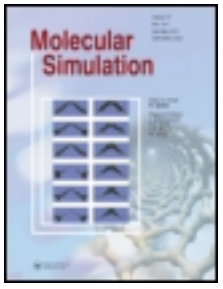

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/gmos20

Polypharmacology and supercomputer-based docking: opportunities and challenges

Sally R. Ellingson[ab], Jeremy C. Smith[bc] & Jerome Baudry[bc]

[a] Genome Science and Technology, University of Tennessee, Knoxville, TN, USA

[b] Oak Ridge National Laboratory, UT/ORNL Center for Molecular Biophysics, Oak Ridge, TN, USA

[c] Department of Biochemistry, Cellular and Molecular Biology, University of Tennessee, Knoxville, TN, USA

Published online: 17 Apr 2014.

To cite this article: Sally R. Ellingson, Jeremy C. Smith & Jerome Baudry (2014): Polypharmacology and supercomputer-based docking: opportunities and challenges, Molecular Simulation, DOI: 10.1080/08927022.2014.899699

To link to this article: http://dx.doi.org/10.1080/08927022.2014.899699

Molecular Simulation, 2014
http://dx.doi.org/10.1080/08927022.2014.899699

Taylor & Francis
Taylor & Francis Group

RECENT ADVANCES IN MOLECULAR BIOLOGY

Polypharmacology and supercomputer-based docking: opportunities and challenges

Sally R. Ellingson[a,b]*, Jeremy C. Smith[b,c] and Jerome Baudry[b,c]*

[a]*Genome Science and Technology, University of Tennessee, Knoxville, TN, USA;* [b]*Oak Ridge National Laboratory, UT/ORNL Center for Molecular Biophysics, Oak Ridge, TN, USA;* [c]*Department of Biochemistry, Cellular and Molecular Biology, University of Tennessee, Knoxville, TN, USA*

(*Received 20 December 2013; final version received 23 February 2014*)

Polypharmacology, the ability of drugs to interact with multiple targets, is a fundamental concept of interest to the pharmaceutical industry in its efforts to solve the current issues of the rise in the cost of drug development and decline in productivity. Polypharmacology has the potential to greatly benefit drug repurposing, bringing existing pharmaceuticals on the market to treat different ailments quicker and more affordably than developing new drugs, and may also facilitate the development of new, potent pharmaceuticals with reduced negative off-target effects and adverse side effects. Present day computational power, when combined with applications such as supercomputer-based virtual high-throughput screening (docking) will enable these advances on a massive chemogenomic level, potentially transforming the pharmaceutical industry. However, while the potential of supercomputing-based drug discovery is unequivocal, the technical and fundamental challenges are considerable.

Keywords: polypharmacology; docking; drug discovery

1. Introduction

Most of today's pharmaceuticals, or 'drugs', are small organic molecules interacting with proteins in the patient's body. Hence, most drug discovery effort aims at identifying, optimising and clinically validating small molecules that have the needed chemical features to bind strongly and specifically to a protein target relevant to a specific medical condition. Polypharmacology is based on the concept that pharmaceuticals may interact with more than one different protein, and even with proteins without similar sequences and/or structures.[1] To illustrate this concept, Yildirim et al. have built a drug–target network (a bipartite graph) of interactions for all known FDA-approved drugs and their targets currently on the market using data from the DrugBank database.[2] Of the 890 approved drugs with known targets used to develop the drug–target network, 89% are linked with verified multiple protein targets. This indicates that the polypharmacological nature of existing pharmaceuticals is more of a rule than an exception and suggests that future discovered molecules will most likely also possess polypharmacological properties as well.

The polypharmacological, promiscuous nature of pharmaceuticals can have both beneficial and detrimental consequences. The former of which can be exploited to, for example, improve drug efficacy and prevent drug resistance.[3] In addition to the ability of chemical compounds to interact with an array of protein targets, many diseases have multiple genetic determinants, individual genetic determinants may be involved in multiple diseases, and protein function and expression are controlled by a regulatory network of other proteins.[4] Understanding of the full network of drug–target interactions and disease and regulatory pathways will permit for repurposing of approved drugs for new applications and, inversely, novel approaches to repurposing already-studied drug targets for new diseases and guidance in discovering new drugs that take advantage of beneficial secondary target interactions while avoiding adverse effects. The fundamental characterisation and exploration of polypharmacological networks have the potential to change the pharmaceutical industry and lead to more drugs on the market that target new diseases, at a reduced cost and with a better understanding of their potential side effects. Doing so will, however, present unique challenges and will necessitate state-of-the art supercomputing capacities to produce the needed data and analyse it efficiently.

