

Review Article

Applying Computational Scoring Functions to Assess Biomolecular Interactions in Food Science: Applications to the Estrogen Receptors

Francesca Spyrakis¹, Pietro Cozzini¹, and Glen Eugene Kellogg²

¹University of Parma, Department of Food Science, Molecular Modelling Laboratory, Parma, Italy

²Virginia Commonwealth University, Department of Medicinal Chemistry & Institute for Structural Biology, Drug Discovery and Development Richmond, Virginia, USA

Corresponding Author: Glen Eugene Kellogg; email: Glen.Kellogg@vcu.edu

Received 3 March 2016; Accepted 7 September 2016

Editor: Yong Li

Copyright © 2016 Francesca Spyrakis et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract. During the last decade, computational methods, which were for the most part developed to study protein-ligand interactions and especially to discover, design and develop drugs by and for medicinal chemists, have been successfully applied in a variety of food science applications [1, 2]. It is now clear, in fact, that drugs and nutritional molecules behave in the same way when binding to a macromolecular target or receptor, and that many of the approaches used so extensively in medicinal chemistry can be easily transferred to the fields of food science. For instance, nuclear receptors are common targets for a number of drug molecules and could be, in the same way, affected by the interaction with food or food-like molecules. Thus, key computational medicinal chemistry methods like molecular dynamics can be used to decipher protein flexibility and to obtain stable models for docking and scoring in food-related studies, and virtual screening is increasingly being applied to identify molecules with potential to act as endocrine disruptors, food mycotoxins, and new nutraceuticals [3–5]. All of these methods and simulations are based on protein-ligand interaction phenomena, and represent the basis for any subsequent modification of the targeted receptor's or enzyme's physiological activity. We describe here the energetics of binding of biological complexes, providing a survey of the most common and successful algorithms used in evaluating these energetics, and we report case studies in which computational techniques have been applied to food science issues. In particular, we explore a handful of studies involving the estrogen receptors for which we have a long-term interest.

Keywords: Docking; Scoring Functions; Computational Chemistry; Nuclear Receptors; Estrogen Receptors; Food Science; Drug Discovery; Endocrine Disruptors; Binding Free Energy; Force Field; HINT

1. Introduction

1.1. Basic principles and concepts. Very extensive computational technology was developed by academic and industrial scientists for use in discovering, developing and validating small molecules as “drugs”, which is, of course, an economic activity that is worth many billions of dollars. These

technologies have been reviewed previously with respect to their translation to food science [1], so extensive background is not necessary here. However, the key facts are, that on the molecular scale, an interaction is an interaction, whether it is a drug identifying and occupying its binding site and setting into motion a series of steps leading to a biological action, or a protein organizing itself into its compact native

and biologically-active form, or the poisoning of an enzyme or receptor by a toxin in a way that mimics another, and considerably more healthy, interaction of that enzyme or receptor. Also, while the food industry *in toto* is certainly a far more potent economic force than “Big Pharma”, it is *much* more distributed – basically from Farm to Table with a highly variable number of intermediaries depending on how much processing the final foodstuff has endured. So, instead of the hundreds to thousands of drug companies and API (active pharmaceutical ingredient) manufacturers, there are millions of for-profit participants in the food industry worldwide. Quality control is a much more difficult proposition in this scenario. Finally, we don’t know what we don’t know in food science because the potential problem areas are almost infinite in number and thus unimaginable.

In this chapter we are focusing on “scoring functions”. These are components of the computational docking/scoring paradigm, i.e., tools to predict the energetics of protein-small molecule (or sometimes protein-protein) binding and the stability of the resulting complexes. The first part of the problem, docking, is a geometric problem – the positioning of a “solid” structure defined by its Cartesian coordinates within a cavity or pocket composed of another solid. This is not an inherently *difficult* problem, it can always be solved by exhaustive methods, but the trick and applied intelligence is to **not** do that, as there are *many* irrelevant solutions. The second part of the problem is physico-chemical: evaluating the energetics involved in the association of two or more molecules resulting in binding.

There is a very rich literature describing the virtues of particular docking/scoring methods that inevitably compare the predictions and/or limits of scoring functions. Success for a docking method is usually defined as its accuracy in recreating crystal structure “poses” by docking the separate components together. A classic review by Hawkins et al. (2008) demonstrates why this is not really a completely useful or robust exercise [6]. Also of importance to note with respect to scoring functions is that accurate estimations of binding free energies may be more relevant than simply recreating crystal structures, but there are a number of potential issues here too [6, 7].

It is also very significant that continuous progress in “omics” coupled with advances in (and emphasis on) structure determination has led to a growing database of structural information for biological macromolecules and multicomponent complexes. Advances in computer power, or more exactly the far reduced cost of the same, have enabled many predictive techniques for biomacromolecular structure. Until recently, perhaps only over the last decade, the primary focus was on modeling a much-simplified version of the system – largely the receptor/enzyme and its small molecule ligand. Flexibility was considered – if at all – for the small molecule and occasionally for a few specific amino acid residues in the protein part. Other factors such as water were generally ignored, even to the point of deleting water molecules that

Table 1: Energies of biomolecular Interactions.

Interaction type	Energy (kJ mol ⁻¹)
Ion-ion [8]	150
Ion-dipole [8]	40-60
Ion-induced dipole [8]	3-15
Dipole-dipole [8]	5-25
Hydrogen bond [8, 9]	10-40
Dipole-induced dipole [8]	2-10
Induced dipole-induced dipole [8]	0.05-4.0
Hydrophobic [10]	10-12

had been experimentally resolved. Clearly, however, the bulk solvent properties, discrete water molecules in and around the active site, metal ions and/or cofactors are part of the blueprint of biomolecular association, and real biological molecules are not rigid.

Excepting the formation of covalent bonds, the forces of biomolecular associations are all, in broad terms, electrostatic. Charge-charge interactions, whether formal or partial and centered on atoms, are known as salt bridges, dipoles and induced dipoles. Hydrogen bonds are simply a special type of electrostatic interaction, and the Van der Waals (London Force) attractions/repulsions arise from induced dipole-dipole interactions. The hydrophobic effect is actually an emergent electrostatic property due to the attraction (largely hydrogen bonding) between water and polar atoms or groups “forcing” hydrophobic atoms or groups together. It is useful and intuitive, however, to explain the observed structures, especially in aqueous solvent, in terms of a “hydrophobic interaction” beyond what is normally expected from a purely Van der Waals perspective. Some of the above weak forces are set out in Table 1 with typical values.

The many complexities in approximating, quantitatively, the vast number of individual steps of biochemical recognition to yield an interaction explains the overarching failure to obtain a precise and rapid estimation of binding free energy. Such a computational tool, if universally applicable, is sometimes referred to as the “Computational Holy Grail” [11, 12]. The basic problem is that we really have an incomplete understanding of the physics and thermodynamics of biomolecular associations [13]. Concomitantly, applying a purely reductionist approach is likely to ignore the inherent complexity, e.g., the hydrophobic effect, driving such processes. Starting with the basics of thermodynamics, the formation of biological complexes should be considered in terms of the standard Gibb’s free energy of binding (ΔG°), which measures the equilibrium of the reaction $P_{aq} + L_{aq} \leftrightarrow PL_{aq}$ in aqueous solution:

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ, \quad (1)$$

where ΔH° is the enthalpic contribution (breaking/formation of hydrogen bonds or specific polar and non-polar contacts)

and $T\Delta S^\circ$ is the entropic contribution (release of “constrained” water to solvent and changes in conformational mobility of molecules). ΔG° is also related to K_a , the association constant:

$$\Delta G^\circ = -RT \ln K_a, \quad (2)$$

where K_a is for the above reaction. In many cases, we use the reverse dissociation constant K_d , defined as $[P_{aq}, L_{aq}]/[PL_{aq}]$.

While the *direct* interactions between molecules are most important for favorable enthalpies of binding, the ubiquitous presence of water also contributes energetically. Prior to interaction, each component of the reaction is solvated, requiring partial or full desolvation during the binding process. Enthalpy gained is the difference between that of the molecular association and those desolvation enthalpies from the involved molecules. The entropy gain arises from the release of the previously bound waters, which is, in turn, compensated by their reorganization into new and ordered hydrogen bonds with the existing waters in bulk. In addition, unfavorable entropic contributions arise from changes in degrees of freedom, both translational and rotational of the molecules. Hydrogen bonding is often considered to be the main driving force in binding energy (see Table 1), but hydrogen bonds are susceptible to disruption by water, which is ubiquitously present in biological systems. Thus, hydrogen bonds not solvent-accessible are stronger than those on the surface. Hydrophobic interactions, on the other hand, are energetically much weaker when evaluated one-by-one, but can be significant when considered as a whole, and are far less influenced by interactions with solvent.

For evaluating interactions we must consider the total free energy of binding ($\Delta G^\circ_{\text{bind}}$), which is the sum of the interaction energy between the interacting molecules ($\Delta G^\circ_{\text{int}}$), the solvation energies of molecule 1, molecule 2 and of the resulting complex ($\Delta G^{\circ, \text{mol1}}_{\text{solv}}$, $\Delta G^{\circ, \text{mol2}}_{\text{solv}}$, $\Delta G^{\circ, \text{complex}}_{\text{solv}}$, respectively), as well as entropic ($T\Delta S^\circ$) and conformational ($\Delta\lambda$) changes in energy [14]:

$$\begin{aligned} \Delta G^\circ_{\text{bind}} = & \Delta G^{\circ, \text{complex}}_{\text{solv}} - \Delta G^{\circ, \text{prot}}_{\text{solv}} \\ & - \Delta G^{\circ, \text{lig}}_{\text{solv}} + \Delta G^\circ_{\text{int}} - T\Delta S^\circ + \Delta\lambda \end{aligned} \quad (3)$$

There is a catch, however: Dill [15] has shown that adding up terms as above ignores the fact that biological interactions are actually concerted events that cannot be simply deconstructed. Even if each of the interaction types are individually calculated with the utmost of care, the resulting sum may not yield an accurate and reliable $\Delta G^\circ_{\text{bind}}$.

1.2. Practical matters. The first requirement for any structure-based computational experiment is reliable structural data, preferably from experiment. This is usually derived from one of three experimental techniques: X-ray diffraction, neutron diffraction and Nuclear Magnetic Resonance (NMR). Alternatively, when an experimental structure

is not known, but significant “homology” exists between the biomolecule of interest and one or more known structures, theoretical comparative modeling technique may be of some use (*vide infra*). To date, X-ray diffraction has been the most productive experimental method because it is relatively inexpensive and not as susceptible to limitations of other methods. While the big advantages of neutron diffraction over X-ray diffraction are higher resolution and the ability to directly observe hydrogen atoms, neutron diffraction is more expensive and rare because neutron sources are only available at accelerator facilities and the technique necessitates somewhat larger crystals, i.e., more material. The primary NMR limitations are that larger (molecular weight) proteins are more difficult to assign and that solubility is often an issue. However, NMR is a solution technique and the resulting structure is thus free of crystallization and packing effects, while also providing a description, albeit limited, of flexibility.

There are few examples of structures derived by homology modeling – when there are modest comparative identities – that have led to comprehensive understanding of structure or creating models as accurate targets for drug discovery, etc. The problem is that many relatively small errors and uncertainties lead, altogether, to fairly large problems over an entire structure. Thus, this approach is best used when there is more direct experimental structural information.

There are a number of reviews [16, 17] describing limitations of structural data for structure-based studies such as drug discovery; Protein crystal structures are not immune from error: a commentary in *Nature* [18] discussed this issue in important protein structures. It is clear that there is a correlation between input structural quality and successful computational studies [6, 19]. Below, we will briefly highlight some of these factors, while defining some important terms (see also Giacobozzo [20]).

Three parameters representing quality of structural data are: resolution, R factor and temperature (B) factor. First, resolution, usually reported in Å, describes the accuracy in data collection, but not in refinement. With current instrumentation this effectively is a measure of the intrinsic quality of the crystal. Resolution is directly related to Bragg’s law ($d = \lambda/2\sin\theta$). It is important to note that, when atoms are closer together than this resolution, they *can* be differentiated. This is because atoms are modeled as spheres, the entire reflection (peak) is fit, and the underlying chemical structure of the molecule is generally known. In other words, resolution does not indicate a standard deviation of atomic position. Resolutions under 2.0 Å are considered good and values close 1.0 Å are termed atomic resolution. Second, the R factor or residual factor represents how well the refined structure explains the observed data. A newer and increasingly used adaptation is R_{free} , which uses a small subset of data not included in the refinement to test the structure model [21]. Third, the B factor, also called the temperature factor, represents a composite of uncertainties

for an atomic position that incorporates effects such as movements of atoms (along Cartesian axes from isotropic or anisotropic thermal motion), and static disorder in the crystal, among others. For reference, a B of around 20 suggests that atomic positions are indeterminate about ± 0.25 Å. Note that B factors are fit simultaneously with atomic positions during refinement. Thus, the B is not an independent and unbiased metric of structure quality. To ameliorate some of the problems with structural data quality, a number of researchers have “curated” test sets: e.g., LPDB [22], PDBbind [23, 24], BindingDB [25], Binding MOAD [26], CCDC/Astex [27], AutoBind [28]. These sets can be used as test beds for parameterizing scoring functions in docking programs.

Even given a good quality structure model, there are a number of preparation issues before computational studies can commence: 1) proper interpretation of atom “type” and hybridization from structural data files – the ubiquitous “PDB format” does not encode multiple bonds; 2) the positions of protein hydrogen atoms are seldom known experimentally – thus, tautomeric states of histidine or ligands or the ionization states of numerous functional groups on the protein or ligand are similarly unknown. These effects, especially for residues located within the protein active site, can significantly impact (even reverse) ligand binding; and 3) while docking/scoring problems appear to be two body interactions (e.g., protein-ligand, protein-protein, protein-DNA, organic host-guest, etc.), there are invariably a number of other actors on the scene, e.g., other proteins, other ligands and water (both discrete and bulk), along with other cofactors, ions, etc. So, to completely estimate binding energy, we must consider *all* interactions between *all* players, including those that may change due to changes in ionization states.

2. Docking and Scoring

Ligand docking is the prototypical version of this technology. Docking is often applied as a computational technique for lead discovery, i.e., part of a virtual screening workflow. Algorithmically, leads are found with two basic approaches: “true” docking algorithms and *de novo* design methods. Docking approaches consider ligands as the whole, while in *de novo* design, the molecules are built sequentially from a seed that is “grown” by adding building blocks of pre-organized organic fragments. An alternative is simultaneously placing multiple fragments that are later linked [29]. The strategies adopted by *de novo* design programs are a separate subject from this review and the interested reader is referred to the literature [30, 31]. Our interest here is on scoring algorithms for docking and energy evaluation, particularly for ligands docked to biomacromolecules. The docking of two macromolecules together is a much more difficult problem [32], and a topic for another day. However, the evaluation of energy should be, in principle, independent of the types

of molecules if they are in the same environment, i.e., the biological media.

