1 to 4.

We have investigated a congeneric series of ligands binding to BRD4-BD1 (we shall from now on refer to this are simply BRD4). This was performed in the context of a blind test of the protocols in collaboration with GlaxoSmithKline [3]. The goal was to benchmark free energy calculations in a realistic drug discovery scenario. In the original study, we investigated chemical structures of 16 ligands based on a single tetrahydroquinoline (THQ) scaffold. These studies employed two different algorithms (simulation protocols), known as TIES and ESMACS [4], both based on multiple simulations of the same system. Drug design projects have limited resources, so initially large numbers of compounds must be cheaply screened to eliminate poor binders (using ESMACS), before more accurate methods (such as TIES) are needed as good binders are refined and improved.

In order to support such investigations, in addition to scale, the protocols must be executed to utilize flexible resource management schemes where based upon intermediate results at runtime, resources can be (re-)allocated between instances of different protocols or systems, for example, when one calculation has converged whilst another has not. Such adaptability makes it easier to manage complex programs where effi-

cient use of resources is required in order to achieve a time to completion of studies comparable to those of high throughput chemistry.

This work is motivated by the real world problem of providing insight from drug candidate data on a time scale that is as short as possible. We demonstrate the use of a framework – HTBAC, designed to support the aforementioned requirements of accurate and rapid drug binding affinity calculations. HTBAC facilitates the execution of the numbers of simulations while supporting the adaptive execution of algorithms. Furthermore, HTBAC enables the selection of experimental parameters during runtime, which can in principle optimize the use of computational resources whilst producing results with a target uncertainty.

II. Methods and Models

In this section we outline the computational methods employed and the physical system (drug candidates) studied. We discuss how computational methods have been co-designed with the software systems to support scalable approaches on the largest supercomputers.

A. Binding Affinity Calculation Protocols

Computing accurate protein-drug binding affinities (also known as binding free energies) requires a simulation technique which captures the chemical detail of the system. MD simulations are the time dependent numerical integration of the classical equations of motion for molecular systems. Application of MD to atomistic models of proteins and their ligands can be used to answer questions about the properties of a specific system often more readily than experiments on the actual system. Free-energy calculations in the framework of MD simulations not only yield quantitative estimates of binding strength but also provide insights into the most important interactions driving the process. Evidence from blind challenge predictions and retrospective validation studies has suggested that molecular dynamics (MD) can now achieve useful predictive accuracy (1 kcal/mol) [5], [6]. This accuracy is sufficient to greatly accelerate lead optimization [7].

Statistical mechanics provides the prescription for calculating such macroscopic quantities as ensemble averages of microscopic states. Traditionally, these macroscopic properties have usually been calculated from the time average of a single "long" duration trajectory. An intuitive and potentially more time efficient method to capture the mixing dynamics required to describe an equilibrium thermodynamic state is the use of an ensemble of separate trajectories. [8]

The major sources of error in free energy calculations are the representation of the system chemistry encoded in the forcefield used, finite sampling and the free energy estimator. Protocols developed in the Coveney labs have obtained accurate and precise results which successfully reproduce experimental binding free energies from a wide range of systems. [9], [3], [10], [11] Comparisons of results obtained for a large set of sequences will provide valuable insights on the importance of choices made in simulation and analysis for the overall accuracy and predictive power of free energy calculations, and facilitate the refinement of our protocols.

Most methods for calculating binding affinities fit into one of two broad classes; so called alchemical and endpoint methodologies. Alchemical free energy calculations employ unphysical ("alchemical") intermediates to calculate changes in free energies between two systems. It is common in these methods to refer to a variable, λ, which describes the path taken to transform one protein sequence (or ligand) into another. Endpoint methods, as the name suggests, consider the difference in energy between bound and unbound structures. To obtain information on the differences in binding affinity of different sequences for a panel of kinase inhibitors requires a deployment of various strategies, incorporating both alchemical and endpoint approaches. In this work we deploy approaches from both of these classes.