2. Beneficial consequences of polypharmacology

Functional genomic studies have shown that most single-gene knockouts have little to no effect on phenotype.[5–7] The robustness of phenotypes can be explained by the existence of redundant protein functions and signalling routes.[8] This suggests that a polypharmacological drug

*Corresponding authors. Email: sellings@utk.edu, jbaudry@utk.edu

may be efficacious because it is modulating multiple components of a disease pathway or multiple pathways relevant to an undesirable phenotype. In addition, when a pharmaceutical targets multiple points in a pathway, if one point develops mutations that cause drug resistance, there remain multiple mechanisms and pathways in which the drug may still act. For instance, fluoroquinolones are prescribed as broad-spectrum antibiotics at concentrations at which its two main targets (bacterial gyrase and topoisomerase IV) are inhibited, even though the inhibition of only one of them is needed to stop bacterial growth, thus preventing antibiotic resistance caused by single mutations of one of the targets.[3] While there is arguably a case for polypharmacological drug design, the pharmaceutical industry still largely relies on a paradigm in which one drug very selectively interacts with one target because a multitarget approach would be much more complex to design and implement. Hence, new drug design methodologies are needed in order to fully take advantage of the polypharmacological nature of drugs.

3. Detrimental consequences of polypharmacology

Interactions between a drug and multiple proteins also result in undesirable side effects and toxicity. Many adverse drug reactions result from drugs interacting with non-therapeutic antitargets.[3] For example, fenfluramine, an anorexigen, was withdrawn from the market because it led to pulmonary hypertension and heart valve damage due to the unwanted activation of serotonin 5-HT_{2B}.[9–11] It has been demonstrated that animal studies during pre-clinical trial may not yield significant indications of these adverse interactions in humans [12], and such adverse effects are generally not discovered until a drug has reached clinical trial or is already on the market. With the number of different proteins in humans and the genetic variations observable in the population, a full understanding of all possible interactions through experiments and clinical testing alone is infeasible, making computational investigations particularly useful and relevant.

4. Repurposing

One way to make use of some of the resources that have been lost to failed drugs is to find ways to utilise previous investments in research for new discoveries. Drug repurposing (also called repositioning or therapeutic switching) allows for drugs that have already been tested and approved as safe to be marketed and used to treat diseases that the drug was not initially developed to treat. This is possible if the intended drug targets are pleiotropic and involved in multiple disease pathologies or if the drug's off-target interactions are relevant in an alternative disease pathway. Drug repurposing is time and cost effective since a great deal of effort has already gone into developing and testing a drug that has subsequently already gone through the approval process. Repurposing may also be a mechanism to obtain pharmaceuticals that treat neglected diseases that would not otherwise create a profitable market for pharmaceutical companies, such as, for instance, in the case of Eflornithine (originally developed as an anti-cancer drug) that was repositioned and successfully used to treat human African trypanosomiasis, a tropical disease.[13] Here again, computational tools that explore the complete polypharmacological space of existing drugs can greatly accelerate the repurposing of approved drugs.

Drug targets can also be repositioned since many drug targets are pleiotropic. This is similar to, and has overlap with, drug repositioning, but can be unique when a drug target for the disease being investigated has not yet been discovered but was previously studied as a relevant target for an alternative disease. As clinical target validation rates are low,[14] computational tools to predict and identify proteins that are involved in a disease pathway, as well as candidate drug targets, are also useful for improving the efficiency of drug discovery.

5. Towards a systems biology (network-based) approach to drug discovery

Network-based approaches have been developed to identify drug targets, both novel and for repositioning. In [14], genes expressed differentially for a disease of interest are overlaid on a molecular interaction network and network analysis methods used to identify drug targets associated with a disease of interest. Since drug targets may highly influence a disease-specific expression response, the combination of (experimental) expression data and knowledge-based data such as molecular interaction networks can give new insights into drug targets. Identified targets can then be used to develop novel drugs for a specific disease. Alternatively, if the identified target is already used in the treatment of another disease, it can be evaluated for target repositioning. In this context, a computational framework, drugCIPHER, has been developed for predicting drug–target interactions and side effects on a genome-wide scale.[14–16] This framework uses both pharmacological space (i.e. drug therapeutic and chemical similarities) and genomic space (i.e. protein–protein interaction networks) to predict new interactions on a large scale. The power of the method, however, is limited by the quality and incompleteness of current protein–protein interaction data needed as inputs for this approach.