At its core, docking for drug discovery is a rational approach with a goal of predicting the structure and the binding energy of a complex given only the structures of the component free ligand and receptor [33]. The primary aim is to correctly predict the most favorable binding conformation and modality of the ligand in the target’s active pocket. Thus, both a good positioning algorithm and a reliable scoring function are the necessary component prerequisites for a successful docking tool [34]. The search step is simply evaluating the accessible conformational space for the interaction between the two molecules, while balancing accuracy, redundancy and speed. The score step is the evaluation of the generated conformations that discriminates between native (correct) and non-native (incorrect) binding models [35]. In theory, the highest scored complex should mimic what was or would be observed experimentally. This is unlikely for a number of reasons: 1) scoring functions are seldom able to find the global minimum energy; 2) optimization algorithms capable of identifying such a minimum are too complex and slow for docking simulations [36, 37]; and 3) the search space is very large with six degrees of translational and rotational freedom for the ligand *and* numerous internal conformational degrees of freedom for both the ligand and the protein [34]. Generally, only a subregion of conformational space and a reduced set of possible binding poses can be effectively explored.

Counterintuitively, in order to increase the probability of finding the global minimum, or at least to try to sample more promising conformational space effectively, constraint strategies can be adopted. The most severe constraints involve treating one or both components as rigid bodies, thus exploring only six degrees of translational and rotational freedom. This constraint forces the assumption that the starting ligand conformation is very similar to the final experimental one, and that there are no conformational adjustments of the protein due to the ligand’s binding [34]. Not surprisingly, there is a lot of evidence that this is a poor approximation, as induced fit effects in adenylate kinase [38], nuclear receptors [39, 40], G-protein-coupled receptors (GPCRs) [41], among other systems led, at a minimum, to ligand conformational flexibility [42], and later for limited exploration of receptor flexibility [37].

The time and resource demands of rigorous scoring functions make them impractical for virtual screening. Again, the approach has been simplification of these functions to more mathematically tractable algorithms, which leads, inevitably, less accurate energy estimations. This should be viewed, however, in the light of knowledge that no current scoring function, regardless of its complexity, is completely accurate. The goal is thus to find the right compromise between computational time and binding prediction accuracy [34].

A particularly useful innovation was the evolution from the representation of the target biomolecule with explicit atoms [43, 44], to molecular surfaces [45], and eventually to regular three-dimensional grids encoding the receptor's physico-chemical properties [46–49]. The efficiency arises from the ability to “look-up” rather than calculate the scores (energies) of different ligand conformations generated by the docking algorithm. Although Platzer and Scheraga reported the use of molecular dynamics (MD) simulations on enzyme-substrate complexes in the early 1970s [50], Kuntz's DOCK algorithm, which uses combinatorial sampling methods, is acknowledged as the first “docking” tool [51–53]. Many other algorithms and protocols have been designed and implemented since, and docking is considered a key technology in computer-aided drug discovery. Considerable debts to the success of docking are the rapidly increasing availability of interesting and important target structure and the even more rapid decrease in the cost of computing.

From the most basic to most complex, docking methods can be classified with respect to the allowed degrees of flexibility inherent in the algorithm, i.e., from: 1) placing rigid ligands in rigid sites, i.e., the classic “lock and key” [51]; 2) placing flexible ligands in rigid sites [54]; 3) placing flexible ligands into “semiflexible” sites [55]; and 4) full flexibility. Algorithms and methods for the latter are being developed, but numerous issues related to torsion sampling and scoring render this a difficult problem. This article will focus on the scoring function part of the process, but we will first present a brief review of docking methods for reference. It also should be noted that this article is not a critical review of these computational technologies. There is already voluminous literature comparing docking tools and scoring functions, e.g., by Foreman et al. [56], Bissantz et al. [57], Bursulaya et al. [58], Kroemer et al. [59], Cole et al. [60], Warren et al. [61], Truchon et al. [62], Englebienne et al. [63], Weill et al. [64], Biesiada et al. [65], Cerqueira et al. [66], Vajda et al. [67], Wang and Lin [68], and Cheng et al. [28].

2.1. Docking algorithms. At the heart of docking algorithms is the method by which orientations of the incoming molecule are searched and generated for interaction with its target. Algorithms that search conformational space are of three general types: systematic, stochastic and deterministic [37]. Systematic or combinatorial search methods [35] calculate a grid of values for each degree of freedom designated for exploration. Thus, as the system complexity increases, with more degrees of freedom, the search likewise becomes more complex – necessitating termination criteria. Docking methods keying on shape complementarity with point-to-point matching or distance geometry algorithms were adequate for rigid-rigid or semirigid-rigid docking. Adding more flexibility led to “place and join” and incremental construction algorithms that avoid combinatorial explosions in searching conformational space [35]. Notable examples of

docking programs that use point complementary matching are: FTDOCK by Gabb, Jackson and Sternberg [69], FRED (Fast Rigid Exhaustive Docking) developed by OpenEye [70, 71], PhDOCK [72] and Ph4DOCK [73], DragHome by Schafferhans and Klebe [74], and Schneidman-Duhovny's PatchDOCK [75]. Distance geometry-based methods adopt a form of shape matching algorithms; programs of note include the earliest incarnations of DOCK (1982) and the Wallqvist and Covell program ADAM [76]. FRED, by using conformer pose libraries built by OMEGA [77] is also exhaustive. Prominent docking programs using incremental construction or “place and join” include: the 1992 Leach and Kuntz (and later) implementations of DOCK [78], Rarey *et al.*'s FlexX [79], SURFLEX by Jain [80], and SLIDE by Leslie Kuhn [81, 82].

Unfortunately, energy minimization techniques do not exhaustively sample conformational space as they progress towards the closest (usually) local minimum. Thus, other strategies to overcome are applied to cross high energy barriers [35]. Alternative approaches – stochastic evaluations of conformational space – are thus often used. Monte Carlo simulations and Genetic Algorithms (GAs) search for global minima of binding free energies evaluated with molecular dynamics or molecular mechanics [83]. Stochastic algorithms also have the advantage of more complete modeling of both ligand and receptor flexibility. However, they are intrinsically random in nature, and issues with stochastic docking methods are related to trouble in reaching convergence, often requiring multiple independent runs [37]. In Monte Carlo (MC) methods, poses are explored by randomly applying bond rotations, translations, and etc. one degree of freedom at a time. Following, each conformation's energy is evaluated. The earliest formulation of Goodsell and Olson's AutoDock (1990) was based on MC/SA methods [84]. Other software using MC methods include: DockVision by Hart and Read [85], MCDOCK by Liu and Wang [86], HADDOCK (High Ambiguity Driven protein-protein Docking) by Dominguez and co-workers [87], PRODOCK by Trosset and Scheraga [88], and Schrödinger's Glide docking and scoring program [89]. Also considered as stochastic search methods are those relying on genetic algorithms that mimic evolution by manipulating structural data parameters or “chromosomes” [90]. GOLD (Genetic Optimization for Ligand Docking) [91] is a particularly well known GA-based program, as is post-1998 AutoDock [92].

In contrast, deterministic methods rely on the architecture of the initial conformation to project the moves for generating the next state model. There are significant drawbacks of these methods in that there is a high probability of being trapped in local minima due to many high energetic barriers [37]. In response, two divergent mathematical approaches have been applied in MD-based docking algorithms: 1) simulated annealing that has been enhanced with “tricks” to overcome such barriers has been successfully applied in approaches such as CDOCKER by Wu et al. [93], Miranker and Karplus's

Multiple-Copy Simultaneous Search (MCSS) [94], and by Di Nola in the MDD (MD Docking) algorithm [95]; 2) heuristic tabu searches are designed to ensure that unexplored space is sampled during docking simulations with a variety of penalty and reward functions. The best known docking method of this type is PRO_LEADS (Ligand Evaluation by Automatic Docking Studies) by Baxter and co-workers [96].

2.2. Scoring functions for docking: overview. The primary goal of this article is to provide the food scientist with some background information and theory (albeit not at an extremely deep level) to evaluate the pros and cons of using computational docking methodology in his or her research. In particular, we are focusing here on the scoring functions used in these computational tools. This section will illuminate several approaches that have been used to design and implement scoring functions associated with docking tools. Note that there are imprecise definitions for scoring functions because their goal is actually somewhat ambiguous: the measure of a binding force, which is not universally defined, is manifested as a simple numeric value. These values are, in the simplest terms, rankings of one ligand pose relative to another.

There is an interesting feature of scoring functions as used in docking that should be mentioned. In many cases, these scoring functions are used for both scoring *after* docking with their most rigorous energetic evaluations, and scoring *while* posing with less sophistication. Depending on the docking algorithm, many intermediate decisions may be required during docking of a single pose, and this type of scoring need not be as detailed and precise as the final comparison of poses. Thus, completely separating discussions of docking algorithms from scoring functions is difficult. However, since there are distinct algorithmic classes for scoring functions, this separation makes sense pedagogically.

There are four primary classes of scoring algorithms: 1) those based on Newtonian molecular mechanics force fields; 2) those, termed semi-empirical, obtained when molecular mechanics terms are supplemented by empirical parameters such as hydrophobic-contact surface area; 3) those based on analyses of training sets that yield predictive equation with QSAR-like descriptors are of the empirical class; and 4) the knowledge-based function class that use potentials of mean force (PMF) data to interpret known structures in terms of rule sets. Finally, because scoring functions are often built from complementary information, consensus scoring has been used to make docking decisions.

2.3. Force field-based methods . At the core of force field-based scoring algorithms are the molecular mechanics force fields. Briefly, force fields are sets of equations and atom-type specific constants that can be applied to energy calculations. In particular, energy is calculated as the sum of electrostatic and van der Waals potentials, combined with intramolecular

distance, angle and torsion contributions, in a form usually like:

$$E_{total} = \sum_{bonds} Kr(r - r_{eq})^2 + \sum_{angles} Ke(\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{V_n}{2} [1 + \cos(n\varphi - \gamma)] + \sum_{i < j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} + \frac{q_i q_j}{\epsilon R_{ij}} \right] \quad (4)$$

For intramolecular terms, K_r is the bond stretching constant (r and r_{eq} are actual and equilibrium bond lengths); K_e is the “bond” stretching constant for the virtual bond between the 1–3 atoms of an angle in a bend (θ and θ_{eq} are the actual and equilibrium “bond” lengths; and similarly, V , n , φ and γ are parameters associated with the energetics of dihedrals. For intermolecular terms, A_{ij} and B_{ij} are the repulsion and attraction parameters of the 6-12 Lennard-Jones potential, while R_{ij} is the distance between atoms i and j ; electrostatically, q_i and q_j are the atomic point charges and ϵ is the dielectric constant. Some well known and widely used molecular mechanics force fields include: AMBER [97], CHARMM [98], DISCOVER [99], GROMOS [100], MMFF [101], MM2 [102], MM3 [103], MM4 [104], OPLS-AA [105] and TRIPOS [106].

Interestingly, force field methods were developed for evaluating gas-phase phenomena and not biomolecular interactions. More emphasis in developing these equations had thus been on the intramolecular “mechanical” components that enable the prediction of molecular structure from the perspective of bond lengths, angles, etc. In contrast, there is less understanding of intermolecular forces, particularly in biological environments, and less attention has been paid to hydrophobic interactions, solvation, and entropic effects. The main advantages of force field-based scoring functions are speed and mathematical precision. Because force field equations have been optimized for energy minimization, which is highly iterative, binding or internal energy estimates for a single state are very rapid. When differentiating amongst multiple docked models, the application of force field tools is very similar to that in energy minimization: energies are calculated on an atom-by-atom basis. These equations, however, are very amenable to grid representations where each pre-computed grid point encodes, for example, polar and non-polar potentials of the target.

Because of their purely Newtonian origin, and the resulting absence of entropic and solvation effects that are known to be significant components of biological interactions, additional and more sophisticated representations of non-covalent terms, e.g., electrostatics have been implemented. While far more computationally-demanding than Coulomb’s law, the Poisson-Boltzmann Equation (PBE) [107] and the Generalized Born (GB) approximation have been successful additions to non-empirical scoring functions [108].

The first scoring function used in DOCK [51] was an adaptation of force field parameters from AMBER, where protein-ligand interactions were calculated with repulsion and hydrogen bonding terms. Meng and colleagues added a PBE-based electrostatic interaction energy function and Lennard-Jones terms to DockScore [47]. Consideration of solvation effects in the function was added by Shoichet, Leach and Kuntz [109], by first calculating the electrostatic potentials and van der Waals interactions with PBE of DelPhi [110] and CHEMGRID [47], respectively; later, Zap [111] electrostatics was implemented. Interaction free energies were then corrected with HYDREN's [112] calculations of electrostatic and non-polar ligand solvation energies. With this scoring approach's inclusion of the HYDREN penalty terms yielded virtual screening results of molecules with smaller size and lower formal charges. DOCK 5 and DOCK 6 include corrections for ligand conformational entropies, desolvation, Generalized Born/surface area (GB/SA) and Poisson-Boltzmann/surface area (PB/SA) solvation scoring, and receptor flexibility and implicit solvent, both with AMBER.

FLOG (Flexible Ligands Oriented on Grid) by Miller and co-workers [113] used a scoring equation that estimated docking energy as a pairwise sum of van der Waals, electrostatic, hydrophobic and hydrogen bonding terms that were size-normalized to avoid scoring problems with large and bulky molecules. The QXP approach (McMartin and co-workers) was based on a MC search protocol with AMBER energies [114]. While rigid body docking is driven by electrostatic potentials and van der Waals forces, superposition is included as a separate term and helps identify the bioactive conformation of flexible ligands [115]. In later versions, additional QXP terms estimated non-bonded interactions, van der Waals, electrostatic energy and contact energies of interactions and the count of hydrophobic contacts [116].

At this point, treating the contribution of solvent requires more discussion. Kollman and co-workers developed the MM/PBSA approach, which is a solvent continuum method combined with classic molecular mechanics. The latter calculates the usual bond, angle, torsion, van der Waals and electrostatic terms. To this was added Poisson-Boltzmann-based solvation free energy and solute entropy derived from analyses of MD trajectories [117]. While Free Energy Perturbation (FEP) methods are more accurate (and more resource intensive), MM/PBSA involves MD simulations for only the free ligand in solution and the protein-bound ligand [118]. Later, Massova *et al.* introduced a Generalized Born (GB) variant that further reduced computational time [119]. A further enhancement was the OWFEG (One Window Free Energy Grid) method of Pearlman *et al.* [120], which is a simplified FEP that performs the analyses on a single MD trajectory. Free energy is calculated at each point of a grid that represents protein structure with three probes (neutral, positively charged and negatively charged). The final G_{score} is the sum of the grid score for each atom in the ligand.

