FACULDADE DE CIÊNCIAS DA UNIVERSIDADE DO PORTO

Integrating a new scoring function into molecular docking for the design of glucosidase inhibitors

Diogo Santos Martins

Programa Doutoral em Química Sustentável

Supervisor: Prof. Maria João Ramos, FCUP

Co-Supervisor: Prof. Arthur J. Olson, TSRI

January 10, 2017

Integrating a new scoring function into molecular docking for the design of glucosidase inhibitors

Diogo Santos Martins

Programa Doutoral em Química Sustentável

January 10, 2017

Abstract

This thesis reports improvements in the field of molecular docking, a methodology widely used in drug discovery projects to predict the affinity of small molecules to proteins (or to other macromolecules), and to understand the structural basis for the affinity. We developed better descriptors for interactions between zinc atoms in metalloenzymes and their ligands, both structurally (tetrahedral coordination shells are explicitly represented) and energetically. We have also worked on an empirical approach for predicting free energies of solvation in water, aiming at its integration with molecular docking to improve the description of desolvation effects.

In a parallel line of research, we explored the influence of conformational fluctuations of enzyme-substrate complexes on the reaction rate. We found that the energy of transition states (relatively to reactants state) oscillates over time on a nanosecond (or faster) timescale, a fact with implications for the calculation of reaction rates.

Resumo

Esta tese relata melhoramentos no âmbito do *docking* molecular, uma metodologia muito usada em projectos de descoberta de fármacos para prever a afinidade de moléculas pequenas com proteínas (ou com outras macromoléculas), e para perceber a base estrutural que explica a afinidade. Nós desenvolvemos descritores melhorados para a interação entre àtomos de zinco em metaloenzimas e moléculas ligantes, tanto do ponto de vista energético como estrutural (esferas de coordenação tetraédricas são representadas explicitamente). Também trabalhamos numa abordagem empírica para prever energias livres de solvatação em água, contemplando a sua integração com o "docking" molecular para melhorar a descrição dos efeitos de dessolvatação.

Numa linha de investigação paralela, exploramos a influência de flutuações conformacionais do complexo enzima-substrato na velocidade de reacção. Descobrimos que a energia dos estados de transição (relativamente ao estado dos reagentes) oscila ao longo do tempo na ordem dos nanosegundos (ou mais rapidamente), um facto com implicações para o cálculo de velocidades de reação.

Acknowledgements

I thank my supervisor Prof. Maria João Ramos for guidance and extraordinary research conditions, but specially for helping me achieve my goals in an independent manner.

I thank my co-supervisor Prof. Arthur J. Olson for doing large efforts to integrate me in San Diego, and for being an outstanding advisor.

Several people contributed extensively to shape my research: Prof. Pedro Alexandrino Fernandes, Prof. Stefano Forli, António J. M. Ribeiro, Eduardo F. Oliveira, Nuno M. F. S. A Cerqueira, Prof. David S. Goodsell and Rui P. P. Neves.

I also thank those I interacted with for many engaging discussions, for technical help and for making the workplace a nice place to work: José Gaspar Pinto, Sílvia Martins, Óscar Passos, Diana Gesto, Cátia Moreira, Ana Rita Calixto, Fabiola Medina, João Coimbra, Tiago Gesto, Sérgio Sousa, Prof. Alexandre Magalhães, Prof. André Melo, Rui Sousa, Henrique Fernandes, Andreia Pereira, Rui Ribeiro, Ana Catarina Barbosa, Inês Simões, Pedro Ferreira, Pedro Paiva, Prashant Jindal, Matilde Viegas, Natércia Brás, Pradeep Ravindranath, Emilio Angelina, Daniel N. Santiago, Michael E. Pique, Oleg Trott, Adam Gardner, Ruth Huey, Ludovic Autin, Prof. Michel Sanner and Peggy Graber.

I thank Fundação para a Ciência e Tecnologia (FCT) for scholarship SFRH/BD/84922/2012, with contributions from the European Social Fund and the Government of Portugal. The financing program is Programa Operacional Capital Humano (POCH).

Preface

This thesis follows the 'paper collection' model, where scientific publications are presented as chapters and describe the most important results obtained during the PhD studies. Here, chapters 3, 4, 5 and 6 correspond to published papers, chapter 7 is a manuscript we plan to submit and chapter 8 describes unpublished results. Introductory chapters cover important concepts in molecular docking (chapter 1) and computational enzymatic catalysis (chapter 2).

Our efforts towards a better scoring function for glycosidase inhibitors — our primary goal when we started — are described in chapter 8. We hypothesized that desolvation effects may be a key descriptor in carbohydrate binding due to their strong interaction with water — a consequence of having a large number of hydroxyl groups. We improved the description of solvation effects with an empirical method based on atomic contributions scaled by surface areas (chapter 4). Notably, this method performs satisfactorily for carbohydrates (details in chapter 8). Our participation in the D3R Grand Challenge 2015 (chapter 5), a challenge where 180 ligands were to be ranked by affinity, aimed at testing our new desolvation terms in a challenging dataset of protein-ligand complexes. Our desolvation descriptors were inefficient, a result we attribute to shape effects on the energetics of confined waters, severely affecting desolvation of binding pockets (see section 1.1.1). In chapter 3 the AutoDock4.2 scoring function was extended to describe the coordination of Zn^{2+} atoms with improved geometry and energetic terms. This work was performed in Art Olson's lab at The Scripps Research Institute, La Jolla.

Chapters 6 and 7 explore the structure of transition states and reactants on varying conformations of the enzyme. It became clear that transition state structures of the glycolysis step are rigid and well defined, a fact with immediate implications for the design of transition state analogues. Furthermore, these studies indicate that fluctuations in the activation barrier of enzyme catalyzed reactions oscillate due to changes in enzyme structure and in intermolecular interactions in the active site. Notably, the timescale of these fluctuations is orders of magnitude faster than turnover rates. These results may aid in clarifying scientific disputes over the origin of catalytic power in enzymes.

Overall, this thesis contributes with knowledge and modeling tools for the development of novel glycosidase inhibitors, and also inhibitors of other enzymes because the novel desolvation descriptors are generic in nature. It also contributes to our understanding of enzymatic catalysis, not only for glucosidases, but in a more general way.

Contents

List of Figures

List of Tables

Abbreviations

HSP90	Heat Shock Protein 90
GIST	Grid Inhomogeneous Solvation Theory
QM	Quantum Mechanics
PDB	Protein Data Bank
D3R	Drug Design Data Resource
GC2015	Grand Challenge 2015
SAMPL	Statistical Assessment of the Modeling of Proteins and Ligands
ROC (curve)	Receiver Operating Characteristic
AUC	Area Under the (ROC) Curve
RMSD	Root Mean Square Deviation
MO	Molecular Orbital (theory)
VB	Valence Bond (theory)
SCF	Self Consistent Field
BO	Born-Oppenheimer (approximation)
TS	Transition State
TST	Transition State Theory
SCF	Self Consistent Field
HF	Hartree-Fock
DFT	Density Functional Theory
EVB	Empirical Valence Bond

Chapter 1

Molecular Docking

Molecular docking refers to methodologies used to guess the pose of a ligand upon binding to a macromolecular receptor of known structure, and to predict the affinity of the resulting complex. There are two key components in molecular docking software: a search algorithm and a scoring function.

The search algorithm samples ligand conformations and positions inside the receptor binding pocket (referred to as binding poses) while the scoring function scores and ranks each pose. Search algorithms are beyond the scope of this thesis and are not discussed. The score of the supposedly correct pose (best ranked) can also be used to predict ligand affinity or to rank a series of compounds. When libraries of compounds are docked for the search for new inhibitors, the term used is virtual screening; an analogy to high throughput screening (HTS). Overall, scoring functions predict both structural (binding pose) and functional (binding affinity) features of ligand binding.

The following sections introduce concepts required for the development of scoring functions, focusing on a rigorous physical description of ligand binding. Our point of view can be contrasted with the use of machine learning algorithms (which have been criticized recently [3]) to build complex models without interpretable physical meaning, i.e. it is impossible to decompose output scores into separate components such as desolvation, electrostatics, and so on. Since ligand-receptor binding is governed by the laws of physics, scoring functions are likely to improve with a better description of the physical processes underlying molecular association.

1.1 Physics of Ligand-Receptor Binding

The affinity of a ligand-receptor complex is dictated by the difference in free energy between bound and unbound states. Experimentally, the affinity is quantified by a dissociation constant (K_i or K_d) which relates to binding free energy (ΔG) by the following equation:

$$\Delta G = RT\,lnK \tag{1.1}$$

where K is the dissociation constant, R is the ideal gas constant and T is the temperature. It is interesting to note that ΔG is itself temperature dependent ($\Delta G = \Delta H - T\Delta S$), so temperature

actually appears twice in the equation. In the simple case where one ligand binds to a single receptor, the free energy can be written as:

$$\Delta G_{bind} = G_{RL} - (G_R + G_L) \tag{1.2}$$

where G_{RL} is the absolute free energy of the solvated complex, G_R is the absolute free energy of the solvated receptor and G_L is the absolute free energy of the solvated ligand. Since computation of absolute free energies is impractical, scoring functions rely on descriptors (or terms) that predict changes (ΔG) in various energy components associated with ligand-receptor binding.

The most obvious components contributing favorably to ΔG_{bind} are those related to intermolecular interactions between ligand and receptor: van der Waals (vdW), electrostatics and H-bonds. These are calculated based on the structure of the bound complex, using a forcefield (described in section 2.4) or knowledge based potentials (see section 1.2). Scoring functions generally have enough terms to describe ligand-receptor interactions (van der Waals, hydrogen bonds and electrostatics) with nearly chemical accuracy (same accuracy as the underlying forcefield). However, binding free energies depend on other energetic components which change significantly from the bound state (G_{RL}) to the unbound state ($G_R + G_L$). These components are difficult to calculate and are related to conformational changes in the receptor (panel B in figure 1.1), and desolvation effects (panels B and C in figure 1.1). The contribution of desolvation and changes in water structure to ΔG_{bind} are described in detail in section 1.1.1. Scoring functions either ignore these physical features of ligand binding (e.g. receptor conformational changes) or rely on an insufficient description of their complex nature (e.g. receptor desolvation). Furthermore, ligand binding may be accompanied by changes in protonation state of both receptor and ligand molecules. If a stable binding mode is restricted to a single protonation state or tautomer, the existence of multiple states in solution is entropically favorable to G_R and/or G_L relatively to G_{RL}. Indeed, tautomers and protonation states play a significant role in solvation free energy [4], and should be considered when modeling binding affinities.

It is important to point out that selecting the correct pose from an ensemble of docked poses may be successful even when ignoring or misrepresenting desolvation and conformational changes of the receptor. If these changes are equivalent among all docked poses of the same ligand, only intermolecular interactions (vdW, electrostatics and H-bonds) are relevant to select correct binding poses. This may explain why scoring functions perform better at selecting poses than at predicting ligand affinities [5]: it is possible that better terms are required for the problematic components described above.

Entropy plays a significant role in the energetics of ligand binding but is neglected in molecular docking because sampling is not performed. Sampling approaches (molecular dynamics and monte carlo) are time consuming and would prohibit the study of large libraries of compounds. Receptor, ligand and water are subject to changes in entropy from unbound to bound states. Upon binding, the ligand is restricted to the shape of the binding pocket, which translates into an entropic penalty disfavoring binding. Scoring functions often use the number of rotatable bonds in

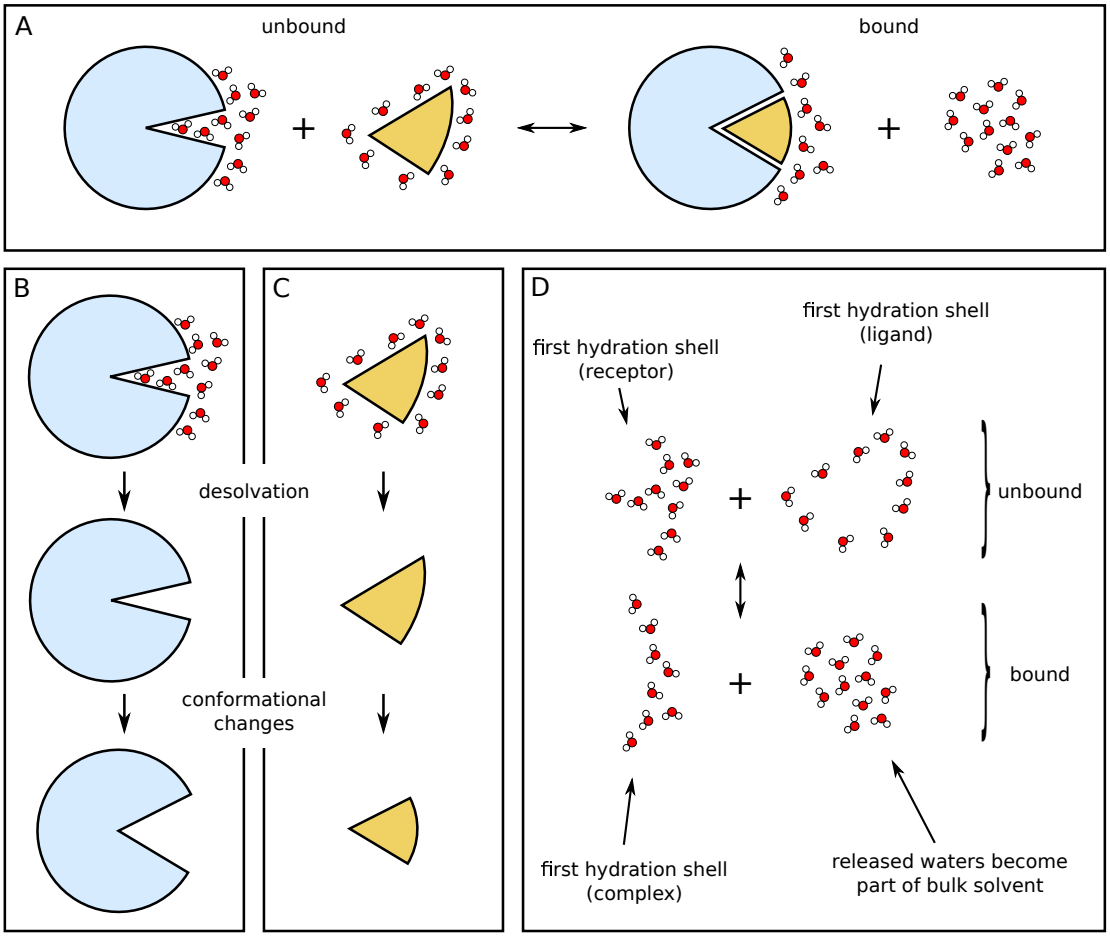


Figure 1.1: End states of ligand binding. (A) Overview of unbound and bound states. (B) Desolvation and conformational changes in the receptor. (C) Desolvation and conformational changes in the ligand (in most cases the ligand is restricted upon binding which translates into an entropic cost disfavoring binding). (D) Solvation shells of receptor and ligand before binding and solvation shell of complex and released waters (n = 12) after binding. (A) Changes in water strcture: in the unbound state 10 explicit water molecules solvate the ligand and another 10 solvate the receptor, while only 8 solvate the bound complex. Thus, the binding process releases 12 waters to bulk solvent. Released waters no longer interact directly (first shell) with neither the ligand nor the receptor.

the ligand to quantify its loss of entropy. Entropy changes in receptor and water are much more difficult to quantify because extensive sampling would be required. In order to be applicable in large scale studies, molecular docking approaches sacrifice the entropic component and rely on a single receptor structure.

1.1.1 Desolvation Effects

Desolvation ΔG^{solv}_{bind} is the change in free energy of solvation (hydration) between bound and unbound states. It makes significant contributions to binding free energies. Desolvation free energy

can be written as the difference between hydration free energies of bound and unbound states: $\Delta G^{solv}_{bind} = \Delta G^{solv}_{RL} - (\Delta G^{solv}_{R} + \Delta G^{solv}_{L})$. Hydration is commonly perceived as an homogenous effect where the exact structure of water is unimportant and the interplay between hydrogen bond acceptors can be averaged into a dielectric constant. This idea has been successfully applied to small solutes because water-water interactions do not change dramatically compared to bulk water. In this section we explain how water-water interactions are likely to play a major contribution to ΔG^{solv}_{bind}.

1.1.1.1 Ligand Desolvation

Hydrophobic effects refer to the aggregation of apolar solutes in water. In the case of ligand binding it specifically refers to the favorable burial of apolar ligands into apolar binding sites. Apolar surfaces interact only weakly with water molecules (mostly vdW), and are unable to counterbalance the cost of creating a cavity in water to accommodate the solute. Cavity creation is disfavorable because strong water-water interactions (H-bonds and electrostatics) are disrupted along the interface between solute and water. Thus, hydrophobic aggregation is driven by maximization of water-water interactions which drives a minimization of solute-water contact area leading to the observed hydrophobic association. The cost of re-organizing waters in contact with apolar molecules can be dominated by either enthalpic or entropic contributions. The following paragraphs describe important features underlying hydrophobicity which are relevant to understand the thermodynamics of ligand binding.

The free energy cost of creating a cavity in water, ΔG_{cav}, depends not only on the extent of interfacial surface but also on the shape of that surface [6]. In the specific case of spherical cavities, the ratio between ΔG_{cav} and surface area (A) depends on the radius of the sphere. For spheres with radius r larger than 10 Å, ΔG_{cav} is directly proportional to the surface area of the cavity, but when $r < 10$ Å, ΔG_{cav} is proportional to the volume of the cavity [7]. In other words, $\Delta G_{cav}/A$ is constant when $r > 10$ Å, but increases with r when $r < 10$ Å. Moreover, cavitation is enthalpically dominated when $r > 10$ Å and entropically dominated when $r < 10$ Å[8]. The dependence of $\Delta G_{cav}/A$ on r for small spheres may appear counter-intuitive but can be rationalized on surface curvature: the smaller the sphere, the higher the curvature, and the easier it is for water molecules to make a stable network of hydrogen bonds around the cavity. One may imagine the cavity imposes restrictions on the network of hydrogen bonds, reducing the number of available configurations for which suitable H-bonds exist, thus explaining the entropic character of ΔG_{cav} for small spheres. As radius increases, the cavity surface becomes larger and flatter, asymptotically approaching zero curvature, and making it impossible for waters to form a hydrogen bonding network. At this point $\Delta G_{cav}/A$ no longer increases with radius because curvature is already similar to that of a plane. Since H-bonds are always restricted (it is not a matter of the number of states with suitable H-bonds) the process is enthalpically dominated. Ordering of waters around small apolar solute were disputed by neutron diffraction experiments [9], but later observed by femtosecond mid-infrared spectroscopy [10].

Comparing spherical with non-spherical cavities provides further insight into the importance of shape. Using molecular dynamics Wallqvist and colleagues compared ΔG_{cav} for a sphere and an oblate ellipsoid of the same volume [11]. Cavity volume was not a suitable predictor because ΔG_{cav} of the sphere differed from that of the ellipsoid. Furthermore, the ratio $\Delta G_{cav}/A$ differed between spherical and elliptic shapes, implying that surface area could not be used to predict ΔG_{cav} either. Thus, a curvature correction was proposed to calculate ΔG_{cav} based on the surface of the cavity. Others have also noted that hydration free energies depend on solute shape [6], in particular the non-polar component [12, 13]. This is in opposition to traditional implicit solvation models where ΔG_{cav} is considered proportional to surface area [14].

In the context of ligand binding, the disappearance of the ligand cavity favors the bound state [15] because water-water interactions are restored. However, the overall contribution of ligand desolvation might be unfavorable because the magnitude of solute-water interactions may exceed that of water-water interactions, i.e. the hydration free energy of the solute is negative, stabilizing G_L (see equation 1.2).

A comprehensive study of several scoring functions [5] showed that the ligand apolar surface area buried upon binding (ΔSAS) can predict binding affinity, displaying better correlation with binding free energy than nearly all tested scoring functions, including notable examples such as Glide-SP/XP [16, 17, 18] GoldScore [19] ChemPLP [20] and ChemScore [21, 22]. It is important to note that ΔSAS is unable to identify the correct binding pose from an ensemble of generated poses, i.e. it lacks docking power, but the importance of desolvation effects in ligand-receptor interactions is indisputable.

1.1.1.2 Receptor desolvation

The major difference between desolvation of ligands and desolvation of binding pockets is that ligands provide a mostly convex surface which moderately perturbs water structure, while binding pockets provide a concave surface which highly confines waters inside [23]. These geometrical features of apolar surfaces have implications for the network of H-bonds formed by nearby waters [24]. Under the confinement of apolar binding sites, waters are unable to make H-bonding networks comparable to those found in bulk water, meaning that binding of a ligand which replaces such confined waters results in an enthalpic gain [25, 26], in opposition with the traditional view where entropy is the sole driver of hydrophobic effects. As is illustrated in figure 1.1, binding free energies strongly depend on the structure of water molecules inside the binding pocket [27]. Molecular dynamics simulations also suggest the existance of binding pockets where waters have an extreme difficulty in making hydrogen bonds with each other, causing the binding pocket to be absent of waters (vacuum) for a fraction of time. In such a scenario, binding is largely favored because vacuum is filled by a ligand [28].

Recent methodologies have been developed to characterize the thermodynamics of solvation based on MD simulations, such as WaterMap [28] from Schrödinger Inc. and Grid Inhomogeneous Solvation Theory (GIST) [29] implemented in cpptraj [30]. SZMAP from OpenEye predicts solvation thermodynamics without actually running MD simulations [31]. These methodologies

inform about the absolute free energies of waters in the binding site, an important component of G_R in equation 1.2. Waters with higher free energies contribute more favorably to ΔG_{bind} after being displaced by a ligand. Furthermore, thermodynamics of waters solvating ligand-receptor complexes can affect affinity: it was the determinant factor to explain selective inhibition against different isoforms of phosphoinositide 3-kinases [32].

The concepts described above have been successfully used to understand receptor-ligand interactions: Kelly and Mancera considered the shape and extent of non-polar surfaces to quantify hydrophobicity [33], Cao and Li improved affinity prediction by using a curvature-dependent surface area model [34] and the Glide-XP scoring function [16] identifies locations in the receptor where an hydrogen bond can be formed but its interaction with a water molecule (in the unbound state) is unfavorable because the water molecule is unable to form its additional complement of hydrogen bonds (compared to bulk water). Of course, such situations arise in non-polar environments with specific shapes.

In this thesis, we developed a new desolvation function (chapter 4) which performed reasonably well for calculating free energies of hydration but displayed no advantage in the context of molecular docking (see chapter 5). Since our new desolvation function had no terms to predict the effect of confinement on water-water interactions, its lack of accuracy inside binding pockets could be anticipated. A recent docking software (rDock [35]) uses a desolvation term similar to ours (based on surface areas) which also neglects water confinement, indicating that these concepts are yet to be recognised for their critical role in ligand binding. The field of molecular docking would likely benefit from a fast method to describe hydration under confinement.

1.2 Overview of Scoring Functions

In a recent publication Liu and Wang [36] used four categories to classify scoring functions, which are:

physics based refers to the use of potentials derived from electronic structure methods or already existing forcefields (e.g. Amber and Charmm). Typically, these potentials describe the energy of pair-wise interactions as a function of interatomic distances, and have a well defined physical meaning: they correspond to van der Waals interactions, electrostatics, hydrogen bonds (which may be included in the electrostatics formalism), desolvation, etc.

knowledge based use of potentials derived from observed contacts between defined atom types, as opposed to using physics based potentials. Statistics are drawn from protein-ligand complexes in the PDB. Frequent pairwise contacts are expected to contribute more to ligand binding. Inevitably, statistical potentials capture stabilizing interactions — such as hydrogen bonds — and may resemble physics based potentials to some extent.

empirical use of linear regression to combine physics or knowledge based potentials. The linear model produces weights (or coefficients) that scale the contribution of each individual term. Generally, less than 10 terms are combined linearly (terms can be descriptors or potentials). Interpretation of weights produced by linear regression is straightforward.

descriptor based some scoring functions are developed by feeding a large number of descriptors into machine learning methods such as neural networks or random forests. Compared with the previous category — empirical scoring functions — descriptor based scoring functions use a much larger number of descriptors and combine them using non-linear models. Descriptors inform about intermolecular interactions established between defined atom types. Generally, it is impossible to understand the learned model because the complexity of the descriptors (and also the complexity of the model itself) do not allow for interpretation in physical terms.

These four categories effectively describe the range of methods employed to build scoring functions. In the original work [36] each scoring function was allocated to a single category, but such an allocation was difficult because most scoring functions are built on concepts from multiple categories. In our oppinion, a 1 to 1 relationship between categories and scoring functions is not possible, e.g. empirical scoring functions always have either physics or knowledge based potentials. Our own work on an improved scoring function for zinc metalloenzymes AutoDock4Zn (chapter 3) is an example of a multi-concept scoring function, as it uses (i) physics based potentials, (ii) knowledge based potentials and (iii) combines them using the 'empirical' approach. The new terms to describe zinc coordination were designed based on analysis of protein-ligand complexes, which makes them 'knowledge based' in nature. This concept was specially relevant to distinguish atom types that coordinate zinc in a tetrahedral geometry from atom types without geometrical preferences. The magnitude of the new potentials was calibrated with the 'empirical' approach, by doing a linear fit to experimental affinities. Finally, the underlying forcefield is the same as in standard AutoDock4, which uses physics based potentials.

1.3 Scoring Function Evaluation

Both the development of scoring functions and the preparation of virtual screening campaigns rely extensively on scoring function evaluation. One of the papers in annex [37] focused on the procedures employed for optimizing virtual screening protocols. Here we discuss four metrics (powers) described in a benchmark study by Li et. al. [5], which are:

docking power the ability to identify the native pose from an ensemble of binding poses. In docking software poses are generated by a search algorithm, but in benchmark studies the ensemble of poses should be generated in prior, eliminating bias from different search algorithms implemented along with different scoring functions.

Docking power is often quantified by the number (fraction) of complexes for which the native pose was given the best score.

scoring power — the correlation between scores (or ΔG_{bind} predictions) and the logarithm of experimental affinities, provided that a correct binding pose is used as input.

ranking power — same as scoring power, but instead of using a correlation coefficient (such as Pearson's R), a rank order correlation coefficient is used (such as Spearman's rho or Kendall's tau). Again, known binding poses are used as input.

screening power — simulates a virtual screening scenario where the aim is to identify active molecules from a large pool of molecules. Binding poses are unknown — scoring functions have to first select the most promising binding poses and also score. The most commonly used metrics are based on ROC curves, such as area under the curve (AUC) and enrichment factors.

There is a fundamental distinction among these metrics: scoring and ranking powers rely on already known binding poses, while docking and screening power require the scoring function to guess the native pose.

In the original publication [5] ranking power was restricted to subsets of ligands that bind the same target, and the reported ranking power was the average over all subsets. We note that this concept may also be applied to scoring power. This approach may reduce errors arising from different experimental setups (it is likely that a single research group studied many different ligands for the same target) and can also alleviate the effect from conformational changes in the receptor and receptor desolvation, which are very difficult to model (see figure 1.1). Whether or not this last aspect is an advantage, depends on the specific purposes of the evaluation.