These above-mentioned examples show instances in which computational tools have been used to orient and facilitate drug discovery and the characterisation of

medically relevant pathways by combining biomedical data, polypharmacological properties of drugs and the recognition that disease phenotypes are the result of an underlying network of interactions. Computational approaches are based on the mining and understanding of the many-to-many relationship between the set of existing (and possible) pharmaceuticals and the set of proteins defining the druggable genome.

In the future, the above-mentioned information will be able to be exploited for purposes beyond the drug discovery and design process, and directly used for patient care in a clinical setting. As described in Ref. [4], an ideal therapeutic strategy would involve an individual screen for each patient that includes their mutations and genomic signature to identify misregulated elements in the underlying network that will be the target of a specialised treatment plan. This would, however, require a full understanding of the polypharmacological profiles of available drugs.

6. Computational docking to study polypharmacology

The previous sections illustrate how the ever-growing wealth of experimental and clinically obtained biological and medical data can be used for knowledge discovery in drug research. In drug discovery, as in most contemporary biology (and indeed as in most contemporary science), another source of data utilised originates from numerical experiments, such as molecular simulations. There are many *in silico* techniques that can be used to study the interaction between a drug candidate and target protein, including extremely computationally intensive approaches of simulating the behaviour of every atom in the protein–ligand complex in solution, and extracting from these simulations thermodynamic quantities, such as protein:ligand binding free energies. For instance, in a recent major computational achievement, the cancer drug dasatinib was simulated to bind in its experimentally determined binding pocket during an unguided molecular dynamics simulation [17] that sampled all possible protein:ligand interactions and described the binding pathway of a pharmaceutical in its protein target at an atomistic level of detail. These techniques are very insightful in determining how a small molecule interacts with its target, but they are too time- and computationally intensive to be used in a high-throughput manner comparable to that used experimentally to identify new hits in libraries of chemicals. Cost-effective methods are needed that can virtually screen a large number of drug–protein complexes quickly. Such a method is virtual docking, an efficient computational process that aims at predicting the bound conformation of a protein–ligand complex and how well it binds through a scoring algorithm.[18,19] Autodock4 [20] and Autodock Vina [21] are two open source and freely available docking tools commonly used in academic pharmaceutical research. Our laboratories developed high-throughput tools utilising these docking engines and the Message Passing Interface (MPI) libraries to efficiently distribute a massive number of docking calculations to supercomputers, namely Autodock4.lga.MPI [22] and VinaMPI,[23] respectively. Docking applications and scoring functions have been compared in reviews.[24–26] The scoring functions commonly used in docking applications use approximations to rapidly estimate protein:ligand binding affinities and the resulting computational efficiency makes these applications useful for virtual high-throughput screens (vHTS) in which millions of drug–protein complexes can be tested quickly (in a matter of days or hours) on sufficiently powerful supercomputers.

In addition to being used for hit discovery (or lead optimisation), vHTS, because of its potential to produce and analyse large amounts of molecular and biological data, can be used to address many of the challenges and opportunities of polypharmacology introduced earlier. For instance, a recent study used docking scores to relate complex drug–protein interaction profiles from DrugBank [2] with effect profiles.[15] The information was combined using correlation and classification methods to generate an effect probability matrix or drug profile, and gives a probability that each drug has any given effect. While powerful, this method is limited by the need to know a priori the effects of the drugs. A tool is needed that can make predictions about possible side effects of novel drugs during the early stages of drug discovery.

