



HHS Public Access

Author manuscript

Bioorg Med Chem. Author manuscript; available in PMC 2017 October 15.

Published in final edited form as:

Bioorg Med Chem. 2016 October 15; 24(20): 4928–4935. doi:10.1016/j.bmc.2016.07.064.

Ensemble-based docking: From hit discovery to metabolism and toxicity predictions

Wilfredo Evangelista^{a,b}, Rebecca L. Weir^{a,b}, Sally R. Ellingson^{b,c,d,†}, Jason B. Harris^{b,c,‡,§},
Karan Kapoor^{b,c,¶}, Jeremy C. Smith^{a,b,c,*}, and Jerome Baudry^{a,b,c,*}

^aDepartment of Biochemistry and Cellular and Molecular Biology, The University of Tennessee, Knoxville, TN, United States

^bUT/ORNL Center for Molecular Biophysics, Oak Ridge National Laboratory, Oak Ridge, TN, United States

^cGenome Science and Technology, University of Tennessee, Knoxville, TN, United States

^dDivision of Biomedical Informatics, College of Medicine, University of Kentucky and Cancer Research Informatics, Markey Cancer Center, Lexington, KY, United States

Abstract

This paper describes and illustrates the use of ensemble-based docking, i.e., using a collection of protein structures in docking calculations for hit discovery, the exploration of biochemical pathways and toxicity prediction of drug candidates. We describe the computational engineering work necessary to enable large ensemble docking campaigns on supercomputers. We show examples where ensemble-based docking has significantly increased the number and the diversity of validated drug candidates. Finally, we illustrate how ensemble-based docking can be extended beyond hit discovery and toward providing a structural basis for the prediction of metabolism and off-target binding relevant to pre-clinical and clinical trials.

Keywords

Drug discovery; Docking; Computational drug discovery; Toxicity; Hit discovery; Lead discovery

1. Introduction

A popular approach to develop drug candidates that are potent and efficient is to rely on structure-based drug discovery, i.e., knowledge of the structure of a protein target, to identify small molecules that possess the desired chemical and structural properties needed to bind to the protein of interest. Docking belongs to a class of structure-based virtual screening approaches and are used to complement and accelerate experimental drug discovery screening campaigns.¹

*Corresponding authors. smithjc@ornl.gov (J.C. Smith), jbaudry@utk.edu (J. Baudry).

†Current address: University of Kentucky, Lexington, KY, United States.

‡Current address: National Center for Computational Toxicology—Research Triangle Park, NC, United States.

§Current address: University of North Carolina, Chapel-Hill, NC, United States.

¶Current address: Georgetown University, Washington, DC, United States.

Docking calculations essentially predict how well a given chemical may bind to a given protein structure, and involve computer programs that perform the following tasks: (i) position a small molecule drug candidate in the (predicted or experimentally known) binding site of the target, (ii) evaluate an interaction energy between the small molecule and its protein environment—varying from an enthalpy-like interaction energy to more accurate and sophisticated binding free energy and (iii) identify those chemicals that are predicted to bind the strongest.¹

In recent efforts to reduce toxicity of drug candidate, attention has been focused on not only predicting binding of drug candidates to the target but also on predicting off-target binding, as toxicity is often the result of off-target binding. Understanding drug discovery at this system-level terms implies that the ideal hit should bind only to certain proteins (to the target, or potentially to other proteins turning a pro-drug into a drug), and to become a solid drug candidate, the initial hit should also not bind to other off-target proteins (to avoid toxicity issues). This is a applicable problem for docking approaches to be used since they are able to sample not only multiple ligands but also multiple protein targets as we will show below.

Virtual screening traditionally follows an *induced fit mechanism* for ligand binding: the chemical to be assayed for its binding energy in a protein is inserted in the protein binding site, and potential structural modifications of the protein *following the binding of the ligand* are evaluated by allowing some degree of flexibility to the protein side chains around the ligand, and sometimes, but more rarely, extending this flexibility to the protein backbone as well such as in the case of MM-PBSA of free energy perturbation calculations.^{2–6} Such a protein-flexibility approach, while more realistic than keeping the protein rigid, requires long calculations that make this approach seldom used in large screening campaigns and reserved to the a detailed binding investigation of a small number of ligands, such as in lead optimization approaches. Alternatively, to quickly describe the dynamical flexibility of the protein receptor around docked ligands, docking scoring functions have been developed that are parametrized to use mostly rigid protein structures and to reproduce experimentally-known binding energies of ligands in as many proteins as possible. These rigid protein approaches usually use soft non-bonded interaction terms that limit steric clashes between a ligand and its protein environment and implicitly represent the overall effect of protein flexibility upon ligand binding.

With increasing computational power becoming available, there have recently been an increasing number of reports that aim at simulating the dynamics of the apo-protein targets, and perform docking in conformers thus sampled, as conceptualized by Lin and co-workers,⁷ and as recently demonstrated in an integrated computational/experimental landscape study.⁸ This *ensemble-based* approach aims at reproducing a *conformational selection mechanism*, where the protein-bound structure is sampled *prior to ligand binding*, and specific conformations are selected by the ligand(s) to form a thermodynamically favored protein:ligand complex of lower global free energy than that of other potential protein:ligand complexes. Technically, this conformational search does not preclude later small-scale protein rearrangements in response to ligand binding, but in practice, the latter,

local induced fit is often omitted in ensemble docking and the same soft scoring functions used.