An important aspect when assessing docking power is to consider the quality and rigidity of the crystallographic structure used as reference. Often, a fragment of the ligand and/or part of the binding site has high mobility and does not exist in a single conformation. In these cases, it is important to bear in mind that coordinates from the PDB are not data — electron density is the data — coordinates are just a model. Temperature factors (b-factors) are a good proxy to infer about the relative rigidity of atoms in the complex: higher values are associated with more flexible atoms (electron density is poorly resolved). In the D3R Grand Challenge 2015 (see chapter 5), one of the six ligands for which participants had to predict poses, had some atoms with high b-factors and the ligand was disqualified because it was found to establish contacts with a second molecule of the receptor due to crystallographic packing.

Ideally, scoring functions would excel by all metrics, but such a scoring function does not exist yet. Therefore, it is important to prioritize metrics in order to make evaluation as procedural as possible. From our own work in chapter 5 and from the work of Li et. al [5] it became clear that docking power should be targeted first, followed by screening power, while ranking/scoring power should receive secondary attention. This is the approach that generates the greatest usefulness, either by guiding the setup of a virtual screening protocol or by identifying the best scoring functions

for a specific target. By 'greatest usefulness' we mean the performance in identifying actives from a large pool of molecules or the accuracy in ranking ligands by affinity. The rationale behind this prioritization of metrics is associated with accurate identification of the correct binding mode: in a real scenario binding poses are unknown for the vast majority of ligands, and accurate scores can only be computed on the correct pose. Indeed docking and screening power were highly correlated [5]. According to the authors, this correlation occurs because identifying the native pose of a ligand among its other incorrect poses (docking power) is similar to distinguishing native poses of binders from incorrect poses of non-binders (as non-binders only have incorrect poses).

A recent study [38] claimed that selecting correct poses is not important to predict binding affinity, by observing no correlation between errors in ΔG and RMSD between docked and x-ray poses. However, since errors in ΔG were large for all RMSD range, the simplest and most straightforward conclusion is that ΔG predictions are independent from pose predictions, which is theoretically unreasonable because intermolecular interactions contribute greatly to G_{RL} in equation 1.2. Interestingly, this paper comes from the group that developed RF-Score, proven by us (chapter 5, ref. [39]) and others [3] to lack docking power. However, in our work ligand ranking improves if RF-Score is used to rescore poses selected by a method with decent scoring power (chapter 5, ref. [39]).

From a physical standpoint, it is hard to find a reason for using different scoring functions for pose selection and re-scoring. The native pose exists because it is more stable than all other possible binding modes. Thus, in principle, one would expect that good docking power brings good scoring power.

1.4 Technicalities of Real Life Application

Scoring functions generally require well modeled structures (both for receptor and ligand) to select native binding poses and to produce meaningful affinity predictions [40, 41]. The term 'well modeled' implies reasonable choices for various aspects: overall receptor conformation and alternate locations, protonation states of both receptor and ligand, components of ligand conformation that are not sampled by the docking search (bond lengths, angles and macrocyclic pucker dihedrals) and partial charges (if considered by the scoring function).

The existence of multiple protonation of titratable protein side chains and/or ligands (tautomers and acid/base equilibria) may complicate the identification of correct binding poses if specific donor/acceptor patterns are involved. For example, imidazole side chains of histidine residues can rotate and adopt two stable protonation states, however the exact conformer and protonation state cannot be determined by crystallography. A priori guessing of bound state is needed when the scoring function or docking engine don't sample these properties. Most docking programs allow residue side chains to rotate during the search, but scoring are not prepared to score different protonation states, i.e. the correct tautomer or protonation state has to be defined in prior.

AutoDock4 and other scoring functions implement directional terms for H-bonds and therefore depend on the position of polar hydrogens. In these cases, it is important to model polar

hydrogens of fixed receptor residues in an appropriate position, or intermolecular interactions will be misrepresented. On the other hand, Vina ignores polar hydrogens: they are used exclusively for the purpose of atom typing. For this reason, Vina lends itself for easier automation of virtual screening protocols as it is more permissive to ambiguities in the definition of protonation states and orientation of polar hydrogens.

Alternate locations are inserted by crystallographers when the electron density suggests that a given group of atoms (generally a side chain) exists in two (or more) conformations, sharing its occupancy among the identified conformations. If these flexible side chains are able to adopt different conformations upon binding of different ligands, the docking protocol should be able to identify the correct conformation of the side chain for any given ligand, either by making the side chain flexible during the search, or by using a different receptor model for each relevant conformation. The same applies to different receptor conformations that involve more dramatic changes beyond simple side chain movement. For example, the binding pocket of Heat Shock Protein 90 (HSP90) adopts different conformations for different ligands, a feature we modeled by the use of two different receptor models (see chapter 5). Importantly, one of the conformations was associated with higher internal energy, and the corresponding docking score was penalized accordingly. Overall, receptor flexibility is one of the most problematic issues in molecular docking as it requires previous knowledge about possible receptor conformations, either by crystallographic or NMR studies or by extensive molecular dynamics symulations. Newer docking programs implement algorithms capable of serching receptor conformations to some extent. As an example, AutoDockFR [42] is efficient at sampling several receptor sidechains simultaneously.

In many protein ligand complexes there are interfacial waters mediating ligand-receptor interactions [43]. When different ligands stabilize different interfacial waters in the same receptor, there are two alternatives: (i) to use different receptor models with different fixed water molecules (as we did in chapter 5), or to use a method that samples water positions during the docking search. Due to the prevalence of this problem there are many tools related to prediction of fixed water molecules [44, 45, 46, 47, 48].

Chapter 2

Computational Enzymatic Catalysis

2.1 Linking Calculations to Experiments

Enzymatic catalysis is arguably one of the most complicated processes to simulate and study with computational models because of the different timescales and levels of theory involved. Even the simplest of enzymes (which requires no co-factors, has no allosteric regulation and performs a single step chemical reaction of first-order kinetics) works in three separate steps: (1) substrate binding, (2) chemical step and (3) product release. The chemical step, when covalent bonds are broken and formed, occurs on the same timescale as molecular vibrations: the transition state lasts less than a picosecond (ps)[]. An electronic structure method (see section 2.3) is required to study this step because the rearrangement of covalent bonds is a rearrangement of electronic structure. On the other hand, substrate binding and product release take somewhere from microseconds to miliseconds to occur and is reasonable well modeled with a molecular mechanics description, as long as extensive conformational sampling is performed. The range of methods required to simulate the action of an enzyme increases our chances of making mistakes, which occur due to inappropriate sampling (insufficient or non-equilibrated MD simulation) and/or incorrect description of chemical interactions (wrong forcefield parameters, inapropriate electronic structure method, etc). For this reason, it is critical to use experimental data to validate our models. However, experimental data is very scarce and does not inform about all steps in the chemical cycle. Consider the follwing rate equation for the simple enzymatic cycle described above:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ES \underset{k_{-2}}{\overset{k_2}{\rightleftarrows}} EP \underset{k_{-3}}{\overset{k_3}{\rightleftarrows}} E + P \tag{2.1}$$

where E denotes free enzyme, S free substrate, P free product, ES enzyme-substrate complex, EP enzyme-product complex, k_1 is the rate of substrate binding and k_i are the rate constants associated with each step in the enzymatic cycle. Ideally, there would be experimental values for all k_i, but often only substrate affinity and k_{cat} are determined. k_{cat} is the combined kinetics of the chemical step and product release:

$$E + S \underset{k_{-1}}{\overset{k_1}{\rightleftarrows}} ES \xrightarrow{k_{cat}} E + P \tag{2.2}$$

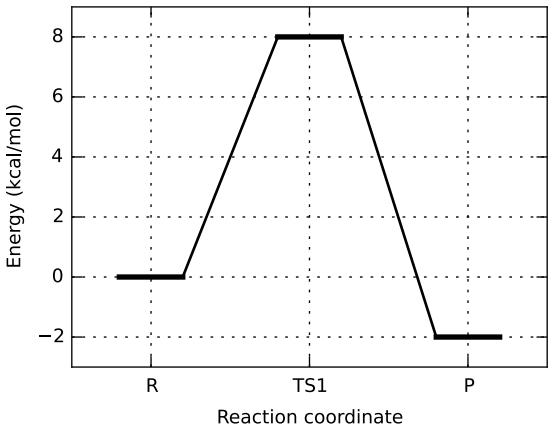


Figure 2.1: Energy surface along conceptual reaction coordinate (toy example). The activation energy is 8 kcal/mol and the reaction energy is -2 kcal/mol.

The grouping of rate constants is a consequence of the inability to measure certain chemical states experimentally: the ES and EP states are indistinguishable and the most informative (yet measureable) phenomena is the increase in product P concentration as the reaction proceeds. In chapters 6 and 7 we studied the chemical step of enzymes and compared our results with k_{cat} values. Under such circumstances k_{cat} defines a maximum possible value for the free energy of activation (see next section).

2.2 Transition State Theory

According to Transition State Theory (TST) chemical reactions proceed through an intermediate structure between reactants and products called the transition state (TS). Structural characterization of transition states depends on the definition of reaction coordinate, which is the minimal set of degrees of freedom associated with breaking and forming bonds, typically expressed as a function of interatomic distances. The transition state is the structure of higher energy along the reaction coordinate has minimum energy along any other coordinate: it corresponds to a saddle point in the energy surface. Characterization of TS structures can take place in both potential and free energy surfaces: the method we employed in chapters 6 and 7 uses potential energy surfaces to find transition states.

The free energy difference between TS and reactants is called the free energy of activation $\Delta G^{\ddagger}$, which is 8 kcal/mol in figure 2.1. The rate of product formation k depends both on $\Delta G^{\ddagger}$ and on temperature, as is predicted by the Eyring equation:

$$k = \frac{k_B T}{h} e^{\frac{\Delta G^{\ddagger}}{RT}} \tag{2.3}$$

where T is temperature, k_B is Boltzmann's constant, h is Planck's constant and R is the ideal

gas constant. The term $\frac{k_B T}{h}$ predicts the number of collisions between reacting particles and $e^{\frac{\Delta G^{\ddagger}}{RT}}$ predicts the number of particles with enought kinetic energy to overcome the transition state. This second term is a consequence of the distribution of kinetic energy as a function of temperature (the Maxwell-Boltzmann distribution), which is applied to condensed matter (such as enzymes) even though it was derived for ideal gases. At first sight temperature always increases reaction rate by (i) increasing the number of collisions in the first term and (ii) increasing the number of particles with enought kinetic energy to surpass the activation barrier. However, if activation entropy is too negative (the number of states associated with the transition state is small), an increase in temperature will increase the free energy barrier because $\Delta G^{\ddagger} = \Delta H^{\ddagger} - T\Delta S^{\ddagger}$. Generally, the reaction rate increases with temperature.

2.3 Electronic Structure Methods

2.3.1 The Behaviour of Electrons

Molecules are made of protons, neutrons and electrons. Protons and neutrons form the nucleus of atoms and are often treated classically in computational chemistry, i.e. the nucleus is described by charge and mass. While this may seem a severe approximation to the quantum nature of protons and neutrons, it is reasonable for the majority of chemical systems. Electrons, on the other hand, cannot be described classically and its behavior is described by a wave function instead. The wave function Ψ is complex-valued and depends on spatial and spin (ω) coordinates: $\Psi(x,y,z,\omega)$ or simply $\Psi(\vec{r},\omega)$.

Born's interpretation of the wave function dictates that the probability of finding an electron in a given part of space is the integral of $|\Psi|^2$ over that space. Since the probability of finding the electron everywhere must be 100%, the following rule applies: $\iiint_{-\infty}^{\infty} |\Psi(x,y,z)|^2 \, dx \, dy \, dz = 1$.

The majority of studies in the field of computational chemistry use the time-independent Schrödinger equation, which ignores the evolution of quantum states over time and also the Born-Oppenheimer (BO) approximation where nuclei are assumed to move orders of magnitude slower than electrons allowing electrons to be permanently equilibrated around nuclei positions. The electronic hamiltonian $\hat{H}_{elec}$ effectively formalizes the BO approximation by ignoring the kinetic energy of nuclei for the calculation of electronic wavefunction:

$$\hat{H}_{elec} = \sum_{i=1}^{n} -\frac{1}{2}\nabla_i^2 + \sum_{i=1}^{n}\sum_{k=1}^{K} \frac{-Z_k}{|\vec{r}_i - \vec{r}_k|} + \sum_{i=1}^{n}\sum_{j>i}^{n} \frac{1}{|\vec{r}_i - \vec{r}_k|} + \sum_{k=1}^{K}\sum_{l>k}^{K} \frac{Z_k Z_l}{|\vec{r}_k - \vec{r}_l|} \tag{2.4}$$

where the first term sums kinetic energies of n electrons, the second term is the electrostatic interaction between n electrons and K nuclei, the third term is the repulsion between electrons and the fourth term is the repulsion between nuclei. Z denotes atomic number. The indices i and j run over electrons while k and l run over nuclei. Denominators $|\vec{r}_a - \vec{r}_b|$ correspond to the distance between particles a and b. Due to the wavefunction description of electrons, the distance between

electrons and other particles is integrated over the probability density $|\Psi(\vec{r})|^2$. Note that $\hat{H}_{elec}$ is expressed in atomic units so the physical constants are unitary.

Electrons are fermions and obey Fermi-Dirac statistics, i.e. two electrons cannot occupy the same quantum state simultaneously. This property is often stated as the Pauli exclusion principle or the antisymmetry requirement, and dictates restrictions to the mathematical description of a multi-electron wavefunction. The alternative to Fermi-Dirac statistics is Bose-Einstein statistics used to describe particles that can occupy the same quantum state (bosons), such as photons and α-particles (He^{2+}). Fermi statistics imply that we cannot express the wavefunction of multiple electrons as a simple product of the individual wavefunctions (known as the Hartree product). The (incorrect) Hartree product for two particles looks like:

$$\Psi(r_1, r_2) = \Psi_1(r_1)\Psi_2(r_2) \tag{2.5}$$

which is incorrect because it does not obey the antisymmetry rule, which dictates that the multi-particle wavefunction changes sign if any two fermions exchange position:

$$\Psi(r_1, r_2) = -\Psi(r_2, r_1) \tag{2.6}$$

A mathematical description that obeys the anti-symmetry rule is accomplished by a Slater determinant:

$$\Psi(r_1, r_2) = \frac{1}{\sqrt{2}} \begin{vmatrix} \Psi_1(r_1) & \Psi_2(r_1) \\ \Psi_1(r_2) & \Psi_2(r_2) \end{vmatrix} = \frac{1}{\sqrt{2}} (\Psi_1(r_1)\Psi_2(r_2) - \Psi_1(r_2)\Psi_2(r_1)) \tag{2.7}$$

which respects the anti-symmetry requirement (equation 2.6). Therefore, it is used in electronic structure methods. This description of the wavefunction complicates the electron repulsion term (third term in eq. 2.4) by making impossible to dissociate one-electron wavefunctions from the total wavefunction. Electron-electron repulsion is the bottleneck in electronic structure calculations.

2.3.2 Basis Sets

Basis sets are pre-calculated solutions for the schrödinger equation. They are based on exact solution of the Schrödinger equation for one electron and are centered on atomic nuclei. In Molecular Orbital (MO) theory, electrons can occupy orbitals that run over entire molecules. In order to build molecular orbitals, atom-centered orbitals are used as building blocks in an approach called Linear Combination of Atomic Orbitals (LCAO). The alternative to MO theory is Valence Bond (VB) theory, where orbitals are confined to individual atoms or covalent bonds as depicted by single Lewis structures (several individual Lewis structures are required to properly describe molecules).

In chapters 6 and 7, we used Pople basis sets which are Gaussian-type orbitals. The numenclature is X-YZg. In this case, 'X' is the number of Gaussian functions for the core orbitals, and 'Y' and 'Z' represent the number of gaussian functions for the inner and outer parts of valence orbitals,

respectively. Since valence orbitals are separated into inner and outer parts the basis set is called double-zeta. An example of a widely used double-zeta basis set is 6-31G. Valence orbitals can be described with further detail by using a triple-zeta approach (X-YZWg). An example would be 6-311G.

Basis sets can be complemented by additional functions that allow molecular orbitals to better adapt to their environments. Diffuse functions are less concentrated on the atomic center and are introduced by the '+' sign, as in 6-31+G for heavy atoms or 6-31++G for both heavy atoms and hydrogens. Polarization functions can also be added for heavy atoms and hydrogens separately: examples are 6-31G(d) which adds one set of d functions to heavy atoms or 6-311G(3df,2p) which adds 3 sets of d functions and one set of f functions to heavy atoms in addition to two sets of p functions to hydrogens.

2.3.3 Hartree-Fock

The Hartree-Fock method approximates the third term in equation 2.4 by calculating the wavefunction of each individual electron in the mean electric field created by the remaining electrons. When the wave function of the i^{th} electron is optimized in the mean field created by all other electrons, the mean field perceived any other electron $j \neq i$ changes. Therefore, one has to iterate several times until all one-electron wavefunctions change no more, i.e. they are consistent with the field they create. For this reason, the Hartree-Fock method is also known as the Self Consistent Field (SCF) method. For the case of closed shell (all orbitals are doubly occupied) the Fock operator for the i^{th} electron is:

$$\hat{F}(i) = -\frac{1}{2}\nabla_i^2 - \sum_{k=1}^{K}\frac{Z_k}{|\vec{r}_i - \vec{r}_k|} + \sum_{j=1}^{N}[2\hat{J}_j(i) - \hat{K}_j(i)] \tag{2.8}$$

where the first term is the kinetic energy of electron i, the second term is the electrostatic interaction between the i^{th} electron and K nuclei, N is the number of doubly occupied orbitals, $\hat{J}$ is the coulomb operator that calculates the electrostatic interaction between the i^{th} electron and the average field created by $j \neq i$ electrons and $\hat{K}$ is the exchange operator that calculates the energy associated with exchange of electrons.

Optimization of the electronic wavefunction aims at finding the lowest energy possible for a given set of nuclear coordinates. According to the variational principle, incorrect wavefunctions always give an energy higher than that associated with the exact wavefunction. Therefore, there is no risk of finding an electronic structure with lower energy than the exact wavefunction (as it would be calculated by the Schrödinger equation).

The difference between Hartree-Fock energy and the exact energy of a single-determinant wavefunction is known as dynamic correlation energy. It corresponds to the error created by optimizing one-electron wavefunctions in the mean electric field of other electrons. Static correlation energy can be recovered by describing the total wavefunction with multiple Slater determinants.

2.3.4 Density Functional Theory

Density Functional Theory (DFT) consists in using electron density instead of an explicit multi-electron wavefunction. Electron density $\rho(\vec{r})$ is simply the density of electrons at each point in space — it's 3-dimensional because $\vec{r} = (x, y, z)$. On the other hand, the exact electronic wavefunction $\Psi(e_1, e_2, ..., e_N)$ is $3N$-dimensional because each electron has three spatial coordinates (plus spin coordinates). The electron density can be computed from the electronic wavefunction:

$$\rho(\vec{r}) = 2\sum_{i=1}^{N} \Psi_i^*(\vec{r})\Psi_i(\vec{r}) \tag{2.9}$$

where N is the number of doubly occupied orbitals. A significant amount of information about individual electrons is neglected when electron density is used. However, Hohenberg and Kohn proved that the wavefunction is a unique functional of the electron density, i.e. it is possible to recover the exact $3N$-dimensional wavefunction from the 3-dimensional electron density. Therefore, all molecular properties that can be computed from the wavefunction can also be computed from the density.

The energy of a molecule in DFT is calculated by a formalism similar to the electronic Hamiltonian (equation 2.4): there is a term for the kinetic energy of electrons, a term for electron-nuclei attraction, a term for nucleus-nucleus repulsion and a term for electron-electron repulsion, all a adapted to deal with electron density instead of a multi-electron wavefunction. In addition, there is an exchange/correlation term V_{XC}, which attempts to predict the energy that arises from electron correlation and from the exchange interaction. These predictions are based on the expected electron correlation/exchange found in a homogenous electron gas. There are different formalisms for calculating V_{XC} energies from the electron gas, leading to a wide variety of DFT functionals. Hybrid functionals — such as B3LYP [49] — incorporate a percentage of Hartree-Fock energy into the V_{XC} term (namely the exchange energy, which has an exact formalism in Hartree-Fock). Because of these approximations, DFT fails for highly correlated systems.

There is a wide variety of DFT functionals which use different variations of the V_{XC} term. It is known that B3LYP is likely to provide reasonable results for calculating various properties in different systems, but benchmark studies are important to identify well suited functionals for particular purposes [50].

2.4 Molecular Mechanics

Molecular mechanics describe the behaviour of molecules in condensed and gas phase using a classic description where each atom is represented by a particle with attributes (partial charge and vdW parameters). The topology of molecules is described by parameters for bonds, angles and dihedrals. In mainstream forcefields such as AMBER [51], GROMOS [52] OPLS [53] the topology of molecules remains unaltered through the simulation (no breaking or forming of covalent bonds).

Exceptions are reactive forcefields such as the empirical valence bond (EVB) [54] approach which has been extensively used to study enzymatic catalysis.

Most forcefields have been validated and parameterized for proteins and DNA. Mainstream forcefields for general organic chemical space are GAFF [55] and OPLS-AA [56]. Partial atomic charges are derived from the electronic structure, tipically with RESP [57] or AM1-BCC [58, 59]. The most used waters models (TIP3P and TIP4P) have been parameterized to reproduce properties of water such as the density profiles over temperature [60].

2.4.1 Non-Bonded Interactions

Non-bonded interactions refer to van der Waals (vdW) and electrostatics. Hydrogen bonds are described within the general framework of electrostatics — the interaction energy depends on the partial charges of the interacting atoms. Older versions of the AMBER forcefield used a 12-10 Lennard-Jones potential, similar to the 12-6 potential used for vdW interactions. The AutoDock scoring function inherited this 12-10 potential and added a directional component which decreases the interaction energy if the orientation between acceptor and donor deviates from the optimal geometry.

Non-bonded interactions are calculated between all pairs of atoms wich are not bound to each other. Atoms separated by two covalent bonds (1-3 interactions) are also discarded. 1-4 interactions (separated by three bonds) are taken into account but are scaled by a factor of 0.5 (tipically).

A 12-6 Lennard Jones potential is used to describe vdW interactions between atoms i and j at distance r_{ij}:

$$E_{ij}^{vdW}(r_{ij}) = \varepsilon_{ij}\left[\left(\frac{R_{ij}}{r_{ij}}\right)^{12} - 2\left(\frac{R_{ij}}{r_{ij}}\right)^{6}\right] \tag{2.10}$$

where ε_{ij} is the well depth (the interaction energy at equilibrium distance) and R_{ij} is the equilibrium distance. The well depth of the interaction between two atoms is the geometric mean of individual atomic parameters $\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j}$ while the equilibrium distance of the interaction is calculated as the aritmetic mean: $R_{ij} = \frac{1}{2}(R_i + R_j)$.

The electrostatic energy between two atoms with partial charges q_i and q_j is calculated by Coulomb's law:

$$E_{ij}^{elec}(r_{ij}) = \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \tag{2.11}$$

where ε_0 is the electric permitivity of vaccum and r_{ij} is the distance between atoms i and j.

2.4.2 Bonded Interactions

Bonded interactions refer to potentials describing atoms separated by 1, 2 or 3 bonds, which are refered to as 1-2, 1-3 and 1-4 interactions, respectively, or alternatively as bonds, angles and dihedrals.

Bonds between two atoms are described by a harmonic potential which takes the form:

$$E_{ij}^{bond}(r_{ij}) = k_{ij}(r_{ij} - R_{ij})^2 \tag{2.12}$$

where r_{ij} is the distance between atoms i and j, R_{ij} is the equilibrium distance for the bond and k_{ij} quantifies the rate of increase in potential energy as the interatomic distance deviates from R_{ij}. Angles are described by an identical formalism:

$$E_{ijk}^{angle}(\theta_{ijk}) = k_{ijk}\left(\theta_{ijk} - \theta_{ijk}^{eq}\right)^2 \tag{2.13}$$

where θ_{ijk} is the angle between atoms i, j and k, θ_{ijk}^{eq} is the equilibrium angle and k_{ijk} quantifies the rate of increase in potential energy as the θ_{ijk} deviates from θ_{ijk}^{eq}.

Dihedral angles describe the energy profile associated with rotation of a bond. Four atoms a, b, c and d are necessary to define a dihedral angle θ_{abcd} — the rotating bond is that between the central atoms b and c. The energy of the dihedral is calculated as follows:

$$E_{abcd}^{dihedral}(\theta_{abcd}) = \frac{k_{abcd}}{2}[1 + cos(n\theta - \gamma)] \tag{2.14}$$

where k_{abcd} is the height of barriers opposing rotation, n is the periodicity (number of local minima over a 360° rotation) and γ sets the position of local minima with respect to 0°. Dihedral angles are also used to describe the planarity of atom with three substituents, in which case they are refered to as 'improper' dihedrals. As an example, the planarity of amide bonds and the pyramidal geometry of amines are described by improper dihedrals.

2.5 QM/MM: ONIOM

Enzymes are large molecules, but the chemical step occurs at a specific location within the active site. Therefore, it is not necessary to calculate the electronic structure of all enzyme atoms. Nevertheless, atoms surrounding the reactive region may contribute to the chemical step by providing a pre-organized electrostatic environment that stabilizes the transition state, and by imposing significant spacial restraits on the reactive region. Thus, it is appealing to use quantum mechanics (QM) exclusively for atoms whose electron structure varies significantly along the reaction, and molecular mechanics (MM) for the remaining atoms.

The ONIOM approach is a scheme to combine QM with MM. ONIOM stands for **O**ur own **N**-layered **I**ntegrated molecular **O**rbital and molecular **M**echanics. Typically, systems are divided in two layers, one that uses a QM method (high layer) and another with a MM description (low layer). The entire system (both layers) is called the 'real system', and the high layer is also known as 'model system'. The ONIOM energy for two layers is calculated as follows:

$$E_{ONIOM} = E_{QM}^{high\ layer} + E_{MM}^{real\ system} - E_{MM}^{high\ layer} \tag{2.15}$$

The boundary between layers correspond to covalent bonds being cut: only one atom remains in the high layer. The missing atom (from the high layer) is replaced by a hydrogen, allowing the calculation of a meaningful electronic structure and therefore a reasonable energy value for the model system ($E_{QM}^{high\ layer}$). The cut between layers is preferably made at single bonds in order to minimize errors.

There are two major variations to calculate the electrostatic interaction between high and low layers: electrostatic embedding and mechanical embedding. In electrostatic embedding atomic partial charges contribute the electrostatic potential where electronic wavefunctions are optimized — the electronic structure is effectively modified by the low layer. The electrostatic interaction between layers appears in the $E_{QM}^{high\ layer}$ term in eq. 2.15. In mechanical embedding, the electronic structure is calculated in vaccum (low layer atoms are absent), and atomic partial charges are assinged to QM atoms based on the calculated electronic structure. The interaction energy is calculated classically and appears in the $E_{MM}^{real\ system}$ term.

Overall, QM/MM approaches are useful to study the importance of electrostatic effects when the system is too large to model with an electronic structure method and also to provide a realistic environment that guarantees that reacting atoms remain in a conformation that does not collide with the enzyme. This also has the property of exposing the conformational landscape of enzymes: in chapters 6 and 7 we used the ONIOM method to demonstrate that the chemical step can only occur at certain enzyme conformations where specific enzyme-substrate interactions take place. This has implications for our understanding of enzyme catalysis because it emphasizes that transition state stabilization varies over time.