Polypharmacology is rationalised in [27] as a result of protein domains serving as drug targets. It is assumed that there are a limited number of domain types that can be combined to form different proteins of different function.[28] This concept implies that drugs bind to multiple proteins because they target a common domain shared between proteins that may otherwise be lacking overall structural and sequence homology. This idea has been used in [29] to identify potential secondary protein targets by looking for binding site similarities. In this work, a workflow which involves molecular docking into a filtered subset of the Protein Data Bank (PDB) [30] was developed to detect polypharmacological targets. The workflow includes (1) sequence homology clustering of all protein chains in the PDB, (2) selection of one representative structure from each cluster to create a subset of the PDB in which each structure is at least somewhat dissimilar, (3) assessment of binding site similarity between potential binding sites in each of the structures in the PDB subset and the known target and (4) docking of the drug candidate into the structures containing similar active sites. This approach has led to the identification of secondary targets for an inhibitor of TbREL1 from *Trypanosoma brucei*, the causative agent of African sleeping sickness. This should

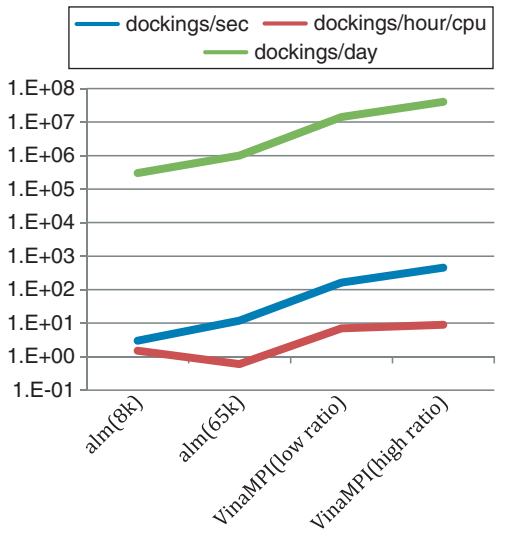

Figure 1. (Colour online) Docking capabilities achieved to date. The *y*-axis is number of dockings per different units (blue line: seconds; red line: hours/cpu; green line: day). The *x*-axis represents different docking technologies and job set-ups. alm(8 k) and alm (65 k) represent Autodock4.lga.MPI ([22]), using the corresponding core counts. VinaMPI(low ratio) represents VinaMPI on 85 k supercomputer cores with a low task-to-worker ratio and VinaMPI (high ratio) represents VinaMPI on 180 k supercomputer cores with a high task-to-worker ratio (See Ref. [1]).

be implemented early in the drug discovery pipeline, before lead optimisation, in order to identify potential undesirable secondary targets and optimise the specificity of the lead molecule. The challenges of this approach are (1) the very high number of docking calculations needed to be performed, (2) the introduction of false-positives due to shortcomings in the docking and scoring algorithms, (3) the dependence on sequence-homology clustering to reduce the number of protein structures to be processed because of computational limitations and (4) the ability to scale this solution to a library of compounds, and not just one candidate compound. To overcome these limitations, vHTS/docking tools are needed that can dock libraries of drug candidates into large numbers of protein structures with reasonable accuracies.

7. Chemogenomic level understanding of polypharmacology

Chemogenomics is the systemic study of the effects of large libraries of drug compounds against a wide variety of macromolecular targets.[31–33] Cerep, a biotechnology company, developed BioPrint, a suite of proprietary data and analysis tools to assist in drug discovery.[34] They provide pharmacological activity data between their library of in-house chemical compounds and a number of protein targets. This binding affinity data can be clustered to identify classes of proteins that interact with similar compounds. This clustering by pharmacological activity is used to identify 'hotspots' of therapeutic and off-target effects of different compounds. The ability to produce such data on a chemogenomic scale would not only be invaluable to the pharmaceutical industry but it would also lead to a better understanding of polypharmacology, and in combination with systems biology, a better understanding of disease pathology and biological mechanisms of diseases.

There are over 21 million commercially available molecules that can be used in screening for drug candidates in the ZINC database.[35] Considering also the chemistry yet to be synthesised, an estimated novemdecillion (10^{60}) small molecules are theorised to exist in the chemical universe.[36] In addition, there are about 1500 human drug targets, representing the intersection of the druggable genome and disease-altering genes [37] and as many as 10,000 ligand-binding domains [38] – without including bacterial or viral protein targets. This creates a super-massive drug discovery space that cannot be explored and validated using experimental screening approaches. Only contemporary supercomputing power has the potential to serve as an exploratory vessel.

