State-of-the-art technology in modern computer-aided drug design

Georgios A. Dalkas®, Dimitrios Vlachakis®, Dimosthenis Tsagkrasoulis, Anastasia Kastania and Sophia Kossida

Abstract

The quest for small drug-like compounds that selectively inhibit the function of biological targets has always been a major focus in the pharmaceutical industry and in academia as well. High-throughput screening of compound libraries requires time, cost and resources. Therefore, the use of alternative methods is necessary for facilitating lead discovery. Computational techniques that dock small molecules into macromolecular targets and predict the affinity and activity of the small molecule are widely used in drug design and discovery, and have become an integral part of the industrial and academic research. In this review, we present an overview of some state-of-the-art technologies in modern drug design that have been developed for expediting the search for novel drug candidates.

Keywords: computer-aided drug design; haptics; 3D virtual environments; 3D printing; flexible docking

INTRODUCTION

It takes ~15 years to develop a new drug from the time it is discovered to when it is available to the market. The average cost to develop each new medicine is estimated to be $0.8–$1 billion, including the cost of thousand of failures: Only 5 in 5000 compounds that enter the preclinical phase make it to human testing, and ultimately only one receives approval. This represents an enormous investment in terms of money, time and human. It includes chemical synthesis, purchase, curation and biological screening of thousand compounds to identify hits followed by their optimization to generate leads which require further synthesis [1]. Furthermore, predictability of animal studies in terms of both efficacy and toxicity is frequently suboptimal. Therefore, it is not surprising that pharmaceutical industry is searching ways to save time, money and resources, with the goal of increasing profit margins. Computer-aided drug design (CADD) has become an essential tool in pharmaceutical companies and in academia [2]. It is estimated that computer modeling and simulations will account for ~20% of pharmaceutical R&D expenditure by 2016 [1].

The availability of faster and cheaper computers and the developments in software and hardware make it possible to simulate complex biological systems with atomic resolution. Continuous improvement in user interface design simplified the user’s interaction to a few mouse clicks and the
development of many modeling softwares is oriented
towards limiting the human intervention. Some of
these computational techniques in computer-aided
drug design include databases, quantitative struc-
ture-activity relationships (QSAR), similarity search-
ing, pharmacophores, homology models, machine
learning, data mining, network and data analysis
tools that use a computer [1].

Current de novo drug design programs need to ad-
dress the following: how to assemble the candidate
compounds; how to evaluate their potential quality
and how to sample the search space effectively [1, 2].
In answering the first of these questions, two
approaches exist in order to build the desired struc-
tures [1]. The first is the linking procedure, which
the algorithm selects the most suitable moieties to
interact with the active site of the protein and then
starts to link them together in a chemically appro-
priate way. The second approach is the growing pro-
cedure, which involves the determination of a group
as a starting point and then the growing of a larger
compound that would fit in the active site and that
would be capable of establishing interactions with it.
The first approach requires that the user must supply
the docking/interaction points of the moieties that
will be used as anchors to the final structure. The
second approach requires that the user selects a suit-
able ‘seed’ as a starting point, and that the user directs
the growing process.

Evaluation of the results obtained could be done
by different methods: receptor-based scoring func-
tions; explicit force-field methods; empirical scoring
functions and knowledge-based scoring functions.
All of these approaches attempt to approximate the
binding free energy [2]. Force fields are computa-
tionally more costly than the other two types of
scoring functions. The last issue related to the
de novo drug design methods is the sampling of the
chemical space associated with the growing com-
pounds. There are many different search algorithms
implemented in the variety of software programs
available, and new ones are continuously developed
[2]. It is also important to mention the nature of the
‘building blocks’ used to construct the structures:
some programs connect single atoms; others use a
library of chemical fragments [2]. The size and the
diversity of the fragment library are fundamental for
a broad search of the chemical space. Moreover, with
these approaches, it is feasible to enrich the com-
pounds with desirable properties and reject the can-
didates with undesirable properties (inactive, toxic,
etc.). Thus, computational approaches are used to
significantly minimize time and resource require-
ments of chemical synthesis and biological testing.

In this article, we review the application of hap-
tics, a technology which takes advantage of the sense
of touch or force feedback by applying forces or vi-
btrations to the user, and its use in the field of
computer-aided drug design; virtual environments
(VEs), which users are not only external observers
of images on a computer screen but are active par-
ticipants on a computer-generated virtual world; 3D
printing, technology that could help in the develop-
ment of new drugs and finally we discuss recent
developments in incorporating the receptor flexibil-
ity in molecular docking simulations.