1) Alchemical Protocol (TIES): Alchemical methods employ MD simulations of unphysical, alchemical intermediate states that attenuate the interactions of the small molecule with its environment. These alchemical intermediate states include both the fully-interacting complex as well as replicas in which the ligand does not interact with its environment, and allow the total free energy of binding—including entropic and enthalpic contributions—to be efficiently computed. Typically, the alchemical path between the states of interest is described by a parameter, λ, which varies between 0 for the initial and 1 for the final state of the transformation of interest. Sampling is then performed at a series of points along this path and the gradient of the energy integrated to calculate the binding affinity. Simulations conducted at a given λ value are said to be sampling a λ window at that point.

The TIES (thermodynamic integration with enhanced sampling) protocol, developed within the Coveney lab, employs ensemble sampling at each λ window to yield reproducible, accurate, and precise relative binding affinities. [3] Based on the direct calculation of ensemble averages, it allows us to determine statistically meaningful results along with complete control of errors. As currently designed, TIES computes the change in binding affinity between two related system (termed 'relative binding affinities').

2) Endpoint Protocol (ESMACS): Computationally cheaper, but less rigorous methods, endpoint methods have been used to directly compute the binding strength of a drug to the target protein from MD simulations (as opposed to differences in affinity).

We have developed an ensemble-based endpoint protocol called ESMACS (enhanced sampling of molecular dynamics with approximation of continuum solvent). The protocol is built on the popular molecular mechanics Poisson–Boltzmann surface area (MMPBSA) [12] method which makes a continuum approximation for the aqueous solvent in order to obtain results on practical timescales. Commonly, MMPBSA analyses are performed on a single MD trajectory, or even a single structure. We have demonstrated the lack of reproducibility of such an approach in both HIV-1 protease and MHC systems, with calculations for the same protein-ligand combina-

tion, with identical initial structure and force field, shown to produce binding affinities varying by up to 12 kcal/mol for small ligands (flexible ligands can vary even more). [10] ESMACS employs MMPBSA to produce ensemble- based, converged and reproducible, determinations of binding free energies (separate ligand and receptor trajectories can also be used to account for adaptation energies). This provides a rapid quantitative approach sensitive enough to determine changes in binding free energies which differentiate susceptible and resistant sequences (typically of the order of 2 kcal/mol).

B. BRD4 System

Initial coordinates for the protein-ligand system were taken from a Protein Data Bank X-ray crystal structure (ID: 4BJX), which contains BRD4 bound to one of the ligands investigated in this study. The other 15 drugs were aligned to the common THQ scaffold of the ligand in the crystal structure to provide initial bound conformations. Preparation of simulation input files were implemented using our automated tool, BAC [13]. Preparation includes (i) parametrization of the compounds (ii) solvation and neutralization of the system and (iii) creation of MD engine configuration files. Parametrization was performed with the widely used AMBER ff99SB-ILDN force field for the protein, TIP3P for water molecules, and the General Amber force field (GAFF) [14] for the ligands. Ligand geometries and point charges were produced using Gaussian 03 at the Hartree–Fock level with 6-31G** basis functions and the restrained electrostatic potential (RESP) module of the AMBER package. All systems were solvated in orthorhombic water boxes with a minimum extension from the protein of 14 Å. The resulting systems contain approximately 40 thousand atoms. Hybrid topologies for use in the TIES protocol were created by combining the individual ligand parameters using the process described in Bhati *et al.* [4].

III. Computational Challenges

As the nature of scientific inquiry and the applications to support that inquiry evolve, there is a critical need to support the execution of scientific workflows on high-performance computing (HPC) infrastructures. This poses three main challenges: (1) scaling the distributed execution of workflows; (2) developing simple and usable workflow systems for HPC resources; and (3) implementing runtime adaptivity.

A. Scalability

Applications composed of multiple tasks with dependences that determine the order of their execution are referred to as 'workflows'. Often times, the structure of the task dependencies is simple and adheres to discernible patterns, even though the individual tasks and their duration are non-trivially distinct. Put together, it is a challenge to support the scalable execution of workflows on HPC resources due to the existing system and runtime software.