We describe here the collaborative contributions of our laboratories in developing computational techniques for ensemble-based (multiple proteins and multiple protein structures) docking, and in the applications of these techniques for hit discovery and for pathway exploration, and we present original results toward predicting the behavior of drug candidates in pre-clinical and clinical trials. Our work ranges from relatively small scale approaches to large-scale, supercomputing-supported, ensemble-based approaches that involve several target protein structures and large chemical databases of drug candidates.

2. Computational methods

2.1. Virtual screening programs

There exists a large number of programs developed for docking, many with well-documented histories of successful application. These programs originate from both academic and commercial laboratories, exemplifying the commercial importance of computational approaches to drug discovery in the pharmaceutical industry. A review of some of these programs and of their respective strengths and challenges has been given elsewhere.^{9,10}

We have used several of these packages with success. However, in this paper we present results obtained using the commercial program MOE (Molecular Operating Environment, Chemical Computing Group, Inc., Montreal, Canada), and the academic program Autodock Vina (A.J. Olson laboratory, The Scripps Research Institute, San Diego, California).¹¹ We used Autodock Vina both in its native distribution and in a parallelized version, called VinaMPI, developed by our laboratories¹² and described below. The MOE program was used in the metabolic activation estrogenization pathway project described here, the Autodock Vina was used in the hit discovery project on modulators of coagulation, and the VinaMPI program was used in the toxicity prediction project.

2.2. Computational resources

The computational resources required to perform docking calculations vary with the scope of a specific screening campaign. While more computational power is always desirable, calculations screening a relatively small number of chemicals, up to a few hundred, on a few structures of a protein target can be achieved in a reasonable time on a modern desktop computer with a few CPU cores and about 200 Gb of hard drive space. We report below one such project that, while ambitious in scope, required relatively low-scale resources to provide a proof of concept in docking applied to multi-protein pathways. In larger, or much larger, screening campaigns, and in particular in the development of future toxicity/potency prediction of drug candidates, considerably more powerful computational resources are required to handle

- i. the sizes of the databases of chemicals to be screened,
- ii. the number of protein structures to be used in ensemble-based docking and,

- iii. the associated large storage and data processing requirements.

Here, we also present work done in our laboratories that used virtual screening approaches on the world's most powerful supercomputers. Our original parallelization of the AutodockVina program was developed on the (now decommissioned) Kraken machine, then the world's most powerful academic supercomputer, and operated by the University of Tennessee, Knoxville, Tennessee. The application research projects were performed on the (again now decommissioned) Jaguar and currently Titan¹³ supercomputers. Jaguar and Titan were, and are, respectively, the most powerful open-science supercomputers in the USA, both operated by the Oak Ridge National Laboratory, Oak Ridge, Tennessee.

2.3. Protein conformer generation

In the results presented below, we have used molecular dynamics (MD) simulations to generate protein conformers on which we performed ensemble-based docking. An MD trajectory is divided into clusters that span the conformers sampled during the MD. In the seminal ensemble-based work of Amaro, McCammon and coworkers, MD simulations of 20 ns were used to sample protein conformations, and these conformations were used to dock ~1800 compounds.¹⁴ In our larger screening campaigns, we have used MD simulations ranging from several hundreds of ns to the microsecond timescale. The MD simulations in our project were performed with the NAMD2 program¹⁵ for atomistic MD simulations, and the Gromacs v.5.0.1^{16,17} and Martini v.2.0 force field^{18,19} for Coarse Grained (CG) MD simulations.

3. Results

3.1. Ensemble-based approaches and computational engineering

In this section we review our work on enabling efficient docking approaches on supercomputers. The primary benefit of using supercomputers is, of course, to be able to run many more docking and MD calculations than on smaller architectures. The docking enables very large, sometimes massive, databases of chemicals to be considered as potential drug candidates, increasing the chemical diversity of the chemicals considered as potential ligands for the targets of interest.

The second reason why being able to run large docking jobs is desirable is that it enables the simulation of the conformational selection mechanism. As illustrated below, using more than one single structure of the protein target very significantly increases the quantity and the diversity of molecules experimentally verified to bind to the protein, when compared to what would be possible using only a single structure of the target. The third reason why using more than one protein structure is desirable is to characterize binding of a drug candidate to other proteins than the target to determine off-target effects i.e., metabolism and toxicity.

The above goals, i.e., (i) increasing the chemical diversity of drug candidates, (ii) using a realistic conformational selection mechanism for docking, and (iii) investigating if chemicals may bind to several different proteins, require increasing computational power as the goals of a screening campaign become more ambitious in scope. The computational power of modern supercomputers in the petaFLOPS (10^{15} floating point operations per second) range

renders possible, in principle, hundreds of thousands to millions of docking calculations per day, i.e., enough to dock very large databases of chemicals against several proteins.

However, having very powerful supercomputers is not enough. To leverage the power of these machines, efficient software is needed. Docking is, in principle, an ‘embarrassingly parallel’ process: docking of a given chemical to a given protein structure is not dependent on the result of the docking of another chemical to the same (or another) protein structure. There is therefore a perception that existing docking engines are, out of the box, efficient enough to be implemented on supercomputers and that ensemble-based docking calculations can be handled by a series of docking instructions in batch. However, major issues in data handling and I/O, as well as realities of queue management policies, render using supercomputers for docking dependent on sophisticated software engineering strategies. These programs have to handle

- i. the generation of the very large amount of data needed in ensemble docking,
- ii. the distribution of the data and instructions to the computing cores performing the docking calculations (or nodes in the case of threaded docking calculations) and retrieval of results, and
- iii. the prioritization and management of the many tasks to be performed.