Chapter 3

AutoDock4$_{Zn}$ An improved AutoDock forcefield for small-molecule docking to zinc metalloproteins

Diogo Santos-Martins, Stefano Forli, Maria João Ramos and Arthur J. Olson

Adapted from ref. [61].

In this work I ran all experiments, analyzed results and wrote parts of the paper.

3.1 Abstract

Zinc is present in a wide variety of proteins, and is important in the metabolism of most organisms. Zinc metalloenzymes are therapeutically relevant targets in diseases such as cancer, heart disease, bacterial infection and Alzheimer's disease. In most cases a drug molecule targeting such enzymes establishes an interaction that coordinates with the zinc ion. Thus, accurate prediction of the interaction of ligands with zinc is an important aspect of computational docking and virtual screening against zinc containing proteins. We have extended the AutoDock forcefield to include a specialized potential describing the interactions of zinc-coordinating ligands. This potential describes both the energetic and geometric components of the interaction. The new forcefield, named AutoDock4$_{Zn}$, was calibrated on a dataset of 292 crystal complexes containing zinc. Re-docking experiments show that the forcefield provides significant improvement in performance in both free energy of binding estimation as well as in root mean square deviation from the crystal structure pose. The new forcefield has been implemented in AutoDock without modification to the source code.

3.2 Introduction

Zinc is present in numerous biological structures, and is found in virtually all aspects of metabolism across multiple species[62]. It can play a structural role as in zinc finger proteins, the most preva-

lent proteins in eukaryotic genomes [63], and is present in all enzyme classes [64], usually in the form of coordinated Zinc(II) or Zn^{2+} ion. Zinc metalloenzymes are therapeutically relevant targets in many diseases, like heart disease[65], cancer[66, 67], bacterial infections[68] and Alzheimer[69, 70]. In most cases, a drug molecule establishes coordination bonds with the zinc ion [71] present in the protein, thus, an accurate description of this interaction is crucial for drug design.

To properly model the zinc coordination interactions, two issues should be addressed: the coordination geometry and the interaction strength.

Most forcefields describe metal coordination using descriptions derived from the original Stote and Karplus nonbonded model[72], where the interaction is described using Lennard-Jones and Coulomb potentials. This description relies on assignment of partial charges [72, 73, 74], and thus accuracy becomes strongly dependent on the choice of charge model. Also, an electrostatic model based on the filled valence orbital of the Zn^{2+} ion fails to explain the prevalence of histidine and cysteine over the more electronegative carboxylate groups of glutamate and aspartate as the most frequent zinc coordinating residues [75, 76, 77, 78].

Moreover, some high potency inhibitors coordinate Zn^{2+} *via* uncharged nitrogens with electron lone pairs, such as those found in sulfonamides [79] and imidazoles [67] (see figure 3.13), that seem to interact more strongly than negatively charged nitro groups[68]. Recently, DFT calculations were used to calibrate a nondirectional zinc coordination forcefield independent of atomic partial charges[80].

Polarization and charge transfer models[81, 82, 83] could provide a more accurate description, although their computational complexity makes them unsuitable for dockings, which typically involve a large number of energy estimations over the course of the calculation.

The coordination geometry issue is addressed differently by bonded and nonbonded models. It has been recently demonstrated that zinc exhibits a strong preference for the tetrahedral geometry, with some of the previously observed variability in coordination spheres being artifactual [78]. Bonded models, such as the Zinc AMBER Force Field [84], describe the tetrahedral coordination with harmonic potentials and angle terms for an explicit bond that provides directionality. Due to the requirement of the explicit bonds where ligands coordinate with zinc, bonded models are not suitable for docking calculations.

In nonbonded models, few forcefields provide directional potentials. Two examples that do are the cationic dummy atom model [66] and the scoring function implemented in FlexX[85]. However, in this latter case, while improving coordination geometry accuracy, no improvements in binding energy prediction were reported.

To be suitable for docking, and virtual screens in particular, modelling the interaction with zinc must provide a description of the geometry that is computationally efficient, and good accuracy in the estimation of the interaction strength.

In this paper, we report the development of a directional, charge-independent model for zinc-coordination forcefield for AutoDock4 which provides higher accuracy than the standard force-

field. Interactions are modeled independently for different atom types, providing specific potentials for each one.

Over the years, AutoDock and its forcefield were modified by us and others to improve scoring of low affinity ligands[86] or to obtain a scoring function tailored to specific targets, like kinases[87]. Other methods added new features, like receptor flexibility models[88, 89], flexible macrocycle docking[90] and docking with waters[44]. The highly customizable architecture of the program allows implementation of substantial modifications relatively easily, and without requiring source code changes.

3.3 Methods

To implement the new zinc-coordination model, first we identified a dataset of high-quality complexes for which experimental affinity values had been determined. The dataset analysis enabled determination of the parameters for geometrical terms that were then calibrated to fit within the AutoDock forcefield. The new forcefield, named AutoDock4$_{Zn}$, was then cross-validated on the dataset.

3.3.1 Dataset creation

To design the new forcefield, a suitable set of zinc metalloprotein-ligand complexes was defined. The ligands cover a wide range of structure diversity and binding affinity, thus providing an optimal calibration set for generic applicability of the forcefield for drug design.

In order to build the dataset, the Binding MOAD [91] was filtered using the following criteria: *a)* presence of at least one zinc ion; *b)* experimentally determined inhibition (K_i) or dissociation (K_d) constants; *c)* no alternate conformation or missing atoms for the ligand and *d)* no alternate side chain conformations in receptor residues within 5 Å from any ligand atom. This filtering led to a set of 510 complexes, which were downloaded from the Protein Data Bank[92].

These complexes were then analyzed to isolate and characterize the zinc coordination geometry within the receptor and its interaction with ligands. Each complex was classified accordingly to the the number of receptor (r) and ligand (l) atoms within coordination distance ($\leq$2.8 Å for sulfur atoms, $\leq$2.5 Å for all others) from the zinc ion[85], and denoted as $Zn_{r,l}$.

A specific treatment was used to analyze the coordination geometry of carboxylic acids. Carboxylic acids from aspartate or glutamate side chains have been described to coordinate zinc mainly with bidentate, monodentate, *syn* or *anti* modes. However, it has been demonstrated that carboxylate groups can adopt any coordination geometry ranging between mono- and bidentate [93, 94], which is poorly described by a discrete classification scheme. To address this issue, carboxylic acid groups on receptors were always considered as monovalent and represented by a weighted average of the position of the two coordinating oxygens (see figure 3.1); the method used to calculate the weighted average for carboxylic acids is described in the Supporting Information.

The distribution of different coordination geometries is summarized in table 3.1. The most represented coordination geometry in our data set is the tetrahedral one ($Zn_{3,1}$, $Zn_{4,0}$), that was

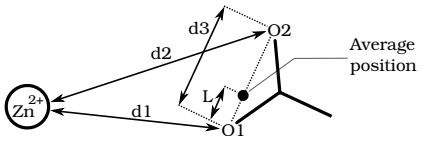


Figure 3.1: Definition of carboxyl group average atom. Details about the methods are reported in Supporting Information.

indeed found to be the most common in biological systems[78]. Other geometries, like five- ($Zn_{4,1}$, $Zn_{3,2}$) and six-coordinated ($Zn_{3,3}$, $Zn_{4,2}$), were also found, but were much rarer.

Complexes where ligands were not directly involved in zinc coordination (*i.e.*, $l = 0$) were discarded. This included also $Zn_{4,0}$ cases, where zinc plays a structural role helping protein folding[76, 75], coordinating four cysteine side chains. Some of the $Zn_{4,0}$ cases were misclassified as $Zn_{2,0}$ because the zinc ion bridges two monomeric units that were split during the analysis process.

Geometries where receptor atoms were not involved, or only partially involved in zinc interactions ($0 \leq r \leq 2$) were also discarded upon visual inspection. In particular, the $Zn_{0,0}$ class contains complexes where Zn is used as an aid in crystallization and has no biological significance, surrounding the protein structures often in large number and at toxic concentrations[78]. Finally, 5 complexes involving serine protease inhibitors from $Zn_{2,2}$ class, were discarded because zinc is known to be recruited transiently as co-inhibitor only, and it is not consistently present in the binding site[95].

This left four coordination classes, $Zn_{3,1}$, $Zn_{3,2}$, $Zn_{3,3}$, $Zn_{4,1}$, $Zn_{4,2}$, resulting in a calibration set of 292 unique complexes. A summary of ligand properties in the set is shown in figure 3.2.

For complexes where the tetrahedral coordination geometry is possible ($Zn_{3,x}$), we analyzed the distribution of the ligand atoms coordinating zinc, using the AutoDock atom types: NA (nitrogen HB acceptor), N (nitrogen non-HB acceptor), OA (oxygen HB acceptor) and SA (sulfur HB acceptor). The ideal zinc tetrahedral geometry was calculated with respect to the averaged position of receptor atoms. The tetrahedral plane was defined as the plane calculated between average coordinating receptor atoms and the zinc atom (figure 3.7).

Table 3.1: Number of $Zn_{r,l}$ classes for each zinc ion found in the initial data set. Complexes with at least one the classes in bold were selected for the final dataset

$Zn_{r,l}$		l			
		0	1	2	3
r	0	14	3	1	0
	1	56	1	0	0
	2	72	2	5	0
	3	57	**244**	**43**	**3**
	4	214	**12**	**15**	0
	5	8	0	0	0

Then, we measured the deviation of ligand atoms from the ideal position in the tetrahedral geometry, defined as the angle between the vector Zn-ligand atom and the tetrahedral plane.

In figures 3.3, 3.4, 3.5 and 3.6 are shown the tridimensional scattering coordinating ligand atoms with respect to the tetrahedral zinc (*a,b*), and their angle deviations (*c*). More details about the alignment method and analysis are reported in Supporting Information.

The deviations analysis showed that nitrogen HB acceptor (NA) is consistently found very close to the ideal position ($> 80\%$ within $\leq 10^{\circ}$, figure 3.3). On the other hand, the placement of nitrogen non-HB acceptor (N), and oxygen (OA) and sulfur (SA) is less well defined, appearing to be dependent solely on the accessibility of the zinc atom in the receptor (figures 3.4, 3.5, and 3.6).

3.3.2 New forcefield

The standard AutoDock forcefield supports several ligand-metal interactions[96]. Similar to all other pairwise interactions in the forcefield, the interaction between ligand atoms and metals contained in the receptor is described mainly by van der Waals (ΔH_{vdW}) and Coulomb electrostatic (ΔH_{elec}) terms, and to a smaller degree, by the desolvation term (ΔG_{desolv}). This approach has several limitations. First, the van der Waals equilibrium distances for the atoms involved in zinc coordination are significantly larger than the coordination distances[74, 80] (*i.e.*, for nitrogen, the vdW equilibrium distance is 2.49 Å, compared to coordination distance of 2.0 Å). Second, due to the lack of a specialized terms for the metal coordination, directionality is not accounted for. Finally, while the electrostatic term is very effective in describing interactions involving partial charges, it makes the energy function highly sensitive to strongly charged groups, such as metals with formal charges. Also, in the Gasteiger[97] charge model used in AutoDock[96], oxygen atoms are systematically assigned a more negative charge than nitrogen and sulfur, thus resulting in the prefered candidates for chelating positively charged metal. While this approach is accurate enough for magnesium ion interactions, it is not sufficient to properly describe zinc coordination preferences.

From our data set analysis, we found that the coordination of zinc requires a specialized treatment, so we modified the standard AutoDock forcefield. The standard forcefield includes the following terms (eq. 3.1)[96]:

$$FEB = W_{vdw}(vdW) + W_{hb}(Hbond) + W_{elec}(Elec) + W_{sol}(Desolv) + W_{tor}(TorsDoF) \qquad (3.1)$$

where the Free Energy of Binding (FEB) is calculated as a sum of van der Waals (vdW), hydrogen bond ($Hbond$), Coulomb electrostatic ($Elec$), desolvation ($Desolv$) and ligand torsional entropy ($TorsDoF$); each term is weighted by a specific value (W_{term}) estimated using a linear regression model.[96] To extend the forcefield, we first disabled the electrostatic potential for zinc by setting its partial charge to zero. Then, the pairwise interactions of each atom types involved in zinc coordination was defined as a new potential energy term. For N, OA and SA atom types, spherical potentials $V_{Zn,N,}$, $V_{Zn,OA}$ and $V_{Zn,SA}$ were defined to reflect the known coordination distances, by

adapting the van der Waals potential in the AutoDock forcefield (eq. 3.2):

$$V_{ij} = \varepsilon_{ij}\left[\left(\frac{r_{ij}}{r}\right)^{12} - 2\left(\frac{r_{ij}}{r}\right)^{6}\right] \tag{3.2}$$

The pairwise equilibrium distance r_{ij} between zinc and N, OA and SA atom types was set to 2.0, 2.1 and 2.25 Å, respectively, and independent ε well-depth values were estimated. Spherical potentials are particularly suitable for accurately reproducing hydroxamate coordination geometries[98, 99].

For the NA type a new directional tetrahedral potential $V_{TZ,NA}$ was defined and the interaction with zinc was splitted in two separate components. The repulsive component is mediated by the zinc atom, while the attractive component is mediated by a new pseudoatom TZ that has been added to the standard forcefield table[100]. The pseudoatom interacts only with NA, therefore no interaction is defined with any other atom type. The pseudoatom is added in the receptor structure for all complexes where the tetrahedral coordination geometry is present, *i.e.*, all $Zn3_{,x}$ classes, where only three receptor atoms are coordinating zinc. The pseudoatom is placed at the unoccupied vertex of the tetrahedral geometry, located at the optimal coordination distance for nitrogen ($r_{ij} = 2.0$ Å) (figure 3.7(a)), and an attractive 12-6 potential with a corresponding ε is defined (figure 3.7(b)). Finally, the zinc-hydrogen pairwise interaction was eliminated to prevent clashes that would interfere with the proper interaction between groups like sulfonamide -NH_2, or hydroxyl, with zinc. This allows ligands to establish the proper coordination interaction independent of the orientation of the hydrogen with respect to the heavy atom.

Therefore, the following potential was added to eq. 3.1:

$$ZincCoord = V_{TZ,NA} + V_{Zn,N} + V_{Zn,OA} + V_{Zn,SA} \tag{3.3}$$

and the FEB becomes the linear combination of the five standard AutoDock terms plus the new zinc coordination pairwise potential.

All modifications to the AutoDock forcefield were made by adapting the forcefield table and parameter files, without source code modifications. The details of the implementation are described in the Supporting Information.

The ε values for eq. 3.3 were then calibrated independently from each other and from the other terms in equation 3.1.

3.3.3 Calibration protocol

The new forcefield was calibrated with an iterative least squares scheme. Initial attempts to calibrate combined terms from eq. 3.1 and 3.3 led to performance degradation in non-zinc complexes. Optimization of different term combination were tried, and best results were obtained by optimizing only terms in eq. 3.3, while keeping the standard terms (eq. 3.1) unmodified.

The calibration protocol consisted of the following steps: *a)* crystallographic structures of the ligands were minimized with the current version of the forcefield, using Solis-Wet local search

implemented in AutoDock[96]; *b)* unweighted terms were calculated from minimized structures; *c)* a regression model was built; *d)* weights from the new regression model were used in the next minimization step. The protocol iterated through steps *c* and *d* five times to achieve convergence; stable weight values were achieved after the first two iterations.

Initial calibration results and cross-validation tests showed that no statistical significance could be achieved for the $V_{Zn,N}$ term. This is likely due to insufficient experimental data for the N atom type. Therefore, standard forcefield term for this interaction (i.e. van der Waals) was restored, while keeping the correct equilibrium radius (2.0 Å) identified in the analysis. Then the calibration was repeated omitting the $V_{Zn,N}$ term from eq. 3.3. Final forcefield weights were selected from the last iteration, with a residual standard error was 2.804 kcal/mol. Final coefficients and extended analysis of the iterative calibration are described in the Supporting Information.

3.4 Results and Discussion

Predictive capabilities of the regression model were assessed with 5-fold cross-validation. The dataset was divided in five bins containing an approximately uniform distribution of ligand atom types coordinating zinc, then re-docking calculations were performed. Cross-validation docking results are summarized in table 3.2. Details on docking preparation and RMSD calculations are available in Supporting information. Reproducing proper metal-coordination geometries and accurate energy estimations are notoriously difficult, especially for zinc[101]. Performance of AutoDock4$_{Zn}$ was evaluated accordingly to three different criteria: FEB estimation error, ligand pose RMSD calculated on all heavy atoms, and deviation from ideal zinc coordination geometry. Overall, the new AutoDock4$_{Zn}$ forcefield performed significantly better than

Table 3.2: Cross-validation of docking performances and FEB estimation accuracy

	FEB error (kcal/mol)			RMSD (Å)		RMSD$_{Zn}$ (Å)	
	<1.0	<2.0	<3.0	<2.0	<2.5	<1.0	<1.5
AutoDock4$_{Zn}$	32%	64%	81%	45%	51%	75%	80%
AutoDock4	18%	34%	53%	36%	42%	33%	46%
Vina	20%	38%	64%	45%	52%	37%	52%

both standard AutoDock and Vina forcefields. These forcefields provide roughly the same prediction errors in FEB estimations, while AutoDock4$_{Zn}$ consistently improved success rate (+50% with $<$ 1.0 kcal/mol) (figure 3.8). The new forcefield also improves RMSD accuracy over the standard AutoDock forcefield in pose prediction accuracy (RMSD), producing results comparable with Vina (figure 3.9). Not surprisingly, a remarkable improvement was achieved in reproducing the proper zinc-coordination geometry (RMSD$_{Zn}$), where AutoDock4$_{Zn}$outperforms the two other forcefields by a large amount (+127% success rate $<$ 1 Å, figure 3.10).

The use of specific potentials for describing the interaction is the main factor responsible for such an improvement, as showed by re-docking experiments.

Table 3.3: Docking performances and FEB estimation accuracy on NEP (1r1j)

	FEB error (kcal/mol)	RMSD (Å)	RMSD$_{Zn}$ (Å)
AutoDock4$_{Zn}$	+0.74	1.21	1.04
AutoDock4	+1.76	4.85	6.38
Vina	+4.68	9.19	4.78

3.4.1 Examples

In some cases, where sulfur is directly involved in coordinating zinc, neither AutoDock4 nor Vina forcefields were able to establish the proper interactions between the ligand and the zinc ion. A key example of is provided by re-redocking results of a potent inhibitor of neutral endopeptidase (NEP)[102] in the crystallographic complex with the PDB Id *1r1j* (figure 3.11). Results are summarized in table 3.3. Both AutoDock4 and Vina predicted zinc to be coordinated by the carboxylate group, resulting in a misalignment of the ligand with respect to the receptor. The AutoDock4$_{Zn}$, on the other hand, predicted the proper coordination by sulfanyl and carbonyl groups and provided more accurate FEB estimation. Similar results were found when re-docking a potent aryl sulfonamide TACE inhibitor in the crystallographic complex with PDB Id *1oi0* (figure 3.12). The increase in docking pose prediction and accuracy in the coordination geometry indentification resulted in a more precise FEB estimation (table 3.4). It must also be noted also that no performance degredation was found when docking compounds without ligand-zinc interactions using the new forcefield. Improvements provided by AutoDock4$_{Zn}$ makes it suitable for virtual screening campaigns involving zinc.

3.5 Conclusions

We extended the standard AutoDock forcefield to include a specialized potential describing interactions of zinc-coordinating ligands. The potential has a description for both energetic and geometric components of the interaction

The new forcefield, named AutoDock4$_{Zn}$, was calibrated on 292 complexes containing zinc using an iterative linear regression model. Re-docking experiments showed that the forcefield provides a considerable improvement in performance when compared to both standard AutoDock4 and Vina forcefields. Improvements are particularly relevant in the accuracy of FEB estimations, as well as in reproducing the proper coordination geometry. In fact, AutoDock4$_{Zn}$ provides a significant advantage when docking zinc-coordinating ligands, as shown in the examples described.

Table 3.4: Docking performances and FEB estimation accuracy on TACE (2oi0)

	FEB error (kcal/mol)	RMSD (Å)	RMSD$_{Zn}$ (Å)
AutoDock4$_{Zn}$	+0.33	2.10	0.65
AutoDock4	+1.56	7.98	11.43
Vina	+2.50	2.34	5.03

Table 3.5: Docking performances and FEB estimation accuracy on farnesyltransferase (1s63)

	FEB error (kcal/mol)	RMSD (Å)	$RMSD_{Zn}$ (Å)
$AutoDock4_{Zn}$	+0.07	0.68	0.34
AutoDock4	+3.52	7.62	7.78
Vina	+2.93	0.50	0.54

Due to the fully-configurable nature of AutoDock4, the new potential was implemented by modifying the standard forcefield tables and few Python helper scripts available at `http://autodock.scripps`

Moreover, the potential itself does not add any overhead to the docking calculation. There is no increase in the search complexity nor in the computational power requirement, therefore docking speed is completely unaffected. Also, performance on dockings not involving zinc-coordination are unchanged. Therefore, accuracy increase and lack of computational overhead make the $AutoDock4_{Zn}$ suitable for virtual screening campaigns, particularly when coordination of zinc is important.

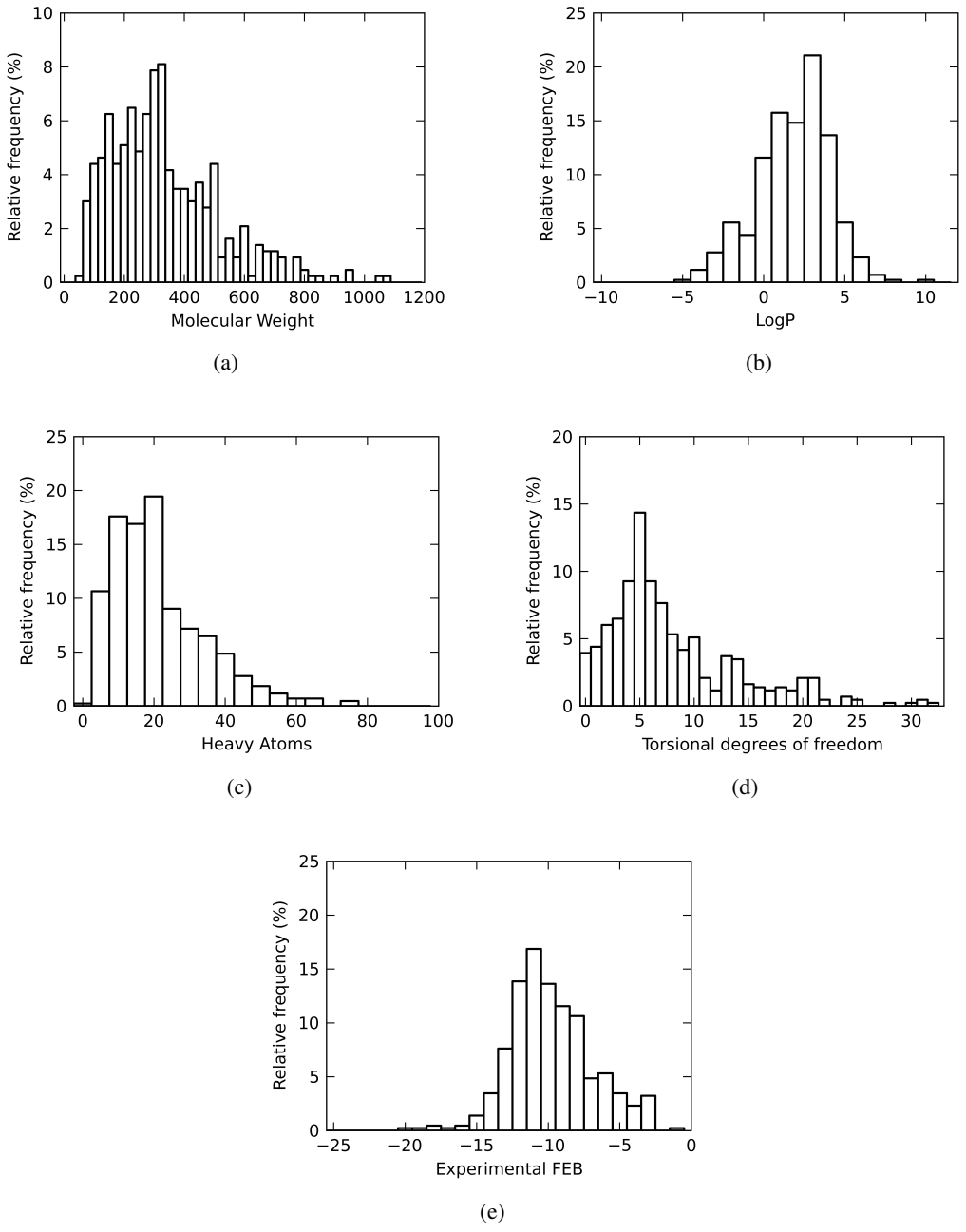
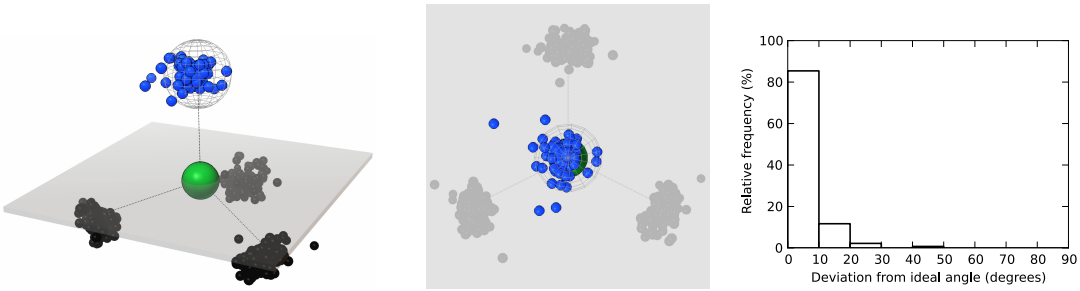
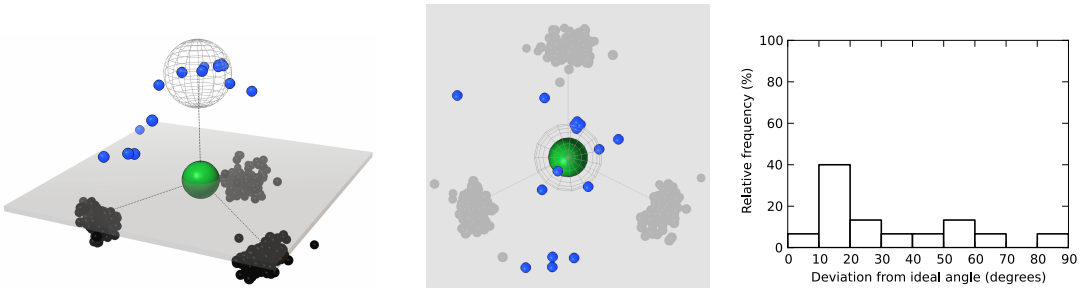


Figure 3.2: Summary of the distributions of ligand properties in the final dataset: molecular weight (*a*), LogP (*b*), number of heavy atoms (*c*), torsional degrees of freedom (*d*), experimental free energy of binding (*e*)



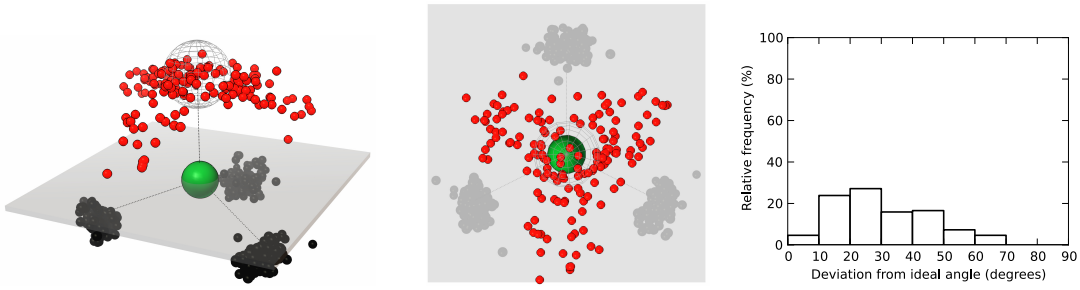
(a) (b) (c)

Figure 3.3: Distribution of 137 NA atom types coordinating zinc: *(a)* perspective projection; *(b)* top view; *(c)* angle histogram. Atoms are shown as spheres: receptor atoms *(black)*, zinc *(green)*; NA atoms *(blue)*. Tetrahedral geometries are colored in *gray*; tetrahedral plane is shown as semitransparent polygon; pseudoatom location is shown as wireframe sphere.