Newer developments include MedusaScore, by Yin *et al.*, a forcefield-based scoring function including van der Waals, solvation and hydrogen bond terms derived from newly parameterized forcefields and algorithms [121]. A key feature of MedusaScore is that it is not trained to fit specific experimental data and thus possesses high transferability.

2.4. Semi-empirical and empirical methods. Semi-empirical methods use empirical or empirically-calibrated energetic terms to calculate or estimate interaction types that are not generally available from molecular mechanics, while empirical scoring functions represent free energies of binding as sums, often empirically weighted by regression, of uncorrelated terms that each account for different types of contributions [122]. There are advantages and disadvantages for both: while allowing contributions for fundamental biological interactions such as hydrogen bonding, solvent effects and hydrophobic interactions, these scoring functions lose some of the supposedly universal "first principles" applicability of the purely physics-based molecular mechanics forcefields.

Typically, semi-empirical scoring functions include electrostatic, hydrophobic, and entropic terms. In the ICM method, electrostatic contributions are represented as the sum of desolvation and Coulombic ΔG s obtained with the Poisson equation [107, 123], hydrophobic contributions from multiplying the solvent accessible surface area (SASA) by the surface tension, and entropy is calculated as $RT \ln[N_s]$ (N_s is the count of conformational states with low-energy found via MC simulations).[124] GOLD [125] combines three terms: hydrogen bonding energy, estimated with a 4-8 (Lennard-Jones-like) potential, and VdW forces for protein-ligand interactions and the internal energy of the ligand are calculated with a 6-12 Lennard-Jones potential. Similarly, an AMBER-based scoring function implemented in early releases of AutoDock used a 6-12 Lennard-Jones potential for vdW interactions, a 12-10 function for hydrogen bonding, and a Coulombic term for electrostatic interactions [126, 127]. AutoDock 3.0 [92] and later use semi-empirical scoring based on Wesson and Eisenberg's thermodynamic cycle [128], and ligand conformational and desolvation contributions from Stouten's pairwise volume-based method [129].

The summation of uncorrelated terms in empirical scoring is without physical foundation, especially with respect to the assumed additivity of such terms [15]. In fact, this paradigm has more similarities to QSAR (Quantitative Structure-Activity Relationships) [130] than to any theoretically sound physics-based method. However, like QSAR, there are pragmatic and interpretative advantages; e.g., the various terms, hydrogen bonds, hydrophobic interactions, ionic contacts, entropic effects, etc., are familiar ones. As in QSAR analyses, functions and weighting coefficients are obtained by fitting parameters to training sets of complexes. Empirical scoring functions are simple and fast, but have several drawbacks largely related to the quality of data,

i.e., experimental binding data or crystallographic structures, in the training sets. SCORE1 by Böhm was probably the first empirical scoring function and was implemented in his *de novo* design program LUDI [131], and later was used as the scoring function of FlexX [132]. Only terms for hydrogen bonding and hydrophobic interactions were part of SCORE1, but terms for losses of internal entropy, deviations from ideal hydrogen bond geometries, the loss of binding energy from restricting the internal degrees of freedom of the ligand when it binds, cation- π interactions, differentiated between neutral and ionic hydrogen bonds, and contributions of water molecules were implemented in the SCORE2, algorithm [133, 134]. The ChemScore function was modeled after SCORE1 with the addition of a hydrogen bond term for water-mediated contacts, ligand flexibility and atom-atom contact estimates of hydrophobic interactions [135]. Following Verdonk's addition of covalent energy terms [136] ChemScore was implemented in GOLD [137], FRED, Sybyl and PLANTS. Wang developed the SCORE algorithm by reconstituting binding affinity as five terms, representing vdW forces, metal-ligand interactions, H-bond energy, desolvation (a form of logP), and deformation, localized on each atom of the ligand [138]. The later SCORE 3.0 [139] includes terms for entropic effects and internal ligand energy.

One trend in scoring function development has been a steady increase in the size of training and test sets used for calibration. This makes sense on a number of levels, particularly related to statistical robustness, but outliers are more difficult to detect and the stories they tell are often not heard. As part of the "Scoring Function Consortium" project [140], Sottriffer recently described the SCFscore algorithm, based on a training set of more than 850 protein-ligand complexes and more than 60 unique descriptors, including some usually absent, such as interactions with aromatic rings. Still missing are descriptors that account for the energetic roles of water molecules, internal strain energy and etc.

HINT (Hydrophobic INteractions) by Kellogg and Abraham [141] is almost completely based on the hydrophobic effect, as measured by $\log P_{o/w}$, the partition coefficient for 1-octanol and water [142, 143] – the free energy of solvent transfer. Each atom's hydrophobic atomic constant, derived by fragmentation of experimentally determined $\log P_{o/w}$ values, is a partial δg and directly related to the atom's propensity of interacting with 1-octanol/water or other hydrophobic or polar atoms. The HINT Score is correlated with experimental free energies of binding to create the scoring function for the dataset of interest. The HINT Score function is optimized for binding free energy and has often been used in docking applications as a post-processing filter.

2.5. Knowledge-based methods. Algorithms of the knowledge-based class capture the steadily evolving "knowledge" that is available in publicly accessible databases, in particular the Protein Data Bank. Mügge, who

was one of the first to organize these data for docking scoring functions, describes potential of mean forces (PMF) scoring functions as statistical potentials that are derived from statistical mechanics [144]. The key statistical assumption is that more commonly observed interactions contribute positively to binding, while less common interactions should be marked as repulsive. Then, with the Boltzmann law, the extensive experimental information for structurally characterized protein-ligand complexes is converted into sets of atom-pair potentials for each pair of interacting atoms. Knowledge-based algorithms have a *physical* equation, but are statistical methods. However, the quality and diversity of the training databases is critical [145]; for example, more rare interactions (like π - π or π -cation) can be poorly predicted.

Wallqvist and co-workers designed their knowledge-based potential function to focus on calculated buried surface areas in complexes by comparing the interface surface area by atom type in complexed and uncomplexed structures [146], which enabled determination of specific atom-atom and residue-residue preferences. The most notable and widely applied PMF smoothed-grained potential was derived by Mügge and Martin from the statistical analysis of protein-ligand complexes found in the PDB [147]. Their list of atom pair potentials was correlated ($n = 77$, $R^2 = 0.61$) to binding free energy with a standard error of $< 2 \log K_i$ units. These PMF potentials have continued to evolve, largely by drastically increasing the sizes of the training sets [148, 149]. For all knowledge-based approaches, since they are calibrated on specific training sets, tunable parameters may require extensive adjustments before these functions can be extrapolated outside their training sets' chemical space.

2.6. Consensus scoring. The survey above of scoring functions included a brief look at several different classes, from Newtonian, first principles-based to semi-empirical, empirical and knowledge-based systems. While all achieve to a large extent their goals of discerning properly positioned ligands in binding sites and some reasonably useful ranking of ligand poses based on estimations of binding free energies, all scoring methods to date rely on approximations, in many cases due to the desire for faster throughput, and these lead to errors. However, it needs to be reiterated that even the most sophisticated methods of free energy estimation are not wholly accurate. In the absence of a single function that is valid for many or all targets [150], a strategy of merging several diverse scoring functions to form a consensus has emerged. This was proposed by Charifson [151] because the combination of diverse functions could potentially compensate for the weaknesses of individual functions, and thus yield better overall scoring. While other strategies such as post-docking minimization [93, 152] or topological filtering [153] The consensus scoring model has been widely accepted [154–156], There are three types of consensus methods: 1) rank by "vote", where only conformations within the top x%

are polled; 2) rank by rank, where the ranking from each function is averaged; and 3) rank by number, where the score value from each function is averaged.

Charifson's rank by vote strategy rescored compounds after docking/virtual screening with thirteen algorithms; the hit lists from each function were compared, and molecules appearing multiple times among the first several positions of each list were selected. A significant improvement in enrichment (discrimination between active and inactive compounds) and fewer false positives were observed in the hitlists filtered in this manner. An interesting observation by Bissanz *et al.*, after analyzing the performances of seven scoring functions, Chemscore, DOCK, FlexX, Fresno, GOLD, PMF and SCORE, with three different docking algorithms on the same systems was that enrichments are largely independent of the method of docking but impacted much more by the scoring methods [57].

The CScore program used a rank by rank strategy [157] (ChemScore, DockScore, GoldScore and PMF) and showed preferences for near-native conformations. Application of three or more functions works better than single or double scoring function methods. Wang *et al.* used a rank by number scheme in developing the X-CScore method [158]. Lastly, Stahl and Rarey [159] appropriated terms from different algorithms and merged them into a single scoring function termed ScreenScore, in order to exploit the strengths of each of the component methods.

3. Unresolved Issues in Docking and Scoring

The key problem with docking and scoring is that "chemistry" is not entirely representable with mathematics. Ligand binding is in many ways a balance between physical and chemical phenomena and the more chemical phenomena are more difficult to represent mathematically. Briefly, we will describe several potential problem areas with the docking and scoring paradigm, along with some discussion of solutions that have been applied in order to address them. Most of what was described above solved reasonably well the problem of creating ligand-biomacromolecule models complexes and scoring those models in terms of their "physical" realism, which is generally based on Newtonian mechanics. However, chemical subtleties that are not accounted for may be the key to identification of the one completely correct model.

The concept of emergent properties from systems biology is also in play. These properties are observed from collections of components acting in concert, but cannot be observed by a reductionist dissection back to the components. The hydrophobic effect is such a property, as hydrophobic species appear to self-associate due to the *actual* self-association of polar (hydrogen bonding) species. This property cannot be gleaned from the physical properties of a single atom. Similarly, chemical "reactions" such as changes in ionization states, resonance structures, or tautomerization are often not parameterized or otherwise represented in current docking

scoring functions. Another factor, which we alluded to several times above, is the flexibility of both the protein and the docked molecule. This is especially relevant in nuclear receptors where active sites are often quite loose. The flexibility problem can be resolved with physics, but at a high cost. Simulating flexibility in docking is a problem that has been attracting significant attention in the past decade. A related issue is that the relationship between molecular models of similar energy is poorly understood. The reflection data underlying crystal structure models are generally collected at very low temperatures – yielding a single conformer, while at biological temperatures many molecular conformations, ionization states, and etc. co-exist. In contrast, structural data collected with solution-phase NMR present a family of structural models, but their effective resolution is relatively poor. In terms of predicting binding modes and for virtual screening, Carlson has suggested that docking to an NMR structure ensemble has advantages over docking to a single structure for a molecular target [160].

In the next several paragraphs some of the issues above are discussed in more detail and solutions currently being employed are presented.

3.1. Hydrophobicity and entropy. Entropy clearly has a significant effect on all chemical processes, and biological binding events are no exception. Entropy is difficult to measure and even more difficult to model with calculations. The simple textbook definition of entropy is that it represents an increase of disorder. This is hard to envision in many cases and the consequences can even be counterintuitive. If two liquids like water and acetone are mixed, it is obvious that a random distribution of one in the other is more disordered: it would be impossible for them to be segregated. On the other hand, changing the mixed liquids to water and octanol would require the creation of highly ordered water cages around the octanol molecules or else they couldn't exist in water. Thus, finding the two liquids in separate layers is the more entropically favored state.

There are a number of effects in protein structure and protein-ligand binding attributable to entropy: 1) the displacement of water molecules from a binding site has the effect of increasing entropy because such solvent molecules have more degrees of freedom and are less ordered; 2) protein folding also has an entropic driving force because more waters are released in a compact protein structure with a hydrophobic core with more polar sidechains exposed to solvent. This explains the, at least partial, relationship between hydrophobic interactions and entropy. Also, calculations and analyses show that the packing of X-ray protein structure models is more "clever" than that of NMR or decoy models, with higher sidechain entropy than even models that are similarly compact [161].

The thermodynamics of ligand binding also involves – in a similar way – entropy, although it is difficult to quantify. Because most scoring functions treat hydrophobic

interactions in simple terms, like quantifying simply with contact surface areas, teasing out the interaction energy associated with entropy is impossible. Even HINT, based on the partitioning, i.e., free energy, of small organic molecules between two solvents is inadequate for identifying and quantifying the many, often small, effects that contribute to entropy. In contrast, free energy perturbation (FEP) methods exploit thermodynamic integration cycles – because differences in free energy are independent of path – and thus estimate binding free energy. Importantly, intermediate states on the path are chosen because of their computational accessibility. Nevertheless, these are difficult and expensive calculations, and at present not really applicable in docking or virtual screening experiments. FEP and related methods are also do not provide notably more accurate free energy predictions than empirical scoring functions that have been calibrated for the data set of interest.

3.2. Modeling water and its biology. Water is always present in the biological environment. It is a key ingredient for protein folding, as a solvent in biological reactions and in supporting protein-protein, protein-DNA and other associations. In protein-ligand associations, water molecules in the active site can mediate (i.e., bridge) interactions between the ligand and protein to change unfavorable interactions to favorable by their appropriate use of two hydrogen bond donors and two acceptors. Water can be thought of as a nano-buffer [162]. Water molecules shape the properties of an active site when they present different steric and electrostatic profile than the desolvated biomacromolecule. Despite several decades of investigation [163, 164], much about water is still poorly understood. While on average there is about one water molecule per residue in a protein, the quality or resolution of the structure has a large effect [165] with one found per residue at ~ 2.0 Å and more (> 1.5) at 1.0 Å resolutions [166]. In some cases, adding water molecules to structural models with computational prediction tools [46, 167–177] is a good strategy, especially if such a water is highly conserved and missing for experimental reasons such as X-ray resolution, etc.

The basic question in docking is determining what water molecules will be displaced or retained. In terms of medicinal chemistry, understanding the roles of water molecules in enzyme active sites can may assist in inhibitor design [178–184]. These roles can be evaluated in terms of a number of criteria such as count of water-involved hydrogen bonds, the crystallographic thermal B factors, accessible surface areas, and, most importantly, their conservation and/or displacement as a result of ligand binding. Experimental structural data are sometimes difficult compare with simulations because single water molecules cannot be observed or tracked experimentally.