To enable docking on supercomputers we have developed the program VinaMPI,¹² an MPI parallelization of the AutodockVina program,¹¹ that scales on tens of thousands of supercomputer cores. This parallelization addresses several key points needed for efficient docking on supercomputers: the handling of an even distribution of work by the computing cores, the handling of communication between cores—including handling of failing cores, and, importantly, the inclusion of the complexity of a given docking calculation in the prioritization and distribution of computing tasks. The complexity of a docking calculation is based on the flexibility of the chemical to be docked: large, flexible chemicals with a high number of rotatable bonds (typically more than 10), and hence requesting a large number of conformers to be generated and docked, are prioritized for docking before smaller, more rigid chemicals. This way, computing nodes are always busy, and it is possible to avoid a situation where jobs finish quickly on cores handling small chemicals and ending up being idle while longer jobs are still running, as illustrated on Figure 1. This load balancing approach is central in our parallelization strategy and enables an efficient use of computing resources.

Importantly, in order to be able to submit a small number of jobs to queues, the handling of several protein structures to be used in ensemble docking is also implemented in VinaMPI. This is a key point in developing massive ensemble-based calculations, as including more than one protein structure in the docking process can lead to I/O issues that essentially prevent the calculations from going forward. The program VinaMPI enables us to perform ~4 million docking calculations in about two hours wall-clock time on the Titan supercomputer, involving several hundreds of thousands of chemicals in several protein structures.²⁰ The program VinaMPI is available freely to the community and can be used on any computing architecture that has policies for submission and execution of batch jobs.

3.2. Ensemble-based approaches to hit discovery²¹ (and references therein)

3.2.1. Based on Kapoor et al., and references therein—In this section we review our results on ensemble-based hit discovery. This approach illustrates the most common application of docking, in which a protein is used in a structure-based approach to computationally identify the molecules from a collection predicted to have the most favorable binding energies, and hence to bind more strongly to the target. This is an approach used in the initial steps of drug discovery, where new hits/leads are needed. In silico screening thus helps to prioritize chemicals for experimental assays, as successfully demonstrated by the Jorgensen lab on many projects.²²

In one project, we aimed at discovering compounds capable of modulating the interactions between the Factor Xa (FXa) and Factor Va (FVa) proteins of the coagulation cascade. Pharmaceuticals exist that inhibit the enzymatic functions of FXa, but they have a very narrow safety profile, and are difficult to use even in hospital settings. The goal of this project was to discover novel molecules that can bind to the surface of FXa and modulate its interactions with FVa, without affecting FXa's enzymatic properties. Our strategy was to identify binding pockets on the surface of FXa, in regions that are hypothesized to interact with its FVa protein partner. A 300 ns MD simulation trajectory was produced for FXa using the NAMD2 engine, and the protein structures in the trajectory were clustered using nearest neighbor RMSD clustering to identify 11 snapshots most representative of the conformations sampled by the protein during the simulation. Potential binding pockets not present on the crystal structure were observed to be formed in some of these structures on the surface of the protein.

A total of 281,128 compounds, originating mostly from the 'clean drug-like' subset of the ZINC database, were used in the docking calculations in 12 FXa protein structures: the 11 structures obtained from the clustering of the MD trajectory, and the crystal structure of the FXa target. On each of these 12 targets, several potential binding sites were investigated 'agnostically' (i.e., without *a priori* favoring one potential binding site over another), resulting in about 3.4 million docking calculations performed using Autodock Vina. The results from docking on the FXa-FVa binding interface indicated that compounds can bind in different sub-binding sites at this interface in different protein MD snapshots. The FXa-FVa interface was divided into ten sub-binding sites based on the clustering of the binding loci for these compounds. The results of this large docking campaign are summarized in Figure 2.

The compounds predicted to bind on the protein:protein interaction surface were ranked based on their calculated docking scores/predicted binding free energies, and their ability to successfully bind to multiple protein conformations while showing selectivity for the binding sites at the FXa-FVa interface. Based on these criteria a total of 535 compounds were considered for experimental validation of the docking calculations. After employing further filters based on compound availability, cost, chemical similarity and the presence of reactive functional groups, 124 compounds (97 compounds originating from the docking calculations and 27 compounds chemically similar to docking hits) were ordered for experimental validation by our collaborators. Experimental work on these compounds, using

a prothrombinase activity assay measuring inhibition of the rate of generation of active thrombin and direct FXa inhibition using FXa chromogenic peptide substrate identified 31 molecules (out of the 124 tested) as non-active site inhibitors, and 10 compounds (out of these 31 inhibitors) exhibited the desired properties of achieving dose-independent partial inhibition of PTase activity in a non-active site-dependent and self-limiting mechanism. Importantly, out of these 10 compounds, 8 would not have been predicted to bind from docking calculations using only the crystal structure, i.e., these compounds scored well in docking calculations in some of the binding sites that were formed only during the molecular dynamics simulation.

The above results illustrate two important aspects of a successful docking campaign for hit discovery: the first aspect is a confirmation that successful docking can be a very useful tool to prioritize compounds for experimental validation. Here, out of 97 compounds suggested from docking to bind to the target protein, 31 were experimentally verified to possess a biological activity related to the regulation of the protein, and 10 compounds exhibited other, further desired biological properties. This is not a direct validation that the compounds of interest do indeed bind to the protein. To do this, target engagement experiments would have to be done. Notwithstanding, the correlation between the docking results and the experimentally-measured biological activity renders likely that these compounds do bind to the protein target with the suggested mechanisms.