Currently, HPC software ecosystem mostly enables strong and weak scaling of applications composed by a single simulation that requires large amount of parallelism. This ecosystem has instead limited support for the concurrent execution of workflows, especially when composed of multiple heterogeneous tasks. Particularly limiting are the need to submit every task to a batch system with long queue waiting time and the limited amount of concurrency and flexibility offered by machine and architecture-specific tools that enable bulk submission.

Multiple workflow systems have emerged in response to this and other problems, each with its own strengths and unique capability but also with specific problems and challenges. In spite of the many successes of workflow systems, there is a perceived high barrier-to-entry, integration overhead and limited flexibility.

B. Simplicity and Usability

Many commonly used workflow systems emerged from an era when the support for distributed computing was fragile, missing features and services. Consistently, workflow systems had a monolithic design that included the end-to-end capabilities needed to execute workflows on heterogeneous and distributed cyberinfrastructures. Further, these workflow systems were typically designed to support large "big science" projects such as those at the LHC [15] or LIGO [16]. There, the same workflow was used by thousands of scientists over many years, justifying the large overhead of developing application workflows, and influencing programming models and interfaces.

However as the nature, number and usage of workflows has evolved so have the requirements: scale remains important but only when delivered with the ability to prototype quickly and flexibly. Further, new performance requirements arise from the need to support concurrent execution of tasks with diverse requirements and relations. For example, when executing multiple homogeneous pipelines of heterogeneous tasks, efficient resource utilization requires to ensure that individual pipelines have similar execution times, minimizing both pipeline-to-pipeline and task-to-task runtime fluctuations.

Together, these factors challenge workflow systems designed to mainly support specific use cases or 'locked-in' end-to-end executions. In the next Section, we discuss the design and implementation of the RADICAL-Cybertools, a set of software building blocks that can be easily composed to design, implement and execute domain specific workflows rapidly and at scale.

C. Adaptivity

Adaptive applications use intermediate data to enable better fidelity in the modeling of complex phenomena that would otherwise require unfeasible amount of computing time. Enabling adaptive capability on HPC systems poses specific challenges in expressibility, instantiation and implementation.

Adaptive applications requires to express both the application workflow and how it should adapt depending on data generated at runtime. The former requires the description of the application task graph, the latter methods to change this task graph. These methods are called at the end of the execution

of a task and, depending on its output, determine the generation of a new portion of the graph. Abort, rollback, proceed, forward and recovery are all examples of such methods [17].

Adaptation of the application at runtime depends on: (i) propagation of adapted task graph to all components; (ii) consistency of the state of task graph between different components; and (iii) minimal overhead. Minimizing overheads is particularly relevant as performing adaptive operations should be irrelevant when compared to the time required by the tasks execution.

IV. Solution

RADICAL Cybertools (RCT), developed by The RADICAL Lab, enables the efficient and dynamic execution of ensembles on heterogeneous computing resources. Different from other runtime systems, RCT decouples the workload execution and resource management details from the expression of the application, which significantly reduces the burden on the end user. RCT has been used extensively to support biomolecular sciences methods and algorithms, e.g., replica exchange, adaptive sampling and high throughput binding affinity calculations.

Here we describe High Throughput Binding Affinity Calculator (HTBAC), which builds upon the RADICAL Cybertools, as the framework solution to support the coordination of the required scale of computations, thereby allowing us to employ thousands of cores at a time.

Most benchmark evaluations of free energy protocols in the literature look at only a small number of systems, drawing inferences from tens or hundreds of runs. HTBAC facilitates studies on unprecedented scales, with the number of systems investigated an order of magnitude larger than published studies, which provides the opportunity to gain invaluable knowledge on the domain of applicability of current MD technologies. In particular, we demonstrate the use of HTBAC to compute the binding affinities of drugs to their target protein, using two simulation protocols with differing levels of rigor and computational cost, ESMACS and TIES.