Consolv [176] applies a K-nearest-neighbors genetic algorithm to identify conserved waters in biomacromolecule-ligand complexes. Its evaluation is based on geometric features that are in turn dependent on the quality of the examined crystallographic data. In practice, each water molecule's environment is characterized: geometrical features known to correlate with water binding suggest whether particular waters are conserved or displaced. While all *conserved* waters are important, and thus should be included modeling and docking, probably only those waters that are bridging or functionally displaceable are essential. Similar to Consolv, WaterScore [185] is based on geometrical parameters (B-factor, solvent-accessible surface area (SASA), number of protein-water contacts) and an energetic term that is the total hydrogen bond energy. WaterScore catalogues water as bound, sterically displaced and (otherwise) displaced. The factors indicated by WaterScore to suggest conservation also seem to be correlated with restricted water mobility. We have proposed a simpler model based on HINT score and Rank [186, 187], where the HINT score is calculated between each water and its site and Rank is only geometry-based, i.e., the potential of each water for making hydrogen bonds. This combination led to a mathematical model that predicts which water molecules will be displaced and which are “relevant”, with up to 90% accuracy. Relevant waters are potentially displaceable but they should be replaced by functional groups that reprise or improve upon the waters' hydrogen-bonding capabilities.

Ideally, docking programs, combined with their scoring functions *should* be able to: 1) insert a ligand into a solvated active site; 2) displace the irrelevant and non-conserved waters, but retain all others; and 3) estimate a reasonable free energy of binding for that ligand, which includes the contributions from the retained water molecules. This goal is not yet achieved, but progress is being made: 1) the FlexX particle algorithm places and continually optimizes water molecules by maximizing the number and quality of hydrogen bonds at the receptor-ligand interface during incremental construction docking simulations [188]; 2) SLIDE uses a desolvation penalty term for displacement to predict the retention or displacement of water molecules upon ligand binding [81, 189]; 3) GOLD [190] makes water molecules individually user switchable and allows them to rotate around their three main axes; 4) a new potential for displaceable water molecules was added to AutoDock by Moitissier *et al.* implemented by combining water grids into a continuum grid [191]; 5) van Dijk and Bonvin proposed a Monte Carlo procedure that applied water-mediated contact propensity to remove extraneous waters after docking completely solvated molecules [192]; and 6) Abel *et al.* described a method called WaterMap that measures the contribution of the solvent to the binding free energy of small molecules in their associated receptors including the effects of the ligand displacing the solvent in atomic detail [193], based on relatively short molecular dynamics simulations. While none of these treatments for water are ideal, they can yield

significant improvement in binding poses and in predicting binding free energies [81, 188, 190, 194–202].

3.3. Ionization and tautomerization. Modeling the protonation state of the system at the active site is another challenge. Since there are several ionizable residues with acidic or basic sidechains, potentially ionizable N- and C-termini, and the acidic or basic functional groups in ligands, there is a lot of room for ambiguity in intermolecular and intramolecular interactions. The ionization states of nearly all functional groups are well understood when they are isolated, and usually their pK_a s are known. However, the pH properties of the microenvironments within or on the surface of proteins and at protein-ligand interfaces are less understood and considerably more difficult to model.

Importantly, crystal structures seldom reveal direct information on the positions of their hydrogen atoms. Only for the highest resolution X-ray structures, at ~ 1.0 Å or better and for neutron diffraction structures are hydrogens observed, and usually those sets are incomplete because the H atom is small. Thus, for ionizable groups, both for proteins and ligands in many environments, it is usually quite difficult to assign their ionization states because such groups are not isolated and influence each other's states. It goes without saying that accurate docking and scoring depends on a definitive assignment of each atom's charge, etc. The tautomerization of specific functional groups on the ligand can also be affected by either the solution pH or the microenvironment pH, which can have a similar effect on the binding of a ligand in the site. Adding to the complexity, the number of model possibilities, i.e., protonation states for each ionizable residue in an ensemble, grows exponentially as the number of these groups increases. An active site commonly has several protonation state models that exist in equilibrium with each other, and a protein-protein interface may have hundreds. Water molecules present in the active site further complicate the problem because of their ability to act as Lewis bases and/or acids. Computational studies on a binding pocket in the engineered cavity of cytochrome c peroxidase [203], showed better docking results when a single conserved water was allowed to move. Crystallographers and computational chemists have been interested in the proper placement and optimization of polar protons for decades. The correction of atom placement, coupled with the assignment of chemically and biologically reasonable protonation states for ionizable residues in proteins is a necessary prerequisite for docking experiments. There are a few web-enabled applications for examining and validating X-ray protein crystal structures such as MolProbity [204], which both places hydrogen atoms and corrects various other errors in structure models.

Computational ionization state evaluation and optimization is a complex problem to which quantum mechanics [205], quantum mechanics–molecular mechanics (QM/MM) [206, 207] approaches have been applied. However, solving

the Poisson-Boltzmann equation to evaluate possible protonation states and hydrogen positions [208–212] has generally been the method of choice, sometimes when combined with MD simulation [213–215]. The speed of PBE solutions has been a handicap for general use as a scoring function, but optimized code such as OpenEye's Zap, and faster computers have been narrowing the gap. The goal of PBE methods is also somewhat orthogonal to docking – in most applications the result is a set of pK_a s for protein residues rather than the set of energetically accessible protonation states. Only a few docking programs tackle the ligand tautomerization problem, e.g., eHITS and GLIDES's LigPrep. For each group, all possible tautomerization and ionization cases are simulated and evaluated, with the best for each selected independent of the others, thus avoiding, rather than dealing with, the inherent combinatorial issues [216]. However, most docking programs require the input of all tautomers by the user. This docking problem has to date received only modest attention, and remains a problem that must be solved before universally reliable docking, scoring or virtual screening will be available.

3.4. Flexibility. All proteins are flexible and this often turns out to be fundamental for their function. Significant conformational flexibility is vital in nuclear receptors for their biological function, as the pharmacology of their ligands involves movements of short α -helix segments at carboxy termini [39, 57, 217]. While structure-based molecular design has been a relatively successful strategy in the past without overly respecting target flexibility, future efforts are not likely to be as fruitful. The current content of the Protein Data Bank (PDB) [218] likely overemphasizes rigid proteins because these are easier to crystallize and analyze. Furthermore, disorder is reduced by collecting diffraction data at low (\sim liquid N_2) temperatures, which is in stark contrast to biology that generally occurs in water at or above room temperature. Thus, even though X-ray crystallography is the best technique to obtain experimental models of biomacromolecules, these models are static and often reveal little about their flexibility –likely to be key for their biological activity. Nuclear Magnetic Resonance (NMR) spectroscopy, however, solves three-dimensional protein structures in solution and thus encodes dynamic flexibility of molecules. NMR also provides an ensemble of conformations with low energy. Currently proteins of 80–100 kD are amenable for studying with NMR [219, 220], and is especially useful for systems that are difficult to crystallize or when the crystals obtained are disordered.

Molecular dynamics (MD) calculations are able to use the structural framework provided by experimental X-ray crystallography and NMR data to extrapolate at least a partial understanding of molecular motions. Thus, MD can generate many protein conformations for docking and/or virtual screening studies [221–223]. Various scales are available from these approaches, usually dictated by the

nature and size of the system in the study. Coarse-grained approaches smooth out the atomistic details in order to make longer time-scale dynamics more accessible, and thus reveal general protein flexibility. Such dynamics can often model protein dynamics, i.e., formation or reshaping of binding pockets, etc., but without atom-level detail, are of insufficient quality for docking. In atomic-level MD, both small atomic fluctuations and large protein movements can be observed, but at higher computational cost. The cost is so large that effective sampling of conformational states is often in question because even several ns is too brief to ensure adequate sampling. In recent years, MD methods have become more ubiquitously accessible, but experimentally determined conformational data remain more desirable.

In simulating docking, integration of detailed flexibility defeats, in a sense, the rationale behind docking – speed and the ability to explore multiple conformations and many putative ligands. However, true protein conformations are often stabilized by the bound ligand, which makes the computational simulation a bit like the chicken and egg. In designing ligands for such targets, it cannot be known beforehand how the target will adapt to binding of a particular ligand. In the past few years, much effort has been expended to simulate active site flexibility in docking. The most conservative methods incorporate flexibility with active-site sidechains, e.g., AutoDock v4 [54, 84, 126, 224] explicitly allows such flexibility, GOLD samples several alternative rotamers for specified sidechains and optimize hydrogen-bond interactions involving these residues [136], and SLIDE rotates sidechains and substructures of ligands to remove unfavorable van der Waals interactions [82]. Simulating sidechain flexibility appears to be most effective when the docking site has *not* been preformed, e.g., into a un-liganded structure. To simulate significant conformational adjustments, an induced fit docking approach such as suggested by Gohlke based on rigidity and elastic network theory [225], can be effective. Induced fit docking simulations support the “conformation selection” concept [226, 227], where observed bound conformations may be predetermined by motions of the protein in the unbound state. Lastly, ensembles of target structures, usually derived from MD, can be used for docking [228]. Ensemble docking places the ligand of interest into each structural snapshot from an MD trajectory, or into a grid that encodes the composite similarity from the ensemble. The disadvantages of ensemble docking are that the ensemble of conformers is seldom complete or even representative, and the shortfalls of scoring functions in representing free energy are magnified here because binding mode differences are often dominated by entropy.

4. Example Case Studies of Estrogen Receptor Toxicology

As stated above, a good understanding of biological interactions (and even better a good estimation of them!), is the basis

for successful medicinal chemistry investigations, even those driven by experimental results. In the same way, application of well-calibrated and understood computational tools can be part of the process for creating and producing safer food and food-related products.

Many current computational investigations are focused on the identification of endocrine-disrupting chemicals (EDCs), which are exogenous compounds capable of interfering with the human endocrine system either by direct or indirect interactions with nuclear receptor proteins. Common targets are often estrogen receptors (ERs) that, once activated or inhibited by interaction with such compounds, called xenoestrogens, alter the normal homeostatic transcription and signaling pathways.

4.1. Docking and scoring. One of the first applications of computational methods in food science was reported by Amadasi *et al.* (2009) [3], with the aim to identify possible xenoestrogens in a set of allowed and common food additives. These were surveyed by virtually screening a food additive list maintained by the Food and Agriculture Organization of the United Nations (FAO) and World Health Organization (WHO) that contained nearly 500 chemicals with their chemical structures. These were each docked/scored in the alpha estrogen receptor using the GOLD algorithm and the in-built GOLD scoring functions as filters. Those compounds passing through these filters were re-scored with HINT. The most “suspicious” molecules, that is, those that better fit the receptor’s binding site and had the highest HINT scores, were selected for experimental testing in an *in vitro* assay. Among the thirteen potential xenoestrogens thus identified for testing, four were already known to exhibit estrogenic activity. The assays of the other nine compounds determined their binding affinities to the receptor and, by inference, their resulting biological effects. In this way, propyl gallate was found to act as an antagonist, while 4-hexylresorcinol as a potent transactivator. Both ligands were active at nanomolar concentrations, as predicted by the *in silico* analysis.

Similar techniques, i.e., GOLD and HINT scoring tools, were applied in investigating the biological effect of the known food contaminant zearalenone and of its sulphonate and glucuronate metabolites [229, 230]. Zearalenone is a known potent micotoxin produced by several *Fusarium* species and can be found in different cereal crops such as barley, wheat, rice and others. Again, it affects estrogen receptor activation by acting as an agonist and inducing gene transcription. *In silico* analyses predicted the molecules to be active towards the alpha estrogen receptor, in good agreement with experimental evidence. These findings further supported the fruitful application of computational technique to support and illuminate experimental results, and in particular to possibly reduce the number of and/or replace animal experiments.

The alpha and beta estrogen receptors were also the targets of an innovative “full-dry” *in silico* method for evaluation of

the potential xenoestrogenic behavior of ellagitannin-derived metabolites. The strategy involved structure-based virtual screening, docking simulations followed by re-scoring, and the results revealed valuable insights into the phase II conjugations (glucuronidation, sulfation, and methylation, occurring *in vivo*) affecting the estrogenicity of these compounds on both α - and β -ER isoforms [231]. Structure-based docking techniques were adopted to predict the possible effects of tioxhantone (TX) derivatives that had been identified as food contaminants, especially in infant formulas, by the European Food Safety Authority. Phase I metabolites for the TX class of compounds were predicted and their binding affinities for the AR ligand-binding pocket were defined using local models based on available information about metabolism and AR activity. The results confirmed that the precursors can bind to the ER targets and suggested some of their metabolites to be possible new disruptors suitable for further *in vitro* and/or *in vivo* toxicological evaluations [232].

With a similar goal, Ng *et al.* combined agonist and antagonist docking models with competitive docking, enabling the evaluation of potential ER biological functional changes due to chemicals binding to the receptor. This also allowed the assessment of a chemical's endocrine disrupting potential. The potential of these applications is impressive, considering their applicability in risk assessment on potential EDCs, but also for the identification of potential agonists and antagonists in drug discovery projects.[233]

Alternariol and alternariol methyl ether, produced by different *Alternaria* species, have been recognized as another set of emerging micotoxins. *In silico* docking, scoring and re-scoring strategies were applied to investigate first whether these molecules could generate genotoxic effects by the activation of DNA topoisomerase I and/or II [229]. These compounds, alternariol and its derivatives, may also have interactions with nuclear receptors such as the alpha and beta estrogen receptors. This possibility should also be assessed. In summary, the fairly simple computational techniques of virtual screening, docking and/or scoring have revealed a lot of experimentally difficult-to-obtain data. We are not yet at the stage where computational procedures can totally replace well-constructed and executed experiments, but just the promising results presented here argue for the importance of including them in risk assessment investigations.

4.2. Molecular dynamics. The purpose of molecular dynamics (MD) methods is to simulate motion, which is a very important aspect for the *very* flexible nuclear receptors like estrogen receptor. Many more of these studies have been published and many more are underway, but we highlight here two recent examples. The best MD studies are incorporated into a comprehensive work that uses conformational information from the dynamics as the starting place for virtual screening, docking, etc. Molecular dynamics techniques were used by Li *et al.* to investigate the binding mechanism of bisphenol A (BPA) to three typical nuclear receptors: ER α ,

ERR γ and PPAR γ . Bisphenol A is commonly present in food packaging and is known to affect the normal function of nuclear receptors at very low doses. The simulations provided structural support for the capability of this xenoestrogen to bind to these receptors, where it appears to mimic the action of the natural hormone and keeps the nuclear receptors in active conformations [234]. These results could form the basis for identification of safer alternatives for BPA and safer packaging.

Dal Palú *et al.* recently described the development of alternative and innovative approaches for simulation of nuclear receptor (NR) intrinsic dynamics. The authors reported a fast strategy based on logic constraint programming and applied it to a study of the flexibility of α -ER alpha. They investigated, in particular, the routes possibly followed by the receptor to transform from the inactive to the agonist- and antagonist-active states. Also, they presented a number of energetically accessible conformations that could possibly be used in structure-based drug design for the identification of new potential binders and disruptors [4]. Their methodology should be reasonably and successfully applied to any other nuclear receptor.