The second lesson of this successful docking campaign is that using an ensemble-based approach has led to a 400% increase in the number of unique compounds with desired properties compared to what would have been discovered using only the crystal structure of the protein target. While computationally more expensive than docking to a single target structure, ensemble docking can indeed lead to a large improvement in the number and diversity of true positives that can be prioritized for experimental validation.

3.3. Ensemble-based approaches to phase-I metabolism²³ (and references therein)

In this section we review our results on docking-based characterization of pro-drug oxidation by P450 cytochrome oxidases and the effect this metabolic activation has on downstream binding to the estrogen receptor. This work is relevant to ensemble docking, albeit on the small scale, in that it involves different protein species, rather than different conformations of the same target. The goal of the work was to predict possible metabolic products of the environmental xenobiotic pollutant PCB-30 (2,4,6-trichlorobiphenyl) after exposure to P450s, and the endocrine disruptive potential of the oxidized metabolic products. PCB-30 is not a pharmaceutical, but it undergoes the same oxidation process that a pharmaceutical would experience in phase-I metabolism, and the metabolites acquire novel biological properties the same way a pro-drug would become bioactive (see Fig. 3).

The parent molecule PCB-30 was first docked to two P450 species: CYP2D6 and CYP3A4, that are the most common forms of cytochrome P450 in humans and responsible for metabolizing about half of known drugs. Primary docking calculations of PCB-30 in P450 models containing the heme's iron-oxo reactive intermediate, as per the P450 model of Baudry,²⁴ were performed with the program MOE using flexible side-chain approaches. PCB-30 can bind in different orientations in active sites of the two P450s, and hence

different carbon atoms can be oxidized. We predicted the most likely metabolites by identifying the carbon atoms of PCB-30 positioned closest to the iron-oxo oxygen atom in the best scoring docked poses. In addition, the predicted reactivity of each carbon was taken into account as per calculated by the SMARTCYP approach as described in 23. For each P450, carbon atoms number 4 and, slightly less favorably, number 3 were predicted to bind close to the reactive iron-oxo's oxygen atom. This suggests that both enzymes will preferably oxidize the parent molecule at the 4 or 3 locations, leading to the 4-OH and 3-OH PCB-30 species. In contrast, the binding modes of PCB-30 involving preferred oxidation at the 2 and 3' locations were found to be only marginally possible.

In secondary docking calculations, the 4-OH, 3-OH and 3'-OH metabolites of PCB-30 were docked to the human alpha estrogen receptor ligand binding domain, again using the MOE docking facility. The 4-OH PCB-30 metabolite was predicted to have a docking binding energy ~3 kcal/mol more favorable than the 3-OH and 3'-OH species, and 4 kcal/mol more favorable than that of the parent molecule, PCB-30.

The above results predict that the 4-OH PCB-30 will be (i) the primary metabolite of PCB-30 upon P450-mediated oxidation, and that (ii) the 4-OH metabolite will bind to the estrogen receptor and exhibit the most estrogenic activity. A following experimental validation was performed by our collaborators as follows: Estrogenic activity of parent chemicals and metabolites were assessed using EC50 values from a bioluminescent yeast assay expressing the human estrogen receptor with a lux reporter system. Metabolites were generated as microsomal reaction mixtures (MRMs) using total and enriched P450 extracts for each specific CYP450 (3A4 and 2D6). GC/MS analysis with purified reference compounds was used to confirm presence of specific primary metabolites (3- and 4-hydroxy species) in each MRM. Experiments confirmed that indeed 4-OH was present as the primary metabolite of PCB-30 from the action of the two P450s, with 3-OH-PCB 30 present as secondary metabolite in lower quantities, and that 4-OH PCB30 exhibits a clear increase in estrogenic action over that of the parent molecule. This case study multi-proteins, ensemble-based docking illustrates the potential use of docking to explore biochemical pathways, such as the bioactivation of the estrogen receptor through phase-I metabolism, and serves as a proof of concept of the potential of docking in the age of system biology.

3.4. Ensemble-based toxicity prediction: prediction off-binding and toward virtual clinical trials

In this section we present original results on *ensemble-based toxicity prediction*, a new direction of our laboratories. Off-target binding is a major reason behind the concerning failure of such a high fraction of drug candidates (up to 90% overall) in pre-clinical and clinical trials.²⁵ When a drug candidate is exposed to proteins beyond its intended target, the drug candidate may—and often does—bind to other proteins in a promiscuous way, diluting its effect to the point of becoming no more active than placebo. This abolishes efficacy. Moreover, a drug candidate may bind strongly to another protein, disrupting, for example, a biochemical pathway not related to that of the diseased state against which the candidate was developed, to the point that that this drug candidate becomes toxic. While binding off target may sometimes be desirable, such as in the case of drug repurposing,²⁰ it remains a major

concern and a major source of failures of drug development programs. Hence, being able to predict the likelihood of a given molecule of engaging in off-target shenanigans is of paramount importance to the entire field of drug discovery and in the developing field of polypharmacology.²⁰ Several predictive statistical and cheminformatics-based approaches have been proposed that can suggest ADME(T) (Adsorption, Distribution, Metabolism, Excretion, Toxicity) properties of drug candidates, and recently a large cheminformatics-based strategy has been used to predict off-target binding for 656 drugs, with a success rate of about 50% in terms of experimental validation of the predictions.²⁶

Just as docking can be used to discover new ligands binding to a given target and to predict binding of substrates and products to enzymes in a biochemical pathway, docking approaches can also be used to predict off-target binding of drug candidates. Following the pioneering work of other groups in this field,^{27–30} we are developing ensemble-based docking as a predictive technology platform to red-flag drug candidates for their potential to bind off-target to proteins known to be involved in most clinical failures. This approach is different from cheminformatics/ligand-based predictions in that it requires no parameters other than those used in the parametrization docking scoring functions, and is hence amenable to use in the case of novel molecules for which little chemical similarity to known drugs is available.