RECENT ADVANCES IN
COMPUTER-AIDED DRUG DESIGN
Haptic technology for drug design
In recent years, there has been increasing interest in
using haptic technology to facilitate the exploration
and analysis of molecular docking. Haptic-based mo-
olecular docking allows developing interactive systems
that could be used in drug design and development.
It is well known that molecular structures are 3D,
and the display screen is 2D, so some tricks are
necessary for 3D visualizations [3]. The mouse is a
2D device, a poor choice for 3D input. The more
input or output are available to our senses, the more
meaningful the modeling process. It could be argued
that in some cases an interactive and informed direct
human intervention in a computer simulation pro-
cess could significantly improve the results obtained
[4]. Haptic interfaces enable human–machine com-
munication through touch, and most commonly, in
response to user movements [5, 6].

A haptic display is composed of three parts: a
force-feedback device that generates computer-
controlled forces to convey to the user a sense of
the natural feel of virtual objects; a physical model
of the object that the user is attempting to perceive
and a haptic renderer that computes the forces
required to comprehend the model. A haptic ren-
derer detects collisions between a probe and a virtual
object to create a response. As a user manipulates a
probe over a molecular model’s surface, collisions are
detected and translated by the modeling or simula-
tion software into reaction forces. The strength of
these forces depends upon the depth of the probe’s
penetration into the molecule’s surface. These forces
are then returned to the user through the haptic device [7].

Brooks et al. [8] developed GROPE-III, the first haptic display system used for molecular visualization. GROPE-III combined a six degree-of-freedom plus grip Argonne Remote Manipulator, normally used for sample manipulation within a sealed laboratory glove box, with a workstation and large screen display. It ran a program called Docker that allowed the researcher to interactively dock a drug molecule within a protein’s active site. The availability of relatively inexpensive and more compact haptic feedback tools, such as Sensable’s PHANToM, has made it possible to build or adapt interactive molecular visualization systems to include haptic interactions. The PHANToM series is one of the most popular force feedback device in the world. Its control software can be easily developed using GHOST, the exclusive programming library for the PHANToM series.

Many groups have been exploring the application of haptic technology to molecular modeling in different academic- and industrially funded projects. Nagata et al. [9] adopted tactile sense technology to approach the protein–ligand docking problem and developed a new algorithm for docking simulations. Most conventional docking softwares are based on numerical differential calculations of the total energy between a protein and a ligand. Nagata et al. utilize the force between a ligand and a protein instead of the total energy. The most characteristic function of the system was its ability to enable the user to ‘touch’ and sense the electrostatic potential field of the protein. The user could scan the surface of a protein using a globular probe, which was given an electrostatic charge, and was controlled by a force feedback device [9]. The electrostatic force between the protein and the probe was calculated in real time and immediately fed back into the force feedback device. The user could easily search interactively positions, where the probe is strongly attracted to the force field. Such positions could be regarded as candidate sites, where functional groups of ligands corresponding to the probe could bind to the target protein.

Bayazit et al. [10] built a system for binding rigid ligands, and Lai-Yuen and Lee [11] have developed a custom haptic force-torque feedback tool and software system for docking flexible ligands that can be used to screen docking candidates and give the user a real-time sense of a ligand’s flexibility. Bidmon et al. [12] built an interactive haptic visualization system for analysis of protein dynamics. By attaching the PHANToM stylus to an atom, the user’s hand is dragged around in space following the motion of the selected atom over time when a protein trajectory is played back. This motion tracking allows the analysis of anisotropic protein dynamics. The user can directly feel the motion in all three dimensions without perceptual issues as there would be for stereoscopic vision. Wollacott and Merz [13] built a system called HAMStER for docking small molecules that offers both visual and haptic feedback of intermolecular forces. Brancale and his co-workers [4] implemented a simple and affordable haptic-based molecular mechanics environment aimed at interactive drug design and ligand optimization, using an easily accessible software/hardware combination by any researcher interested in this field.

Stocks et al. [14] developed the haptic system HaptiMol ISAS. It allows the user to explore the accessible surface of biomolecules using a 3D input device to investigate the shape and water accessibility of the biomolecular. The process of exploration was able to determine regions on the surface that are accessible to the solvent by using a water molecule as the probe. A navigation cube was used to visualize the explored molecular surface and the navigation cube could be automatically scaled to fit the workspace of the haptic device. In another approach proposed by Subasi and Basdogan [15], a rigid ligand molecule manipulated by the user was placed into a rigid protein to search the binding cavity while the molecular interaction forces were transmitted to the user via a haptic device for guidance. They developed a new visualization concept, Active Haptic Workspace, for the efficient exploration of the large protein surface in high resolution using a haptic device having a small workspace. The final configuration of the ligand inside the cavity was calculated off-line through time-stepping molecular dynamics (MD) simulations. The simulations were continued until the ligand molecule arrives to the lowest energy configuration.