HTBAC is not limited to these protocols as additional protocols can be expressed and implemented easily with the HTBAC user-facing API, however for demonstration we focus on these ensemble protocols HTBAC has demonstrated sizable execution and performance of the ESMACS [18] and TIES [19] protocols on leadership class machines including NCSA Blue Waters. For example, we demonstrated how HTBAC scales almost perfectly to hundreds of concurrent multi-stage pipelines for the TIES protocol in Fig. 2.

Both ESMACS and TIES have been successfully used to predict binding affinities quickly and accurately. In their standard non-adaptive forms TIES is approximately 2.5 times more expensive than ESMACS. The ligands investigated here are closely related so it could be expected that the greater accuracy of TIES would be required to differentiate them. However, in a drug design scenario many drug candidates would need to be investigated meaning that optimizing the execution time while retaining (or improving) accuracy is desirable. Given the very

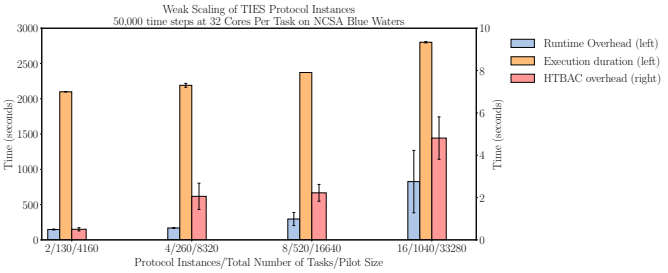

Fig. 2. Weak scaling properties of HTBAC. We investigate the weak scaling of HTBAC as the ratio of the number of protocol instances to resources is kept constant. Overheads of HTBAC framework (right), and RCT overhead (left) and total execution time TTX (left) for experimental configurations investigating the weak scaling of TIES. We ran two trials for each protocol instance configuration. Error bars in TTX in 2 and 8-protocol runs are insignificant.

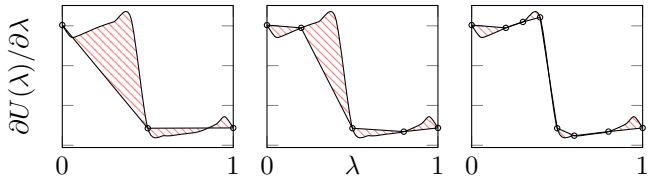

Fig. 3. Adaptive quadrature of the function $f(\lambda) = \partial U(\lambda)/\partial\lambda$ in the interval $[0, 1]$ using the trapezoidal rule. The figures from left to right show consecutive levels of recursion and bisection of intervals. The true integration error is the difference between the interpolated function and the actual function (shaded area). If the error on an interval is larger than a threshold, the interval is bisected.

large number of drug candidates, it is imperative to gain maximum insight into potential candidate compounds using time and resources efficiently. This provides one clear motivation for the use of adaptive methods which minimize the compute time used whilst producing binding free energy estimates meeting pre-defined quality criteria (such as convergence or statistical uncertainty below a given threshold).

Typically, datasets will involve ligands with a wide range of chemical properties which can impact not only the time to convergence, but the type of sampling required to gain accurate results. In general, there is no way to know before running calculations exactly which setup of calculation is required for a particular system. Even within closely related ligands, such as the THQ based BRD4 ligands studied here, the $\partial U/\partial\lambda$ curveλ integrated in TIES varies between physical systems (drugs). Consequently, the number (or the exact location) of the λ windows that will most impact the calculation are not known *a priori*, and change between systems. As multiple simulations must be run for each window, sampling with a very high frequency is expensive and impractical. Furthermore, adaptive placement of λ windows is likely to better capture the shape of the $\partial U/\partial\lambda$ curve, leading to more accurate and precise results for a given computational cost (see figure 3).

In order to support such investigations, in addition to being scalable, HTBAC is enhanced to support flexible resource reallocations schemes where resources can be moved between simulations run using different protocols or systems, for example, when one calculation has converged whilst another has

not. This adaptability makes it easier to manage complex programs where efficient use of resources is required in order to achieve a time to completion of studies comparable to those of high throughput chemistry.