8. Limitations of computational docking for chemogenomic level polypharmacology

The power of vHTS/docking to be successfully used for hit discovery has been demonstrated in many studies involving relatively small-scale projects (low number of targets, relatively low number of drug candidates). [15,29,32,39–41] With today's computational power and docking technologies such as developed in our laboratories,[23] several millions of compounds can be virtually screened in a single day. Figure 1 shows the evolution of docking capabilities achieved by our laboratories to date. Our docking technology Autodock4.lga.MPI, based on Autodock4 was able to perform 300,000 dockings in a 24 h period while utilising 8 k processing cores.[22] By increasing the core count to 65 k, the performance per core is reduced but this method successfully screened 1 million compounds in a 24- h period.[26,42] Our more recently developed VinaMPI approach focused on the ability to scale the docking program Autodock Vina at larger core counts. In benchmarks, this code ran on 3/4th of the Kraken supercomputer (i.e., on 85 k cores) with a continued decrease in time-to-completion of the job.[23] Recent improvements on the task-to-worker ratio mean that we estimate that nearly 40 million compounds can be screened on the Department of Energy's Titan Supercomputer, presently the most powerful supercomputer in the USA, using 180 k cores in a 24 -h period.

However, *in silico* vHTS still has significant drawbacks. The scoring algorithms do not always generate

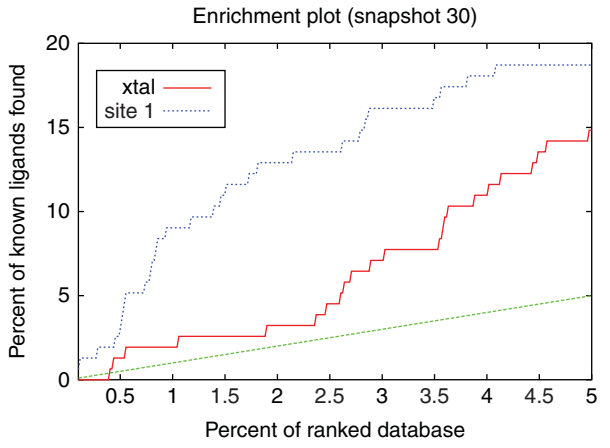

Figure 2. (Colour online) The percent of active compounds identified in the ranked compound library after docking. Red: enrichment using the crystal structure; blue: enrichment using a snapshot obtained from a molecular dynamics trajectory; green: random enrichment.

scores that correlate well enough with experimentally measured binding affinities.[43] When millions of compounds are being processed, the number of false-positives can be in the thousands. In addition, to reduce the computational complexity of the problem, protein structures are usually kept rigid, or mostly rigid, which essentially limits these numerical experiments to the investigation of – at best – an 'induced fit' binding mechanism or – at worst – a 'lock-and-key' over-simplification of protein:ligand binding mechanisms. Approaches that sample efficiently the dynamic flexibility of many protein targets are needed to investigate 'conformational selection' binding mechanisms in which drug candidates bind in a 'selected' protein conformation otherwise accessible at room temperature.

Areas for potential improvement in scoring functions include more advanced potential energy models and better incorporation of solvent effects and configuration entropy. Another approach is to perform more computationally rigorous free energy methods on top scoring vHTS docked compounds to generate more accurate scores and weed out false-positives. Reviews that address these directions include Refs [43–46].

The dynamics of protein targets, controlling many biological processes such as molecular recognition and catalytic activity, may be obtained from molecular dynamics simulations. While all atom simulations of large proteins are very computationally expensive, the ability to efficiently model active-site flexibility can greatly improve virtual docking and indeed allow for a 'conformational selection' binding mechanism to be included in the virtual screening process. This has been conceptualised in [47] and has led to successful applications in which potential drug candidates were identified that would not have been found through traditional virtual screenings using only a static, experimentally solved structure of the protein, as demonstrated in [48]. These alternative conformations can also be used to find novel binding sites not existing in the crystal structure.[49] When dealing with large chemical databases of potential drug candidates, our laboratories have also observed that the use of selected snapshots from a molecular dynamics simulation of a protein target leads to significantly improved database enrichment over that obtained using only a static (crystal) structure [see Figure 2 for an example using human tyrosine-protein kinase c-src (PDB ID 2SRC) and its set of ligands and decoys from the Directory of Useful Decoys [50]]. However, the derivation of a method for extracting snapshots that represent conformational states relevant to drug binding is still an active area of research.