A list of 44 proteins has been proposed to be used in safety panels by four drug discovery companies (Astra Zeneca, GlaxoSmithK-line, Novartis and Pfizer), based on a consensus on which proteins are most often associated with off-target binding and clinical failures.³¹ Out of these proteins 35 (80% of the total) are membrane proteins, seven are enzymes and two are nuclear receptors. Of the 35 membrane proteins 24 are GPCRs, 3 are transporters, and 7 are ion channels. 25 of the membrane proteins have a (sometimes incomplete) structure deposited in the Protein data bank.

We have focused a first set of calculations on one of these 44 proteins: the human adenosine A2A receptor, a GPCR with a (relatively) short extra membrane loop and available structures amenable to the building of a protein model. This protein is found in CNS tissues, and its function is to regulate pain, cerebral blood flow, basal ganglia functions, respiration and sleep. Adverse side effects due to off-target drug-induced activation are pain, asthma, seizures and other neurological disorders).

Lists of ligands and decoys are available in the DUD-E database for the protein. The PDB entries 3QAK and 2YDV, two X-ray crystallographic structure resolved at 2.6 Å and 2.71 Å resolution, respectively, and both co-crystallized with an agonist molecule, were used and combined to build a complete model of the protein, by using segments that were present in one structure and not in the other one and in particular to obtain a complete structure for extracellular loops. The protein was included in a membrane model, and explicit solvation and counter-ions were added (see Fig. 4). A 1 μs microsecond coarse-grained molecular dynamics simulation of the model using the Gromacs v5.0.1 program and Martini v.2.0 force field was performed on the Moldyn cluster at the Center for Molecular Biophysics and the Newton High Performance cluster at the University of Tennessee. The coarse-grained

trajectory was mapped back to an atomistic model using python and shell scripts from www.cg.martini.nl and adapted for this case.

A cluster analysis was performed on the trajectory and 32 snapshots were selected using a 2.0 Å cut-off for the pairwise RMSD value of all MD frames based on the entire protein structure, to represent the range of conformations sampled by the protein during the molecular dynamics simulation. A set of 11,733 molecules were docked in each of the 32 snapshots and in the crystal structure of the protein using our VinaMPI program. The docking box included the entire channel and cytoplasmic and extra-cytoplasmic proteins regions outside of the membrane. This set of 11,733 molecules was obtained from the DUD-E database entry corresponding to this protein, and contains 844 known ligands, and 10,889 molecules chemically similar to the 844 ligands and labeled as 'decoys'. These 11,733 molecules were then ranked according to their calculated docking scores. In the case of the ensemble-docking using the 32 snapshots, all the docking results were combined together and only the top-score of any given molecule of all 32 scores (one for each conformers) was used in the ranking, following the protocol of Ellingson.²⁰

The known ligands represent only ~7% of the compounds in the database used in the docking. When using the crystal structure to perform the docking, only 58 known ligands are ranked amongst the 10% top-scoring molecules (i.e., amongst the top 1174 compounds), which is close to random. However, when using the ensemble docking approach, a total of 343 unique known ligands are ranked amongst the 10% top-scoring molecules. Assuming that all decoys indeed do not bind to the protein, about a third of the top 1174 molecules, if assayed experimentally, would have been identified as ligands. Using an ensemble-based strategy has hence increased the rate of 'known ligand' prediction by ~600% with respect to what can be achieved using only the crystal structure, comparable to the FXa:FVa case described above. About 40% of the known ligands would have hence been correctly predicted to be capable of binding to this protein target. Had these 844 chemicals been drug candidates developed against another protein target, off-binding to adenosine A2A receptor would have correctly been predicted for nearly half of them.

Figure 5 shows that most of the correctly-identified ligands select essentially 5 protein conformers, which corresponds to about one sixth of the representative protein conformations. In the present case, it was already known what compounds were 'ligands' and what compounds were 'decoys', and we could identify *a posteriori* the MD snapshots most-often selected by validated ligands. In contrast, in the case of novel proteins and/or of novel compounds docked to a protein, one does not know *a priori* which of the protein conformers will be selected. There is a great interest in identifying protein conformers selected by ligands for binding. We have proposed an approach based on pooling together docking scores obtained in different snapshots, as used in the present case, and selecting the best scores for each compound amongst all the protein conformers.²⁰ Other approaches are being developed and proposed.³² This will undoubtedly continue to be of importance as ensemble-based docking approaches continue to be implemented.

4. Discussion and conclusions

Together, the above results illustrate the potential of ensemble-based docking approaches for increasing the quality and the scope of structure-based drug discovery. Simulating in silico a conformational selection mechanism is much more computationally expensive than using a single structure, but contemporary computer power allows such an approach, especially when bringing massively-parallel supercomputer power into play. Multiple protein-based docking can also be useful in a more traditional induced-fit approach on more modest computational resources to explore limited biochemical pathways. The latter approach, when combined with conformational selection and supercomputing, will certainly play an increasing role in obtaining a predictive understanding of metabolomics, in toxicity prediction and in providing a structural rationale for a systems approach to the biochemical functioning of cells. With the large quantity of data being generated, it will become vital to continue to develop data-mining techniques to process the docking results obtained on several conformers of a protein, and in particular to identify—if this is at all possible—what protein conformations will be selected by what chemicals.