The novel contributions of HTBAC are: (i) Unprecedented throughput: it allows the concurrent screening for drug binding affinities of multiple compounds at unprecedented scales, both in the number of candidates and resources utilized; (ii) Agile selection of different binding affinity protocols: HTBAC supports inter-protocol adaptivity, leading to resources being assigned at runtime to the "optimal" protocol (as determined by accuracy for given computational cost); (iii) Support for intra-protocol adaptivity, which provides the efficient execution of individual protocols.

A. Performance Metrics

TIES is rigorous but computationally expensive and has a limited range of applicability; ESMACS is approximate but can be employed across any set of ligands at lower computational cost. Both protocols are designed to simulate a large range of mutations. A single protocol instance represents a unique physical system. Each protocol instance contains a pipeline which maps to a sequence of simulations steps which include minimization, equilibration and production MD simulation. These simulation pipelines are replicated, where replicas differ only by their parameter configurations, namely initial velocities, which are randomly generated and assigned by the MD engine at the start of execution. ESMACS consists of 25 replicas i.e. 25 pipelines, while TIES consists of 13 λ windows, which are additional sampling parameters, and 5 replicas for a total of 65 pipelines. The additional pipelines in TIES contribute to the greater computational cost. Both protocols run for a total of 6 ns simulation durations. ESMACS produces 3.5 GB/system (24 MB/ns) while TIES produces 10 GB/system (24 MB/ns). Each simulation step in TIES and ESMACS requires 32 cores. Protocols run approximately 10-12 hours, depending on the physical system and the number of timesteps provided by the user.

There are several measures of performance that are relevant. The most pertinent is the weak scaling property which demonstrates the ability of HTBAC to solve large number of drug candidates in essentially the same amount of time (as the resources increase). To this effect, we investigated weak scaling behavior for screening sixteen drug candidates concurrently using thousands of multi-stage pipelines on more than 32,000 cores on NCSA Blue Waters (as shown in Figure 2). Similar scaling has been demonstrated on other platforms such as Titan for different protocols.

V. Impact of Solution

The flexibility provided by HTBAC to run adaptive workflows offers huge advantages scientifically. First, the intra-protocol adaptivity allows the automated optimization of calculations to ensure that results are obtained with known precision across systems which may exhibit very different behavior (for example levels of 'roughness' in the $\partial U/\partial\lambda$ curve in TIES). This has a significant impact of the reliability of comparisons between runs. The ability to switch between protocols on the other hand offers a mechanism through which 'cheaper' approximate methods (such as ESMACS) can be used to scan large regions of chemical space, whilst more accurate and 'expensive' ones are employed to investigate areas of specific interest (TIES). This maps directly onto processes such as hit to lead optimization in drug discovery and could be of particular use in investigating activity cliffs. This is a phenomena where small chemical changes provide large differences in drug efficacy. If changes are detected using an approximate method it is important to verify that they come from real chemical effects and not simply inaccuracies in the computational algorithm employed.

The scale enabled by HTBAC also has an impact on the potential for scientific discovery using free energy calculations. Most studies in the literature are limited to the investigation of tens of protein-ligand complexes. In order to establish the validity of particular combinations of forcefield and simulation protocol and quantify uncertainties much larger campaigns are needed. Our ensemble has already provided evidence that the variability in single runs is sufficient to swamp true differences between systems of interest. The combined need for both large numbers of systems and multiple repeats of each one produces a requirement for middleware to manage huge numbers of simulations.

HTBAC allows the concurrent screening for drug binding affinities of multiple compounds at unprecedented scales, both in the number of candidates and resources utilized. Specifically, we investigated weak scaling behavior for screening sixteen drug candidates concurrently using thousands of multi-stage pipelines on more than 32,000 cores on NCSA Blue Waters. This permits a rapid time-to-solution that is essentially invariant with respect to the calculation protocol, size of target system and number of ensemble simulations. In addition, HTBAC enabled the adaptive execution of the TIES protocol providing greater convergence (i.e., lower errors) for a given amount of computational resources.