Heyd and Birmanns [16] presented an interactive global docking (IGD) approach, which combines the best features of non-interactive exhaustive search techniques and purely interactive visualization methods. It provides the user with visual feedback about global docking scores, steric clashes and good protein–protein interfaces. In addition, haptic rendering could be employed to further enhance the interaction with the user. The additional information
supplied during IGD allows researchers not only to rely on their personal knowledge of the system but also to draw on objective, software-generated, fitting information. Stone et al. [17] developed a system termed Interactive MD (IMD), which permits manipulation of molecules in MD simulations with real-time force feedback and graphical display. Impartation was achieved through an efficient socket connection between a visualization program and a MD program.

A visual haptic-based biomolecular docking system for helix–helix docking research was proposed by Sourina et al. [18] and implemented this application in e-learning [19]. In other works, the same group developed a prototype system HmolDock (Haptic-based Molecular Docking) [20] of biomolecular docking and proposed an improved haptic rendering algorithm for with torque force [21], where the user can experience six degree-of-freedom (DOF) haptic force-torque feedback during the process. For the 6-DOF haptic-based molecular docking, Daunay and Micelli [22] developed a haptic-based molecular docking system which provides haptic feedback for a flexible ligand–protein docking and obtains a stable manipulation based on wave variables. An important limitation of this system was the size of protein molecules.

In project Combination of Sensorimotor Renderings for the Immersive Analysis of Results (CoRSAIRe), Férey et al. [23] designed an immersive and multimodal application, where virtual reality (VR) devices, such as the 3D mouse and haptic devices, were used to interactively manipulate two proteins to explore possible docking solutions. During this exploration, visual, audio and haptic feedbacks were combined to render and evaluate chemical or physical properties of the current docking configuration.

Considering the level of user interaction, it is possible to imagine the impact that the haptic technology could have on drug design, not only on de novo approaches but also on lead optimization and other aspects of CADD. The haptic technology appears to be another competent tool in the field of molecular docking for developing more interactive systems that could be used in computer-aided drug design [24].

**3D virtual reality environment**

The interactive inspection of 3D models is the base of a big number of applications in computer graphics, where most of the time is spent on providing the system with proper user input for exploring a geometric model from different viewpoints [25]. VEs are a new state-of-the-art human–computer interaction technology which allows a user to interact with a computer-simulated environment [26]. Most current VR environments are primarily visual experiences, displayed either on a computer screen or through stereoscopic displays. Some simulations include sensory information, such as sound through speakers. Users can interact with a VE or a virtual artifact either through the use of standard input devices such as keyboards and mice or through multimodal devices [26].

Today’s VR systems consist, at a minimum, of the following components: (a) Graphics rendering units: The computer hardware to compute the virtual scene and render it to a frame buffer, ready to be sent to a display device. This is typically a high-end graphics PC. (b) 3D stereo display units: Serve as the interface from the computer to the user. These used to be projectors and screens, but with the advent of 3D flat panel LCD or plasma displays these are more and more common. (c) Tracking system: Serves as the interface from the user to the computer [27]. Consumer graphics computers became powerful enough to compete with high-end, specialized graphics mainframes and significant developments have been happening with display and tracking technology. Tracking systems used to be either wireless and very expensive or tethered and still expensive. Today, wireless optical tracking systems are available at a cost similar to a high-end PC. And with the advent of consumer 3D TVs, VR displays have finally made it to the consumer market [27].

In the last decade, many groups have created numerous software applications for VEs. These VR environments allow scientists to explore worlds as small as nanoparticles and as big as the cosmos. For example, the researcher can move over a strand of DNA and look around him. Researchers at the California Institute for Telecommunications and Information Technology (Calit2) have built a number of novel 3D VR display systems over the last years. The most notable ones are the StarCAVE [28], the NexCAVE [29] and the AESOP wall. In addition, the most advanced classic CAVE to date, called Cornea, was installed in mid-2009 at the King Abdullah University of Science and Technology (KAUST) by Mechdyne Corporation and co-located with a new-generation of unique VR systems, designed also by DeFanti’s team. Other manufacturers
of CAVE-like systems include Christie, Barco, N.V., Eon Reality, Inc. and Visbox, Inc. [29].