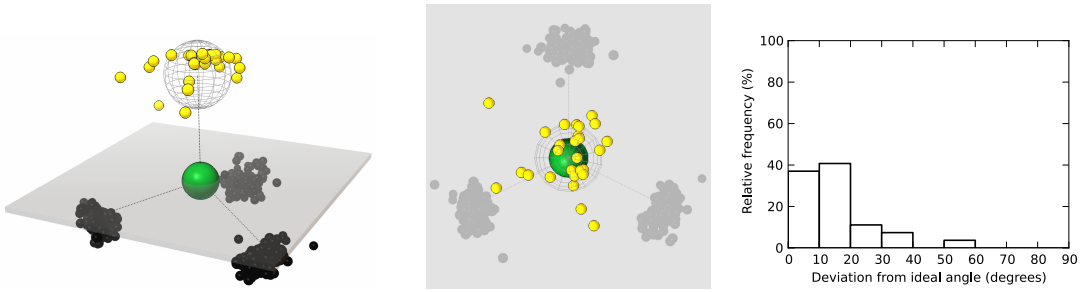
(a) (b) (c)

Figure 3.4: Distribution of 15 N atom types coordinating zinc: *(a)* perspective projection; *(b)* top view; *(c)* angle histogram. Atoms are shown as spheres: receptor atoms *(black)*, zinc *(green)*; N atoms *(blue)*. Tetrahedral geometries are colored in *gray*; tetrahedral plane is shown as semitransparent polygon; pseudoatom location is shown as wireframe sphere.



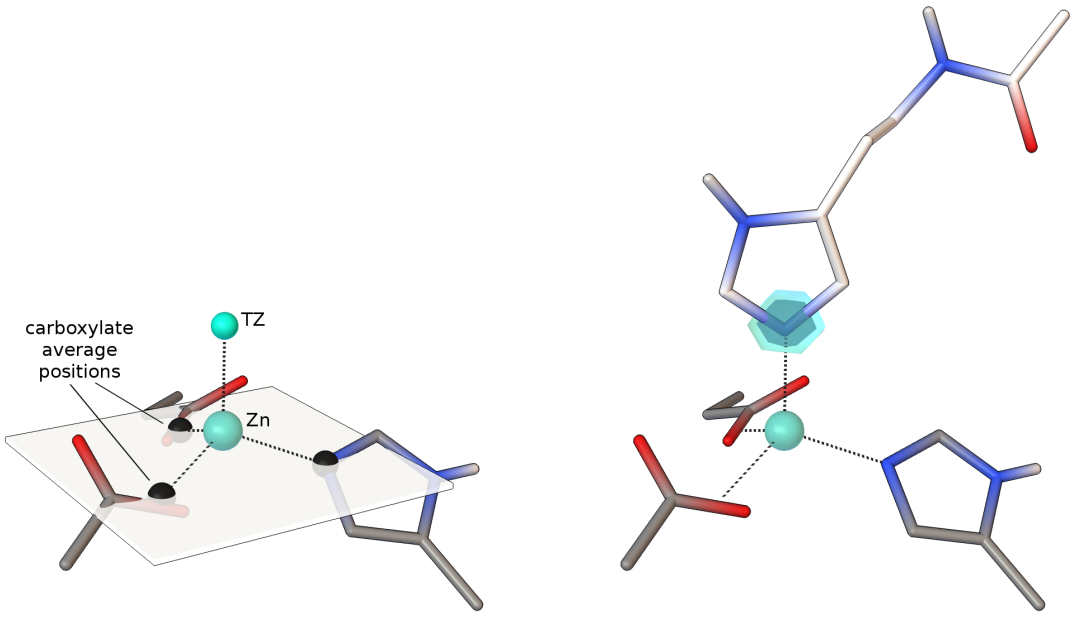
(a) (b) (c)

Figure 3.5: Distribution of 151 OA atom types coordinating zinc: *(a)* perspective projection; *(b)* top view; *(c)* angle histogram. Atoms are shown as spheres: receptor atoms *(black)*, zinc *(green)*; OA atoms *(red)*. Tetrahedral geometries are colored in *gray*; tetrahedral plane is shown as semitransparent polygon; pseudoatom location is shown as wireframe sphere.



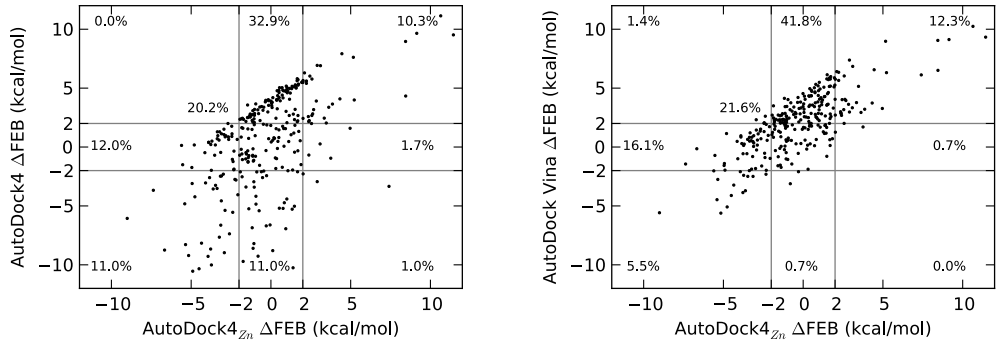
(a) (b) (c)

Figure 3.6: Distribution of 27 SA atom types coordinating zinc: *(a)* perspective projection; *(b)* top view; *(c)* angle histogram. Atoms are shown as spheres: receptor atoms *(black)*, zinc *(green)*; SA atoms *(yellow)*. Tetrahedral geometries are colored in *gray*; tetrahedral plane is shown as semitransparent polygon; pseudoatom location is shown as wireframe sphere.



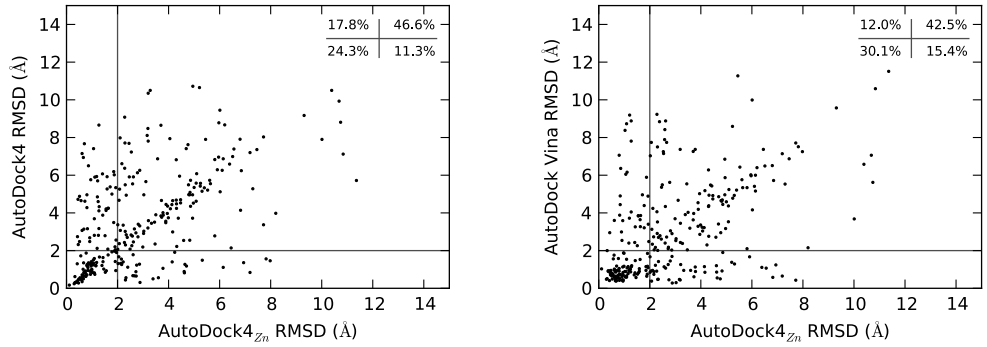
(a) (b)

Figure 3.7: Tetrahedral zinc geometry. *(a)* Ligand and receptor atoms are shown as sticks colored by atom type. The tetrahedral plane defined by three receptor atoms (*black* spheres) is determined. The TZ pseudoatom is located at unoccupied corner of the ideal tetrahedral geometry. Coordination geometry is calculated on weighted average oxygen positions from carboxylic side chains. *(b)* The potential for atom type NA (nitrogen acceptor) is shown as iso-contour surfaces (*cyan*).



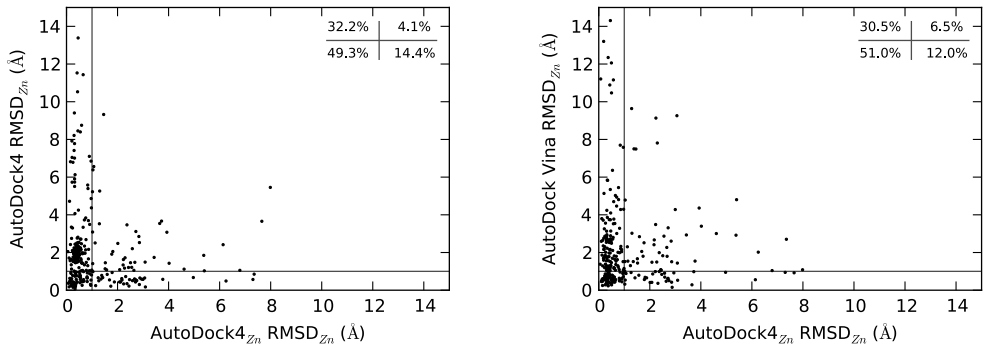
(a) (b)

Figure 3.8: Comparison of FEB prediction errors of the new forcefield with (*a*) standard AutoDock4 forcefield and (*b*) AutoDock Vina



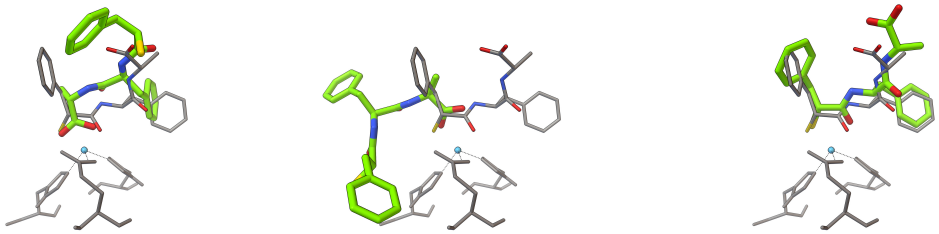
(a) (b)

Figure 3.9: Comparison of RMSD error of the new forcefield with (*a*) standard AutoDock4 forcefield and (*b*) AutoDock Vina



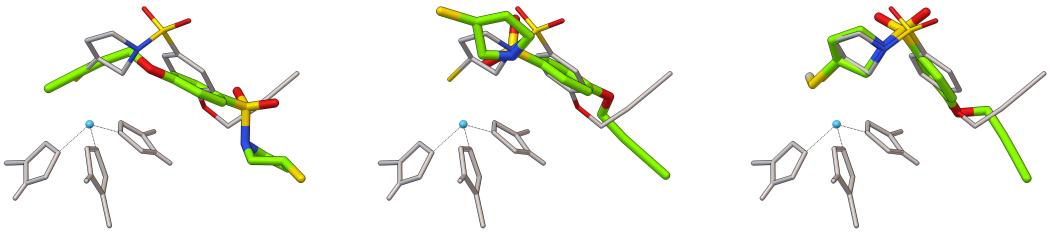
(a) (b)

Figure 3.10: Comparison of RMSD error on zinc coordination geometry of the new forcefield with (*a*) standard AutoDock4 forcefield and (*b*) AutoDock Vina



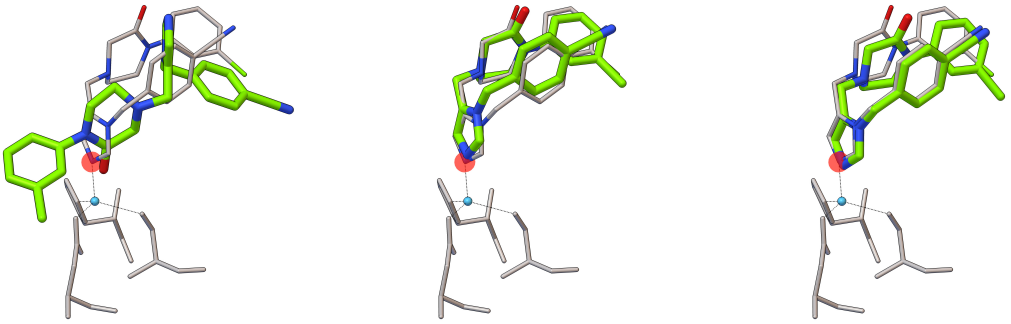
(a) AutoDock4 (4.85 Å) (b) AutoDock Vina (9.19 Å) (c) AutoDock4$_{Zn}$ (1.21 Å)

Figure 3.11: Comparinson of re-docking accuracy with 1r1j using (*a*) standard AutoDock4, (*b*) AutoDock Vina and (*c*) AutoDock4$_{Zn}$ forcefields (RMSD are shown in parentheses). Zinc-coordinating residues and experimental ligand pose are shown as thin *gray* sticks; docked poses are shown as *green* thick sticks. Hydrogens are not shown for sake of clarity.



(a) AutoDock4 (7.98 Å) (b) AutoDock Vina (2.34 Å) (c) AutoDock4$_{Zn}$ (2.10 Å)

Figure 3.12: Comparison of re-docking accuracy of 2oi0 using (*a*) standard AutoDock4, (*b*) AutoDock Vina and (*c*) AutoDock4$_{Zn}$ forcefields (RMSD are shown in parentheses). Zinc-coordinating residues and experimental ligand pose are shown as thin *gray* sticks; zinc is *cyan*; docked poses are shown as *green* thick sticks. Hydrogens are not shown for sake of clarity.



(a) AutoDock4 (7.62 Å) (b) AutoDock Vina (0.5 Å) (c) AutoDock4$_{Zn}$ (0.68 Å)

Figure 3.13: Comparison of re-docking accuracy of 1s63 using (*a*) standard AutoDock4, (*b*) AutoDock Vina and (*c*) AutoDock4$_{Zn}$ forcefields (RMSD are shown in parentheses). Zinc-coordinating residues and experimental ligand pose are shown as thin *gray* sticks; zinc is *cyan*; docked poses are shown as *green* thick sticks; the location and the optimal radius of the TZ pseudoatom potential is shown as semi-transparent sphere (*red*). Hydrogens are not shown for sake of clarity.

Chapter 4

Calculation of distribution coefficients in the SAMPL5 challenge from atomic solvation parameters and surface areas

Diogo Santos-Martins, Pedro Alexandrino Fernandes and Maria João Ramos

Adapted from ref. [103]

In this work I ran the experiments, analyzed results and wrote most of the paper.

4.1 Preface

This paper results from our efforts to improve the desolvation term in AutoDock. Drug molecules display a wide variety of chemical groups, thus desolvation terms should be able to quantify the interaction of different groups with water. For this reason, we worked on the prediction of free energies of hydration ΔG^{solv}_{water} of the 642 compounds in the FreeSolv-0.32 database [104, 105]. Importantly, we were able to predict the free energy of solvation of carbohydrates with satisfactory accuracy (see section 8.4), an important result as most glucosidase ligands are or resemble carbohydrates. Furthermore, the number of fitted parameters (atom types) is much smaller in this work than in other publications, making our model less likely to overfit.

4.2 Abstract

In the context of SAMPL5, we submitted blind predictions of the cyclohexane/water distribution coefficient (D) for a series of 53 drug-like molecules. Our method is purely empirical and based on the additive contribution of each solute atom to the free energy of solvation in water and in cyclohexane. The contribution of each atom depends on the atom type and on the exposed surface area. Comparatively to similar methods in the literature, we used a very small set of atomic parameters: only 10 for solvation in water and 1 for solvation in cyclohexane. As a result, the method is protected from overfitting and the error in the blind predictions could be reasonably estimated.

Moreover, this approach is fast: it takes only 0.5 seconds to predict the distribution coefficient for all 53 SAMPL5 compounds, allowing its application in virtual screening campaigns. The performance of our approach (submission 49) is modest but satisfactory in view of its efficiency: the root mean square error (RMSE) was 3.3 log D units for the 53 compounds, while the RMSE of the best performing method (using COSMO-RS) was 2.1 (submission 16). Our method is implemented as a Python script available at https://github.com/diogomart/SAMPL5-DC-surface-empirical.

4.3 Introduction

The free energy of solvation ΔG^{solv} can be separated in (i) cavitation free energy and (ii) solute-solvent interaction free energy. The cavitation free energy corresponds to the cost of disrupting solvent-solvent interactions in order to create a cavity that accommodates the solute. Solute-solvent interactions include van der Waals interactions and electrostatic interactions. Hydrogen bonds can be treated separately or within the general framework of electrostatic interactions. The assumption that cavitation and solute-solvent interaction free energies are additive provides a simple framework where the balance between these two terms rationalizes observed phenomena. For example, the hydrophobic effect observed for apolar solutes in water results from the high cost of forming a cavity (which includes the entropic penalty associated with constrained water molecules) and lack of counterbalancing strong solute-water interactions.

The solute-solvent interaction energy is mostly determined by the first layer of solvent molecules and by exposed solute atoms, simply because atoms in close proximity make the largest vdW and electrostatic contribution (charged buried atoms, such as in transition metal complexes, may be exceptions to this general rule). Moreover, if the solvent has hydrogen bond donors/acceptors, only exposed solute atoms can participate in hydrogen bonds with solvent molecules. For this reason, computational definitions of surface area have found application in the calculation of solute-solvent interactions, either by applying rigorous electrostatic formalisms as in the Poisson-Boltzman equation, or simply to estimate the contribution of different solute atoms in empirical models, as is the case of the present work.

The free energy of solvation of a molecule can be predicted as the sum of the individual contribution of each solute atom, weighted by its exposed surface area and by an atomic solvation parameter associated with its atom type [106]. Despite its simplicity, this formalism has been reported multiple times in scientific publications. Table 4.1 provides a comparison of published models used to calculate the free energy of solvation in water. Most studies employ a large number of parameters allowing the model to adhere very well to experimental data. In the work of Boyer et. al. [107] a total of 84 parameters were fitted, leading to a mean absolute error (MAE) of 1.41 kcal/mol. Other publications report extremely low MAE's achieving 0.54 kcal/mol by Wang et. al. [108] and 0.65 kcal/mol by Hou et. al. [109]. However, using a large number of parameters relatively to the size of the training set makes the model susceptible to overfitting. Ooi et. al. [110] fitted 7 parameters using only 22 molecules for training, and reported an extremely low root mean square error (RMSE) for compounds in the training set (RMSE = 0.32 kcal/mol) but a

Table 4.1: Comparison of quality of fit (training errors) for ΔG^{solv}_{water} for several models found in the literature and for the one proposed in this work. SAS stands for Solvent Accessible Area and SES stands for Solvent Excluded Area. MAE stands for Mean Absolute Error and RMSE is the Root Mean Squared Error, both presented in kcal/mol.

	solvent radius (Å)	type of surface	partial charges	fitted parameters	dataset size	MAE	RMSE
Ooi 1987 [110]	1.4	SAS	-	7	22	—	0.32
Wang 2001 [108]	0.6	SAS	-	54	401	0.54	0.79
Hou 2002 [109]	0.5	SAS	-	58	415	0.65	0.75
Boyer 2012 [107]	1.4	SAS	RESP	84	596	1.41	—
This work	1.5	SES	Gasteiger-Marsili	10	642	1.25	1.69

significantly larger error when the model was tested on molecules outside the training set (RMSE = 2.0 kcal/mol). For this reason, in this work we used a reduced number of parameters and a large training set.

These models have been used to predict the solvation free energy of different solute conformations [110]. This is possible because surface areas effectively capture the solvent exposure of each solute atom, preventing shielded atoms (e.g. after intramolecular hydrogen bonding) from contributing to the hydration free energy of the conformer. Moreover, atomic solvation parameters and surface areas have also been used to calculate partition coefficients [111] and aqueous solubilities [112], and have been integrated into both molecular dynamics [113] and molecular docking [114].

The existence of approximate but computationally inexpensive methods enables the large scale prediction of free energies of solvation. The question is: how does the performance of empirical models compare to more physical models? In the previous edition of the SAMPL challenge (SAMPL4), one of the blind predictions of the hydration free energies was an empirical model that performed almost as well as the more physically grounded methods [115, 116]. Instead of using surface areas to estimate solvent exposure of solute atoms, the proximity of other solute atoms from the atom of interest was taken into account. A total of 34 atom types were defined, but the total number of fitted parameters was 102 (68 parameters were used to describe shielding effects and quantify solvent exposure of solute atoms).

In this work we built empirical models to predict the free energy of solvation of organic compounds in water and cyclohexane, where the contribution of each solute atom is weighted by its exposed surface area and an atomic solvation parameter specific to its atom type. Then, we used these models to predict the cyclohexane/water distribution coefficient of 53 SAMPL5 molecules (depicted in figure S2) for which experimental log D values have been calculated [117, 118]. Despite the reduced number of atomic solvation parameters (10 for the free energy of solvation in water and 1 for the free energy of solvation in cyclohexane), our method performed reasonably. Unsurprisingly, more physical methods made better logD predictions (see the SAMPL5 overview paper [4], but the computational efficiency of our approach makes it valuable for large scale applications.

Table 4.2: Van der Waals radii used in this work.

element	vdW radius (Å)	element	vdW radius (Å)
H	1.20	P	1.80
C	1.70	S	1.80
N	1.55	Cl	1.75
O	1.52	Br	1.85
F	1.47	I	1.98

4.4 Methods

Cyclohexane/water distribution coefficients (D) were calculated from the free energies of solvation in each solvent, according to the following equation:

$$logD = \frac{\Delta G^{solv}_{water} - \Delta G^{solv}_{cyclohexane}}{2.303RT} \tag{4.1}$$

where T is the temperature ($293K$) and R is the ideal gas constant ($1.9872\ cal\ mol^{-1}\ K^{-1}$). The chosen temperature value is approximate: some training molecules had their solvation free energies determined at 298K while others were studied at 293K. The following sections describe the calculation of free energies of solvation in water and in cyclohexane.

Throughout this work, the following simplifications were adopted: (i) molecules were used in the conformation provided by SAMPL5 organizers and no conformational sampling was performed; (ii) only a single protonation state was considered for each molecule, corresponding to a neutral state, and ignoring different tautomeric states.

4.4.1 Free energy of solvation in water (hydration)

The free energy of hydration (ΔG^{solv}_{water}) was calculated as the sum of atomic contributions over all solute atoms. The contribution of an individual atom depends on the atomic solvation parameter, and on its solvent exposure:

$$\Delta G^{solv}_{water} = \sum_{i}^{N} W_i \times S_i \tag{4.2}$$

where N is the number of solute atoms, W_i is the atomic solvation parameter of the i^{th} atom and S_i is the solvent exposure of the i^{th} atom. Solvent exposure was calculated either as the solvent accessible surface (SAS) area or the solvent excluded surface (SES) area, computed with MSMS [119] using a solvent probe radius of 1.5 Å. The van der Waals radii for solute atoms are listed in table 4.2. The difference between SAS and SES is illustrated in figure 4.1.

In an alternative formalism, we included atomic partial charges for the calculation of free energies of hydration, using the Gasteiger-Marsili [120] method implemented in Openbabel 2.3.2 [121, 122]. The contribution of partial charges is also weighted by the solvent exposure of each

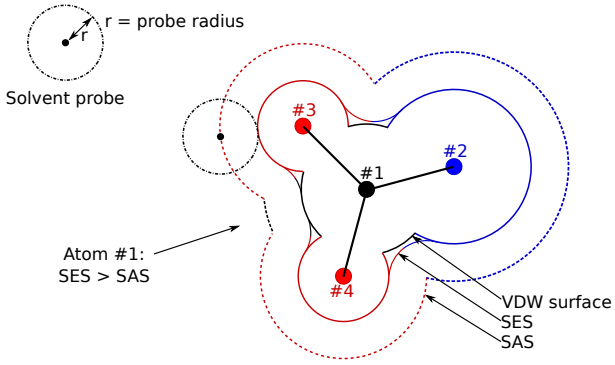


Figure 4.1: Solvent accessible surface (SAS) and solvent excluded surface (SES). SES and SAS are both computed by rolling the probe sphere over the van der Waals surface of the molecule. The SAS is determined by the center of the probe, while the SES is determined by the surface of the probe. The SAS is generally larger than the SES, but the SES of buried atoms can be larger that the SAS. In this example, atom #1 is only solvent accessible on the left side between atoms #3 and #4, where its SES is larger than its SAS.

solute atom, and is implemented by an additional term relatively to equation 4.2:

$$\Delta G^{solv}_{water} = \sum_{i}^{N} W_i \times S_i + Q \sum_{i}^{N} |q_i| S_i \tag{4.3}$$

where q_i is the partial charge of the i^{th} atom and Q is the weight factor for the contribution of partial charges to hydration free energy. Equations 4.2 and 4.3 provide two alternative models, including or excluding atomic charges.

4.4.1.1 Training procedure

Atomic solvation parameters (W_i and Q) were fitted by the least squares method to reproduce the experimental free energy of hydration of 642 compounds in the FreeSolv-0.32 database [104, 105], using the R software package [123]. Some atom types displayed poor statistical significance and were manually set to zero. This is either because there are few molecules in the training set containing these atom types or because they are often buried (e.g. phosphorous in phosphate groups). In different models (SES or SAS, with or without partial charges), the excluded atom types varied. The formalism presented in equations 4.2 and 4.3 lacks a term to explicitly describe the cost of creating a cavity in water to accommodate the solute. However, since the atomic solvation parameters are fitted to experimental free energies of solvation, the cost of cavity formation is implicitly incorporated into the atomic solvation parameters.

4.4.1.2 Atom types

We devised a simple atom typing scheme that resulted in a reduced number of atomic solvation parameters. Atom types indicate three attributes: (i) the element, (ii) aromaticity (iii) the possibility of making hydrogen bonds with solvent waters. Both the aromaticity and hydrogen bonding

were predicted by Openbabel 2.3.2. The resulting atom types are shown in table 4.3. There are two types for hydrogen atoms (polar and apolar hydrogens), two types of carbon (aromatic and non-aromatic carbon), four types of nitrogen (aromatic / non-aromatic, able / unable to accept hydrogen bonds). Oxygen has a single atom type, thus all oxygens are typed as O. All oxygens in the FreeSolv-0.32 database and in the SAMPL5 set are considered H-bond acceptors by Openbabel 2.3.2. The remaining elements are composed of a single atom type each. Chemical groups which are H-bond donors, such as hydroxyl groups and amines, rely on the presence of polar hydrogens (HD) to describe their H-bond donor properties.

4.4.2 Free energy of solvation in cyclohexane

For training the model to predict $\Delta G^{solv}_{cyclohexane}$ we used a total of 18 compounds with experimental values [124]. These compounds (figure S3) are sidechain analogues of the 20 naturally occurring aminoacids except glycine and proline. Due to the reduced number of compounds used for training the model, we opted to fit only a single parameter: the SES area of the molecule. The free energy of solvation in cyclohexane is then calculated as:

$$\Delta G^{solv}_{cyclohexane} = W_c \times A \tag{4.4}$$

where W_c is the fitted parameter (using the least squares method) and A is the SES area of the solute, using a solvent probe radius of 1.5 Å. The value of W_c and the quality of the model are discussed in the results section.