5. Conclusions

In this last section, we make a few general comments and then discuss in more detail the relevance of docking, and in particular scoring functions, to receptors likely targeted by food and feed molecules as, for instance, nuclear receptors. Docking had rather modest beginnings, and perhaps its inventors could never have imagined the current broad applications of their concept. The docking paradigm has become an almost universal tool for computational chemistry applied to drug discovery and in probing other small molecule-biomacromolecule interactions. There are numerous success stories, many of which have been presented in this article's references.

A fundamental question that might be asked is, are there multiple "correct" solutions to a docking problem? Both NMR and MD suggest that there are more often than not multiple, energetically indistinct, conformers. Remember, the crystal "structure" is also a model that is *only* special because it fits the experimental electron density [6]. That model is unlikely to be unique, and other conformers, tautomers, variants of ionization states, and etc. may also fit the electron density well. Even the wrong molecule may have been modeled into the density, because of a mistake or after an in-situ reaction. Thus, believing the top scoring candidate model is the only one of value, while dumping others with similar score, may lead to erroneous results.

It is important to keep in mind that running a docking program is an experiment, albeit computational. To succeed, the assumptions and conditions of the experiment should be evaluated and the correct controls should be run. We have usually viewed any computational result with suspicion

until another approach, preferably experimental, can be applied to the problem. Concomitantly, because running the same experiment the same way will yield the same result (precisely), to evaluate uncertainty in docking, different methods should be applied, such as changes in conditions, parameters, docking program, or especially scoring function. In “wet” experiments the observed results are not from a single, discrete molecule interacting with another, but from many molecules of the first species interacting with many of the other. Should we assume that the one-to-one reactions we model with docking calculations are representative of experiment? The only consolation is that all docking algorithms rely on some degree of randomness to drive their calculation engines. The best rule of thumb is to not treat computational results as “answers” to problems, but to see them clues or hypothesis generators for designing experiments.

In our perspective, *in silico* experiments, such as virtual screening, docking and scoring, and etc. represent invaluable tools to investigate the possible interactions of chemical small molecules or biological molecules with relevant receptors. Nuclear Receptors have become increasingly relevant targets of study for both pharmaceuticals and food chemicals, from hit to lead compounds, from xenostrogens, to nutraceuticals, mycotoxins, inks for food packaging, etc. However, a rapidly emerging concern, for environmental, ethical as well as financial reasons, is to reduce the frequency and number of *in vitro* and *in vivo* tests. Thus, *in silico* experiments have become mandatory for the investigation of large libraries of compounds in order to predict their possible interactions with the mentioned targets. Follow up experiments are then much more focused. Finally, as computational approaches to these problems are beginning to enter the regulatory domain, computational practices will have to be well established and properly applied, and the user (and regulatory officials) must be conscious of the pros and cons of the various tools and methodologies, and above all avoid the use of such software as a “magic black box”.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

The authors acknowledge the inspiration of Prof. Donald J. Abraham (VCU) and Prof. Andrea Mozzarelli (U. Parma), who sparked our collaboration because of their mutual interest in structural biology and drug discovery.

References

- [1] *From Medicinal Chemistry to Food Science: A Transfer of In Silico Methods Applications*, Nova Science Publisher, 2015.
- [2] T. Hartung and S. Hoffmann, Food for thought... on *in silico* methods in toxicology, *Altex*, **26**, no. 3, 155–166, (2009).
- [3] A. Amadasi, A. Mozzarelli, C. Meda, A. Maggi, and P. Cozzini, Identification of xenoestrogens in food additives by an integrated *in silico* and *in vitro* approach, *Chemical Research in Toxicology*, **22**, no. 1, 52–63, (2009).
- [4] A. Dal Palu, F. Spyraakis, and P. Cozzini, A new approach for investigating protein flexibility based on Constraint Logic Programming. The first application in the case of the estrogen receptor, *European Journal of Medicinal Chemistry*, **49**, 127–140, (2012).
- [5] A. Vedani, M. Dobler, Z. Hu, and M. Smiesko, OpenVirtualToxLab—A platform for generating and exchanging *in silico* toxicity data, *Toxicol Lett*, **232**, 519–532, (2015).
- [6] P. C. D. Hawkins, G. L. Warren, A. G. Skillman, and A. Nicholls, How to do an evaluation: Pitfalls and traps, *Journal of Computer-Aided Molecular Design*, **22**, no. 3-4, 179–190, (2008).
- [7] F. Spyraakis, A. Amadasi, M. Fornabaio, D. J. Abraham, A. Mozzarelli, G. E. Kellogg, and P. Cozzini, The consequences of scoring docked ligand conformations using free energy correlations, *European Journal of Medicinal Chemistry*, **42**, no. 7, 921–933, (2007).
- [8] P. W. Atkins, *General Chemistry*, Zanichelli, Bologna, 1992.
- [9] G. A. Jeffrey and W. Saenger, *Hydrogen bonding in biological structures*, Springer-Verlag, Berlin, 1994.
- [10] M. Nishio, *The CH-Pi interactions: evidence, nature and consequences*, Wiley-VCH, New York, 1998.
- [11] H. Gohlke and G. Klebe, Approaches to the Description and Prediction of the Binding Affinity of Small-Molecule Ligands to Macromolecular Receptors, *Angewandte Chemie International Edition*, **41**, no. 15, 2644–2676, (2002).
- [12] D. A. Pearlman and P. A. Kollman, The calculated free energy effects of 5-methyl cytosine on the B to Z transition in DNA, *Biopolymers*, **29**, no. 8-9, 1193–1209, (1990).
- [13] J. R. H. Tame, Scoring functions: A view from the bench, *Journal of Computer-Aided Molecular Design*, **13**, no. 2, 99–108, (1999).
- [14] I. Muegge, Effect of Ligand Volume Correction on PMF Scoring, *Journal of Computational Chemistry*, **22**, no. 4, 418–425, (2001).
- [15] K. A. Dill, Additivity Principles in Biochemistry, *The Journal of Biological Chemistry*, **272**, 701–704, (1997).
- [16] A. M. Davis, S. J. Teague, and G. J. Kleywegt, Application and Limitations of X-ray Crystallographic Data in Structure-Based Ligand and Drug Design, *Angewandte Chemie International Edition*, **42**, no. 24, 2718–2736, (2003).
- [17] M. A. DePristo, P. I. W. De Bakker, and T. L. Blundell, Heterogeneity and inaccuracy in protein structures solved by X-ray crystallography, *Structure*, **12**, no. 5, 831–838, (2004).
- [18] C.-I. Bränden and T. Alwyn Jones, Between objectivity and subjectivity, *Nature*, **343**, no. 6260, 687–689, (1990).
- [19] V. Mohan, A. C. Gibbs, M. D. Cummings, E. P. Jaeger, and R. L. DesJarlais, Docking: Successes and challenges, *Current Pharmaceutical Design*, **11**, no. 3, 323–333, (2005).
- [20] C. Giacobuzzo, H. L. Monaco, D. Viterbo, F. Scordari, G. Gilli, G. Zanotti, and M. Catti, *Fundamentals of crystallography*, Oxford University Press, New York, 1992.
- [21] A. T. Brunger, Free R value: A novel statistical quantity for assessing the accuracy of crystal structures, *Nature*, **355**, no. 6359, 472–475, (1992).

- [22] O. Roche, R. Kiyama, and C. L. Brooks, Ligand-Protein DataBase: Linking Protein-Ligand Complex Structures to Binding Data, *Journal of Medicinal Chemistry*, **44**, no. 22, 3592–3598, (2001).
- [23] R. Wang, X. Fang, Y. Lu, and S. Wang, The PDBbind Database: Collection of Binding Affinities for Protein-Ligand Complexes with Known Three-Dimensional Structures, *Journal of Medicinal Chemistry*, **47**, 2977–2980, (2004).
- [24] R. Wang, X. Fang, Y. Lu, C. Y. Yang, and S. Wang, The PDBbind Database: Methodologies and Updates, *Journal of Medicinal Chemistry*, **48**, 4111–4119, (2005).
- [25] X. Chen, Y. Lin, and M. K. Gilson, The binding database: Overview and user's guide, *Biopolymers*, **61**, no. 2, 127–141, (2001).
- [26] L. Hu, M. L. Benson, R. D. Smith, M. G. Lerner, and H. A. Carlson, Binding MOAD (Mother of All Databases), *Proteins: Structure, Function and Genetics*, **60**, no. 3, 333–340, (2005).
- [27] J. W. M. Nissink, C. Murray, M. Hartshorn, M. L. Verdonk, J. C. Cole, and R. Taylor, A new test set for validating predictions of protein-ligand interaction, *Proteins: Structure, Function and Genetics*, **49**, no. 4, 457–471, (2002).
- [28] T. Cheng, Q. Li, Z. Zhou, Y. Wang, and S. H. Bryant, Structure-based virtual screening for drug discovery: A problem-centric review, *AAPS Journal*, **14**, no. 1, 133–141, (2012).
- [29] H. J. Bohm, New computational approaches to predict protein-ligand interactions, in *In Structure-based ligand design*, R. Mannhold, H. Kubinyi, and H. Timmerman, Eds., **6**, Wiley-VCH: Weinheim, 1998.
- [30] M. A. Murcko, Recent advances in ligand design methods, in *In Reviews in Computational Chemistry*, Lipkowitz, K. a. B. DB, Ed., **11**, p. 66, Wiley-VCH, John Wiley and Sons, Inc, New York, 1997.
- [31] D. Clark, C. W. Murray, and J. Li. Current issues, in de novo molecular design, in *In Reviews in Computational Chemistry*, Lipkowitz, K. a. B. DB, Ed., **11**, 67–125, Wiley-VCH, John Wiley and Sons, Inc, New York, 1997.
- [32] H. I. Parikh and G. E. Kellogg, Intuitive, but not simple: Including explicit water molecules in protein-protein docking simulations improves model quality, *Proteins: Structure, Function and Bioinformatics*, **82**, no. 6, 916–932, (2014).
- [33] J. M. Blaney and J. S. Dixon, A good ligand is hard to find: Automated docking methods, *Perspectives in Drug Discovery and Design*, **1**, no. 2, 301–319, (1993).
- [34] R. D. Taylor, P. J. Jewsbury, and J. W. Essex, A review of protein-small molecule docking methods, *Journal of Computer-Aided Molecular Design*, **16**, no. 3, 151–166, (2002).
- [35] C. A. Sotriffer, G. Klebe, M. Stahl, and H. J. Bohm, Docking and Scoring Functions/Virtual Screening, in *In Burger's Medicinal Chemistry and Drug Discovery*, D. J. Abraham, Ed., **1**, 281–331, John Wiley Sons, Inc, 2003.
- [36] I. Muegge and I. Rarey, Small molecule docking and scoring, in *In Reviews Comput Chem*, K. Lipkowitz and D. Boyd, Eds., **17**, p. 61, Wiley-VCH, John Wiley and Sons, Inc, New York, 2001.
- [37] N. Brooijmans and I. D. Kuntz, Molecular Recognition and Docking Algorithms, *Annual Review of Biophysics and Biomolecular Structure*, **32**, 335–373, (2003).
- [38] K. A. Henzler-Wildman, M. Lei, V. Thai, S. J. Kerns, M. Karplus, and D. Kern, A hierarchy of timescales in protein dynamics is linked to enzyme catalysis, *Nature*, **450**, no. 7171, 913–916, (2007).
- [39] C. Hellal-Levy, J. Fagart, A. Souque, J.-M. Wurtz, D. Moras, and M.-E. Rafestin-Oblin, Crucial role of the H11-H12 loop in stabilizing the active conformation of the human mineralocorticoid receptor, *Molecular Endocrinology*, **14**, no. 8, 1210–1221, (2000).
- [40] X. Hu and M. A. Lazar, Transcriptional Repression by Nuclear Hormone Receptors, *Trends in Endocrinology & Metabolism*, **11**, no. 1, 6–10, (2000).
- [41] A. E. Brady and L. E. Limbird, G protein-coupled receptor interacting proteins: Emerging roles in localization and signal transduction, *Cellular Signalling*, **14**, no. 4, 297–309, (2002).
- [42] P. Cozzini, G. E. Kellogg, F. Spyralis, D. J. Abraham, G. Costantino, A. Emerson, F. Fanelli, H. Gohlke, L. A. Kuhn, G. M. Morris, M. Orozco, T. A. Pertinhez, M. Rizzi, and C. A. Sotriffer, Target flexibility: An emerging consideration in drug discovery and design, *Journal of Medicinal Chemistry*, **51**, no. 20, 6237–6255, (2008).
- [43] C. Levinthal, S. J. Wodak, P. Kahn, and A. K. Dadivanian, Hemoglobin interaction in sickle cell fibers I: Theoretical approaches to the molecular contacts, *Proceedings of the National Academy of Sciences of the United States of America*, **72**, no. 4, 1330–1334, (1975).
- [44] S. J. Wodak, M. De Crombrughe, and J. Janin, Computer studies of interactions between macromolecules, *Progress in Biophysics and Molecular Biology*, **49**, 29–63, (1987).
- [45] R. Norel, S. L. Lin, H. J. Wolfson, and R. Nussinov, Shape complementarity at protein-protein interfaces, *Biopolymers*, **34**, no. 7, 933–940, (1994).
- [46] P. J. Goodford, Aldose reductase inhibitors: a potential new class of agents for the pharmacological control of certain diabetic complications, *Journal of Medicinal Chemistry*, **28**, no. 7, 849–857, (1985).
- [47] E. Meng, B. Shoichet, and I. Kuntz, Automated docking with grid-based energy evaluation, *Journal of Computational Chemistry*, **13**, no. 4, 505–524, (1992).
- [48] M. Totrov and R. Abagyan, Flexible protein-ligand docking by global energy optimization in internal coordinates, *Proteins: Structure, Function and Genetics*, **29**, no. 1, 215–220, (1997).
- [49] J. Y. Trosset and H. A. Scheraga, Reaching the global minimum in docking simulations: A Monte Carlo energy minimization approach using Bezier splines, *Proceedings of the National Academy of Sciences of the United States of America*, **95**, 8011–8015, (1998).
- [50] K. E. Platzter, F. A. Momany, and H. A. Scheraga, Conformational energy calculations of enzyme-substrate interactions. I. Computation of preferred conformations of some substrates of -chymotrypsin, *International Journal of Peptide and Protein Research*, **4**, no. 3, 187–200, (1972).
- [51] I. D. Kuntz, J. M. Blaney, S. J. Oatley, R. Langridge, and T. E. Ferrin, A geometric approach to macromolecule-ligand interactions, *Journal of Molecular Biology*, **161**, no. 2, 269–288, (1982).
- [52] D. T. Moustakas, P. T. Lang, S. Pegg, E. Pettersen, I. D. Kuntz, N. Brooijmans, and R. C. Rizzo, Development and validation of a modular, extensible docking program: DOCK 5, *Journal of Computer-Aided Molecular Design*, **20**, no. 10-11, 601–619, (2006).
- [53] C. M. Oshiro, I. D. Kuntz, and J. S. Dixon, Flexible ligand docking using a genetic algorithm, *Journal of Computer-Aided Molecular Design*, **9**, no. 2, 113–130, (1995).
- [54] G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, and A. J. Olson, Automated docking using