These developments fit into a wider vision in which the use of flexible and responsive computational protocols produce accurate, precise and reproducible estimates of the free energy of binding with meaningful error bars. Not only would this allow for wider uptake of computational techniques in industrial settings but opens up possibilities of using these technologies in clinical decision support scenarios. By creating a 'digital twin', where the target protein is based on the real patients genetic sequence, a specific individuals response to different treatments could be predicted.

VI. Analysis of Solution

Our previous work deploying both ESMACS and TIES has typically involved comparing 10 to 20 computed binding affinities. In the original BRD4 inhibitor study, conducted in collaboration by UCL and GlaxoSmithKline [3], 16 drugs were investigated. All drugs were studied using ESMACS and 12 TIES transformations were performed. The non-adaptive

protocols required approximately 10k and 25k core hours per system for ESMACS and TIES respectively, for a total of $\sim$460k core hours for the whole study. Without HTBAC, each system was run by hand. This scale of study is only appropriate for retrospective analysis showing the potential of the methods involved. For *in silico* approaches to have a real impact in industrial scenarios much larger numbers of systems must be run, which is not practical without appropriate middleware. HTBAC satisfies this need by providing a logically programmable interface which facilitates the routine of running studies requiring sustained usage of millions of core hours per day.

The drug design process involves the filtering of millions of compounds to smaller number of 'hits' that bind the target protein and then further refinement to 'leads' that form the basis of candidate drugs. Encapsulation of the ESMACS and TIES protocols in HTBAC means that these protocols can be efficiently applied to the middle and end of this process. Adaptive functionality means that HTBAC can ensure efficient use of the resources depending on the properties of the ligands under investigation. This applies not only to the tuning of calculations (e.g. the use of adaptive quadrature) but also at a more strategic level where lower fidelity results could still be used to inform resource reallocation. For example, if we find that a particular ligand binds poorly compared to others in a study we could halt those simulations even if the size of the binding affinity difference had not fully converged. Careful design of decision points in TIES, based on simulation duration and λ window placement, could perhaps double the number of predictions for a given computational cost.

Additionally, the flexibility supported by HTBAC provides a platform to run the varied simulation campaigns needed to stress test protocols and forcefields in order to further refine our modeling approaches. HTBAC provides the functionality needed to address various computation requirements presented by drug screening campaigns, including facilitating the use of models in which accuracy can be traded for a reduction in computational cost; or inversely, by enabling tighter calculations with efficient resource utilization.

VII. Impact and Result

We demonstrate how scalable, accurate and rapid binding affinity calculation using HTBAC can enable effective clinical decision making by showing performance and scale of the number of drug candidates screened as a function of the number of core-hours. In addition, we show additional functionality of HTBAC to enable the adaptivity of intra-protocol execution, thereby providing greater convergence (i.e., lower errors) for a given amount of computational resources. As such, HTBAC advances binding affinity calculation to support scale and optimize for time-to-insight for investigative drug screening computational campaigns.

References

[1] Natalie H Theodoulou, Nicholas CO Tomkinson, Rab K Prinjha, and Philip G Humphreys. Clinical progress and pharmacology of small molecule bromodomain inhibitors. *Current Opinion in Chemical Biology*, 33(Supplement C):58 – 66, 2016. Chemical genetics and epigenetics * Molecular imaging.

[2] Paul Bamborough, Hawa Diallo, Jonathan D. Goodacre, Laurie Gordon, Antonia Lewis, Jonathan T. Seal, David M. Wilson, Michael D. Woodrow, and Chun-wa Chung. Fragment-based discovery of bromodomain inhibitors part 2: Optimization of phenylisoxazole sulfonamides. *Journal of Medicinal Chemistry*, 55(2):587–596, 2012. PMID: 22136469.

[3] S. Wan, A. P. Bhati, S. J. Zasada, I. Wall, D. Green, P. Bamborough, and P. V. Coveney. Rapid and reliable binding affinity prediction of bromodomain inhibitors: a computational study. *J. Chem. Theory Comput.*, 13(2):784–795, 2017.

[4] A. P. Bhati, S. Wan, D. W. Wright, and P. V. Coveney. Rapid, accurate, precise and reliable relative free energy prediction using ensemble based thermodynamic integration. *J. Chem. Theory Comput.*, 13(1):210–222, 2017.