4.5 Results and Discussion

In the following sections, we discuss (i) the model for predicting free energies of solvation in water, (ii) the model for predicting free energies of solvation in cyclohexane and (iii) the blind prediction of cyclohexane/water distribution coefficients (logD) for SAMPL5 compounds.

4.5.1 Prediction of free energy of solvation in water

Atomic solvation parameters were fitted using the least squares method to reproduce the experimental free energies of hydration of 642 molecules in the FreeSolv-0.32 database [105, 104]. In order to evaluate the benefit of using partial charges calculated by a fast method (Gasteiger-Marsili) and also to test different surfaces (SES and SAS) to quantify solvent exposure of solute atoms, four sets of atomic solvation parameters were derived. The resulting parameters are reported in table 4.3. The quality of prediction using SES areas and including atomic charges is depicted in figure 4.2.

The magnitude and sign of atomic solvation parameters quantifies the contribution of each atom type to the free energy of hydration. More negative values indicate a larger favorable contribution to the hydration free energy. However, the contribution of an individual atom is weighted

Table 4.3: Atomic solvation parameters used in the calculation of the free energy of hydration, fitted to experimental data in the Freesolv-0.32 database using the least squares approach. These parameters correspond to W_i in equations 4.2 and 4.3 and to Q in eq. 4.3. Zeroed parameters were set manually due to poor statistics.

		Atomic Solvation Parameters (W_i) (cal mol^{-1} Å^{-2})			
atom type	desription	SES (eq. 4.3)	SAS (eq. 4.3)	SES (eq. 4.2)	SAS (eq. 4.2)
H	Hydrogen (apolar)	+11.2 (±1.6)	+3.2 (±0.7)	+8.1 (±2.3)	-2.0 (±0.5)
HD	Hydrogen (polar)	-193.5 (±13.3)	-45.7 (±5.1)	-303.3 (±13.0)	-95.5 (±4.4)
C	Carbon	0	+21.2 (±5.8)	-44.6 (±9.8)	0
A	Carbon (aromatic)	-12.6 (±3.1)	0	-40.7 (±3.0)	-22.4 (±2.2)
N	Nitrogen	0	0	+144.9 (±38.4)	0
NA	Nitrogen (aromatic)	-626.6 (±65.4)	-465.2 (±52.5)	-621.5 (±71.5)	-497.6 (±59.0)
NH	Nitrogen (H-bond acc.)	-128.7 (±22.4)	-24.6 (±8.3)	-130.9 (±24.4)	-47.7 (±9.1)
NHA	Nitrogen (arom./acc.)	-185.6 (±21.7)	-98.6 (±11.9)	-244.3 (±22.8)	-124.9 (±13.3)
O	Oxygen (H-bond acc.)	-42.2 (±7.2)	-10.9 (±2.8)	-108.8 (±4.5)	-46.8 (±2.2)
F	Fluorine	+72.4 (±6.5)	+38.9 (±2.8)	+31.7 (±5.7)	+11.3 (±2.6)
P	Phosphorus	0	0	0	0
S	Sulfur	0	0	0	-15.7 (±5.1)
Cl	Chlorine	+13.9 (±2.8)	+8.4 (±1.3)	-7.1 (±2.1)	-5.1 (±1.0)
Br	Bromine	0	0	0	0
I	Iodine	0	0	0	0
Weight factor for Gasteiger charges (Q)		-246.9 (±22.1)	-162.0 (±9.2)	—	—
Training RMSE (kcal mol^{-1})		1.69	1.76	1.80	1.99

by its surface area, explaining the larger magnitude of solvation parameters obtained using SES areas (the SAS is always larger than the SES except for highly buried atoms, as is exemplified for atom #1 in figure 4.1).

Atom types capable of making Hydrogen bonds with water have the most negative solvation parameters (HD, O, NH, NHA). This means that solute-solvent hydrogen bonds can be captured by atomic solvation parameters. The coefficient Q from equation 4.3 scales the contribution of partial charges and also has a large negative value, indicating that the contribution of electrostatic interactions have also been incorporated in the parameters. This observations are consistent with a physically meaningful model, with a straightforward interpretation of atomic solvation parameters. It is important to note that inclusion of partial charges in the model (equation 4.3) decreases the magnitude of the atomic solvation parameter of atoms involved in hydrogen bonds (HD, O, NH, NHA) by about four-fold if SES areas are used and three-fold if SAS areas are used. This means that Gasteiger-Marsili charges are able to describe a significant part of solute-solvent hydrogen bonds.

One particular atom type, NA (aromatic nitrogen that does not accept H-bonds) displays a more negative solvation parameter than atom types involved in hydrogen bonds, which is hard to rationalize. This is partially explained by the low exposure of NA atoms, which are shielded by three substituent groups in a planar geometry, making solvent contacts possible only in small surface patches above and below the plane of the aromatic ring. However, even considering their low

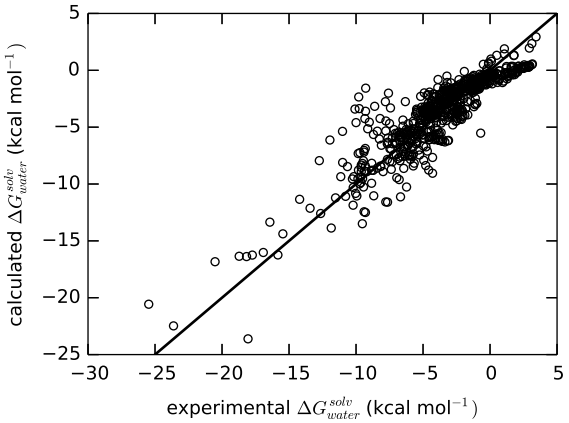


Figure 4.2: Prediction of hydration free energies for molecules in the training set using SES areas and including partial charges.

solvent exposure, NA atoms can make significant contributions: for cyanuric acid (the molecule from FreeSolv-0.32 database with the largest SES area associated with NA atoms), NA atoms contribute with almost -8 kcal/mol to the hydration free energy. For comparison, the contribution from hydrogen bond donors/acceptors and from partial charges is about -15.3 kcal/mol for cyanuric acid, and the free energy of hydration is overestimated by -5.6 kcal/mol. These observations suggest that the parameter for NA has overfitted. We'll return to this discussion in view of the results obtained in blind logD prediction for SAMPL5 compounds containing NA atoms.

Among the four sets of parameters derived to predict ΔG^{solv}_{water}, the quality of the fit was slightly better (lower RMSE) with the use of SES areas and the inclusion of partial charges. Thus, this model was used to make blind predictions of the cyclohexane/water logD for compounds in the SAMPL5 set.

4.5.2 Prediction of free energy of solvation in cyclohexane

Free energies of solvation in cyclohexane were predicted using equation 4.4, in which a single parameter W_c is multiplied by the SES area of the solute to obtain $\Delta G^{solv}_{cyclohexane}$. From a physical point of view, we are assuming that the free energy of solvation is directly proportional to the solute area. This assumption is reasonable because the dielectric constant of cyclohexane is very low ($\varepsilon = 2.02$), and van der Waals interactions constitute the largest contribute to intermolecular stabilization. Using a set of 18 molecules, W_c was fitted to -36 cal mol^{-1} Å^{-2}. The quality of the fit is depicted in figure 4.3, and has a RMSE of 1.02 kcal/mol. On an additional set of 91 molecules from ref.[14], the RMSE is 1.07 kcal/mol (see figure S1 and table S1). The model systematically underestimates free energies of solvation for more negative values and overestimates more positive values. This bias could be fixed by introducing an intercept term B in equation 4.4 and transforming it into $\Delta G^{solv}_{cyclohexane} = W_c \times A + B$. However, the presence of an intercept term B would mean that a molecule with no surface area would have an interaction with cyclohexane, which is physically unreasonable. For this reason, we decided to avoid the use of an intercept.

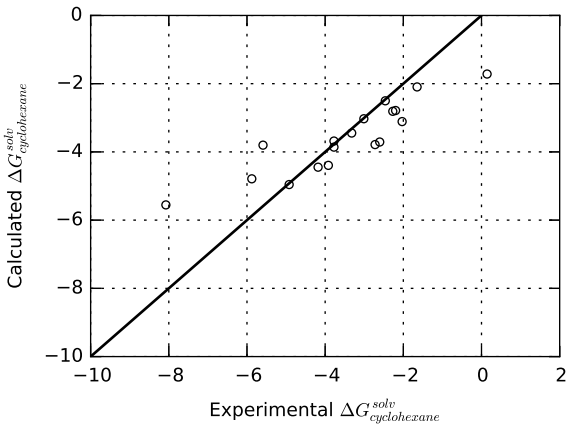


Figure 4.3: Prediction of solvation free energies in cyclohexane for molecules in the training set.

Moreover, in a retrospective analysis, we used an intercept in the model to predict $\Delta G^{solv}_{cyclohexane}$ but this showed no improvement in the prediction of logD values for SAMPL5 compounds, and showed a systematic bias for larger molecules, indicating that equation 4.4 without intercept is more appropriate to calculate $\Delta G^{solv}_{cyclohexane}$.

4.5.3 Prediction of logD for SAMPL5 compounds

The prediction of logD values for compounds in the SAMPL5 challenge was based on the free energies of solvation in water and on cyclohexane (equation 4.1). The free energy of solvation in water was calculated using SES areas and partial charges (equation 4.3). The model that predicts the free energy of solvation in cyclohexane consists of a single coefficient multiplied by the SES area of the molecule (equation 4.4). Our predictions are reported in table S2.

Figure 4.4 compares the calculated and experimental logD values for the 53 SAMPL5 molecules. While a correlation is readily observable the model exaggerates the magnitude of the predictions, both for negative and positive valued logD's. In other words, if the calculated logD was scaled by a factor of about 0.3, the predictions would approach the equality line. The key parameters to describe the quality of the prediction are the Pearson's correlation coefficient of 0.58, the Kendall rank order correlation coefficient of 0.42, a mean signed error of -1.06 (1.42 kcal/mol in ΔG^{solv} units), a mean absolute error (MAE) of 2.57 (3.45 kcal/mol) and a root mean square error (RMSE) of 3.27 (4.39 kcal/mol). These values correspond to modest prediction of logD values. The Kendall rank order correlation coefficient (0.42) also indicates modest performance in ranking the compounds.

The largest outlier is adenosine (ID: SAMPL5_074) which is predicted to have a logD value of -14.1 while the experimental value is -1.9. Analysis of the predictions submitted by other participants revealed a systematic bias towards more negative values. This may indicate a problem with the experimantal value of this molecule, or the existence of a phenomena that lies outside the scope of the modeling techniques, such as the formation of adenosine dimers in cyclohexane, satisfying

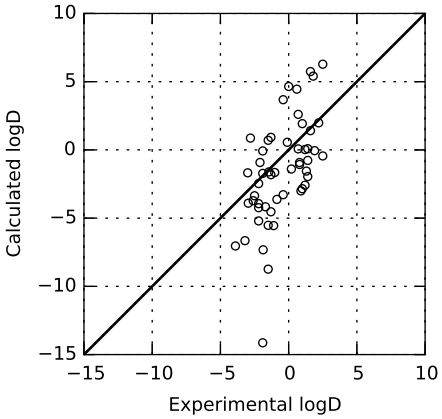


Figure 4.4: Blind prediction of cyclohexane/water logD values for SAMPL5 compounds.

a significant number of hydrogen bond donors/acceptors. However, this is a merely speculative explanation for the systematic deviation of the predictions. If SAMPL5_074 is removed, the RMSE decreases from 3.27 (4.39 kcal/mol in ΔG^{solv} units) to 2.84 (3.80 kcal/mol), and the MAE reduced from 2.57 (3.45 kcal/mol) to 2.39 (3.20 kcal/mol).

4.5.3.1 Discussion of the NA atom type

The atomic solvation parameter for NA in the ΔG^{solv}_{water} model is the most negative among all fitted parameters, which is suspicious in view of the smaller magnitude of other parameters associated with strong interactions with water: hydrogen bonds and atomic charges. As is depicted in figure 4.5, our blind predictions on SAMPL5 compounds confirmed the suspicions: a larger contribution from NA atoms is indeed associated with a biased prediction of logD values towards distribution of the solute in water. In view of these results, we concluded that the atomic solvation parameter for atom type NA has overfitted. It is important to note that the aberrant NA parameter does not explain all errors in our model: molecules in which NA is absent still present large deviations from the experimental value (see fig. 4.5). The value of this analysis lies in the identification of an error created by a machine learning approach through interpretation of the physical meaning of solvation parameters.

In an attempt to explain the aberrant NA parameter, we performed three retrospective (non-blind) experiments, in which the cyclohexane/water logD was calculated for the 53 SAMPL5 molecules using a modified set of parameters to calculate the solvation free energy in water:

1. Set $W_{\mathrm{NA}} = 0$ without change to any other atomic parameter.

2. Set $W_{\mathrm{NA}} = 0$ and re-calibrate the remaining parameters using all 642 molecules in the FreeSolv-0.32 database.

3. Set $W_{\mathrm{NA}} = 0$ and re-calibrate the remaining parameters using all molecules in the FreeSolv-0.32 database except those that contain NA (19 out of 642 molecules contain NA).

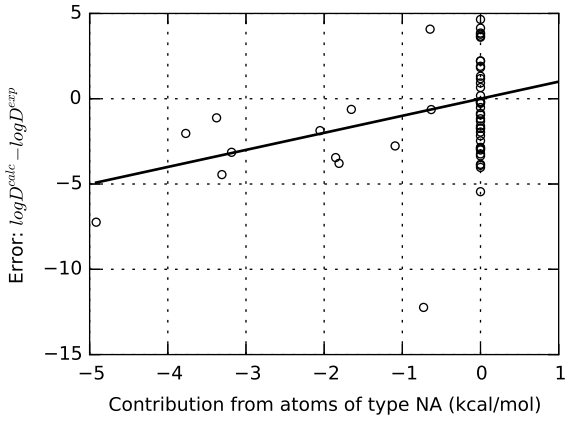


Figure 4.5: Error in the blind prediction of cyclohexane/water logD values is associated with the contribution from NA atoms.

The performance of the new sets of solvation parameters were evaluated in (i) the full SAMPL5 set (n = 53), (ii) all compounds except SAMPL5_074 (n = 52), and (iii) the subset of SAMPL5 compounds that do not contain NA atoms (n=40). We note that SAMPL5_074 contains NA and is excluded in subset (iii).

The results are reported in table 4.4. Overall, retrospective experiment 2 increased the error, while experiments 1 and 3 decreased the error relatively to the original set of parameters (table 4.3). We speculate that NA containing molecules are implied in the origin of the aberrant NA parameter during the fitting process. Setting $W_{\mathrm{NA}} = 0$ without change to other solvation parameters (experiment 1) lowers the error, but forcing $W_{\mathrm{NA}} = 0$ while allowing other parameters to re-optimize (experiment 2) increases the error. Excluding NA containing molecules from the training set (experiment 3) also decreased the error. It is possible that specific chemical features in training set molecules containing NA introduce a bias in the NA parameter in order to reproduce the free energy of hydration. For exaple, the existence of multiple tautomeric states in NA containing molecules (e.g. cyanuric acid), or the induction of a dipole moment in aromatic rings by the presence of a nitrogen atom instead of a carbon atom are complex physical properties beyond the scope of the present model. Thus, the NA parameter optimizes to a meaningless value because it

Table 4.4: RMSE (log D units) evaluated on SAMPL5 compounds using updated atomic solvation parameters from retrospective experiments.

	Evaluation set		
	All SAMPL5 (n = 53)	Except 074 (n = 52)	NA free (n = 40)
Experiment 1	2.97	2.52	2.63
Experiment 2	3.33	2.82	2.92
Experiment 3	2.89	2.45	2.55
Submission #49	3.27	2.84	2.63

is prevalent in molecules that happen to contain chemical features that the present model is unable to describe.

4.5.3.2 Estimating the model error

In the submission of results to SAMPL5, participants were asked to estimate the uncertainty of the predictions. We estimated the uncertainty of our model based on the training RMSE of predictions of solvation free energy in water (1.68 kcal/mol) and cyclohexane (1.02 kcal/mol), which accumulate to 2.7 kcal/mol. We rounded up this value to 3 kcal/mol because the compounds in the SAMPL5 set are larger and chemically more diverse than those used for fitting parameters. According to equation 4.1, 3 kcal/mol correspond to 2.24 logD units. Our logD predictions displayed a RMSE of 3.27 and a mean absolute error (MAE) of 2.6, which is higher than the estimated error. If the compound with ID SAMPL5_074 is excluded (this compound was systematically predicted to have a lower logD by other SAMPL5 participants), the RMSE lowers to 2.84 and the MAE to 2.39. which is not far from our RMSE estimate of 2.24. Overall, the errors in the blind challenge were higher than we have anticipated, but the error estimate is reasonable.

4.6 Conclusions

In this work, we employed an empirical model based on atomic solvation parameters and on the surface area of exposed solute atoms to predict the free energies of solvation in two solvents: water and cyclohexane. This approach was used to make blind predictions of the cyclohexane/water distribution coefficients of 53 molecules in the context of the SAMPL5 challenge. Our predictions were not among the best performing methods in the challenge, but can be considered satisfactory in view of its speed: it takes an average of 0.01 seconds per molecule.

The most striking feature of this work relatively to similar studies is the reduced number of atomic solvation parameters. Tipically, the number of atom types ranges between 30 and nearly 100, but here we have fitted parameters for only 10 atom types (for predicting free energies of hydration). Thus, our model can capture only simple features of the solute-solvent interaction, such as hydrogen bonds, but in compensation has a straightforward interpretation of the physical meaning of atomic solvation parameters and is less susceptible to overfitting. As a result, the error in the blind predictions is only slightly higher than the errors obtained in the fitting stage.

Chapter 5

Interaction with specific HSP90 residues as a scoring function: Validation in the D3R Grand Challenge 2015

Diogo Santos-Martins

Adapted from ref. [39].

5.1 Abstract

Here is reported the development of a novel scoring function that performs remarkably well at identifying the native binding pose of a subset of HSP90 inhibitors containing aminopyrimidine or resorcinol based scaffolds. This scoring function is called PocketScore, and consists of the interaction energy between a ligand and three residues in the binding pocket: Asp93, Thr184 and a water molecule. We integrated PocketScore into a molecular docking workflow, and used it to participate in the Drug Design Data Resource (D3R) Grand Challenge 2015 (GC2015). PocketScore was able to rank 180 molecules of the GC2015 according to their binding affinity with satisfactory performance. These results indicate that the specific residues considered by PocketScore are determinant to properly model the interaction between HSP90 and its subset of inhibitors containing aminopyrimidine or resorcinol based scaffolds. Moreover, the development of PocketScore aimed at improving docking power while neglecting the prediction of binding affinities, suggesting that accurate identification of native binding poses is a determinant factor for the performance of virtual screens.

5.2 Introduction

Molecular docking is a tool for modeling the interaction between small molecules and macromolecules. It is used to predict binding poses and the affinity of ligands. Because of its widespread use and recognized value, it is of the utmost importance to keep the scientific community aware of

its limitations [125, 126, 127]. For this purpose, blind challenges are invaluable because participants make predictions without access to the solutions, thus preventing unintentional bias towards reproduction of the correct results. Moreover, data is typically donated by pharma companies, reflecting the ligand diversity explored to modulate clinically relevant targets. The outcomes of blind challenges are thus the best assessment of the true value of molecular docking (and other methodologies) in drug design for specific targets.

In this work we developed a molecular docking approach specific to Heat Shock Protein 90 (HSP90) and report our participation in the Grand Challenge 2015 organized by the Drug Design Data Resource (D3R) [128]. In this challenge, participants were asked to rank 180 molecules according to their potency as HSP90 inhibitors. IC50 values ranged five orders of magnitude, from low nanomolar to micromolar, and there were over 50 inactive molecules. Participants were also asked to predict the binding pose of a subset of six molecules.

Molecular docking programs rely on two key tools: a search algorithm and a scoring function. The search algorithm samples a large number of binding poses, which constitute tentative solutions for a match to the native pose. The scoring function evaluates all generated poses to identify the best pose (the one expected to reproduce the experimental binding mode). The score of the best pose is used to predict ligand affinity. Therefore, scoring functions perform two tasks: binding mode prediction, and affinity prediction. If the selected pose is incorrect, either because the search algorithm failed to generate the native pose or because the scoring function has selected it incorrectly, the prediction of binding affinity is performed on a wrong pose. Consistent selection of native binding poses is important to support subsequent prediction of ligand affinities.

The concepts of "docking power" and "screening power" are routinely used to benchmark the performance of different scoring functions [129, 5] Docking power is the performance in identifying experimental binding modes from an ensemble of binding poses generated beforehand to guarantee that all scoring functions are evaluated on the same ground, i.e. independently of specific search algorithms implemented in different software packages. "Scoring power" is the correlation between experimental binding affinity and scores produced based on native poses. In a virtual screening scenario, active and inactive molecules are both docked and scored with the aim of identyfing active molecules. Unsurprisingly, Li et. al. [5] observed a positive correlation between docking power and the ability in identyfing active molecules from a set of compounds (screening power). On the other hand, there was no correlation between scoring power and screening power.

These results suggest that reproducing native poses is required to succeed in identifying active molecules while correlating with binding affinities is of secondary importance. It is possible for a scoring function to display excellent scoring power if native binding poses are available, and simultaneously display poor screening power because incorrect poses are scored with unrealistically favourable scores [5].

Other studies reported similar results by demonstrating that docking methods with satisfactory docking power can effectively discriminate between active and inactive molecules, despite having lack of scoring power [130, 131]. Another study added further confirmation of the greater importance of docking power over scoring power, by reporting that scoring functions that lack docking

power show poor screening power, despite being able to predict ligand affinity with exceptional accuracy [3]. Due to the importance of pose selection (docking power), recent scoring functions were developed by aiming exclusively at improved docking power [132, 133].

In this work, we developed a scoring function that displays enhanced docking power for a subset of HSP90 inhibitors, and integrated it in a molecular docking workflow to rank the 180 molecules in the D3R Grand Challenge 2015 set. This scoring function is called "PocketScore" because it quantifies the interaction between a ligand and a small subset of residues in the binding site of HSP90, while ignoring most atoms in the binding site. Despite its simplicity, the results were satisfactory when compared to well established scoring functions (AutoDock Vina [134]) and to predictions submitted by other participants.

5.3 Methods

Our docking protocol consists in three distinct steps, which are illustrated in figure 5.1. In the first step, we generate an ensemble of binding poses for a given input ligand. We expect at least one pose to match the native binding mode. In the second step, we identify which pose is more likely to be correct. This is performed by scoring each pose with a scoring function; the pose with the best score is selected. In the third and final step, we re-score the pose selected in step 2 with yet another scoring function. We use the score produced at this step to predict ligand affinity. This workflow allows us to use different scoring functions for docking (pose selection) and scoring (predicting ligand affinity). The performance of the virtual screen depends on the exact methodology used in each step: how ensembles of binding poses are generated, how one pose is selected, and how the selected pose is rescored.

5.3.1 Training Set Compilation

The training set consisted of HSP90 X-ray structures co-crystallized with several different inhibitors. We considered the 81 HSP90 structures used in the study associated with the Directory of Useful Decoys - Enhanced (DUD-E) [135], available at the DUD-E website [136]. The DUD-E provides several sets of active molecules and decoys that can be used to evaluate virtual screening methods. Some PDB entries were removed from the dataset, reducing its size from 81 to 69 complexes. We removed entries which are co-crystals of the same ligand and correspondingly display the same binding pose, in order to remove duplicated information. We have also removed ligands that display more than one binding mode, either in different PDB entries, or in different chains in the same PDB entry. Although it is physically reasonable for a ligand to possess multiple (stable) binding modes, we prefered to remove these complexes in order to avoid the complexity of dealing with multiple experimental poses for the same ligand. One large macrocylic ligand was removed because AutoDock Vina is unable to sample conformations of cyclic structures. Structures 4yky and 4ykr were provided to D3RGC participants and were included in the training set. The final selection of PDB IDs is listed as a reference[137].

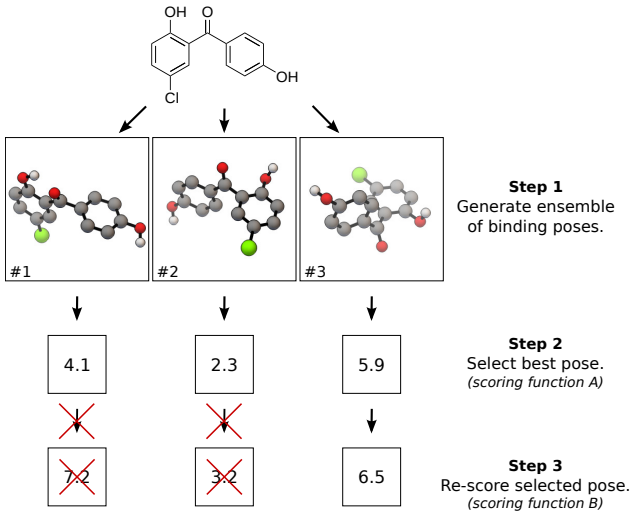


Figure 5.1: Workflow implemented in this study. In the first step, an ensemble of binding poses is generated for the input ligand. In the second step, poses #1, #2 and #3 are ranked by scoring function A, with scores, 4.1, 2.3 and 5.9, respectively. This leads to selection of pose #3 as most likely to match the native binding mode. Then, in step 3, the selected pose is re-scored by a second scoring function (B), leading to the final score of 6.5 that is used to predict ligand affinity. In the text, we would refer to the protocol illustrated here as scoring-function-B//scoring-function-A, to denote the specific combination of scoring functions for re-scoring//pose selection.

5.3.2 Generation of binding poses

The first step of the virtual screening workflow is the generation of an ensemble of binding poses. This step is considered successful if at least one pose in the ensemble is close to the native structure. AutoDock Vina was used to generate binding poses, using default global search exhaustiveness (8) and maximum number of output poses (9). The free energy of binding is calculated by Vina for each pose, but these values are ignored at this step. This allows us to subsequently use an external scoring function (other than that implemented in AutoDock Vina), to rank all poses in the ensemble, and directly assess the performance of that scoring function in identifying native poses - the docking power.

We used ligands from the training set to evaluate if generated ensembles contained a pose close to the experimental structure. We adjusted features of the receptor model to increase the number of ligands for which successful ensembles were generated. We did not increase the number of returned poses by Vina, as others have also evaluated sampling of poses regardless of scores, and suggested that increasing the number of output poses is not as beneficial as changing the underlying procedure for generating poses [138]. Specifically, we tested different conformations of HSP90 and included various water molecules in the receptor. A detailed analysis of different HSP90 conformations and water molecules interacting with known inhibitors is provided below. This analysis supported the choice of conformations and waters to be tested. By combining four conformations with twelve water configurations, we built a total of 48 receptor models. The evaluation of all these receptors for pose generation is described in the results section.