- a Lamarckian genetic algorithm and an empirical binding free energy function, *Journal of Computational Chemistry*, **19**, no. 14, 1639–1662, (1998).
- [55] C. N. Cavasotto and R. A. Abagyan, Protein Flexibility in Ligand Docking and Virtual Screening to Protein Kinases, *Journal of Molecular Biology*, **337**, no. 1, 209–225, (2004).
- [56] K. W. Foreman, A. T. Phillips, J. B. Rosen, and K. A. Dill, Comparing search strategies for finding global optima on energy landscapes, *Journal of Computational Chemistry*, **20**, no. 14, 1527–1532, (1999).
- [57] C. Bissantz, G. Folkers, and D. Rognan, Protein-based virtual screening of chemical databases. 1. Evaluation of different docking/scoring combinations, *Journal of Medicinal Chemistry*, **43**, no. 25, 4759–4767, (2000).
- [58] B. D. Bursulaya, M. Totrov, R. Abagyan, and C. L. Brooks III, Comparative study of several algorithms for flexible ligand docking, *Journal of Computer-Aided Molecular Design*, **17**, no. 11, 755–763, (2003).
- [59] R. T. Kroemer, A. Vulpetti, J. J. McDonald, D. C. Rohrer, J.-Y. Trosset, F. Giordanetto, S. Cotesta, C. McMartin, M. Kihlén, and P. F. W. Stouten, Assessment of docking poses: Interactions-based accuracy classification (IBAC) versus crystal structure deviations, *Journal of Chemical Information and Computer Sciences*, **44**, no. 3, 871–881, (2004).
- [60] J. C. Cole, C. W. Murray, J. W. M. Nissink, R. D. Taylor, and R. Taylor, Comparing protein-ligand docking programs is difficult, *Proteins: Structure, Function and Genetics*, **60**, no. 3, 325–332, (2005).
- [61] G. L. Warren, C. W. Andrews, A.-M. Capelli, B. Clarke, J. LaLonde, M. H. Lambert, M. Lindvall, N. Nevins, S. F. Semus, S. Senger, G. Tedesco, I. D. Wall, J. M. Woolven, C. E. Peishoff, and M. S. Head, A critical assessment of docking programs and scoring functions, *Journal of Medicinal Chemistry*, **49**, no. 20, 5912–5931, (2006).
- [62] J.-F. Truchon and C. I. Bayly, Evaluating virtual screening methods: Good and bad metrics for the “early recognition” problem, *Journal of Chemical Information and Modeling*, **47**, no. 2, 488–508, (2007).
- [63] P. Englebienne, H. Fiaux, D. A. Kuntz, C. R. Corbeil, S. Gerber-Lemaire, D. R. Rose, and N. Moitessier, Evaluation of docking programs for predicting binding of Golgi α -mannosidase II inhibitors: A comparison with crystallography, *Proteins: Structure, Function and Genetics*, **69**, no. 1, 160–176, (2007).
- [64] N. Weill, E. Therrien, V. Campagna-Slater, and N. Moitessier, Methods for docking small molecules to macromolecules: A user’s perspective. 1. The theory, *Current Pharmaceutical Design*, **20**, no. 20, 3338–3359, (2014).
- [65] J. Biesiada, A. Porollo, P. Velayutham, M. Kouril, and J. Meller, Survey of public domain software for docking simulations and virtual screening, *Human Genomics*, **5**, no. 5, 497–505, (2011).
- [66] N. M. F. S. A. Cerqueira, D. Gesto, E. F. Oliveira, D. Santos-Martins, N. F. Brás, S. F. Sousa, P. A. Fernandes, and M. J. Ramos, Receptor-based virtual screening protocol for drug discovery, *Archives of Biochemistry and Biophysics*, **582**, article no. 6986, 56–67, (2015).
- [67] S. Vajda, D. R. Hall, and D. Kozakov, Sampling and scoring: A marriage made in heaven, *Proteins: Structure, Function and Bioinformatics*, **81**, no. 11, 1874–1884, (2013).
- [68] J. C. Wang and J. H. Lin, Scoring Functions for Prediction of Protein-Ligand Interactions, *Curr Pharm Des*, **19**, 2174–2182, (2013).
- [69] H. A. Gabb, R. M. Jackson, and M. J. E. Sternberg, Modelling protein docking using shape complementarity, electrostatics and biochemical information, *Journal of Molecular Biology*, **272**, no. 1, 106–120, (1997).
- [70] M. R. McGann, H. R. Almond, A. Nicholls, J. A. Grant, and F. K. Brown, Gaussian docking functions, *Biopolymers*, **68**, no. 1, 76–90, (2003).
- [71] G. B. McGaughey, R. P. Sheridan, C. I. Bayly, J. C. Culberson, C. Kreatsoulas, S. Lindsley, V. Maiorov, J.-F. Truchon, and W. D. Cornell, Comparison of topological, shape, and docking methods in virtual screening, *Journal of Chemical Information and Modeling*, **47**, no. 4, 1504–1519, (2007).
- [72] D. Joseph-McCarthy, B. E. Thomas IV, M. Belmarsh, D. Moustakas, and J. C. Alvarez, Pharmacophore-based molecular docking to account for ligand flexibility, *Proteins: Structure, Function and Genetics*, **51**, no. 2, 172–188, (2003).
- [73] J. Goto, R. Kataoka, and N. Hirayama, Ph4Dock: Pharmacophore-based protein - Ligand docking, *Journal of Medicinal Chemistry*, **47**, no. 27, 6804–6811, (2004).
- [74] A. Schafferhans and G. Klebe, Docking ligands onto binding site representations derived from proteins built by homology modelling, *Journal of Molecular Biology*, **307**, no. 1, 407–427, (2001).
- [75] D. Schneidman-Duhovny, Y. Inbar, R. Nussinov, and H. J. Wolfson, PatchDock and SymmDock: Servers for rigid and symmetric docking, *Nucleic Acids Research*, **33**, no. 2, W363–W367, (2005).
- [76] A. Wallqvist and D. G. Covell, Docking enzyme-inhibitor complexes using a preference-based free-energy surface, *Proteins: Structure, Function and Genetics*, **25**, no. 4, 403–419, (1996).
- [77] <http://www.eyesopen.com/products/applications/omega.html>.
- [78] A. R. Leach and I. D. Kuntz, Conformational analysis of flexible ligands in macromolecular receptor sites, *J Comput Chem*, **13**, 730–748, (1992).
- [79] M. Rarey, B. Kramer, T. Lengauer, and G. Klebe, A fast flexible docking method using an incremental construction algorithm, *Journal of Molecular Biology*, **261**, no. 3, 470–489, (1996).
- [80] A. N. Jain, Ligand-Based Structural Hypotheses for Virtual Screening, *Journal of Medicinal Chemistry*, **47**, 947–961, (2004).
- [81] V. Schnecke and L. A. Kuhn, Virtual screening with solvation and ligand-induced complementarity, *Perspectives in Drug Discovery and Design*, **20**, 171–190, (2000).
- [82] M. I. Zavodszky and L. A. Kuhn, Side-chain flexibility in protein-ligand binding: The minimal rotation hypothesis, *Protein Science*, **14**, no. 4, 1104–1114, (2005).
- [83] M. Totrov and R. Abagyan, Drug-Receptor Thermodynamics: Introduction and Applications, in *R. Abagyan. Drug-Receptor Thermodynamics: Introduction and Applications*, p. 603, John Wiley E Sons: Chichester, 2001.
- [84] D. S. Goodsell and A. J. Olson, Automated docking of substrates to proteins by simulated annealing, *Proteins: Structure, Function and Genetics*, **8**, no. 3, 195–202, (1990).
- [85] T. N. Hart and R. J. Read, A multiple-start Monte Carlo docking method, *Proteins: Structure, Function and Genetics*, **13**, no. 3, 206–222, (1992).
- [86] M. Liu and S. Wang, MCDOCK: A Monte Carlo simulation approach to the molecular docking problem, *Journal of Computer-Aided Molecular Design*, **13**, 435–451, (1999).

- [87] C. Dominguez, R. Boelens, and A. M. J. J. Bonvin, HADDOCK: A protein-protein docking approach based on biochemical or biophysical information, *Journal of the American Chemical Society*, **125**, no. 7, 1731–1737, (2003).
- [88] J.-Y. Trosset and H. A. Scheraga, PRODOCK: Software package for protein modeling and docking, *Journal of Computational Chemistry*, **20**, no. 4, 412–427, (1999).
- [89] S. D. Schrödinger Inc., CA 92122.
- [90] G. Jones, P. Willett, R. C. Glen, A. R. Leach, and R. Taylor, Development and validation of a genetic algorithm for flexible docking, *Journal of Molecular Biology*, **267**, no. 3, 727–748, (1997).
- [91] G. Jones, P. Willett, and R. C. Glen, A genetic algorithm for flexible molecular overlay and pharmacophore elucidation, *Journal of Computer-Aided Molecular Design*, **9**, no. 6, 532–549, (1995).
- [92] G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, and A. J. Olson, Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function, *Journal of Computational Chemistry*, **19**, no. 14, 1639–1662, (1998).
- [93] G. Wu, D. H. Robertson, C. L. Brooks III, and M. Vieth, Detailed analysis of grid-based molecular docking: A case study of CDocker—A CHARMM-based MD docking algorithm, *J Comput Chem*, **24**, no. 13, 1549–1562, (2003).
- [94] A. Miranker and M. Karplus, Functionality maps of binding sites: A multiple copy simultaneous search method, *Proteins: Structure, Function and Genetics*, **11**, no. 1, 29–34, (1991).
- [95] A. Di Nola, D. Roccatano, and H. J. C. Berendsen, Molecular dynamics simulation of the docking of substrates to proteins, *Proteins: Structure, Function and Genetics*, **19**, no. 3, 174–182, (1994).
- [96] C. A. Baxter, C. W. Murray, D. E. Clark, D. R. Westhead, and M. D. Eldridge, Flexible docking using Tabu search and an empirical estimate of binding affinity, *Proteins: Structure, Function and Genetics*, **33**, no. 3, 367–382, (1998).
- [97] D. Case, J. Berryman, R. Betz, D. Cerutti, T. Cheatham, T. Darden, R. Duke, T. Giese, H. Gohlke, A. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. Lee, S. LeGrand, P. Li, T. Luchko, R. Luo, B. Madej, K. Merz, G. Monard, P. Needham, H. Nguyen, I. Omelyan, A. Onufriev, D. Roe, A. Roitberg, R. Salomon-Ferrer, C. Simmerling, W. Smith, R. Swails, R. Walker, J. Wang, R. Wolf, X. Wu, D. York, and P. A. Kollman, (2015), University of California, San Francisco.
- [98] B. Brooks, R. Bruccoleri, B. Olafson, D. States, S. Swaminathan, and M. Karplus, CHARMM: A program for macromolecular energy, minimization, and dynamics calculations, *J Comput Chem*, **4**, no. 2, 187–217, (1983).
- [99] U. Dinur and A. Hagler, New Approaches to Empirical Force Fields, in *In Reviews in Computational Chemistry*, Lipkowitz, K. a. B. DB and VCH., Eds., **2**, 99–164, New York, 1991.
- [100] J. E. H. Koehler, W. Saenger, and W. F. van Gunsteren, A molecular dynamics simulation of crystalline α -cyclodextrin hexahydrate, *European Biophysics Journal*, **15**, no. 4, 197–210, (1987).
- [101] T. A. Halgren, Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94, *Journal of Computational Chemistry*, **17**, no. 5-6, 490–519, (1996).
- [102] N. L. Allinger and H. Von Voithenberg, Conformational analysis-131. A molecular mechanics treatment of phosphines I Paper 130, N.L. Allinger, J. Am. Chem. Soc. 99, 8127 (1977), *Tetrahedron*, **34**, no. 6, 627–633, (1978).
- [103] N. L. Allinger, Y. H. Yuh, and J.-H. Lii, Molecular mechanics. The MM3 force field for hydrocarbons. 1, *Journal of the American Chemical Society*, **111**, no. 23, 8551–8566, (1989).
- [104] J. P. Bowen and N. L. Allinger, In *Rev Comput Chem*, in *In Rev Comput Chem*, K. Lipkowitz and D. Boyd, Eds., **2**, 81–97, VCH Publishers, New York, 1991.
- [105] G. Kaminski and W. L. Jorgensen, Performance of the AMBER94, MMFF94, and OPLS-AA force fields for modeling organic liquids, *Journal of Physical Chemistry*, **100**, no. 46, 18010–18013, (1996).
- [106] ., Tripos Inc., St. Louis MO, 1995.
- [107] B. Honig and A. Nicholls, Classical electrostatics in biology and chemistry, *Science*, **268**, no. 5214, 1144–1149, (1995).
- [108] D. Bashford and D. A. Case, Generalized Born Models of Macromolecular Solvation Effects, *Annu Rev Phys Chem*, **51**, 129–152, (2000).
- [109] B. K. Shoichet, A. R. Leach, and I. D. Kuntz, Ligand solvation in molecular docking, *Proteins: Structure, Function and Genetics*, **34**, no. 1, 4–16, (1999).
- [110] A. Nicholls and B. Honig, A rapid finite difference algorithm, utilizing successive over-relaxation to solve the Poisson–Boltzmann equation, *J Comput Chem*, **12**, 435–445, (1991).
- [111] J. A. Grant, B. T. Pickup, and A. Nicholls, A smooth permittivity function for Poisson–Boltzmann solvation methods, *J Comput Chem*, **22**, no. 6, 608–640, (2001).
- [112] A. A. Rashin, Hydration phenomena, classical electrostatics, and the boundary element method, *Journal of Physical Chemistry*, **94**, no. 5, 1725–1733, (1990).
- [113] M. D. Miller, S. K. Kearsley, D. J. Underwood, and R. P. Sheridan, FLOG: A system to select 'quasi-flexible' ligands complementary to a receptor of known three-dimensional structure, *Journal of Computer-Aided Molecular Design*, **8**, no. 2, 153–174, (1994).
- [114] C. McMartin and R. S. Bohacek, QXP: Powerful, rapid computer algorithms for structure-based drug design, *Journal of Computer-Aided Molecular Design*, **11**, no. 4, 333–344, (1997).
- [115] C. McMartin and R. S. Bohacek, Flexible matching of test ligands to a 3D pharmacophore using a molecular superposition force field: Comparison of predicted and experimental conformations of inhibitors of three enzymes, *J Comput Aided Mol Des*, **9**, 237–250, (1995).
- [116] M. Mezei, Efficient Monte Carlo sampling for long molecular chains using local moves, tested on a solvated lipid bilayer, *Journal of Chemical Physics*, **118**, no. 8, 3874–3879, (2003).
- [117] P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case, and T. E. Cheatham III, Calculating structures and free energies of complex molecules: Combining molecular mechanics and continuum models, *Accounts of Chemical Research*, **33**, no. 12, 889–897, (2000).
- [118] B. Kuhn and P. A. Kollman, A ligand that is predicted to bind better to avidin than biotin: Insights from computational fluorine scanning, *Journal of the American Chemical Society*, **122**, no. 16, 3909–3916, (2000).
- [119] I. Massova and P. A. Kollman, Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding, *Perspectives in Drug Discovery and Design*, **18**, 113–135, (2000).