9. Conclusions

The promiscuous (polypharmacological) nature of drugs can be exploited to both repurpose existing drugs and design better, more effective drugs. However, the search space of all drug possibilities and protein targets is too large to thoroughly explore experimentally. Efficient and accurate computational methods for exploring this space could revolutionise the pharmaceutical industry. In this regard, virtual docking holds great promise as a lynchpin of the future drug-repurposing pipeline. As advances are made in docking and scoring methods, the combination of the massive amount of interaction information that can be generated via simulation and extreme computational power available with supercomputers with ever-growing sources of genomic, disease and drug profile data will pave the way for a new generation of pharmaceutical discovery and personalised medicines.[51]

Acknowledgements

This work was partially funded by D.O.E – Oak Ridge National Laboratory grants number 13-1277 and 13-1278 to S.R.E and J.B. Contract grant sponsor: National Institute for Computational Science, (NICS); contract grant number: TG-MCA08X032; Contract grant sponsor: NIH; contract grant number: 1KL2RR031974.

References

[1] Keiser MJ, Roth BL, Armbruster BN, Ernsberger P, Irwin JJ, Shoichet BK. Relating protein pharmacology by ligand chemistry. Nat Biotechnol. 2007;25(2):197–206.
[2] Knox C, Law V, Jewison T, Liu P, Ly S, Frolkis A, Pon A, Banco K, Mak C, Neveu V, Djoumbou Y, Eisner R, Chi Guo A, Wishart DS. DrugBank 3.0: a comprehensive resource for "omics" research on drugs. Nucleic Acids Res. 2011;39(Database issue):D1035–D1041.
[3] Peters J-U. Polypharmacology – foe or friend? J Med Chem. 2013;56(22):8955–8971.
[4] Boran A, Iyengar R. Systems approaches to polypharmacology and drug discovery. Curr Opin Drug Discov Dev. 2010;13(3):297–309.

[5] Zambrowicz BP, Sands AT. Modeling drug action in the mouse with knockouts and RNA interference. Drug Discov Today Targets. 2004;3(5):198–207.

[6] Winzeler EA. Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. Science. 1999;285 (5429):901–906.

[7] Giaever G, Chu AM, Ni L, Connelly C, Riles L, Véronneau S, Dow S, Lucau-Danila A, Anderson K, André B, Arkin AP, Astromoff A, El-Bakkoury M, Bangham R, Benito R, Brachat S, Campanaro S, Curtiss M, Davis K, Deutschbauer A, Entian K-D, Flaherty P, Foury F, Garfinkel DJ, Gerstein M, Gotte D, Güldener U, Hegemann JH, Hempel S, Herman Z, Jaramillo DF, Kelly DE, Kelly SL, Kötter P, LaBonte D, Lamb DC, Lan N, Liang H, Liao H, Liu L, Luo C, Lussier M, Mao R, Menard P, Ooi SL, Revuelta JL, Roberts CJ, Rose M, Ross-Macdonald P, Scherens B, Schimmack G, Shafer B, Shoemaker DD, Sookhai-Mahadeo S, Storms RK, Strathern JN, Valle G, Voet M, Volckaert G, Wang C-Y, Ward TR, Wilhelmy J, Winzeler EA, Yang Y, Yen G, Youngman E, Yu K, Bussey H, Boeke JD, Snyder M, Philippsen P, Davis RW, Johnston M. Functional profiling of the *Saccharomyces cerevisiae* genome. Nature. 2002;418(6896):387–391.

[8] Kitano H. Towards a theory of biological robustness. Mol Syst Biol. 2007;3:137(1–7). doi:10.1038/msb4100179.

[9] Connolly H, Crary J. Valvular heart disease associated with fenfluramine–phentermine. N Engl J Med. 1997;337(9):581–588.

[10] Rothman RB, Baumann MH, Savage JE, Rauser L, McBride A, Hufeisen SJ, Roth BL. Evidence for possible involvement of 5-HT2B receptors in the cardiac valvulopathy associated with fenfluramine and other serotonergic medications. Circulation. 2000;102(23):2836–2841.

[11] Hutcheson JD, Setola V, Roth BL, Merryman WD. Serotonin receptors and heart valve disease – it was meant 2B. Pharmacol Ther Els B.V. 2011;132(2):146–157.

[12] Olson H, Betton G, Robinson D, Thomas K, Monro A, Kolaja G, Lilly P, Sanders J, Sipes G, Bracken W, Dorato M, Van Deun K, Smith P, Berger B, Heller A. Concordance of the toxicity of pharmaceuticals in humans and in animals. Regul Toxicol Pharmacol. 2000;32(1):56–67.