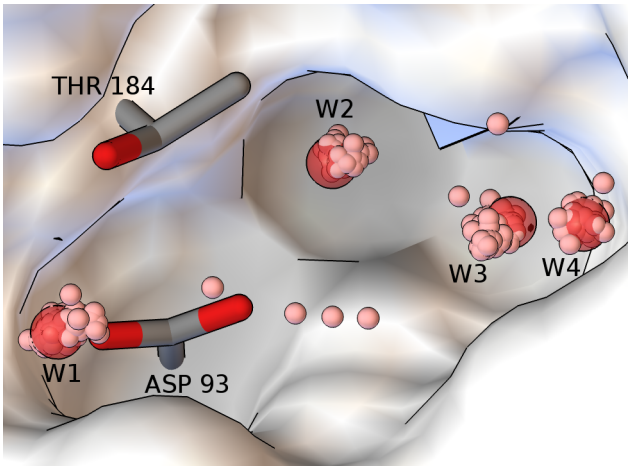


Figure 5.2: Conserved water molecules in the binding site of HSP90. Superimposition of crystallographic structures from the training set reveals important water sites that mediate protein-ligand interactions. Each pink sphere represent a crystallographic water. Red spheres labeled W1 through W4 indicate the location of the four water sites considered for building receptor models.

Upon evaluation of all 48 receptors, we finally settled on using two receptors: 1uyg conformation with waters W2 and W3 (1uyg/W2+3) and the 1yc3 conformation with waters W1 and W3 (1yc3/W1+3). More details on conformations and location of waters are provided below. With the use of two receptors, the size of sets of binding poses increases to a maximum of 18 (up to 9 poses are produced by AutoDock Vina for each receptor). These two receptors were used to generate poses in subsequent assessments of docking power, screening power and ranking power.

5.3.2.1 Modelling fixed water molecules

The training set was investigated for the presence of conserved waters lying in the interface between a ligand and the binding pocket of HSP90. Superimposition of structures from the training set revealed the presence of four conserved waters, as is illustrated in figure 5.2. These waters are labeled W1, W2, W3 and W4. The stability of these four waters has been confirmed by molecular dynamics simulations [139]. Other molecular docking studies included these waters in the receptor model [140, 141]. Since ligands can make hydrogen bonds with these water molecules, but can also displace them upon binding, it is unclear which water molecules should be kept to build an ideal receptor model. There is a total of sixteen possible combinations resulting from the inclusion or exclusion of each of the four waters, but visual analysis of complexes in the training set suggested that W4 is only present if W3 is also present. This restriction reduces the number of possible combinations to twelve. Together with different conformations of the binding pocket, these twelve water configurations were used to build receptor models.

5.3.2.2 Conformation of the binding pocket

With the aim of building realistic receptor models able to accommodate ligands that induce different conformations in HSP90[142], we analysed the variety of HSP90 bound states. All structures

in the training set were aligned to 2JJC coordinates (downloaded from the D3R website), using the "super" command in Pymol and considering all protein atoms. Visual inspection revealed that residues 100-124 adopt a variety of conformations. A clustering protocol was devised to select representative structures: first, the RMSD between alpha carbons of residues 100-124 was calculated between all possible pairs of structures in the training set, using the coordinates aligned by the full protein (coordinates were aligned with Pymol using 2JJC as reference, and RMSD values were subsequently calculated with a Python script). This resulted in a symmetric square matrix of size $N \times N$, where N is the size of the training set. This matrix is a "distance matrix" as it quantifies dissimilarity between pairs of structures, and it is referred to as the "pairwise RMSD matrix". We used the complete linkage algorithm using the pairwise RMSD matrix as input, producing a dendrogram that aided in the selection of representative structures (figure 5.3). Various clusters were visually identified, having low RMSD values between any pair of structures inside each cluster. There are three large clusters at the center of the RMSD matrix, and three smaller ones at the bottom right corner. We considered the three smaller clusters as a single cluster, leading to a total of four clusters. Within each cluster, the structure that displayed best performance for virtual screening in the DUD-E study was selected as representative. The data about performance of each structure is available online at the DUD-E website [143]. Four structures were used to build receptor models, associated with the PDB ID 1uyg, 1yc3, 1yc4 and 2cct. The conformation of residues 100 to 124 in these representative structures is highlighted in figure 5.4. The relative flexibility of each residue in the selected segment (100 to 124) is illustrated in figure 5.5.

5.3.3 Ranking poses and re-scoring

In steps 2 and 3 of the workflow (figure 5.1), scoring functions are used to select the best pose from the ensemble of generated poses and to subsequently re-score the selected pose, producing the final score that predicts ligand affinity. It is possible to use the same or different scoring functions for steps 2 and 3. We use a double slash notation to refer to a specific combination of scoring functions, e.g. scoring-function-B//scoring-function-A denotes the use of of scoring function B for re-scoring and scoring function A for pose selection. Three scoring functions were used: the one implemented in AutoDock Vina (referred to as Vina), PocketScore (described in detail below), and a scoring function reported to have excellent scoring power: RF-Score-3[144].

When we optimized steps 2 and 3, step 1 had already settled on the use of two receptors to generate ensembles of binding poses: 1uyg/W2+3 and 1yc3/W1+3. It has been argued that the helical conformation of HSP90 (1uyg) has higher internal energy, creating a need for a compensatory term to make scores associated with the helical conformation comparable with other conformations [45, 145]. In this work, we denote the value of the helical penalty applied to poses docked in the helical conformation (1uyg) inside brackets: Vina($+x$) or RF-Score-3($-x$). The penalty is positive for Vina, and negative for RF-Score-3, because Vina outputs a free energy prediction while RF-Score-3 produces a positive valued score.

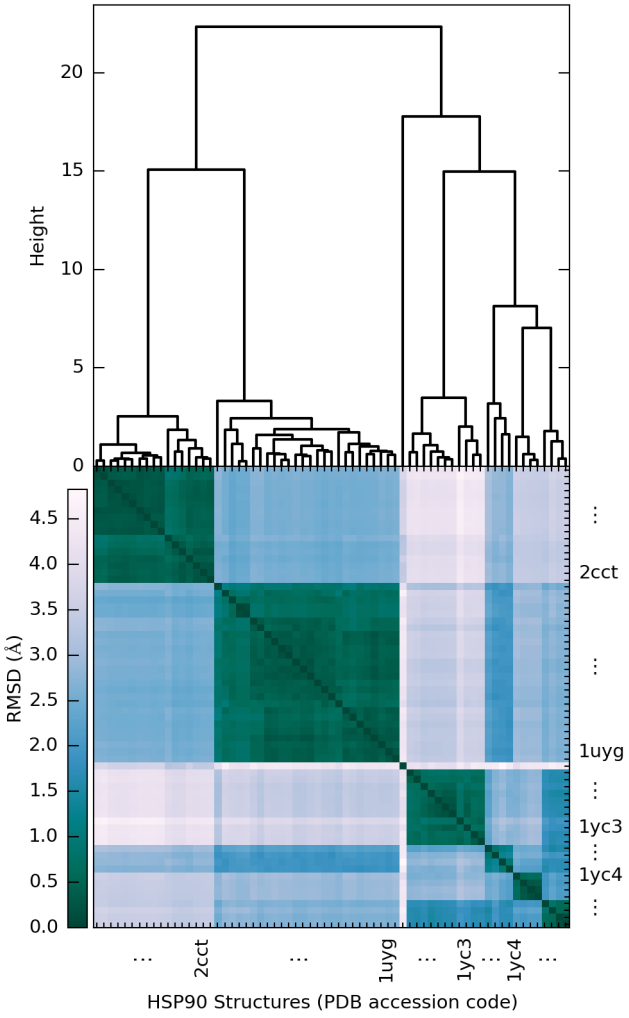


Figure 5.3: Conformational analysis of residues 100 to 124 in HSP90 structures from the training set. The upper panel shows the dendrogram produced by complete linkage of the pairwise RMSD matrix, which is depicted in the lower panel. The x-axis is shared between panels. The size of the RMSD matrix is 67×67 (3hyz and 3k98 were excluded due to missing atoms in the region of interest). Labels are omitted for all structures except those selected as representative. The three larger clusters are represented by 2cct, 1uyg, and 1yc3, and the three smaller clusters that appear in the bottom right corner of the RMSD matrix are represented by 1yc4.

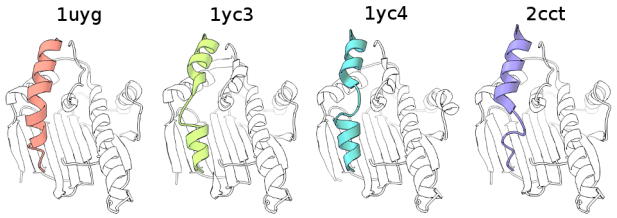


Figure 5.4: Representative conformations of HSP90 in the training set, highlighting the flexible region (residues 100-124). In structure 1uyg, the flexible region adopts an alpha-helical conformation, and the binding pocket is larger.

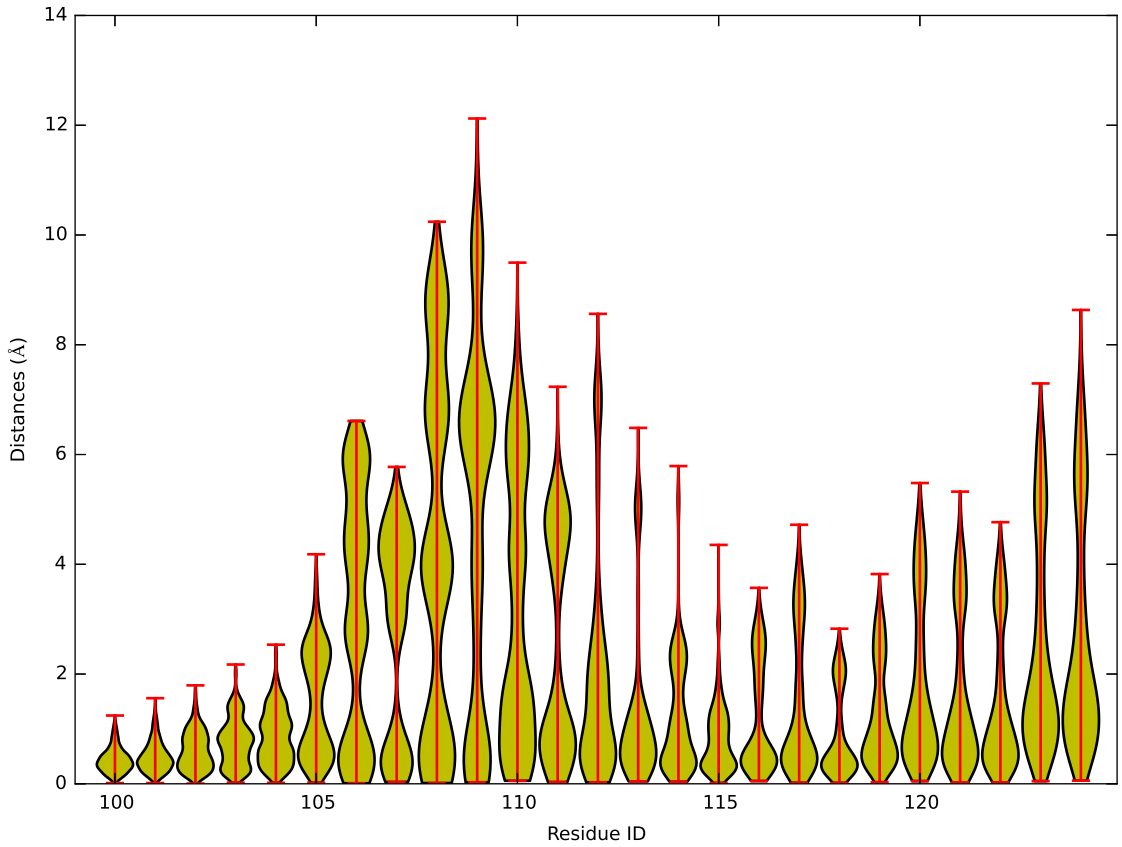


Figure 5.5: Distribution of distances between alpha-carbons in all possible pairs of structures in the training set. Structures were aligned beforehand with the "super" command in Pymol to 2JJC coordinates using all atoms.

5.3.3.1 PocketScore

PocketScore is the interaction score between a ligand and three residues in HSP90: Asp83, Thr186 and W2. This interaction energy is calculated using a `-score_only` calculation with Vina, using a receptor model containing only the three residues of interest. Instead of using the full Vina scoring function, the torsional term was excluded, and the interaction energy is calculated as the sum of the remaining terms (gauss1, gauss2, repulsion, hydrophobic and hbond) multiplied by their default weights. The torsional term was excluded because the focus was on the interaction between the ligand and the residues of interest. PocketScore was created after realizing (through visual analysis) that most HSP90 inhibitors bind through hydrogen bonds with both Asp93 and Thr184, making it easy to identify poses that reproduce the most common binding mode we found in the training set.

5.3.3.2 Consensus Scoring Function 1 - CSF1

During the development of the present work, we observed that PocketScore had better docking power (ability to identify native binding poses) than the scoring function in Vina or RF-Score-3. This implies that a poor PocketScore value is probably associated with an incorrect pose. In order to prevent Vina and RF-Score-3 from re-scoring incorrect poses, we implemented a consensus scoring function that is controlled by PocketScore and uses the score calculated by a secondary scoring function Z only if PocketScore is favourable (more negative values):

$$CSF1(Z) = \begin{cases} PocketScore + Z, & \text{if } PocketScore \leq -1.25 \\ PoceketScore, & \text{otherwise} \end{cases} \tag{5.1}$$

where Z denotes the secondary scoring function which is either Vina or RF-Score-3. If Z is RF-Score-3, the value of Z is multiplied by -1 to maintain all produced scores negative. If PocketScore is favourable, CSF1 returns the sum of PocketScore with the external scoring function. This is typically a large negative value controlled mostly by Z, since Vina and RF-Score-3 produce larger values (in magnitude) than PocketScore). When PocketScore is unfavourable, CSF1 prevents the secondary scoring function Z from contributing to the score, and simply returns the value from PocketScore, which is smaller in magnitude than those of Vina or RF-Score-3.

5.3.3.3 Consensus Scoring Function 2 - CSF2

CSF2 implements the same concept as CSF1, but instead of using a sharp threshold at a fixed value, the contribution of the external scoring function Z is weighted by a factor w that changes linearly between 0 and 1 as the value of PocketScore increases in magnitude from -1.25 to -1.5:

$$F2(Z) = PocketScore + w \times Z \tag{5.2}$$

where w is calculated as follows:

$$w = \begin{cases} 1.0, & \text{if } PocketScore < -1.5 \\ 0, & \text{if } PocketScore > -1.25 \\ -4 \times PocketScore - 5, & \text{if } -1.5 \leq PocketScore \leq -1.25 \end{cases} \tag{5.3}$$

Again, if Z is RF-Score-3, the value of Z is multiplied by -1 to maintain scores returned by CSF2 negatively valued.

5.3.4 Technical details

Openbabel [121] was used to generate 3D coordinates for ligands in the D3RGC set, and to protonate ligands from both the training set and the D3RGC set at pH=7. Ligands in the DUD-E set are provided with 3D coordinates and already protonated. Receptors were also protonated by Openbabel at pH=7. No further preparation steps were done to the proteins as the quality of the crystal structures was considered satisfactory. Strucure 1uyg has alternate locations for one sidechain atom in ILE 203 and alternate location A was chosen (the choice of alternate location is irrelevant because this residue is 15 Åaway from the active site.) Openbabel was also used for converting molecules between different file formats.

For the purpose of generating sets of binding poses, AutoDock Vina was used with default parameters: exhaustiveness = 8 and num_modes = 9. The size of the search space covers the entire binding pocket and is defined by a box centered on coordinates $x = 2$, $y = 8$, $z = 22.5$, and with size (in) $x = 20$, $y = 20$, $z = 30$. These coordinates are only meaningful for HSP90 structures aligned to structures provided by the D3R. The structure with PDB ID 4ykr is aligned to the same reference frame and is publicly available.

Root mean square deviations (RMSD) between docked ligands and crystallographic structures were calculated with a Python script that makes use of Pybel[122] (Python bindings for Openbabel) to identify symmetric atoms, providing symmetry corrected RMSD values. For example, rotation of a phenyl ring by 180 degrees results in a RMSD of zero because the atoms that change position are equivalent. All hydrogen atoms are ignored in RMSD calculations.

Calculation of the Kendall rank correlation coefficient (Kendall Tau) was performed with the Python module scipy [146], which implements the Tau-b version of Kendall's Tau. We report positive Kendall's Tau values when ligand scores and potency correlate in the expected order. Specifically, when more negative values predict more potent ligands, the scores were evaluated against log10(IC50). If larger positive values predict stronger ligands (e.g. RF-Score-3), the scores are evaluated against -log(IC50).

5.4 Results and Discussion

In this work, we designed a molecular docking workflow that uses distinct scoring functions for the search algorithm, selection of binding pose, and re-scoring the correct pose. This workflow is

schematized in figure 5.1. In typical docking workflows, these three steps are served by the same scoring function. Our approach provides a framework to optimize each step individually, making use of different scoring functions for different steps of the workflow.

In the following sections, we start by reporting results for pose generation, which corresponds to step 1 of the workflow (figure 5.1). Then we discuss the docking power power of different scoring functions, which depends on both steps 1 and 2. Finally, we report the performance of several variations of our protocol for identifying active molecules in the DUD-E HSP90 set and for ranking the 180 molecules in the D3R Grand Challenge 2015 set, both of which involve application of the workflow in its entirety.

5.4.1 Generating binding poses

The first of the three steps in the implemented workflow is the generation of binding poses. For any particular ligand, our goal was to generate an ensemble of poses in which at least one pose displays the native binding mode. In order to achieve this goal, a total of 48 receptors were built and evaluated to assess their performance in generating native binding poses for all 69 ligands in the training set. The number of receptors results from 4 different HSP90 conformations and from 4 water molecules that were modeled in 12 different configurations ($4 \times 12 = 48$). Details about these receptor conformations and water molecules are provided in methods section. We used AutoDock Vina to generate an ensemble of up to 9 poses for each ligand in each receptor model. An ensemble of poses is considered successful if at least one pose has a RMSD under 2 from the crystallographic ligand. Scores calculated by Vina for each pose are ignored at this step.

The number of ligands from the training set for which successful sets of poses were generated by each receptor model is reported in table 5.1. The helical conformation (1uyg) proved more efficacious than the other tested conformations independently of water configurations. This is explained by the larger binding pocket of 1uyg, which accommodates the native pose of several ligands while other conformations pose steric clashes. Additionally, ligands that bind in other HSP90 conformations can also be docked in the 1uyg binding pocket. The best results using a single receptor model are obtained using 1uyg with waters W2 and W3 (1uyg/W2+3). This receptor can generate successful sets of poses for 58 of the 69 ligands in the training set.

Analysis of table 5.1 reveals a pattern about water molecules: receptors including water W4 perform significantly worse than all other receptors. This occurs because a large number of HSP90 inhibitors contain a resorcinol group which displaces water W4, as is illustrated in figure 5.8. Receptors containing water W4 are therefore unable to accommodate the native binding pose of ligands that contain resorcinol. There is another common scaffold in HSP90 inhibitors that does not displace water W4, the aminopyrimidine group, and involves a hydrogen bond with water W4. This binding mode is illustrated in figure 5.9. Interestingly, receptors lacking water W4 are able to generate native poses for aminopyrimidine derivatives, indicating that this hydrogen bond is not required to reproduce the correct binding mode of these ligands.

After analysing ensembles of poses generated by single receptors, we investigated whether merging poses from two distinct receptors is beneficial. Figure 5.6 shows which receptors generate

Table 5.1: Generation of successful ensembles of poses by different receptor models. All 69 ligands from the training set were re-docked in each of the 48 tested receptor models. Up to 9 binding poses were generated by Autodock Vina for each ligand in each receptor. An ensemble of poses is successful if at least one pose is within 2 from the crystallographic structure. A perfect receptor model would generate successful ensembles for all 69 ligands. The number of successful ensembles generated by each of the 48 receptors is displayed according to receptor properties: rows indicate which waters were included, and columns indicate the used conformation. Values reported here are not to be confused with docking power, which is the ability of a scoring function to identify the correct binding pose.

	Conformation			
Waters	1uyg	1yc3	1yc4	2cct
dry	52	35	28	21
W1	52	36	31	26
W2	53	33	35	20
W3	55	35	33	14
W1+2	55	34	36	23
W1+3	56	39	31	22
W2+3	58	34	32	16
W1+2+3	57	36	33	18
W3+4	29	11	9	7
W1+3+4	31	14	11	5
W2+3+4	31	12	11	5
W1+2+3+4	31	12	12	4

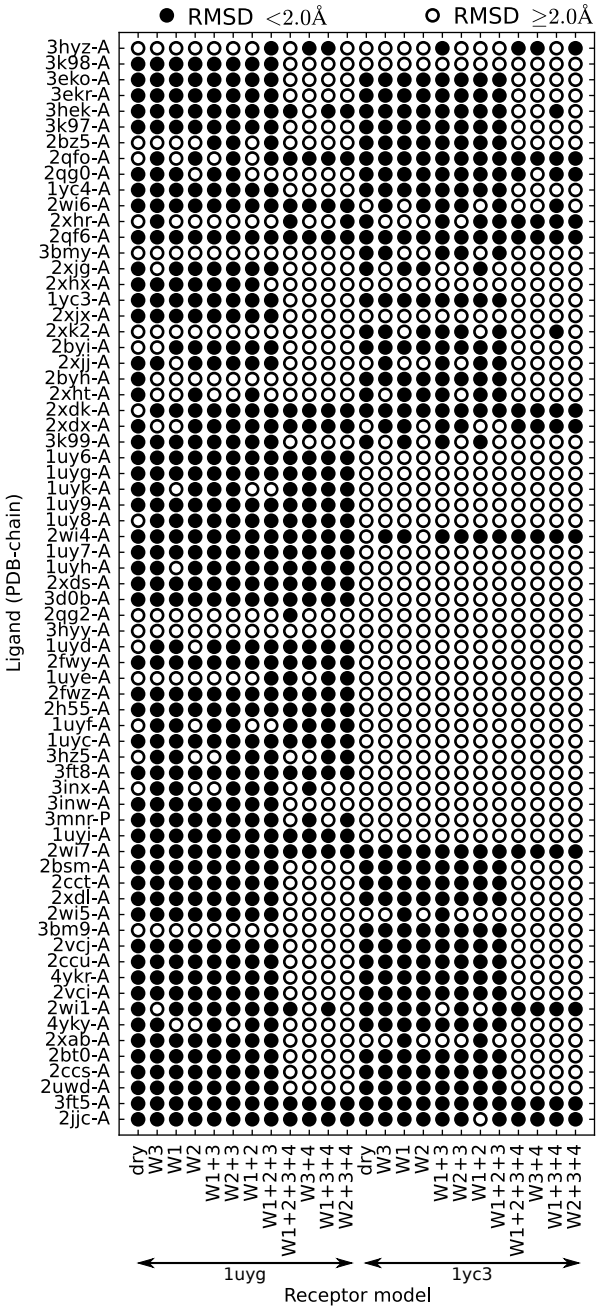


Figure 5.6: Success in generating the correct binding pose, by receptor model (x-axis) and for each ligand in the training set (y-axis). Up to 9 poses are considered for each receptor-ligand pair, using AutoDock Vina. A filled circle is used if at least one binding pose has a RMSD < 2 from the experimental binding mode.

successful ensembles of poses for which ligands. Analysis of this figure revealed that the receptor in the 1yc3 conformation with waters W1 and W3 (referred to as 1yc3/W1+3) can generate correct poses for eight ligands that can not be properly docked in the 1uyg/W2+3 receptor, which is the best performing standalone receptor. These eight ligands are associated with PDB IDs 4yky, 3bm9, 2xht, 2byh, 2xk2, 3bmy, 2xhr and 3hyz. Thus, adding poses generated in receptor 1yc3/W1+3 to poses generated in 1uyg/W2+3 increases the number of successful sets of poses from 58 to 66. These two receptors combined only fail for 3 out of the 69 ligands in the training set (1uye, 3hyy and 2qg2). Due to this high success rate, the protocol settled on the use of these two receptors for step 1, and the remaining steps of the work have been evaluated on poses generated with 1yc3/W1+3 and 1uyg/W2+3.

5.4.2 Docking Power Evaluation

The second step in the workflow concerns the selection of the correct pose from the ensemble of poses generated in stage 1. We rank poses in the ensemble by doing a single point calculation with a given scoring function. The best ranked pose is predicted to be correct by that particular scoring function and is selected. For a series of ligands with experimental binding structure, the number of selected poses that match the experimental structure is a measure of the docking power of a given scoring function. This approach allows us to compare different scoring functions independently from search algorithms because the poses are generated beforehand [5].

Three scoring functions were tested for this task: Autodock Vina (referred to as Vina), RF-Score-3 and PocketScore. The ensembles of poses are generated using two receptors: 1yc3/W1+3 and 1uyg/W2+3. Poses docked in the 1uyg/W2+3 receptor are penalized to compensate for the higher internal energy associated with this conformation [45, 145]. We use the nomenclature Vina($+x$) and RF-Score-3($-x$) to denote the use of a penalty value x. A larger penalty causes preference to poses docked in the 1yc3 conformation.

Two separate sets are used to evaluate docking power: the training set of 69 complexes, and a subset of six ligands from D3RGC. Prediction of the binding pose of these 6 ligands was requested to participants during stage 1 of the D3RGC (experimental structures were disclosed afterwards).

Table 5.2 summarizes the docking power of tested scoring functions. For ligands in the training set, PocketScore is able to identify a correct binding pose for 52 ligands, with a RMSD under 2, representing about $3/4$ of the training set. The performance of Vina depends on the penalty to poses docked in the helical conformation, with Vina(+.5) and Vina(+1.) providing the best results: 40 ligands correctly redocked. RF-Score-3 can only identify 11 correct poses. The poor docking power of RF-Score-3 is coherent with the methodology used in its development. RF-Score-3 has been designed to correlate with ligand potency based on native poses, and was not exposed to the problem of discriminating incorrect from native poses.

With respect to the subset of six ligands from the D3RGC set, PocketScore again displays the best performance, identifying a correct pose for 5 of the six ligands. The performance of Vina is strongly dependent on the penalty applied to poses docked in the helical conformation. An increasing penalty (0, +0.5, +1 and +2) leads to selection of an increasing number of correct poses

Table 5.2: Docking Power Results (Training set and Stage 1 of D3RGC). Vina, RF-Score-3 and PocketScore were tested for their ability in selecting correct poses. For ligands in the training set, the number of correctly selected poses is provided. For the subset of six ligands from the D3RGC, explicit RMSD values are reported for each ligand. RMSD values under 2 are highlighted in bold.

	RMSD(< 2)	D3RGC ligand ID - RMSD()					
Pose selection	Training Set (n=69)	40	44	73	164	175	179
Vina	36 (51.4%)	6.5	9.8	6.6	6.5	6.4	**1.9**
Vina(+.5)	40 (57.1%)	**1.**	9.8	6.6	6.5	6.4	**0.8**
Vina(+1.)	40 (57.1%)	**1.**	9.8	**1.5**	6.5	**0.4**	**0.8**
Vina(+2.)	37 (52.9%)	**1.**	9.2	**1.5**	**0.9**	**0.4**	**0.8**
RF-Score	3 (4.3%)	6.5	9.8	6.6	6.2	6.4	3.9
RF-Score(-.5)	10 (14.3%)	5.6	9.8	6.6	5.5	6.4	3.3
RF-Score(-1.)	11 (15.7%)	5.6	6.1	**1.5**	5.5	6.	3.3
RF-Score(-2.)	11 (15.7%)	5.6	6.1	**1.5**	5.5	6.	3.3
PocketScore	52 (74.3%)	**1.**	2.9	**0.8**	**1.4**	**0.4**	**0.6**

(1, 2, 4 and 5). This correlation between the helical penalty and the success rate of Vina is a consequence of the binding mode of these six ligands: all of them bind in a conformation similar to 1yc3.