[5] David L. Mobley Michael R. Shirts and John D. Chodera. Alchemical free energy calculations: Ready for prime time? *Annu. Rep. Comput. Chem.*, 3:41–59, 2007.

[6] Lingle Wang, Yujie Wu, Yuqing Deng, Byungchan Kim, Levi Pierce, Goran Krilov, Dmitry Lupyan, Shaughnessy Robinson, Markus K. Dahlgren, Jeremy Greenwood, Donna L. Romero, Craig Masse, Jennifer L. Knight, Thomas Steinbrecher, Thijs Beuming, Wolfgang Damm, Ed Harder, Woody Sherman, Mark Brewer, Ron Wester, Mark Murcko, Leah Frye, Ramy Farid, Teng Lin, David L. Mobley, William L. Jorgensen, Bruce J. Berne, Richard A. Friesner, and Rober Abel. 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.*, 137:2695–2703, 2015.

[7] Michael R. Shirts, David L. Mobley, and Scott P. Brown. Free energy calculations in structure-based drug design. In *Structure Based Drug Design*. Cambridge University Press, 2009.

[8] P. V. Coveney and S. Wan. On the calculation of equilibrium thermodynamic properties from molecular dynamics. *Phys. Chem. Chem. Phys.*, 18:30236–30240, 2016.

[9] D. W. Wright, B. A. Hall, O. A. Kenway, S. Jha, and P. V. Coveney. Computing clinically relevant binding free energies of HIV-1 protease inhibitors. *J. Chem. Theory Comput.*, 10(3):1228–1241, 2014.

[10] S. Wan, B. Knapp, D. W. Wright, C. M. Deane, and P. V. Coveney. Rapid, precise, and reproducible prediction of peptide–MHC binding affinities from molecular dynamics that correlate well with experiment. *J. Chem. Theory Comput.*, 11(7):3346–3356, 2015.

[11] S. Wan and P. V. Coveney. Rapid and accurate ranking of binding affinities of epidermal growth factor receptor sequences with selected lung cancer drugs. *J. R. Soc. Interface*, 8(61):1114–1127, 2011.

[12] I. Massova and P.A. Kollman. Computational alanine scanning to probe protein-protein interactions: A novel approach to evaluate binding free energies. *J. Am. Chem. Soc.*, 121(36):8133–8143, 1999.

[13] S. K. Sadiq, D. W. Wright, S. J. Watson, S. J. Zasada, I. Stoica, and P.V. Coveney. Automated Molecular Simulation Based Binding Affinity Calculator for Ligand-Bound HIV-1 Proteases. *J. Chem. Inf. Model.*, 48(9):1909–1919, 2008.

[14] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, and D. A. Case. Development and testing of a general Amber force field. *J. Comput. Chem.*, 25(9):1157–1174, 2004.

[15] Amos Breskin and Rüdiger Voss. *The CERN large hadron collider: accelerator and experiments*. Cern Geneva, 2009.

[16] William E Althouse, Michael E Zucker, et al. LIGO: The laser interferometer gravitational-wave observatory. *Science*, 256(5055):325, 1992.

[17] Wil MP van der Aalst and Stefan Jablonski. Dealing with workflow change: identification of issues and solutions. *Computer systems science and engineering*, 15(5):267–276, 2000.

[18] Jumana Dakka, Matteo Turilli, David W. Wright, Stefan J. Zasada, Vivek Balasubramanian, Shunzhou Wan, Peter V. Coveney, and Shantenu Jha. High-throughput binding affinity calculations at extreme scales. *Computational Approaches for Cancer Workshop at SuperComputing17*, 2017.

[19] Jumana Dakka, Kristof Farkas-Pall, Vivek Balasubramanian, Matteo Turilli, David W Wright, Shunzhou Wan, Stefan Zasada, Peter V Coveney, and Shantenu Jha. Rapid, concurrent and adaptive extreme scale binding free energy calculation. 2017. https://arxiv.org/abs/1801.01174.