- [120] D. A. Pearlman and P. S. Charifson, Improved scoring of ligand-protein interactions using OWFEG free energy grids, *Journal of Medicinal Chemistry*, **44**, no. 4, 502–511, (2001).
- [121] S. Yin, L. Biedermannova, J. Vondrasek, and N. V. Dokholyan, MedusaScore: An accurate force field-based scoring function for virtual drug screening, *Journal of Chemical Information and Modeling*, **48**, no. 8, 1656–1662, (2008).
- [122] H. Bohm and M. Stahl, *Medicinal Chemistry Research*, **9**, 445–462, (1999).
- [123] M. Davis and J. A. McCammon, Electrostatics in biomolecular structure and dynamics, *Chem Rev*, **90**, 509–521, (1990).
- [124] M. Schapira, M. Totrov, and R. Abagyan, Prediction of the binding energy for small molecules, peptides and proteins, *Journal of Molecular Recognition*, **12**, no. 3, 177–190, (1999).
- [125] G. Jones, P. Willett, and R. C. Glen, Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation, *Journal of Molecular Biology*, **245**, no. 1, 43–53, (1995).
- [126] G. M. Morris, D. S. Goodsell, R. Huey, and A. J. Olson, Distributed automated docking of flexible ligands to proteins: Parallel applications of AutoDock 2.4, *Journal of Computer-Aided Molecular Design*, **10**, no. 4, 293–304, (1996).
- [127] J. O. Hirschfelder, C. F. Curtiss, and R. B. Bird, *Molecular theory of gases and liquids*, John Wiley, New York, 1964.
- [128] L. Wesson and D. Eisenberg, Atomic solvation parameters applied to molecular dynamics of proteins in solution, *Protein Science*, **1**, no. 2, 227–235, (1992).
- [129] P. F. W. Stouten and C. Frommel, Effective solvation term based on atomic occupancies for use in protein simulations, *Molecular Simulation*, **10**, no. 2-6, 97–120, (1993).
- [130] S. M. Green and G. R. Marshall, 3D-QSAR: a current perspective, *Trends Pharmacol Sci*, **16**, 285–291, (1995).
- [131] H. J. Bohm, The computer program LUDI: A new method for the de novo design of enzyme inhibitors, *J Comput Aided Mol Des*, **6**, 61–78, (1992).
- [132] M. Rarey, S. Wefing, and T. Lengauer, Placement of medium-sized molecular fragments into active sites of proteins, *Journal of Computer-Aided Molecular Design*, **10**, no. 1, 41–54, (1996).
- [133] H. J. Bohm, The development of a simple empirical scoring function to estimate the binding constant for a protein-ligand complex of known three-dimensional structure, *J Comput Aided Mol Des*, **8**, 243–256, (1994).
- [134] H. J. Bohm, Prediction of binding constants of protein ligands: A fast method for the prioritization of hits obtained from de novo design or 3D database search programs, *J Comput Aided Mol Des*, **12**, 309–323, (1998).
- [135] M. D. Eldridge, C. W. Murray, T. R. Auton, G. V. Paolini, and R. P. Mee, Empirical scoring functions: I. The development of a fast empirical scoring function to estimate the binding affinity of ligands in receptor complexes, *Journal of Computer-Aided Molecular Design*, **11**, no. 5, 425–445, (1997).
- [136] M. L. Verdonk, J. C. Cole, M. J. Hartshorn, C. W. Murray, and R. D. Taylor, Improved protein-ligand docking using GOLD, *Proteins: Structure, Function and Genetics*, **52**, no. 4, 609–623, (2003).
- [137] C. De Graaf, C. Oostenbrink, P. H. J. Keizers, T. Van Der Wijst, A. Jongejan, and N. P. E. Vermeulen, Catalytic site prediction and virtual screening of cytochrome P450 2D6 substrates by consideration of water and rescoring in automated docking, *Journal of Medicinal Chemistry*, **49**, no. 8, 2417–2430, (2006).
- [138] R. Wang, L. Liu, L. Lai, and Y. Tang, SCORE: A new empirical method for estimating the binding affinity of a protein-ligand complex, *Journal of Molecular Modeling*, **4**, no. 12, 379–394, (1998).
- [139] J. Pei, Q. Wang, Z. Liu, Q. Li, K. Yang, and L. Lai, PSI-DOCK: Towards highly efficient and accurate flexible ligand docking, *Proteins: Structure, Function and Genetics*, **62**, no. 4, 934–946, (2006).
- [140] C. A. Sotriffer, P. Sanschagrin, H. Matter, and G. Klebe, SFCscore: Scoring functions for affinity prediction of protein-ligand complexes, *Proteins: Structure, Function and Genetics*, **73**, no. 2, 395–419, (2008).
- [141] G. E. Kellogg, S. F. Semus, and D. J. Abraham, HINT: A new method of empirical hydrophobic field calculation for CoMFA, *Journal of Computer-Aided Molecular Design*, **5**, no. 6, 545–552, (1991).
- [142] C. Hansch and A. J. Leo, *Substituent constants for correlation analysis in chemistry and biology*, John Wiley and Sons, Inc, New York, 1979.
- [143] D. J. Abraham and A. J. Leo, Extension of the fragment method to calculate amino acid zwitterion and side chain partition coefficients., *Proteins*, **2**, no. 2, 130–152, (1987).
- [144] I. Muegge and M. Rarey, Small molecule docking and scoring, *Reviews in Computational Chemistry*, **17**, 1–60, (2001).
- [145] S. Liu, C. Zhang, H. Zhou, and Y. Zhou, A physical reference state unifies the structure-derived potential of mean force for protein folding and binding, *Proteins: Structure, Function and Genetics*, **56**, no. 1, 93–101, (2004).
- [146] A. Wallqvist, R. L. Jernigan, and D. G. Covell, A preference-based free-energy parameterization of enzyme-inhibitor binding. Applications to HIV-1-protease inhibitor design, *Protein Science*, **4**, no. 9, 1881–1903, (1995).
- [147] I. Muegge and Y. C. Martin, A general and fast scoring function for protein-ligand interactions: A simplified potential approach, *Journal of Medicinal Chemistry*, **42**, no. 5, 791–804, (1999).
- [148] I. Muegge, Y. C. Martin, P. J. Hajduk, and S. W. Fesik, Evaluation of PMF scoring in docking weak ligands to the FK506 binding protein, *Journal of Medicinal Chemistry*, **42**, no. 14, 2498–2503, (1999).
- [149] I. Muegge, PMF scoring revisited, *Journal of Medicinal Chemistry*, **49**, no. 20, 5895–5902, (2006).
- [150] E. Krovat, T. Steindl, and T. Langer, Current Computer-Aided Drug Design, *Current Computer-Aided Drug Design*, **1**, 93–102, (2005).
- [151] P. S. Charifson, J. J. Corkery, M. A. Murcko, and W. P. Walters, Consensus scoring: A method for obtaining improved hit rates from docking databases of three-dimensional structures into proteins, *Journal of Medicinal Chemistry*, **42**, no. 25, 5100–5109, (1999).
- [152] E. Perola, W. P. Walters, and P. S. Charifson, A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance, *Proteins: Structure, Function and Genetics*, **56**, no. 2, 235–249, (2004).
- [153] M. Stahl and H. J. Bohm, Development of filter functions for protein-ligand docking, *J Mol Graph Model*, **16**, 121–32, (1998).
- [154] S. Gruneberg, M. T. Stubbs, and G. Klebe, Successful Virtual Screening for Novel Inhibitors of Human Carbonic Anhydrase: Strategy and Experimental Confirmation, *Journal of Medicinal Chemistry*, **45**, 3588–3602, (2002).

- [155] J.-C. Mozziconacci, E. Arnoult, P. Bernard, Q. T. Do, C. Marot, and L. Morin-Allory, Optimization and validation of a docking-scoring protocol; application to virtual screening for COX-2 inhibitors, *Journal of Medicinal Chemistry*, **48**, no. 4, 1055–1068, (2005).
- [156] Z. Zhou, A. K. Felts, R. A. Friesner, and R. M. Levy, Comparative performance of several flexible docking programs and scoring functions: Enrichment studies for a diverse set of pharmaceutically relevant targets, *Journal of Chemical Information and Modeling*, **47**, no. 4, 1599–1608, (2007).
- [157] R. D. Clark, A. Strizhev, J. M. Leonard, J. F. Blake, and J. B. Matthew, Consensus scoring for ligand/protein interactions, *Journal of Molecular Graphics and Modelling*, **20**, no. 4, 281–295, (2002).
- [158] R. Wang, L. Lai, and S. Wang, Further development and validation of empirical scoring functions for structure-based binding affinity prediction, *Journal of Computer-Aided Molecular Design*, **16**, 11–26, (2002).
- [159] M. Stahl and M. Rarey, Detailed analysis of scoring functions for virtual screening, *Journal of Medicinal Chemistry*, **44**, no. 7, 1035–1042, (2001).
- [160] K. L. Damm and H. A. Carlson, Exploring Experimental Sources of Multiple Protein Conformations in Structure-Based Drug Design, *J Am Chem Soc*, **129**, 8225–8235, (2007).
- [161] J. Zhang and J. S. Liu, On Side-Chain Conformational Entropy of Proteins, *PLoS Comput Biol*, **2**, p. e168, (2006).
- [162] M. H. Ahmed, F. Spyraakis, P. Cozzini, P. K. Tripathi, A. Mozzarelli, J. N. Scarsdale, M. A. Safo, and G. E. Kellogg, Bound water at protein-protein interfaces: Partners, roles and hydrophobic bubbles as a conserved motif, *PLoS ONE*, **6**, no. 9, Article ID e24712, (2011).
- [163] J. A. Rupley and G. Careri, Protein hydration and function, *Advances in Protein Chemistry*, **41**, 37–172, (1991).
- [164] H. Frauenfelder, P. W. Fenimore, and B. H. McMahon, Hydration, slaving and protein function, *Biophysical Chemistry*, **98**, no. 1-2, 35–48, (2002).
- [165] P. A. Karplus and C. Faerman, Ordered water in macromolecular structure, *Current Opinion in Structural Biology*, **4**, no. 5, 770–776, (1994).
- [166] O. Carugo and D. Bordo, How many water molecules can be detected by protein crystallography? *Acta Crystallogr D Biol Crystallogr*, **55**, 479–483, (1999).
- [167] C. Barillari, J. Taylor, R. Viner, and J. W. Essex, Classification of water molecules in protein binding sites, *Journal of the American Chemical Society*, **129**, no. 9, 2577–2587, (2007).
- [168] D. N. A. Boobbyer, P. J. Goodford, P. M. McWhinnie, and R. C. Wade, New hydrogen-bond potentials for use in determining energetically favorable binding sites on molecules of known structure, *Journal of Medicinal Chemistry*, **32**, no. 5, 1083–1094, (1989).
- [169] L. Ehrlich, M. Reczko, H. Bohr, and R. C. Wade, Prediction of protein hydration sites from sequence by modular neural networks, *Protein Engineering*, **11**, no. 1, 11–19, (1998).
- [170] A. T. Garcia-Sosa, R. L. Mancera, and P. M. Dean, WaterScore: a novel method for distinguishing between bound and displaceable water molecules in the crystal structure of the binding site of protein-ligand complexes, *J Mol Model*, **9**, 172–182, (2003).
- [171] L. Jiang, B. Kuhlman, T. Kortemme, and D. Baker, A "solvated rotamer" approach to modeling water-mediated hydrogen bonds at protein-protein interfaces, *Proteins: Structure, Function and Genetics*, **58**, no. 4, 893–904, (2005).
- [172] M. A. Kastenzholz, M. Pastor, G. Cruciani, E. E. J. Haaksma, and T. Fox, GRID/CPCA: A new computational tool to design selective ligands, *Journal of Medicinal Chemistry*, **43**, no. 16, 3033–3044, (2000).
- [173] Z. Li and T. Lazaridis, Thermodynamic Contributions of the Ordered Water Molecule in HIV-1 Protease, *J Am Chem Soc*, **125**, p. 6636, (2003).
- [174] D. G. Lloyd, A. T. Garcia-Sosa, I. L. Alberts, N. P. Todorov, and R. L. Mancera, The effect of tightly bound water molecules on the structural interpretation of ligand-derived pharmacophore models, *J Comput Aided Mol Des*, **18**, 89–100, (2004).
- [175] W. R. Pitt, J. Murray-Rust, and J. M. Goodfellow, AQUARIUS2: Knowledge-based modeling of solvent sites around proteins, *Journal of Computational Chemistry*, **14**, no. 9, 1007–1018, (1993).
- [176] M. L. Raymer, P. C. Sanschagrin, W. F. Punch, S. Venkataraman, E. D. Goodman, and L. A. Kuhn, Predicting conserved water-mediated and polar ligand interactions in proteins using a k-nearest-neighbors genetic algorithm, *Journal of Molecular Biology*, **265**, no. 4, 445–464, (1997).
- [177] S. M. Roe and M. M. Teeter, Patterns for prediction of hydration around polar residues in proteins, *Journal of Molecular Biology*, **229**, no. 2, 419–427, (1993).
- [178] I. L. Alberts, N. P. Todorov, and P. M. Dean, Receptor flexibility in de novo ligand design and docking, *Journal of Medicinal Chemistry*, **48**, no. 21, 6585–6596, (2005).
- [179] M. R. Arkin, M. Randal, W. L. DeLano, J. Hyde, T. N. Luong, J. D. Oslob, D. R. Raphael, L. Taylor, J. Wang, R. S. McDowell, J. A. Wells, and A. C. Braisted, Binding of small molecules to an adaptive protein-protein interface, *Proceedings of the National Academy of Sciences of the United States of America*, **100**, no. 4, 1603–1608, (2003).
- [180] P. Cozzini, M. Fornabaio, A. Marabotti, D. J. Abraham, G. E. Kellogg, and A. Mozzarelli, Free energy of ligand binding to protein: Evaluation of the contribution of water molecules by computational methods, *Current Medicinal Chemistry*, **11**, no. 23, 3093–3118, (2004).
- [181] H. Gohlke and G. Klebe, Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors, *Angewandte Chemie - International Edition*, **41**, no. 15, 2644–2676, (2002).
- [182] C. S. Poornima and P. M. Dean, Hydration in drug design. 3. Conserved water molecules at the ligand-binding sites of homologous proteins, *J Comput Aided Mol Des*, **9**, 521–531, (1995).
- [183] C. S. Poornima and P. M. Dean, Hydration in drug design. 2. Influence of local site surface shape on water binding, *J Comput Aided Mol Des*, **9**, 513–520, (1995).
- [184] C. S. Poornima and P. M. Dean, Hydration in drug design. 1. Multiple hydrogen-bonding features of water molecules in mediating protein-ligand interactions, *J Comput Aided Mol Des*, **9**, 500–512, (1995).
- [185] A. T. Garcia-Sosa, R. L. Mancera, and P. M. Dean, WaterScore: a novel method for distinguishing between bound and displaceable water molecules in the crystal structure of the binding site of protein-ligand complexes, *J Mol Model*, **9**, 172–182, (2003).
- [186] A. Amadasi, J. A. Surface, F. Spyraakis, P. Cozzini, A. Mozzarelli, and G. E. Kellogg, Robust classification of "relevant" water molecules in putative protein binding sites, *Journal of Medicinal Chemistry*, **51**, no. 4, 1063–1067, (2008).