[13] Croft SL. Public–private partnership: from there to here. Trans R Soc Trop Med Hyg. 2005;99(Suppl 1):S9–S14.

[14] Emig D, Ivliev A, Pustovalova O, Lancashire L, Bureeva S, Nikolsky Y, Bessarabova M. Drug target prediction and repositioning using an integrated network-based approach. PLoS One. 2013;8 (4):e60618(1–17). doi:10.1371/journal.pone.0060618.

[15] Simon Z, Peragovics Á, Vigh-Smeller M, Csukly G, Tombor L, Yang Z, Zahoránszky-Köhalmi G, Végner L, Jelinek B, Hári P, Heténi C, Bitter I, Czobor P, Málnási-Csizmadia A. Drug effect prediction by polypharmacology-based interaction profiling. J Chem Inf Model. 2012;52(1):134–145.

[16] Zhao S, Li S. Network-based relating pharmacological and genomic spaces for drug target identification. PLoS One. 2010;5(7):e11764 (1–10). doi:10.1371/journal.pone.0011764.

[17] Shan Y, Kim ET, Eastwood MP, Dror RO, Seeliger MA, Shaw DE. How does a drug molecule find its target binding site? J Am Chem Soc. 2011;133(24):9181–9183.

[18] Shoichet BK. Virtual screening of chemical libraries. Nature. 2004;432(7019):862–865.

[19] Werner T, Morris MB, Dastmalchi S, Church WB. Structural modelling and dynamics of proteins for insights into drug interactions. Adv Drug Deliv Rev Els B.V. 2012;64(4):323–343.

[20] Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ. Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. J Comput Chem. 1998;19(14):1639–1662.

[21] Trott O, Olson A. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem. 2010;31:455–461.

[22] Collignon B, Schulz R, Smith JC, Baudry J. Task-parallel message passing interface implementation of Autodock4 for docking of very large databases of compounds using high-performance supercomputers. J Comput Chem. 2011;32(6):1202–1209.

[23] Ellingson SR, Smith JC, Baudry J. VinaMPI: facilitating multiple receptor high-throughput virtual docking on high-performance computers. J Comput Chem. 2013;34(25):2212–2221.

[24] Warren GL, Andrews CW, Capelli A-M, Clarke B, LaLonde J, Lambert MH, Lindvall M, Nevins N, Semus SF, Senger S, Tedesco G, Wall ID, Woolven JM, Peishoff CE, Head MS. A critical assessment of docking programs and scoring functions. J Med Chem. 2006;49(20):5912–5931.

[25] Moitessier N, Englebienne P, Lee D, Lawandi J, Corbeil CR. Towards the development of universal, fast and highly accurate docking/scoring methods: a long way to go. Br J Pharmacol. 2008;153:S7–S26.

[26] Ellingson SR, Dakshanamurthy S, Brown M, Smith JC, Baudry J. Accelerating virtual high-throughput ligand docking: current technology and case study on a petascale supercomputer. Concurr Comput Pract Ex. 2013. doi:10.1002/cpe.3070.

[27] Moya-García A, Ranea JAG. Insights into polypharmacology from drug-domain associations. Bioinformatics. 2013;29(16): 1934–1937.

[28] Kummerfeld SK, Teichmann SA. Protein domain organisation: adding order. BMC Bioinform. 2009;10:39. doi:10.1186/1471-2105-10-39.

[29] Durrant JD, Amaro RE, Xie L, Urbaniak MD, Ferguson MAJ, Haapalainen A, Chen Z, Di Guilmi AM, Wunder F, Bourne PE, McCammon JA. A multidimensional strategy to detect polypharmacological targets in the absence of structural and sequence homology. PLoS Comput Biol. 2010;6(1):e1000648. doi:10.1371/journal.pcbi.1000648.

[30] Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. Nucleic Acids Res. 2000;28(1):235–242.

[31] di Bernardo D, Thompson MJ, Gardner TS, Chobot SE, Eastwood EL, Wojtovich AP, Elliott SJ, Schaus SE, Collins JJ. Chemogenomic profiling on a genome-wide scale using reverse-engineered gene networks. Nat Biotechnol. 2005;23(3):377–383.