The StarCAVE [28] is a room-sized immersive VR system with ~10 ft diameter. The user wears polarized glasses and stands in the center of an array of 15 screens, each driven by two projectors for passive stereo. A cluster of 18 high-end graphics PCs renders 3D images on 34 HD (high definition) projectors (1920–1080 pixels each) with Nvidia Quadro 5600 graphics cards. They use passive stereo, so the user has to wear glasses with polarizing filters. Users of the StarCAVE can interact with the visuals on the 360° display—by pointing a ‘wand’ that makes it easy to fly through the 3D images and zoom in or out. For surround sound output, they have a 5.1 channel surround sound system. The StarCAVE represents the third generation of surround-VR rooms. The classic CAVE was conceived and designed in 1991 by DeFanti and Sandin, who at the time were professors and co-directors of the Electronic Visualization Laboratory (EVL) at the University of Illinois at Chicago (UIC) [28, 29].

The group also created a VR application to view data sets of the Protein Data Bank (PDB) from the Research Collaboratory for Structural Bioinformatics (RCSB). The application can display the 3D macromolecular structures from the PDB in any of our VEs. Using the program ImmersivePDB [27], the viewer can move through and around a structure projected in VR. Structures can be compared to one another, they can be automatically aligned and a variety of visualization modes can be selected from. The software has an interface that makes a connection to the RCSB PDB Web site to download and display files. Both single user and collaborative modes are supported.

CORNEA [29], a Mechdyne-designed, one of the brightest and highest resolution CAVE environments at the moment in the world and one of only two 6-sided installations in the world to employ 24 digital 4K projectors (four per wall) to create an impressive 200 million pixel stereoscopic display. Each of Cornea’s six ~3 m-by-3 m screens displays ~15 megapixels/eye (there is overlap for blending between top and bottom projector sets, so each screen is not quite twice 2K lines). This amounts to a total resolution of ~90 megapixels/eye (of course, since one cannot see behind one’s head, at least half of the 90 megapixels/eye is unseen by any viewer at any time). Cornea makes it possible for users to record their VR experiences, both visually and aurally [29]. The facility can stream 32 channels of high-definition audio and video from Cornea to KAUST’s Interactive Media Room, where it can be recorded and archived. This capability provides a needed tool for scientists to collaborate with their counterparts at institutes that lack VEs. Mechdyne engineers designed and implemented this recording capability.

Researchers of Weill Cornell Medical College have leveraged this technology to help guide them to explain how cocaine and dopamine bind at the neurotransmitter site in the transporter molecule. The Weill Cornell CAVE is powered by eight Christie Mirage 3-chip DLP projectors with active stereo capability. The Christie projectors deliver a resolution of 1920 x 1920 (3.68 megapixels) per wall—334% higher resolution than most previous CAVEs. The results are superior 3D images that set new standards in molecular modeling and other avenues of biomedical research. Dr Harel Weinstein stated that with this technology they are able to investigate images at the molecular and cellular level with clarity and accuracy that was previously not achievable.

The next frontier for 3D printing: drug design

The evolving technology of 3D printing or additive manufacturing has enabled researchers to print inexpensively and quickly 3D objects. 3D printing technology is currently being studied by biotechnology firms and academia for possible use in drug design and in tissue engineering applications where organs and body parts are built using inkjet techniques. All of these technologies utilize a layer-by-layer build-up of the physical part with some method of support for overhangs in the vertical build direction. The great advantage of these methods is that nearly any shape can be created, limited only by the imagination and the structural integrity of the building material [30]. A number of different types of solid printers are on the market, utilizing materials ranging from cornstarch to metal and enabling the production of parts with various physical and mechanical properties. Solid printers that can produce full-color parts are now available [30].

Gillet et al. [30] use 3D printers to create physical models of molecules and assemblies. They developed
software that can add virtual overlays such as electrostatic potential and ligands to the hand-held models. The software composites a view of the physical model seen by a video camera with computer-generated overlays and projects the results on the computer screen.

Rather than paying experts to fabricate glassware and specialized reactors, researchers in the future may be able to design their own reaction vessels with the push of a few buttons, according to an article in Nature Chemistry [31]. This research group led by Lee Cronin and based at Glasgow University in Scotland has reportedly achieved a breakthrough in 3D printing, bringing the wonders of 3D printing into drug development [31]. Professor Cronin and his team using specialized open-source design software and low-cost 3D printer ready available on the consumer market have created ‘reactionware vessels’, essentially custom-designed polymer gels that houses chemical reactions for organic and inorganic synthesis. By adding other chemicals, such as those needed in a reaction, or catalysts, to the reactionware vessels as they are printed by the 3D printer, the group made the vessel itself part of the reaction process.