- [187] A. Amadasi, F. Spyraakis, P. Cozzini, D. J. Abraham, G. E. Kellogg, and A. Mozzarelli, Mapping the energetics of water-protein and water-ligand interactions with the "natural" HINT forcefield: Predictive tools for characterizing the roles of water in biomolecules, *Journal of Molecular Biology*, **358**, no. 1, 289–309, (2006).
- [188] M. Rarey, B. Kramer, and T. Lengauer, The particle concept: Placing discrete water molecules during protein-ligand docking predictions, *Proteins: Structure, Function and Genetics*, **34**, no. 1, 17–28, (1999).
- [189] R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis, and P. S. Shenkin, Glide: A New Approach for Rapid, Accurate Docking and Scoring. I. Method and Assessment of Docking Accuracy, *Journal of Medicinal Chemistry*, **47**, no. 7, 1739–1749, (2004).
- [190] M. L. Verdonk, G. Chessari, J. C. Cole, M. J. Hartshorn, C. W. Murray, J. W. M. Nissink, R. D. Taylor, and R. Taylor, Modeling water molecules in protein-ligand docking using GOLD, *Journal of Medicinal Chemistry*, **48**, no. 20, 6504–6515, (2005).
- [191] N. Moitessier, E. Westhof, and S. Hanessian, Docking of aminoglycosides to hydrated and flexible RNA, *Journal of Medicinal Chemistry*, **49**, no. 3, 1023–1033, (2006).
- [192] A. D. J. van Dijk and A. M. J. J. Bonvin, Solvated docking: Introducing water into the modelling of biomolecular complexes, *Bioinformatics*, **22**, no. 19, 2340–2347, (2006).
- [193] R. Abel, T. Young, R. Farid, B. J. Berne, and R. A. Friesner, Role of the active-site solvent in the thermodynamics of factor Xa ligand binding, *Journal of the American Chemical Society*, **130**, no. 9, 2817–2831, (2008).
- [194] C. De Graaf, P. Pospisil, W. Pos, G. Folkers, and N. P. E. Vermeulen, Binding mode prediction of cytochrome P450 and thymidine kinase protein-ligand complexes by consideration of water and rescoring in automated docking, *Journal of Medicinal Chemistry*, **48**, no. 7, 2308–2318, (2005).
- [195] F. Österberg, G. M. Morris, M. F. Sanner, A. J. Olson, and D. S. Goodsell, Automated docking to multiple target structures: Incorporation of protein mobility and structural water heterogeneity in autodock, *Proteins: Structure, Function and Genetics*, **46**, no. 1, 34–40, (2002).
- [196] P. Pospisil, T. Kuoni, L. Scapozza, and G. Folkers, Methodology and problems of protein-ligand docking: Case study of dihydroorotate dehydrogenase, thymidine kinase, and phosphodiesterase 4, *Journal of Receptors and Signal Transduction*, **22**, no. 1-4, 141–154, (2002).
- [197] J.-M. Yang and C.-C. Chen, GEMDOCK: A Generic Evolutionary Method for Molecular Docking, *Proteins: Structure, Function and Genetics*, **55**, no. 2, 288–304, (2004).
- [198] A. Afanasyeva, S. Izmailov, M. Grigoriev, and M. Petukhov, AquaBridge: A novel method for systematic search of structural water molecules within the protein active sites, *Journal of Computational Chemistry*, **36**, no. 26, 1973–1977, (2015).
- [199] A. Bortolato, B. G. Tehan, M. S. Bodnarchuk, J. W. Essex, and J. S. Mason, Water network perturbation in ligand binding: Adenosine A2A antagonists as a case study, *Journal of Chemical Information and Modeling*, **53**, no. 7, 1700–1713, (2013).
- [200] K. Elokely, P. Velisetty, L. Delemotte, E. Palovcak, M. L. Klein, T. Rohacs, and V. Carnevale, Understanding TRPV1 activation by ligands: Insights from the binding modes of capsaicin and resiniferatoxin, *Proceedings of the National Academy of Sciences of the United States of America*, **113**, no. 2, E137–E145, (2016).
- [201] P. Sharma, A. Shukla, K. Kalani, V. Dubey, S. K. Srivastava, S. Luqman, and F. Khan, Water Molecules Increases Binding Affinity of Natural PI3K γ Inhibitors Against Cancer, *Curr Comput Aided Drug Des*, **11**, 304–320, (2016).
- [202] F. Spyraakis and C. N. Cavasotto, Open challenges in structure-based virtual screening: Receptor modeling, target flexibility consideration and active site water molecules description, *Arch Biochem Biophys*, **583**, 105–119, (2015).
- [203] R. Brenk, S. W. Vetter, S. E. Boyce, D. B. Goodin, and B. K. Shoichet, Probing molecular docking in a charged model binding site, *Journal of Molecular Biology*, **357**, no. 5, 1449–1470, (2006).
- [204] I. W. Davis, L. W. Murray, J. S. Richardson, and D. C. Richardson, MolProbity: Structure validation and all-atom contact analysis for nucleic acids and their complexes, *Nucleic Acids Research*, **32**, W615–W619, (2004).
- [205] W. Zhong, J. P. Gallivan, Y. Zhang, L. Li, H. A. Lester, and D. A. Dougherty, From ab initio quantum mechanics to molecular neurobiology: A cation- π binding site in the nicotinic receptor, *Proc Natl Acad Sci U S A*, **95**, 12088–12093, (1998).
- [206] R. B. Murphy, D. M. Philipp, and R. A. Friesner, A Mixed Quantum Mechanics/Molecular Mechanics (QM/MM) Method for Large-Scale Modeling of Chemistry in Protein Environments, *Journal of Computational Chemistry*, **21**, no. 16, 1442–1457, (2000).
- [207] J. Gao and X. Xia, A priori evaluation of aqueous polarization effects through Monte Carlo QM-MM simulations, *Science*, **258**, no. 5082, 631–635, (1992).
- [208] J. Antosiewicz, J. A. McCammon, and M. K. Gilson, Prediction of pH-dependent properties of proteins, *Journal of Molecular Biology*, **238**, no. 3, 415–436, (1994).
- [209] D. Bashford and M. Karplus, pK_a's of ionizable groups in proteins: Atomic detail from a continuum electrostatic model, *Biochemistry*, **29**, no. 44, 10219–10225, (1990).
- [210] E. Demchuk and R. C. Wade, Improving the continuum dielectric approach to calculating pK_as of ionizable groups in proteins, *Journal of Physical Chemistry*, **100**, no. 43, 17373–17387, (1996).
- [211] L. I. Krishtalik, A. M. Kuznetsov, and E. L. Mertz, Electrostatics of proteins: Description in terms of two dielectric constants simultaneously, *Proteins: Structure, Function and Genetics*, **28**, no. 2, 174–182, (1997).
- [212] A.-S. Yang, M. R. Gunner, R. Sampogna, K. Sharp, and B. Honig, On the calculation of pK_a(s) in proteins, *Proteins: Structure, Function and Genetics*, **15**, no. 3, 252–265, (1993).
- [213] G. J. Bartlett, C. T. Porter, N. Borkakoti, and J. M. Thornton, Analysis of catalytic residues in enzyme active sites, *Journal of Molecular Biology*, **324**, no. 1, 105–121, (2002).
- [214] R. E. Georgescu, E. G. Alexov, and M. R. Gunner, Combining conformational flexibility and continuum electrostatics for calculating pK_as in proteins, *Biophysical Journal*, **83**, no. 4, 1731–1748, (2002).
- [215] H. W. T. Van Vlijmen, M. Schaefer, and M. Karplus, Improving the accuracy of protein pK_a calculations: Conformational averaging versus the average structure, *Proteins: Structure, Function and Genetics*, **33**, no. 2, 145–158, (1998).
- [216] Z. Zsoldos, D. Reid, A. Simon, S. B. Sadjad, and A. P. Johnson, eHiTS: A new fast, exhaustive flexible ligand docking system, *Journal of Molecular Graphics and Modelling*, **26**, no. 1, 198–212, (2007).

- [217] G. Krauss, *Biochemistry of signal transduction and regulation*, in G. Krauss. *Biochemistry of signal transduction and regulation*, p. 541, Wiley-VCH, 2003.
- [218] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, and P. E. Bourne, The Protein Data Bank, *Nucleic Acids Research*, **28**, no. 1, 235–242, (2000).
- [219] R. Horst, G. Wider, J. Fiaux, E. B. Bertelsen, A. L. Horwich, and K. Wüthrich, Proton-proton Overhauser NMR spectroscopy with polypeptide chains in large structures, *Proceedings of the National Academy of Sciences of the United States of America*, **103**, no. 42, 15445–15450, (2006).
- [220] A. Grishaev, V. Tugarinov, L. E. Kay, J. Trehwella, and A. Bax, Refined solution structure of the 82-kDa enzyme malate synthase G from joint NMR and synchrotron SAXS restraints, *Journal of Biomolecular NMR*, **40**, no. 2, 95–106, (2008).
- [221] J.-H. Lin, A. L. Perryman, J. R. Schames, and J. A. McCammon, Computational drug design accommodating receptor flexibility: The relaxed complex scheme, *Journal of the American Chemical Society*, **124**, no. 20, 5632–5633, (2002).
- [222] J. A. McCammon, *Biochim Biophys Acta*, **1754**, 221–224, (2005).
- [223] J. R. Schames, R. H. Henchman, J. S. Siegel, C. A. Sotriffer, H. Ni, and J. A. McCammon, Discovery of a Novel Binding Trench in HIV Integrase, *Journal of Medicinal Chemistry*, **47**, no. 8, 1879–1881, (2004).
- [224] R. Huey, G. M. Morris, A. J. Olson, and D. S. Goodsell, Software news and update a semiempirical free energy force field with charge-based desolvation, *Journal of Computational Chemistry*, **28**, no. 6, 1145–1152, (2007).
- [225] A. Ahmed and H. Gohlke, Multiscale modeling of macromolecular conformational changes combining concepts from rigidity and elastic network theory, *Proteins: Structure, Function and Genetics*, **63**, no. 4, 1038–1051, (2006).
- [226] A. Macchiarulo, I. Nobeli, and J. M. Thornton, Ligand selectivity and competition between enzymes in silico, *Nature Biotechnology*, **22**, no. 8, 1039–1045, (2004).
- [227] C.-J. Tsai, S. Kumar, B. Ma, and R. Nussinov, Folding funnels, binding funnels, and protein function, *Protein Science*, **8**, no. 6, 1181–1190, (1999).
- [228] M. Zentgraf, J. Fokkens, and C. A. Sotriffer, Addressing protein flexibility and ligand selectivity by "in situ cross-docking", *ChemMedChem*, **1**, no. 12, 1355–1359, (2006).
- [229] L. Dellafiora, G. Galaverna, C. Dall'Asta, and P. Cozzini, Hazard identification of cis/trans-zearalenone through the looking-glass, *Food and Chemical Toxicology*, **86**, 65–71, (2015).
- [230] V. A. Ehrlich, L. Dellafiora, J. Mollergues, C. Dall'Asta, P. Serrant, M. Marin-Kuan, E. Lo Piparo, B. Schilter, and P. Cozzini, Hazard assessment through hybrid in vitro/in silico approach: The case of zearalenone, *Altex*, **32**, no. 4, 275–286, (2015).
- [231] L. Dellafiora, P. Mena, P. Cozzini, F. Brighenti, and D. Del Rio, Modelling the possible bioactivity of ellagitannin-derived metabolites. in silico tools to evaluate their potential xenoestrogenic behavior, *Food and Function*, **4**, no. 10, 1442–1451, (2013).
- [232] T. Ginex, C. Dall'Asta, and P. Cozzini, Preliminary hazard evaluation of androgen receptor-mediated endocrine-disrupting effects of thioxanthone metabolites through structure-based molecular docking, *Chemical Research in Toxicology*, **27**, no. 2, 279–289, (2014).
- [233] H. W. Ng, W. Zhang, M. Shu, H. Luo, W. Ge, R. Perkins, W. Tong, and H. Hong, Competitive molecular docking approach for predicting estrogen receptor subtype α agonists and antagonists, *BMC Bioinformatics*, **15(Suppl 11)**, p. S4, (2014).
- [234] L. Li, Q. Wang, Y. Zhang, Y. Niu, X. Yao, and H. Liu, The Molecular Mechanism of Bisphenol A (BPA) as an Endocrine Disruptor by Interacting with Nuclear Receptors: Insights from Molecular Dynamics (MD) Simulations, *PLoS One*, (2015).