Evolution of the methodology outlined here can be envisaged, and should take into account the inexorable march towards the exascale in supercomputing power. The use of massively-parallel supercomputing power to exhaustively search protein conformational space using MD may be contemplated, and this, then could be combined with Markov-State modeling to form a complete thermodynamic description of the metastable state space of apo-proteins.³³ Within the confines of the conformational selection approach, this would need to be performed only once per protein. A library of accessible conformations would then be established, to which large databases of drugs could be docked and the results parsed to determine both target binding and off-target effects.

Acknowledgments

The results presented here describe the work performed computationally by laboratories in the UT/ORNL Center for Molecular Biophysics. We express our profound gratitude to our experimental collaborators (co-authors on the respective publications) who have performed the experimental work on two of these projects: the Center for Environmental Biotechnology at the University of Tennessee for the PCB-30 estrogenization pathway project, and Shifa Biomedical and the laboratory of Dr. Cynthia Peterson at the University of Tennessee for the FXa:FVa project. This work was financially supported by the University of Tennessee Governor Chair program for J.C.S. and by a start up grant from the Department of Biochemistry and Cellular and Molecular Biology to J.B.; as well as by the National Institutes of Health 1R43HL11426 for the FXa:FVa project to J.B., and by the National Institutes of Health 1KL2RR031974 and the Department of Energy for the development of VinaMPI to J.B. and J.C.S.

References and notes

1. Meng X-Y, Zhang H-X, Mezei M, Cui M. *Curr Comput Aided Drug Des.* 2011; 7:146. <http://dx.doi.org/10.1038/nature13314.A>. [PubMed: 21534921]
2. Chipot, C.; Pohorille, A. *Free Energy Calculations.* 2007. <http://dx.doi.org/10.1007/978-3-540-38448-9>
3. Limongelli V, Bonomi M, Parrinello M. *Proc Natl Acad Sci.* 2013; 110:6358. <http://dx.doi.org/10.1073/pnas.1303186110>. [PubMed: 23553839]
4. Wang L, Wu Y, Deng Y, Kim B, Pierce L, Krilov G, Lupyan D, Robinson S, Dahlgren MK, Greenwood J, Romero DL, Masse C, Knight JL, Steinbrecher T, Beuming T, Damm W, Harder E, Sherman W, Brewer M, Wester R, Murcko M, Frye L, Farid R, Lin T, Mobley DL, Jorgensen WL,