It is important to note that the optimal helical penalty for Vina varies between ligand datasets. For the training set, the optimal value is +0.5 or +1, while for the subset of six D3RGC ligands the optimal value is +2. On the other hand, PocketScore displays consistent performance in different datasets without any parameter adjustments. Overall, PocketScore has the highest docking power, followed by Vina with intermediate performance, and finally RF-Score-3 which displays the weakest docking power.

5.4.3 Screening Power in the DUD-E HSP90 set

The DUD-E test set for HSP90 consists of 125 molecules with known activity and 4942 molecules presumed to be inactive (decoys). Ideally, scores produced for active molecules would be significantly better than those of decoys. The screening power metric employed is the area under the ROC curve (AUC), which can be interpreted as the probability of ranking an active molecule better than an inactive [147]. A random prediction yields an AUC of about .5 while a perfect method would produce an AUC of 1.0. The DUD-E paper reported an AUC of 0.69 for this test set, using DOCK 3.6 [148].

To evaluate screening power ligands must be ranked by an affinity prediction, or more generally, by a score. Thus, all three steps of our molecular docking workflow are involved (figure 5.1). With respect to step 1, ensembles of poses are generated with receptors 1yc3/W1+3 and 1uyg/W2+3. Regarding pose selection and re-scoring (steps 2 and 3), we explicitly evaluated the combined screening power of pairs of scoring functions.

Table 5.3 displays AUCs for the DUD-E HSP90 test set using different scoring functions for both pose selection and re-scoring. The bottom half of the table, spanning from the 5th to the

Table 5.3: Area under the ROC curve (AUC) in the DUD-E HSP90 test set, as a function of both pose selection (columns) and pose re-scoring (rows).

	Pose Selection				
Re-scoring	Vina	Vina(+0.5)	RF-Score	RF-Score(-.5)	PocketScore
Vina	.32	.33	.31	.45	.51
Vina(+0.5)	.37	.38	.30	.45	.50
RF-Score	.43	.42	.45	.50	.53
RF-Score(-0.5)	.52	.52	.47	.50	.53
PocketScore	.82	.85	.67	.72	.84
CSF1(Vina)	.81	.84	.66	.72	.83
CSF2(Vina)	.82	.85	.67	.72	.84
CSF1(RF-Score)	.81	.84	.67	.72	.83
CSF2(RF-Score)	.82	.85	.67	.72	.84

9th rows, displays larger AUC values ranging from .66 to .85. AUCs in the first four rows do not exceed .53, which is typical of random predictions. The last five rows correspond to the use of PocketScore as is (fifth row) or as an implicit argument for consensus scoring functions CSF1 and CSF2, suggesting PocketScore as the best scoring function for re-scoring binding poses. If PocketScore is used for re-scoring, the best results are obtained when poses are selected by either PocketScore, Vina or Vina(+0.5).

PocketScore only takes three residues in the receptor into account (Asp93, Thr184 and water W2), ignoring most of the binding pocket. PocketScore values only report if ligands are able to establish hydrogen bonds with these three residues. The distribution of PocketScore values (both pose selection and re-scoring) for actives and decoys is reported in figure 5.7.

5.4.4 Ranking Power - D3R Grand Challenge 2015

The D3R Grand Challenge involved ranking a set of 180 molecules according to their inhibitory activity against HSP90. Here, we report the ranking power of our workflow on this challenging set. Different scoring functions are used for pose selection and re-scoring (steps 2 and 3 in figure 5.1). Ensembles of binding poses are generated using receptors (1yc3/W1+3 and 1uyg/W2+3). The double slash notation Scoring-function-B//scoring-function-A denotes the use of scoring function A for poses selection and scoring function B for pose scoring. The Kendall rank order coefficient (Kendall's Tau) was used to measure ranking power. Results are summarized in table 5.4.

Better screening performance is achieved in the D3RGC when we use PocketScore for pose selection (corresponding to the last column in table 5.4). If we use PocketScore to select poses, a statistically significant ranking of ligand affinity is achieved almost independently of the scoring function that re-scores selected poses (with the exception of RF-Score-3//PocketScore and RF-Score-3(-1)//PocketScore).

The combination Vina//Vina corresponds to using AutoDock Vina for virtual screening, and achieves a Kendall Tau of 0.11 with statistical significance. The use of a penalty for poses docked

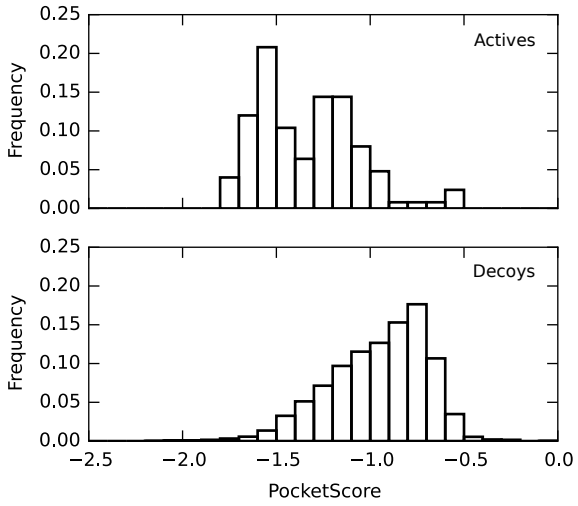


Figure 5.7: Distribution of PocketScore values for actives and decoys in the DUD-E HSP90 set. PocketScore was used for both pose selection and re-scoring.

Table 5.4: Kendall rank correlation coefficient between calculated scores and IC50 values for ligands in the D3R Grand Challenge. Different scoring functions were tested for pose selection (columns) and pose re-scoring (rows). *p-value < 0.05, **p-value < 0.01, ***p-value < 0.001.

	Pose Selection				
Re-scoring	Vina	Vina(+0.5)	RF-Score	RF-Score(-1)	PocketScore
Vina	0.11*	0.10*	0.07	0.21***	0.15**
Vina(+0.5)	0.11*	0.13*	0.08	0.21***	0.12*
RF-Score	0.06	0.06	0.08	0.09	0.07
RF-Score(-1.0)	0.07	0.07	0.08	0.09	0.05
PocketScore	0.03	0.01	0.01	0.07	0.22***
F1(Vina)	0.03	0.01	0.01	0.08	0.21***
F2(Vina)	0.04	0.01	0.01	0.07	0.23***
F1(RF-Score)	0.04	0.01	0.01	0.08	0.22***
F2(RF-Score)	0.04	0.01	0.01	0.07	0.25***

in the helical conformation Vina(+0.5)//Vina(+0.5) only increases Kendall's Tau to 0.13. These results are an important reference because they correspond to the usage of AutoDock Vina.

Surprisingly, Vina//RF-Score-3(-1) and Vina(+0.5)//RF-Score-3(-1) also perform well, with Kendall's Tau values of 0.21. This is unexpected because all redocking tests with the training set indicate that RF-Score-3 is unable to select correct poses. We speculate that this particular combination of scoring functions leads to cancellation of errors. It is important to remind the reader that this combination of scoring functions had no screening power in the DUD-E HSP90 set (table 5.3).

PocketScore//PocketScore is sufficient for a statistically significant rank order correlation with ligand affinity, achieving a Kendall's Tau of 0.22. This indicates that PocketScore captures important interactions of a large number of HSP90 inhibitors. Similar approaches have been developed for other targets in which known interactions are used to process molecular docking results [149].

The best ranking performance is achieved with the use of CSF2(RF-Score-3)//PocketScore, displaying a Kendall's Tau of 0.25. Therefore, RF-Score-3 can improve the ranking power of PocketScore//PocketScore but only if used as an argument for a consensus scoring function: RF-Score-3//PocketScore displays no ranking power at all. The behavior of CSF2 is controlled by PocketScore. If PocketScore predicts a good ligand, the value from RF-Score-3 is used. If the value from PocketScore is low, the value from RF-Score-3 is ignored. The better performance of CSF2(RF-Score-3)//PocketScore relative to RF-Score-3//PocketScore is consistent with the low docking power of RF-Score-3. It is likely that RF-Score-3 scores incorrect poses with excessively favourable scores. By allowing RF-Score-3 to only re-score poses which are likely to be correct, we are able to effectively make use of the improved scoring power (correlation with binding affinity) of RF-Score-3.

5.4.4.1 Domain of applicability of PocketScore

We observed that satisfactory performance of PocketScore is associated with specific ligand scaffolds. We divided the 180 ligand from the D3R Grand Challenge 2015 HSP90 set by visual inspection in three categories: ligands containing a resorcinol group adjacent to a hydrogen acceptor group (figure 5.8), ligands containing an aminopyrimidine group (figure 5.9) and ligands that do not contain either the resorcinol/Acc or the aminopyrimidine scaffold (figure 5.10). Both the resorcinol/Acc and the aminopyrimidine scaffolds are associated with satisfactory ranking of ligands by PocketScore//PocketScore, with Kendall's Tau values of 0.22 (figure 5.11) and 0.25 (figure 5.12) , respectively. However, ligands that do not contain any of these two scaffolds are not satisfactorily ordered by PocketScore//PocketScore, achieving a Kendall's Tau of 0.1 without statistical significance (figure 5.13). We speculate that these ligands bind HSP90 in a different site that does not involve the residues considered by PocketScore, and PocketScore is of no value for docking and scoring such ligands. Notably, most ligands in the training set contain either the aminopyrimidine or the resorcinol/Acc scaffolds, consistent with the specialization of PocketScore for these classes of molecules.

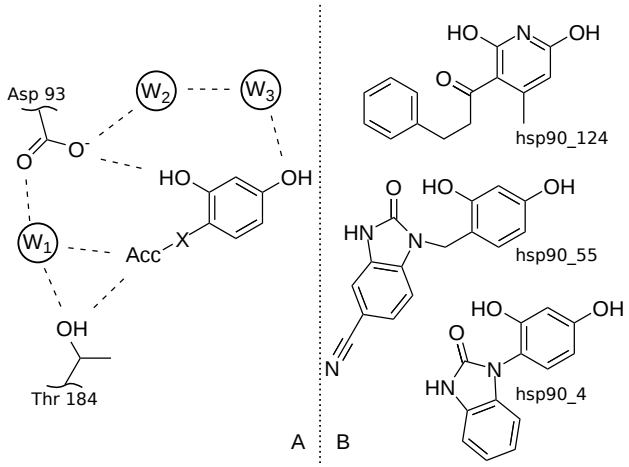


Figure 5.8: Binding mode of the resorcinol/H-bond acceptor scaffold. Panel A represents the structure of the scaffold with the resorcinol group and the H-bond acceptor group (Acc) separated by a linker (-X-). Hydrogen bonds involving the protein and the scaffold are represented by dashed lines. Waters W1 and W3 establish H-bonds with the scaffold, while W4 is displaced. This binding mode was observed in structures from the training set. A total of 58 ligands from the D3RGC HSP90 set contain this scaffold. The interaction pattern depicted here are a partial match to a pharmacophoric model developed for HSP90 inhibitors [1]. This scaffold is discussed extensively in ref [2]. Panel B illustrates three molecules with different liker (-X-) sizes. In ligands hsp90_4 and hsp90_55, Acc is a benzimidazolone, a frequent group in ligands from the D3RGC HSP90 set.

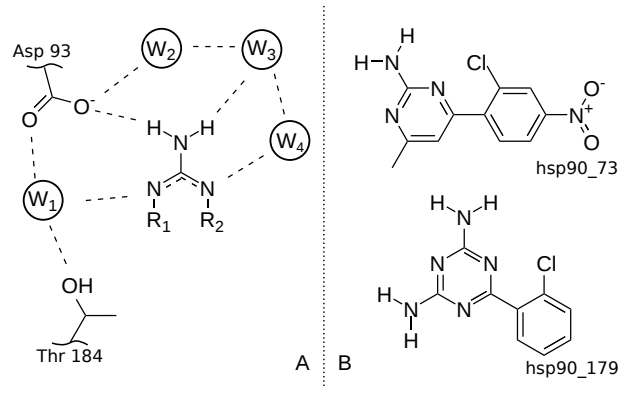


Figure 5.9: Binding mode of aminopyrimidine derivatives. Panel A illustrates the aminopyrimidine scaffold and the hydrogen bonds established in the binding site. Contrarily to the resorcinol/H-bond acceptor scaffold illustrated in figure 5.8, water W4 is not displaced and establishes hydrogen bonds with ligands. This binding mode was observed in structures from the training set. A total of 59 ligands from the D3RGC HSP90 set contain this scaffold. Panel B illustrates two molecules containing this scaffold.

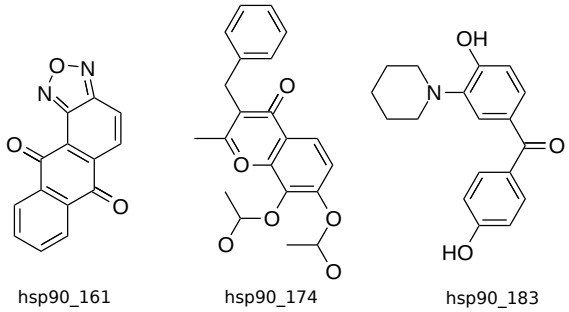


Figure 5.10: Ligands from the D3RGC HSP90 set that do not contain any particular scaffold. A total of 63 ligands do not contain either the resorcinol/Acc scaffold (figure 5.8) or the aminopyrimidine group (figure 5.9).

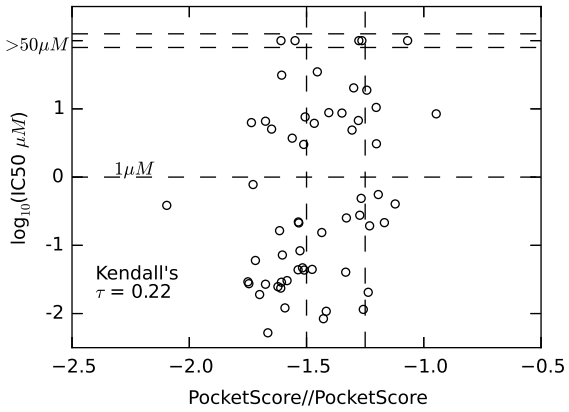


Figure 5.11: PocketScore vs. activity for GC2015 ligands containing the resorcinol/Acc scaffold (figure 5.8). Horizontal dashed lines provide visual guidance and vertical dashed lines correspond to PocketScore cutoffs in consensus scoring functions.

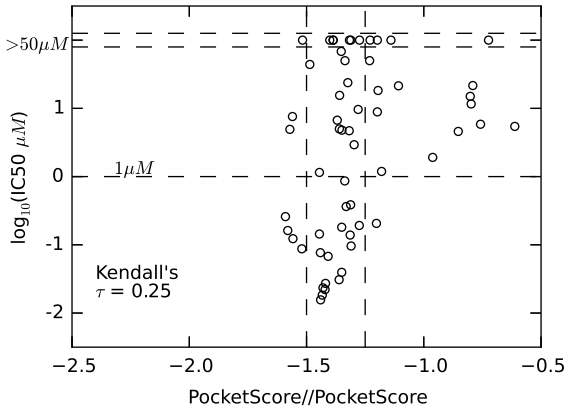


Figure 5.12: PocketScore vs. activity for GC2015 ligands containing aminopyrimidine (figure 5.9). Horizontal dashed lines provide visual guidance and vertical dashed lines correspond to PocketScore cutoffs in consensus scoring functions.

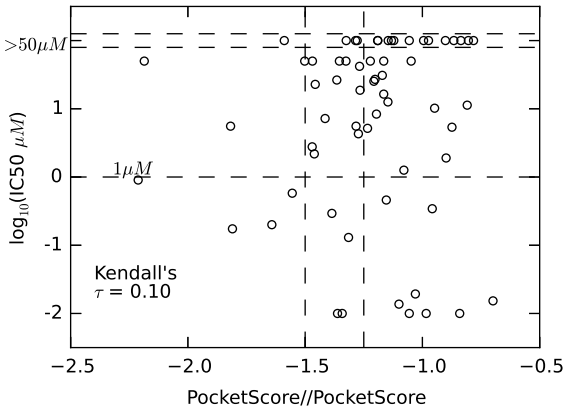


Figure 5.13: PocketScore vs. activity for GC2015 uncategorized ligands, i.e. not containing either the aminopyrimidine (figure 5.9) or the resorcinol/Acc (figure 5.8) scaffolds. Horizontal dashed lines provide visual guidance and vertical dashed lines correspond to PocketScore cutoffs in consensus scoring functions.

5.5 Conclusions

In the context of our participation in the D3R Grand Challenge 2015, we designed PocketScore, a scoring function that quantifies the interaction energy between a ligand and three specific residues in the binding site of HSP90. To assess the value of PocketScore in a molecular docking workflow, we tested the performance of various scoring functions to identify native binding poses (docking power) from an ensemble of poses generated beforehand with AutoDock Vina. PocketScore displayed greater docking power than other scoring functions, specifically RF-Score-3 and AutoDock Vina, demonstrating the importance specific interactions to appropriately model binding of ligands to HSP90.

We have also tested PocketScore for virtual screening applications, which require the calculation of scores to rank ligands by their predicted binding affinity. In our workflow, ligand scores are explicitly calculated by a pair of scoring functions: scoring-function-A//scoring-function-B, where B identifies the best binding pose from the ensemble of generated poses, and A produces the final score by re-scoring the selected pose. In the DUD-E HSP90 test, PocketScore//PocketScore was able to score active compounds significantly better than decoys, resulting in an AUC of 0.84. This result demonstrates that PocketScore is not only able to select native binding poses, but can also be used to discriminate binders from non-binders for molecules containing resorcinol and aminopyrimidine scaffolds. In the D3R Grand Challenge 2015, the importance of docking power became evident as combinations of scoring functions that use PocketScore for pose selection provided better predictions for the ranking of ligands. The best results were obtained CSF2(RF-Score-3)//PocketScore, where CSF2 is a consensus score that adds the value of RF-Score-3 to the output only if PocketScore is favourable, thus preventing RF-Score-3 from re-scoring poses that are likely to be incorrect.

Overall, this study corroborates previous findings on the importance of docking power to the

success of virtual screening campaigns [5]. It also demonstrates that specific residues on the binding pocket of HSP90 are important to model the binding of small molecules to HSP90, and that other generic scoring functions (AutoDock Vina and RF-Score-3), may underestimate these important interactions specific to a subset of HSP90 inhibitors.

Chapter 6

Enzymatic Flexibility and Reaction Rate: A QM/MM Study of HIV-1 Protease

António J. M. Ribeiro, Diogo Santos-Martins, Maria J. Ramos and Pedro A. Fernandes

Adapted from ref. [150].

In this work I ran part of the experiments and participated in results analysis.

Keywords: Enzymatic catalysis; QM/MM; Instantaneous Disorder; Transition State Theory; Protein Dynamics.

6.1 Abstract

The relevance of conformational fluctuations on enzyme rates has been a matter of debate for decades. Single molecule experiments have detected variations on the catalytic rates between different enzyme molecules, and within the same enzyme molecule, in a time scale larger than turnover. Computational methods can detect different energy barriers, induced by thermal conformational fluctuations, at a microscopic timescale, several orders of magnitude faster than the turnover rate of the fastest enzyme. Others have observed these barrier fluctuations, but few computational studies have dissected them in detail and tried to understand their origins and consequences. For this purpose we studied the first step of the reaction catalyzed by HIV-1 Protease, starting from 40 different conformations. We found activation free energies ranging from 14.5 to 51.3 kcal mol^{-1} . The calculated apparent barrier is 16.5 kcal mol^{-1} , which is very close to the experimental value of 15.9 kcal mol^{-1} for product release. These fluctuations are determinant to the overall rate, and are correlated to specific structural changes. The effect of each enzymatic conformation on the stabilization of the transition state can be explained by the electrostatic interaction of every protein residue with the flow of net electronic density (negative charge) from the reactants to the transition state.

6.2 Introduction

Enzyme structures fluctuate over time on a multidimensional free-energy landscape [151, 152, 153, 154]. Even at equilibrium, a broad enzymatic state such as the "enzyme-substrate complex" is a blend of innumerable interchanging conformations. This complexity extends to catalysis: there are countless possible transition state geometries connecting reactant conformations to product conformations. Conventional experimental kinetic studies overlook this diversity, because they measure properties of ensembles of enzymes, which are averaged over time and over moles of molecules. Enzymatic rates from single molecule experiments, however, present both what has been coined as static disorder (rate differences on different enzymatic molecules, assumed to be due to rate variations much slower than turnover) and dynamic disorder (rate differences on the same enzymatic molecule along a time close to turnover)[155]. Initially, it was thought that folding fluctuations slower than the time of the experiment caused static disorder, [156, 157] while faster structural fluctuations caused dynamic disorder [157, 158]. It is now accepted that intrinsic structural differences, such as the existence of post-translation modifications, or of truncated or partially folded enzymes, [159, 160] account for most cases of static disorder. As for dynamic disorder, it was shown that it is not a universal characteristic, as there are proteins that do not display it. [161, 160, 162]. Theoretical modelling of enzymes has been vital to the progress of the field of enzymatic catalysis. Most notably, computational methods are the only systematic approach to describe the reaction pathways followed by enzymes at an atomic level [163, 164, 165, 166, 167]: Enzymatic reactions can be modeled today for the whole enzyme with QM/MM methods, and chemical accuracy. There are a growing number of theoretical studies that focus on the link between structural flexibility and its effect on enzymatic catalysis. One of the main observations is that the calculated activation energies are dependent on the chosen initial structure for the reactants [168, 169, 170, 171, 172, 173, 174, 175, 176]. Depending on the exact methodology or enzyme, differences on the barrier between 5 kcal mol^{-1} [168, 170], and more than 30 kcal mol^{-1} [169, 174] have been found. Some explanations for the relation between the differences in the structure and in the barriers have been proposed, [168, 169, 170] but they are mostly structural considerations, like specific distances between active center atoms, and hence specific to the enzymes in question. In this paper we explored in greater detail the electrostatic interactions of non-catalytic residues with the active center, and found that, in conjunction with different active site conformations, they are the main cause for instantaneous barrier fluctuations. In principle, this kind of analysis can be applied to any enzyme.

For the sake of clarity, we introduce the concept of instantaneous disorder. It is impossible to reconcile the dispersion in the barriers obtained by computational methods with the dispersion obtained from experimental data on dynamic disorder due to the altogether different time scales considered. Dynamic disorder reflects variability in conformations that occur in a timescale larger than k_{cat}. Computational methods, on the other side, have only access to much smaller time scales. Instantaneous disorder is then the instantaneous fluctuations in the enzyme structure with a time scale much smaller than k cat , which leads to different activation barriers. We picture a free energy

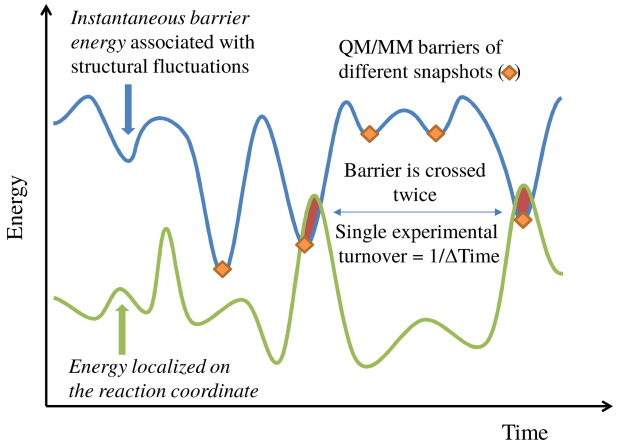


Figure 6.1: A model for the catalytic landscape of enzymes and its relation with QM/MM results.

profile for the reaction that is constantly changing, with innumerable different possible energy barriers that change both quantitatively (different energies) and qualitatively (different paths). The barriers affected by instantaneous disorder cannot be measured experimentally because most of them are never crossed. Figure 6.1 is illustrative of this model. We consider two independent motions: the movement of all the residues in the enzyme, which leads to different energy barriers; and the movement of the atoms in the reaction coordinate. The reaction can only take place when the energy localized on the reaction coordinate (bottom line) is enough to overcome the barrier provided by the enzymatic conformation at that time (top line). The effect of the enzyme scaffold in the barrier is given by its instantaneous interactions (mostly electrostatic) with the active center. Variations in the enzyme structure affect the barrier only through changes in these instantaneous interactions. It is obvious that these two motions are not truly independent but, for the sake of the argument, it is only necessary that the existent correlation is not the cause of the observed rate.

In this work we used 40 conformations taken from a MD (molecular dynamics) simulation of HIV-1 protease to obtain 40 free energy profiles of its catalytic mechanism with a QM/MM methodology. HIV-1 protease was chosen as model enzyme due to its small size and simplicity: two identical chains of 99 residues each compose protease, [177, 178] and it has only two (aspartic) catalytic residues [179, 180, 177, 181, 182, 183, 184, 185, 178]. Furthermore, its mechanism is well established from experimental [179, 180, 181, 182, 183, 184, 185, 186, 187] and theoretical [188, 189, 190] standpoints, as well as from studies on other aspartic proteases [191, 192, 193, 194, 195]. To limit the MD sampling to productive conformations we fixed a hydrogen bond between the protonated oxygen of Asp25B and the carbonyl oxygen of the scissile peptide bond. This hydrogen bond is also found in an ensemble of unconstrained structures, but its occupancy is quite low. The entropic cost of such constraint was calculated recently to be 4.6 kcal mol^{-1} at physiological temperature [196]. This contribution was added to all the barriers in order to compare them directly with the experimental turnover, that constitutes an upper limit for the chemical step, as product release is rate limiting. The active center of each structure was inspected to identify geometric parameters important to the barrier. Furthermore, to understand the effect

of every protease residue in catalysis, we employed a method where each residue was removed from both reactant and transition state structures. By comparing the barriers in the "deleted" systems with the barriers in the complete enzyme, we were able to define exactly the contribution of the 198 residues to catalysis in each of the 40 models . We observed a highly heterogeneous catalysis landscape with activation energies ranging from 14.5 to 51.3 kcal mol^{-1} . We found that different reaction paths and different conformations in the active center led to different barriers, as expected. However, even initial structures that led to the same mechanism and had very similar active center geometries presented large barrier heterogeneity. The main cause behind the dispersion in these barriers is the oscillation of the electrostatic interactions (coming from Coulomb and van der Waals forces) between the active center and the rest of the enzyme due to thermal conformational fluctuations. Our results show consistently that residues that help the flow of negative charge towards the scissile peptide bond will decrease the barrier, while residues that hinder such flow of negative charge density increase the barrier. An analysis of the different positioning of these residues in the snapshots is enough to explain why barriers oscillate so much. Further studies will be necessary to assess if this conclusion can be generalized to more enzymes. It is tempting to hypothesize that the effect will be important in enzymes where the charge distribution is very different in the reactants and transition state structure, whereas it will be less important in free-radical reactions, for example.