[32] Rognan D. Chemogenomic approaches to rational drug design. Br J Pharmacol. 2007;152(1):38–52.

[33] Faulon J-L, Misra M, Martin S, Sale K, Sapra R. Genome scale enzyme-metabolite and drug–target interaction predictions using the signature molecular descriptor. Bioinformatics. 2008;24 (2):225–233.

[34] Krejsa CM, Horvath D, Rogalski SL, Penzotti JE, Mao B, Barbosa F, Migeon JC. Predicting ADME properties and side effects: the BioPrint approach. Curr Opin Drug Discov Dev. 2003;6 (4):470–480.

[35] Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG. ZINC: a free tool to discover chemistry for biology. J Chem Inf Model. 2012;52(7):1757–1768. doi:10.1021/ci3001277.

[36] Reymond J-L, Awale M. Exploring chemical space for drug discovery using the chemical universe database. ACS Chem Neurosci. 2012;3(9):649–657.

[37] Hopkins AL, Groom CR. The druggable genome. Nat Rev Drug Discov. 2002;1(9):727–730.

[38] Bailey D, Zanders E, Dean P. The end of the beginning for genomic medicine. Nat Biotechnol. 2001;19:207–209.

[39] Kinnings SL, Liu N, Buchmeier N, Tonge PJ, Xie L, Bourne PE. Drug discovery using chemical systems biology: repositioning the safe medicine Comtan to treat multi-drug and extensively drug resistant tuberculosis. PLoS Comput Biol. 2009;5(7):e1000423. doi:10.1371/journal.pcbi.1000423.

[40] Jenwitheesuk E, Horst JA, Rivas KL, Van Voorhis WC, Samudrala R. Novel paradigms for drug discovery: computational multitarget screening. Trends Pharmacol Sci. 2008;29(2):62–71.

[41] Wei D, Jiang X, Zhou L, Chen J, Chen Z, He C, Yang K, Liu Y, Pei J, Lai L. Discovery of multitarget inhibitors by combining molecular docking with common pharmacophore matching. J Med Chem. 2008;51(24):7882–7888.

[42] Ellingson SR, Dakshanamurthy S, Brown M, Smith JC, Baudry J. Accelerating virtual high-throughput ligand docking: screening one million compounds using a petascale supercomputer. Proc Third Int Work Emerg Comput Methods Life Sci. 2012;33–38. doi:10.1145/2483954.2483961.

[43] Gilson MK, Zhou H-X. Calculation of protein–ligand binding affinities. Annu Rev Biophys Biomol Struct. 2007;36:21–42.
[44] Pohorille A, Jarzynski C, Chipot C. Good practices in free-energy calculations. J Phys Chem B. 2010;114(32):10235–10253.
[45] Shirts MR, Mobley DL, Chodera JD. Alchemical free energy calculations: ready for prime time? Annu Rep Comput Chem. 2007;3(07):41–59.
[46] Michel J, Essex JW. Prediction of protein–ligand binding affinity by free energy simulations: assumptions, pitfalls and expectations. J Comput Aided Mol Des. 2010;24(8):639–658.
[47] Lin J-H, Perryman AL, Schames JR, McCammon JA. Computational drug design accommodating receptor flexibility: the relaxed complex scheme. J Am Chem Soc. 2002;124 (20):5632–5633.
[48] Amaro RE, Schnaufer A, Interthal H, Hol W, Stuart KD, McCammon JA. Discovery of drug-like inhibitors of an essential RNA-editing ligase in *Trypanosoma brucei*. Proc Natl Acad Sci USA. 2008;105(45):17278–17283.
[49] Wassman CD, Baronio R, Demir Ö, Wallentine BD, Chen C-K, Hall LV, Salehi F, Lin D-W, Chung BP, Wesley Hatfield G, Richard Chamberlin A, Luecke H, Lathrop RH, Kaiser P, Amaro RE. Computational identification of a transiently open L1/S3 pocket for reactivation of mutant p53. Nat Commun. 2013;4:1407.
[50] Huang N, Shoichet BK, Irwin JJ. Benchmarking sets for molecular docking. J Med Chem. 2006;49(23):6789–6801.
[51] Fernald GH, Capriotti E, Daneshjou R, Karczewski KJ, Altman RB. Bioinformatics challenges for personalized medicine. Bioinformatics. 2011;27(13):1741–1748.