- Berne BJ, Friesner RA, Abel R. *J Am Chem Soc.* 2015; 137:2695. <http://dx.doi.org/10.1021/ja512751q>. [PubMed: 25625324]
5. Whalen KL, Chang KM, Spies MA. *Mol Inf.* 2011; 30:459. <http://dx.doi.org/10.1002/minf.201100014>.
6. Whalen KL, Chau AC, Spies M. *A Chem Med Chem.* 2013; 8:1681. <http://dx.doi.org/10.1002/cmcd.201300271>. [PubMed: 23929705]
7. Lin JH, Perryman AL, Schames JR, McCammon JA. *J Am Chem Soc.* 2002; 124:5632. <http://dx.doi.org/10.1021/ja0260162>. [PubMed: 12010024]
8. Fischer M, Coleman RG, Fraser JS, Shoichet BK. *Nat Chem.* 2014; 6:575. <http://dx.doi.org/10.1038/nchem.1954>. [PubMed: 24950326]
9. Warren GL, Andrews CW, Capelli AM, Clarke B, LaLonde J, Lambert MH, Lindvall M, Nevins N, Semus SF, Senger S, Tedesco G, Wall ID, Woolven JM, Peishoff CE, Head MS. *J Med Chem.* 2006; 49:5912. <http://dx.doi.org/10.1021/jm050362n>. [PubMed: 17004707]
10. Moitessier N, Englebienne P, Lee D, Lawandi J, Corbeil CR. *Br J Pharmacol.* 2008; 153:S7. <http://dx.doi.org/10.1038/sj.bjp.0707515>. [PubMed: 18037925]
11. Trott O, Olson AJ. *J Comput Chem.* 2010; 31:455. <http://dx.doi.org/10.1002/jcc.19499576>
12. Ellingson SR, Smith JC, Baudry J. *J Comput Chem.* 2013; 34:2212. <http://dx.doi.org/10.1002/jcc.23367>. [PubMed: 23813626]
13. Oak Ridge National Laboratory. Titan, US Dep Energy. 2012. <https://www.olcf.ornl.gov/titan/>
<https://www.olcf.ornl.gov/titan/>
14. Amaro RE, Schnaufer A, Interthal H, Hol W, Stuart KD, McCammon JA. *Proc Natl Acad Sci USA.* 2008; 105:17278. <http://dx.doi.org/10.1073/pnas.0805820105>. [PubMed: 18981420]
15. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kalé L, Schulten K. *J Comput Chem.* 2005; 26:1781. <http://dx.doi.org/10.1002/jcc.20289>. [PubMed: 16222654]
16. Berendsen HJC, van der Spoel D, van Drunen R. *Comput Phys Commun.* 1995; 91:43. [http://dx.doi.org/10.1016/0010-4655\(95\)00042-E](http://dx.doi.org/10.1016/0010-4655(95)00042-E).
17. Abraham, MJ.; Murtola, T.; Schulz, R.; Páll, S.; Smith, JC.; Hess, B.; Lindahl, E. *SoftwareX.* 2015. p. 19 <http://dx.doi.org/10.1016/j.softx.2015.06.001>
18. Marrink S, De Vries A, Mark A. *J Phys Chem.* 2004; 750 <http://dx.doi.org/10.1021/jp036508g>.
19. Marrink SJ, Risselada HJ, Yefimov S, Tieleman DP, De Vries AH. *J Phys Chem B.* 2007; 111:7812. <http://dx.doi.org/10.1021/jp071097f>. [PubMed: 17569554]
20. Ellingson SR, Smith JC, Baudry J. *Mol Simul.* 2014; 7022:1. <http://dx.doi.org/10.1080/08927022.2014.899699>.
21. Kapoor K, McGill N, Peterson CB, Meyers HV, Blackburn MN, Baudry J. *J Chem Inf Model.* 2016; 56:535. <http://dx.doi.org/10.1021/acs.jcim.5b00596>. [PubMed: 26848511]
22. Jorgensen WL. *Acc Chem Res.* 2009; 42:724. <http://dx.doi.org/10.1021/ar800236t>. [PubMed: 19317443]
23. Harris JB, Eldridge ML, Sayler G, Menn FM, Layton AC, Baudry J. *Environ Toxicol Chem.* 2014; 33:1615. <http://dx.doi.org/10.1002/etc.2595>. [PubMed: 24687371]
24. Baudry J, Li W, Pan L, Berenbaum MR, Schuler MA. *Protein Eng.* 2003; 16:577. <http://dx.doi.org/10.1093/protein/gzg075>. [PubMed: 12968075]
25. Hay M, Thomas DW, Craighead JL, Economides C, Rosenthal J. *Nat Biotechnol.* 2014; 32:40. <http://dx.doi.org/10.1038/nbt.2786>. [PubMed: 24406927]
26. Lounkine E, Keiser MJ, Whitebread S, Mikhailov D, Hamon J, Jenkins JL, Lavan P, Weber E, Doak AK, Côté S, Shoichet BK, Urban L. *Nature.* 2012; 486:361. <http://dx.doi.org/10.1038/nature11159>. [PubMed: 22722194]
27. Reardon S. *Nature.* 2013; 503:449. <http://dx.doi.org/10.1038/503449a>. [PubMed: 24284710]
28. Wallach I, Jaitly N, Lilien R. *PLoS One.* 2010; 5 <http://dx.doi.org/10.1371/journal.pone.0012063>.
29. Labute MX, Zhang X, Lenderman J, Bennion BJ, Wong SE, Lightstone FC. *PLoS One.* 2014; 9 <http://dx.doi.org/10.1371/journal.pone.0106298>.

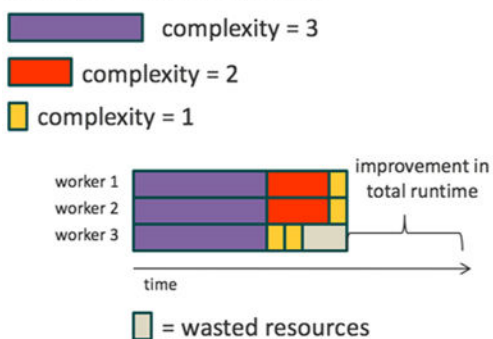
30. Chang RL, Xie L, Xie L, Bourne PE, Palsson B. PLoS Comput Biol. 2010; 6 <http://dx.doi.org/10.1371/journal.pcbi.1000938>.
31. Bowes J, Brown AJ, Hamon J, Jarolimek W, Sridhar A, Waldron G, Whitebread S. Nat Rev Drug Disc. 2012; 11:909. <http://dx.doi.org/10.1038/nrd3845>.
32. Huang Z, Wong CF. J Chem Inf Model. 2016; 56:21. <http://dx.doi.org/10.1021/acs.jcim.5b00299>. [PubMed: 26651874]
33. Noe F, Horenko I, Schütte C, Smith JC. J Chem Phys. 2007; 126 <http://dx.doi.org/10.1063/1.2714539>.

Example

Screen a library of 9
compounds against one target
protein using 3 workers

Workload = 9 tasks
Tasks per worker = 3 tasks

Tasks have varying complexity

Example

Screen a library of 9
compounds against one
target protein using 3 workers

total_complexity = 17
#workers = 3
average_complexity = 5.67

Figure 1. Illustration of (top) idle cores leading to wasted computational resources in docking and (bottom) improved usage of computational resources with an equal distribution of complexity strategy.

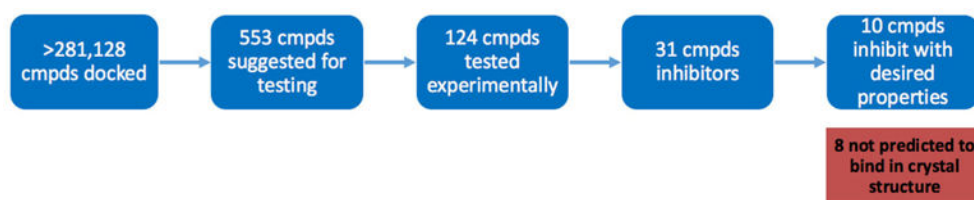


Figure 2.
Flowchart of docking results for the FXa:FVa hit discovery project.