The use of 3D printers makes it possible to fabricate fully customized reactionware vessels, which gives researchers much better control over the reactions that take place within. As they stated in the article, this approach constitutes a relatively cheap, automated and reconfigurable chemical discovery platform that makes techniques from chemical engineering accessible to typical synthetic laboratories [31]. The team has successfully synthesized three previously unreported drug compounds and then controls the outcome of a fourth compound by manipulating the chemical composition of the reactionware that was holding it. By altering the chemical composition of the polymer gel, it is possible to change the catalyst for the whole process.

Combining 3D printing technology and augmented reality can also help researchers design more effective drugs. Arthur Olsen’s Molecular Graphics Lab at the Scripps Research Institute in La Jolla, CA models biological viruses and attempts to figure out what kind of proteins and ligands can latch onto them and see which might inhibit or disable them [32]. With 3D printing technology, they created accurate plastic models of virus segments and the potential drug molecules, and then they attached augmented-reality tags to them so that computer vision can help the researchers find the optimal fit.

**Teaching an old horse new tricks: improved molecular docking**

Molecular docking is playing an increasingly important role in lead discovery and design. The initial ‘lock-and-key’ theory of ligand binding, in which a motionless receptor was thought to accommodate a molecule without undergoing any conformational rearrangements, has been discarded in favor of binding models that account the random flexibility of receptors and the small molecules [33, 34]. Most docking software packages simulate the different conformations of the ligands inside the receptor’s binding pocket. Concerning the receptor, the docking algorithms predict an incorrect binding pose for ~50–70% of all ligands when only a single and rigid receptor conformation is considered [35]. However, due to its large number of degrees of freedom, there are some limitations to the use of the explicit flexibility of the receptor [35, 36]. Moreover, it is known that receptors, such as proteins and DNA, are naturally flexible molecular systems and this flexibility is essential for their functions.

There are many alternative ways to incorporate the receptor flexibility in molecular docking simulations [35–38] and the challenge of docking a flexible ligand to a receptor with varying degrees of flexibility has been addressed by a few research groups. Amaro et al. [39] applied a rigorous characterization of local and global binding effects and the computational efficiency was improved by reducing the receptor ensemble to a representative set of conformations. Machado et al. [40] used an ensemble of different receptor conformations derived from a MD simulation trajectory and they executed a series of molecular docking simulations considering in each one a receptor snapshot derived from the MD trajectory. For a limited number of start configurations, it is possible to combine docking with MD or Monte Carlo simulations. This allows, in principle, for full atomic flexibility or flexibility restricted to relevant parts of the proteins during docking. The HADDOCK program employs MD simulations including ambiguous restraints to drive the partner structures towards the approximately known interface [41]. Guilbert and James [42] developed a program, MORDOR (MOlecular Recognition with a Driven dynamics OptimizeR), which allows induced-fit binding of small molecule ligands with RNA targets via flexible-receptor, flexible-ligand docking. In this work, they describe a new procedure for docking ligands to RNA, which includes the
flexibility of the receptor over the course of the binding search using molecular minimization techniques. This enables an induced fit of the ligand and receptor. The principle of the method is basically to add an atomic root-mean-square-deviation (rmsd) constraint energy term to the potential energy, which drives the ligand to probe the surface of the receptor using energy minimization.

CONCLUSION
Considering the level of user interaction, it is possible to imagine the potential impact that these technologies could have on drug design, not only on de novo approaches but also on lead optimization and other aspects of computer-aided drug discovery. Computational techniques would allow the researcher to use the information obtained from these approaches in the design and optimization of novel potential drugs. More importantly, considering the affordability of the computer hardware, the devices used and the implementation of such approaches in a user friendly, freely available software package, these technologies become easily accessible to the average medicinal chemist and compatible with research or teaching purposes.

Key Points
- The application of haptics, a technology which takes advantage of the sense of touch or force feedback by applying forces or vibrations to the user, and its use in the field of drug design.
- Virtual environments (VE), where users are not only external observers of images on a computer screen but are active participants on a computer-generated virtual world.
- 3D printing, technology useful for the development of new drugs.
- Recent developments in incorporating the receptor flexibility in molecular docking simulations.

FUNDING
The EDGE (National Network for Genomic Research) EU; Greek State co-funded Project (09SYN–13–901 EPAN II Co-operation Grant); European Union (European Social Fund—ESF); Greek national funds through the Operational Program ‘Education and Lifelong Learning’ of the National Strategic Reference Framework (NSRF)—Research Funding Program, Thales and Investing in knowledge society through the European Social Fund.

References