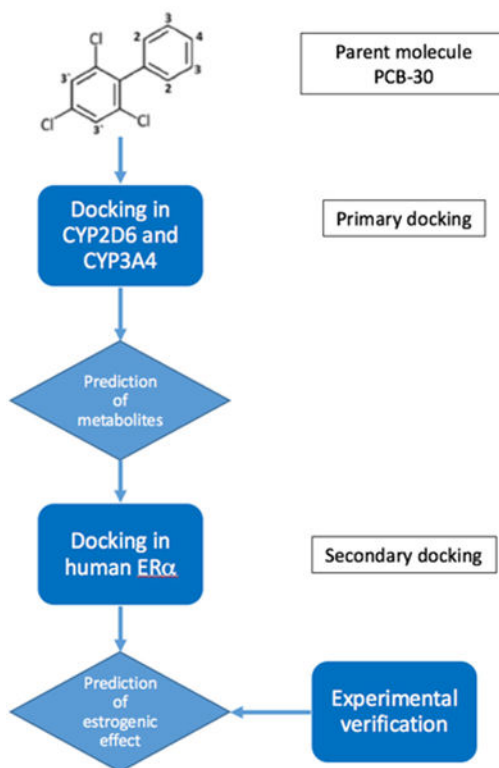


Figure 3. Flowchart of pathway exploration in the PCB-30 estrogenization project.

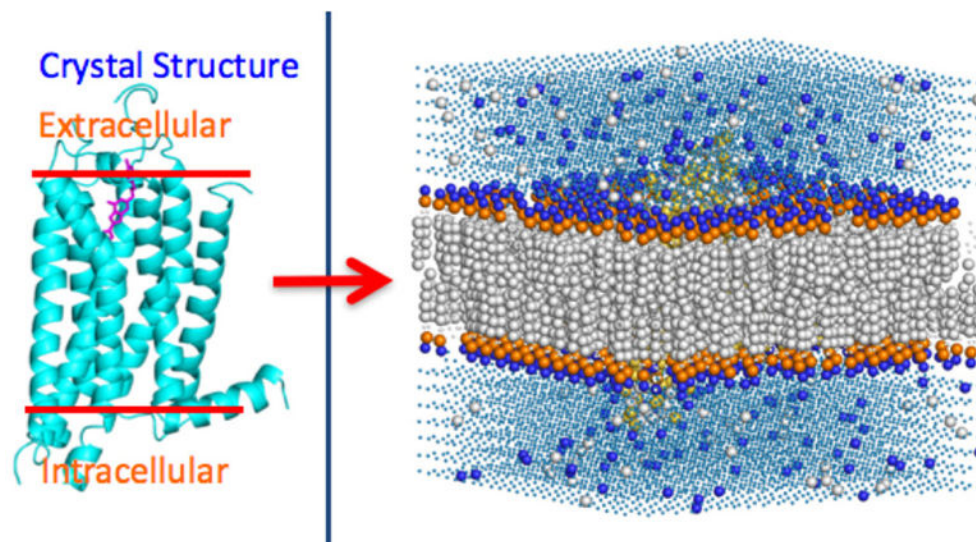


Figure 4. Simulated coarse-grained model of the protein embedded in a membrane system, showing hydration (thin blue points) and counter-ions (van der Waals spheres in the solution phase).

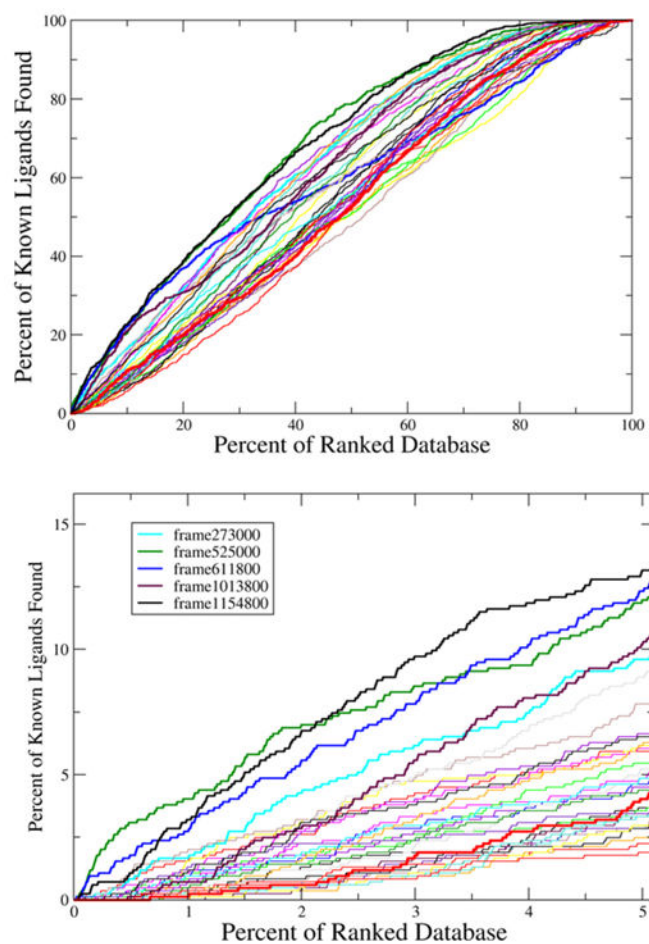


Figure 5. Enrichment factors for each of the 32 snapshots and for the crystal structure (thick red line). Top: for the entire docked database; Bottom: zoom-in on the top 5% of the docked database.