6.3 Methods

The overall computational protocol in this work followed these steps: a) Modeling of the enzyme-substrate complex from the 4HPV PDB structure; b) Molecular dynamics simulations (10+5 ns) to stabilize the modeled structure; c) Three independent MD simulations (total of 80 ns) to sample the conformational space of the system; d) QM/MM calculation of the reactants and transition state of the first step of the reaction for 40 different initial structures, resulting from the previous molecular dynamics simulations, by unconstrained geometry optimization of both stationary points at the ONIOM(B3LYP/6-311+g(2d,2p)|AMBER) level; e) Assessment of the influence of every residue in the activation energy for each one of the 40 structures. The protease model was built from the X-ray structure 4HPV. 47 This structure contains the entire HIV-1 protease bond to the substrate-based inhibitor Ac-Thr-Ile-Nle-[CH 2 -NH]-Nle-GIn- Arg.amide. We modeled the inhibitor into the Ac-Thr-Ile-Met-[CO-NH]-Met-GIn-Arg.amide substrate by changing the [CH2-NH] group to an amide [CO-NH] group, and the Nle (Norleucine) residues to methionine residues. In addition, we added the catalytic water into the active center, and protonated Asp25B. This residue is experimentally known to be protonated to fulfill its role in the catalytic mechanism. This model contains 3232 atoms. The modeling task was done in the GaussView software [197].

In order to equilibrate the modeled structure we did a 10 ns molecular dynamics simulation without any structural restrictions. During this simulation, the active center adopted a conformation that was not adequate to the catalytic reaction. To force the protein to adopt the required conformation we constrained the distance between the catalytic hydrogen atom of Asp25B and the

carbonyl oxygen atom of the substrate with an harmonic potential having an equilibrium length of 1.8 Å and a force constant of 50 kcal $mol^{-1}Å^{-2}$. The free energy cost of restraining the sampling to this subspace has been calculated before to be 4.6 kcal mol^{-1} [196]. This contribution was added to all barriers. We ran another 5 ns MD with this new Hamiltonian. From the last structure of this 5 ns MD, three more simulations were launched with different seeds for the initial velocities, in order to provide sampling space for the subsequent work. Two of these simulations ran for 24 ns, and the other for 32 ns, to a total of 80 ns. The catalytic water is known to occupy only transiently the active site. Our MD simulations confirmed this experimental observation, as the catalytic water diffused away from the active site after a few tenths of ns. Any time that the catalytic water diffused from the active site we stopped the simulations and started new simulations with the catalytic water inside the active site again. That is why we made three simulations, two with 24 ns and a third with 32 ns. These were the times during which the catalytic water remained in the active site.

Forty QM/MM models of the enzyme substrate complex were defined from 40 structures of the sampling dynamics equally spaced in time (i.e. with time intervals of 2 ns). For the ONIOM models, all water molecules were removed, except for the catalytic and structural ones. There were no water molecules on the inside of the protein except for these two. The solvent was removed from the calculations deliberately, and PCM calculations with a set of dielectric constants probed the effect of the environment (see below). Protease binds and cleaves two very large substrate polyproteins (Gag and Gag-Pol), much larger that protease itself (Gag-Pol has about 1500 residues). Most of the protease becomes buried in the substrate protein. Despite the enormous scientific and pharmacologic relevance of this complex, it has remained elusive to crystallography so far. It is unknown how much of protease is exposed to solvent and how much is buried in the binding partner. Additionally, protease acts after viral assembly, in a densely packed environment. To estimate the possible magnitude of environment effects, we performed single point energy calculations with an implicit solvation model (ONIOM-PCM), using different values for the dielectric constant (4, 10 and 80). As expected, higher dielectric constants lead to more pronounced changes in the calculated activation barrier, but the differences (in average 0.64 ± 0.57 kca mol^{-1} withε=4, 0.78 ± 0.72 kcalmol^{-1} with ε=10, and 0.81 ± 1.08kcal mol^{-1} with ε=80) are not significant when compared to the range of activation barriers that is under study. Since the changes in activation barriers induced by solvation are negligible, and the real environment is not really known in detail, we report values corresponding to the enzyme without solvent. The effect of the environment in all calculated barriers is reported in the Supporting information (Table S1 and Figure S1).

The reaction path was studied in the same manner for all models. We started by optimizing the reactants structure and subsequently scanning the reaction coordinate (i.e. the distance between the oxygen of the water nucleophile and the carbon of the peptide bond). The structure with the highest energy in the scan was identified and used as a guess to freely optimize the transition state. A frequency calculation was done to confirm the nature of the optimized structure. To obtain a reactants structure in the same relative minimum as the transition state, we performed an IRC (intrinsic reaction coordinate) calculation in the reactants direction for 10 steps. Instead of pro-

longing the IRC calculation all the way to the reactants, the last structure from the IRC was taken and freely optimized, while making sure that no structural rearrangements (independent from the reaction coordinate) occurred in this last optimization. The different active site conformations were firstly analyzed with the aim of finding structural features correlated with the size of the barriers. This kind of analysis was used before to identify geometric descriptors correlated with the magnitude of the barriers [172]. Here, we limited our analysis to bi-dimensional linear regression models, and used four interatomic distances as descriptors: (1) the hydrogen bond between Asp25A and the catalytic water, (2) the distance between the oxygen of the catalytic water and the carbonyl carbon in the peptide bond, (3) the hydrogen bond between the protonated Asp25B and the catalytic water and (4) the hydrogen bond between Asp25B and the oxygen in the carbonyl group. Details about the performance of all regression models are provided in the Supporting Information.

In a second part of the analysis of the results, we employed a procedure to assess the influence of each MM layer residue in the activation energy of the reaction. For each of the optimized structures of reactants and transition states we did a series of single point energy calculations where we removed every single protein residue, substrate residue or structural water (a second water molecule is known to be present in the protease active center, which has only a structural role). This gives a total of 201 calculations for each pair (reactant and transition state) of optimized structures (196 protein residues, 4 substrate residues and the water molecule, all part of the low layer), or 402 calculations for each barrier; since we end up with 39 productive structures we did more than 15000 single point energy calculations in total. The influence of each residue in the barrier is given by the difference between the native barrier and the barrier calculated without the residue. A positive number means that the deleted residue increases the barrier, while a negative number means that the deleted residue decreases the barrier (it stabilizes the transition state more than it stabilizes the reactants). Note that the purpose of this procedure is to calculate the potential energy effect of each residue on the activation energy of the wild type enzyme. It is a conceptual quantity that cannot be measured experimentally. We are not trying to calculate the free activation energy of mutated enzymes, which is a different physical quantity. Our purpose is to understand if residues stabilize or destabilize the barrier at the specific TS conformation, in absolute terms.

To measure reference distances between residues and the active site, we considered the following geometric points: the oxygen atom for the nucleophilic water; the amide carbonyl carbon for the substrate; the average atomic positions for neutral residues; the guanidine carbon for arginine residues; the nitrogen side chain for lysine residues; and the carboxylate carbon for aspartate and glutamate residues.

6.3.1 Molecular Dynamics Simulations Details

A total of 9960 water molecules were added to the protein in a rectangular box of 88 Å X 67 Å X 71 Å. At least 12 Å were left between the surface of the protein and the face of the box. Explicit van der Waals interactions were truncated at 10 Å and the Coulombic interactions were calculated with the PME method, with the real part also truncated at 10 Å [198]. A time step of 1 fs was used

in simulations where the distance between the side chain proton of Asp25B and the oxygen of the peptide bond was constrained with a harmonic potential. For simulations without this restriction we used the SHAKE algorithm [199] and a time step of 2 fs. An initial warm-up dynamics of 100 ps (from 0 to 300 K) was done in the canonical ensemble (NVT). The production dynamics ran in the isothermal-isobaric ensemble (NPT) with the Langevin thermostat and isotropic position scaling, at 300 K and 1 bar. Parameters from the AMBER03 force field [200] were used for all the amino acids in the system, including the protein and peptide substrate. TIP3P water molecules [60] were used for the catalytic, structural and solvent water molecules. Molecular mechanics simulations were done with the AMBER10 software [201].

6.3.2 ONIOM model details

For the QM/MM calculations we divided the system into two layers. The high layer comprises the side chain of the two catalytic Aspartate residues, Asp25A and Asp25B, the water nucleophile, and 12 atoms of the substrate, as seen in figure 6.2. We deliberately used a very small QM layer in this study, for two reasons. The first is that the division represents the conceptual division between the "reacting atoms" and the "environment". It is much simpler and theoretically clear to define the "environment" at the MM level because MM allows us to clearly isolate the contributions of each atom or residue to the barrier. Second, as we are studying the reaction dozens of times, with transition state optimizations and IRC calculations (and more than 15,000 single point energy calculations), the use of a large QM layer would limit the number of conformations we could explore, which is the focus of the study. Besides, the choice of studying HIV-1 protease was made purposefully on this basis; it is a small protein, with only two catalytic residues and with a very well-known and undisputed catalytic mechanism. A larger QM layer would increase the accuracy of the energies, but would not change them meaningfully. The MM layer includes all the remaining protein and substrate atoms, as well as the structural water molecule. The interaction between the layers was treated with the electrostatic embedding scheme. Comparable models have been used in the past to study the reaction mechanism of HIV- 1 protease and similar enzymes [188, 189, 190, 195] The QM layer was optimized with the B3LYP [49, 202] density functional and the 6-31G(d) basis-set [203], while the MM layer was treated with the parm96 force field [51] as implemented in the GAUSSIAN09 program. Sautet and co-workers have shown that B3LYP properly reproduces the geometries and energies of the first transition state of the HIV-1 protease reaction given by MP2 and CCSD(T) [190]. The difference in activation energy between B3LYP/6-311++G(2d,2p) and CCSD(T)/6-311++G(d,p) levels of theory is only -0.6 kcal mol^{-1}. The effect of the (D3) dispersion correction [204] was also calculated. Its contribution to the barrier was quite small (1.0 kcal mol^{-1} in average). Given the magnitude of the correction and the fact that B3LYP was shown to be excellent in reproducing the barrier for this transition state we haven't included the D3 correction in the results. The dispersion correction for every barrier is shown in the SI (Table S1 and Figure S2).

Zero point energies and entropic corrections were also calculated, in order to obtain free energies of activation. The 39 barriers of the native enzyme were recalculated with the 6-311+g(2d,2p)

Table 6.1: Free energies of activation (kcal mol^{-1}) and rate constant (s^{-1}) for the three mechanisms found and experimental data. A correction of 4.6 kcal mol^{-1} relative to the MD constraint is included in all the values

Mechanism	Barrier range	Average barrier	Aparent barrier	k_{cat}
A.1	14.5 - 38.2	27.5 $\pm$6.6	16.2	24
A.2	23.2 - 41.1	31.9 $\pm$5.5	24.7	2.3 x 10^{-5}
B	38.5 - 51.3	44.9 $\pm$9	38.9	2.3 x 10^{-15}
Total	14.5 - 51.3	29.6 $\pm$7.5	16.5	15
Experimental [207]				41 $\pm$6

basis set. All ONIOM [205] calculations were done with GAUSSIAN09 [206].

6.4 Results and Discussion

6.4.1 The fluctuations of the free activation energies

Among the 40 initial reactant structures studied, only one was non-productive. In this case, the optimization of the reactants led to a structure where Asp25B is making a hydrogen bond with Asp25A, instead of making it with the substrate. For the remaining 39 initial structures, we calculated the activation energies for the first step of the catalytic reaction of HIV-1 protease. We divided the barriers in three categories (A.1, A.2 and B) based on three clearly different conformations adopted by the active center. The first two categories (A.1 and A.2) are essentially the same mechanism, while B is a completely different path. Scheme 1 depicts these different paths, which are described later, and Table 1 summarizes their kinetic data. For variant A.1 of mechanism A, the average of the barriers is 27.5 kcal mol^{-1} with a standard deviation (SD) of 6.6 kcal mol^{-1} . The apparent barrier (i.e. the barrier that would be observed experimentally from a macroscopic population of enzymes in these initial states and proportion) was calculated using the transition state theory from the average of the k cat values. Its value amounts to 16.2 kcal mol^{-1}. For the variant A.2 of mechanism A, the average of the barriers is 31.9 kcal mol^{-1} with and an SD of 5.5 kcal.mol^{-1} . The apparent barrier for this mechanism is 24.7 kcal mol^{-1} -1 . There are only 2 structures that followed mechanism B, one with a barrier of 38.5 kcal mol^{-1} , the other with a barrier of 51.3 kcal mol^{-1}. The average of these two barriers is 44.9 kcal mol^{-1} $^{-1}$ the apparent barrier is 38.9 kcal mol^{-1}. The overall barrier average for all structures is 29.6 kcal mol^{-1}, with a standard deviation of 7.5, and the overall apparent barrier is 16.5 kcal mol^{-1} . The experimental value for the overall free energy barrier of HIV-1 protease with this substrate, 15.9 kcal mol^{-1}, 61 which constitutes an upper limit for the chemical step, is in agreement with the overall apparent barrier.

The results shown here, and previously seen in other simulations as well, indicate that the activation barriers of enzymes fluctuate significantly, and are often related to structural flexibility. We call these fluctuations instantaneous disorder. Protease, in less than 100 ns, presents many states with very different activation barriers. The observed (or apparent) k cat is given by equation 1, where P i is the probability of finding the enzyme in the reactant state i, and $\Delta G_i^{\circ\ddagger}$ is the

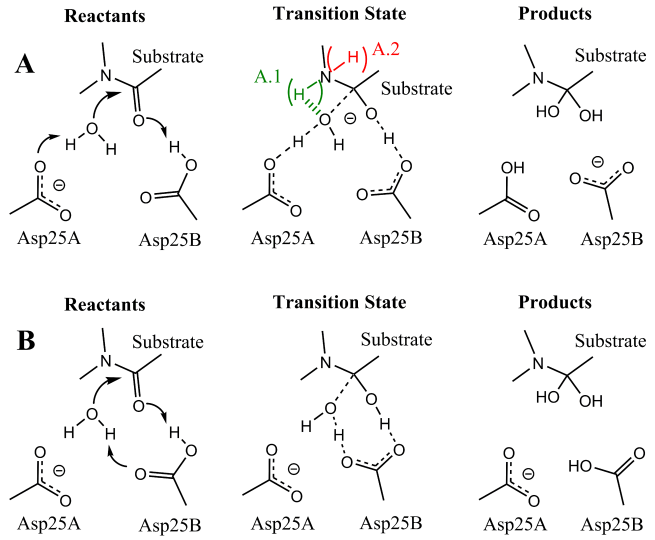


Figure 6.2: Different mechanisms and configurations adopted by the active center of protease. Part A: The water nucleophile attacks the peptide bond and gives a proton to Asp25A. Depending on the configuration of the peptide bond when it loses the planarity, the nucleophile can be more or less stabilized by the highlighted hydrogen (configurations A.1 and A.2). This mechanism is the most commonly describbed in the literature. Part B: The water nucleophile attacks the peptide bond, but this time it gives its proton to Asp25B. In this unfavorable reaction path Asp25A loses its catalytic role.

corresponding barrier free energy.

$$k_{cat} = \kappa \frac{k_B T}{h} \sum_{i=1}^{n} P_i e^{\frac{-\Delta G_i^{\circ\ddagger}}{RT}} \tag{6.1}$$

For the purpose of this work, we assumed that all reactants have equivalent probabilities. The probability of each barrier is accounted indirectly by the number of times the barrier appears in our 40 structures. Note that averaging the k_{cat} in this manner is not the same as averaging the activation energies and calculating k_{cat} from the average barrier. The k_{cat} calculated from the average of the barriers has no physical meaning, and cannot be associated with any experimental value. The general model that explains these results is depicted in figure 6.1. The enzyme goes through many conformations, which are associated with different barriers (instantaneous disorder). Independently, the energy accumulated on the reaction coordinate also fluctuates. The reaction occurs when the energy localized on the reaction coordinate is higher than the current instantaneous barrier. The height of this barrier is essentially determined by the structure of the enzyme as it is in the reactants state, since the enzyme has no time to go through large rearrangements within the timescale of a molecular vibration: the time the chemical reaction takes to occur.

The observation of these large fluctuations in the catalytic barriers is not new. In the case of protease, a decrease in the activation barrier from 50 kcal mol^{-1} to 20 kcal mol^{-1} was previously reported, as the substrate approximates the catalytic aspartates [208]. We obtain a similar effect when we compare mechanism B to mechanism A. Additionally, HIV-1 protease is able to undergo

relatively large structural rearrangements in the flaps region [209], which also must affect catalysis. On a microsecond timescale, these flap movements are associated mostly with substrate entry and exit [210, 211, 212, 213], and are important to understand the full catalytic cycle of HIV protease, in particular because product exit is thought to be rate-limiting in physiological conditions. However, since in the present work we are mostly concerned with the motions that directly affect the catalytic step, we focused instead on conformational fluctuations occurring when the substrate is bound to the protease, and the flaps are in the closed state. In other enzymes, such as ketosteroid isomerase, small changes in active site hydration (the entry of two additional water molecules around the catalytic Asp38), driven by a protein conformational change that closes/opens the active site, induced a raise in the free energy barrier of around 20 kcal/mol [214]. Variations up to 17 kcal mol^{-1} were also observed in the P450 catalyzed epoxidation and hydroxylation of propene and cyclohexene [215]. In this case, the origin of the variation seems to stem from the multiple possible substrate orientations, a factor that is expected to be particularly relevant when promiscuous enzymes bind small substrates. In the case of fatty acid amide hydrolase, 36 different barriers were derived, with a free energy span of about 11 kcal/mol [172]. In that work the authors carried out a successful statistical analysis that identified the specific residues/interactions that were responsible for most of the observed differences.

There has been a long-standing debate about the relationship between enzyme dynamics and enzyme catalysis [216]. We do not consider the fluctuations we observed here, 'dynamic effects', in the sense that they are not dynamically coupled to movements along the reaction coordinate. We consider that the movement along the reaction coordinate is much faster than the movement along other orthogonal directions in the PES. The roughness of the time-dependent PES and the subsequent instantaneous barriers arise from a thermal equilibrium distribution. The conformations explored here are just specific sub-states of the whole macroscopic ensemble. Additionally we rationalize their effect through the TST without any role of the transmission factor.

The effect of tunneling will be small compared to the fluctuations in activation free energy that are originated by the different enzyme conformations. However, this effect will be different from conformation to conformation, as the height and the width of the barriers changes, and will be more relevant for the lower barriers, that also contribute more to the observed activation free energy.

6.4.2 The effect of conformational fluctuations in catalysis

After analyzing the 40 structures of reactants and transition states, we found three factors that account for the fluctuations:

1. Different reaction mechanisms (figure 6.2). Two structures follow a mechanism other than the path described in the literature for HIV-1 protease (path B). In this mechanism, Asp25A loses its role in the reaction, and the proton of the nucleophile goes to the free oxygen atom of Asp25B instead. Since the oxygen atom of Asp25B is less electronegative than the oxygen atom of Asp25A, the proton of the nucleophile is less stabilized in the transition state, which becomes more energetic;

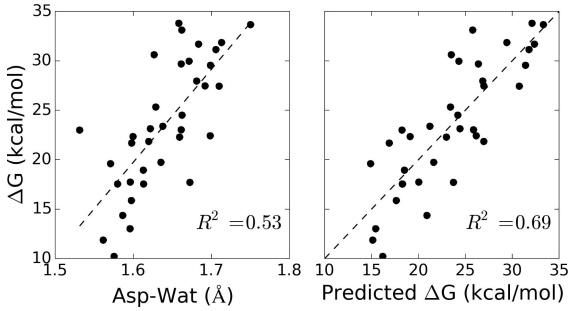


Figure 6.3: Correlation between activation barriers and key interatomic distances. The plot on the left side shows the computed activation barriers as a function of the shortest distance between a proton of the catalytic water and a carboxylic oxygen of Asp25A ('Asp-Wat' distance). On the right side, the x-axis represents the predicted activation barrier using linear regression with two explanatory variables: the 'Asp-Wat' distance and the distance between the oxygen in the catalytic water and the carbonyl carbon in the peptide bond ('Wat-Pep' distance).

2. Different active center conformations. The hydrogen atom of the scissile amine bond is found in two conformations. In the A.1 position, the hydrogen is making a hydrogen bond with the attacking hydroxide. This interaction stabilizes the negative charge that builds up in the hydroxide in the TS. On the A.2 position the hydroxide faces the lone electron pair of the amide nitrogen, a repulsive arrangement that leads to higher activation energies; We also found that simple geometrical descriptors, such as key interatomic distances, correlate well with differences in the activation barriers [208, 172, 214]. We have chosen four obvious distances directly related to the reaction coordinate (definitions for these distances are provided in the methods section). We built all possible regression models using up to two explanatory variables (distances). Figure 6.3 illustrates the performance of the best one- dimensional model and the best two-dimensional model, both providing satisfactory predictive power. The single variable displaying the highest correlation with the activation barrier is the distance corresponding to the hydrogen bond between the catalytic water and Asp25A 'Asp-Wat'. From the chemical viewpoint this result is intimately associated with the reaction coordinate: it represents the importance of abstracting a proton from the water in order to allow the attack on the carbonyl group. The best bi-dimensional regression model added the 'Wat-Pep' distance (the distance between the water oxygen and the carbonyl carbon) to the already discussed Asp-Wat, increasing R^2 from 0.53 to 0.69. This second distance completes the chemical path linking Asp25A and the carbonyl group in the peptide bond. It is worthy to note that the largest change in charge distribution along the reaction coordinate is precisely between Asp25A and the peptide bond (see figure 6.5).

The results shown in figure 6.3 indicate that simple bi-dimensional regressions are insufficient to explain quantitatively the wealth of different barriers that protease shows within a few ns. Instead of introducing more explanatory variables into the model, we looked for the physical sources of these fluctuations and analyzed the contribution of each individual residue to the energy barrier.

3. Conformational fluctuations in the rest of the enzyme structure that affect the barrier through electrostatic interactions. The sign and size of each residue contribution to the barrier (stabilization

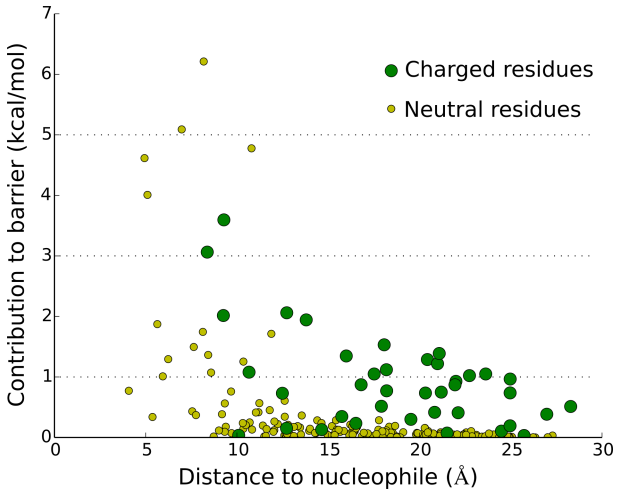


Figure 6.4: Averaged contribution of each residue for the reaction barrier (absolute values), plotted against the (average) distance of the residue to the nucleophilic water oxygen.

of destabilization) can be explained by looking at the flow of negative charge along the reaction path (see figure 6.5). The rest of the discussion will focus on this effect. Before addressing the problem of the residues' contributions to the activation energy fluctuations it will be instructive to focus first on explaining how the residues affect the catalytic barriers. The results we obtained in this respect are extremely intuitive and consistent. To begin with, it is fairly noticeable that residues near the active center have a greater influence on catalysis than the more distant ones (figure 6.4).

The decay is very fast to neutral residues, but less accentuated to charged ones. The effect of charged residues is still meaningful even at distances of 20 Å from the active center. As for neutral residues, after 10 Å the influence is already negligible. As it is clear, most of the contribution (either positive or negative) of neutral residues on the activation energy comes from the first layer of residues around the active center. The interpretation on charged residues is more complicated. Even if we admit that ignoring a 1 kcal mol^{-1} contribution for a single residue at 20 Å of the active center is acceptable, ignoring dozens of such contributions is not prudent. All these residues are moving and their charge is affecting the barrier, the dispersion in the values of individual barriers could be significant, even if these movements do not affect the average barrier. Since the apparent barrier is dominated by contributions of these transient smaller barriers, these fluctuations are of the upmost importance.

The contribution of each residue towards stabilizing or destabilizing the TS is easy to rationalize if we explain it against the charge transfer that takes place when the system evolves from the reactants to the transition state. As seen on figure 6.5, in the reactants, the negative charge on the active center is localized on Asp25, while in the transition state the charge is delocalized through the active center but centered on the substrate and hydroxide ion. The net transfer of negative charge is then upwards and rightwards. It is expected, and we will show just that, that residues that help the rearrangement of electronic density towards the transition state configuration will

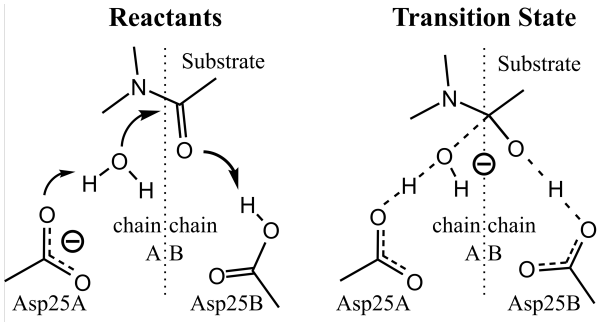


Figure 6.5: The flow of negative charge that happens along the reaction coordinate.

lower the activation barrier and residues that stabilize negative charge in Asp25A will increase the barrier. In other words, if we consider the orientation shown in figure 6.5, positive residues in the upper right corner stabilize the transition state, and in the lower left stabilize the reactants. The opposite is true for negative residues. This interpretation is also extensive to neutral residues, if we consider their partial charges. If the positive partial charge is closer to the active center than the negative partial charge, the residue will behave as if it was a positive residue, and vice-versa. The effect of neutral residues will decay faster than the effect of charged residues, because the partial charges are small and add to zero, with a counterbalancing charge almost at the same distance from the active center. The neutral residues that significantly affect the barrier and their average contribution to the activation energy are depicted in figure 6.6. A negative number means that the residue lowers the activation energy, while a positive value means that the residue increases the activation energy. Leu24 and the structural water molecule are the residues that decrease the barrier to a greater extent. They do it by opposite effects: the structural water molecule stabilizes the transition state with a positive partial charge near the peptide bond. Instead, Leu24 destabilizes the reactants by having its carbonyl group near the side chain of Asp25. The other three residues increase the barrier in a similar (but opposite) way, by having the hydrogen atom of the peptide bond very close to Asp25A. Gly27A and Thr24B establish a hydrogen bond with the Asp25A carboxylate. These five residues are extremely close to the active center and have a very marked impact on the activation energies. To be more confident on the magnitude of these values, it would be desirable to use a model where these residues are included in the high layer of the ONIOM calculations, but there is no reason to think that such approach would change the conclusions in a meaningful way. Moreover, if we wanted to use a higher theoretical level we would have to decrease the sampling to compensate for the additional computational cost, and that trade does not seems to be advantageous in the context of this study.

We now extend the analysis to all the residues (neutral and charged) that affect the barrier by more than 0.5 kcal mol^{-1} . In figure 6.7, these residues are represented in sticks, while the rest of the protein is represented in ribbons. All the information on this figure is coherent with the interpretation we have been outlining. Neutral residues that affect the barrier significantly (in yellow) form a tight core around the active center (The contributions of all neutral residues are included in the SI). Positive residues that are closer to the peptide bond (dark blue) stabilize

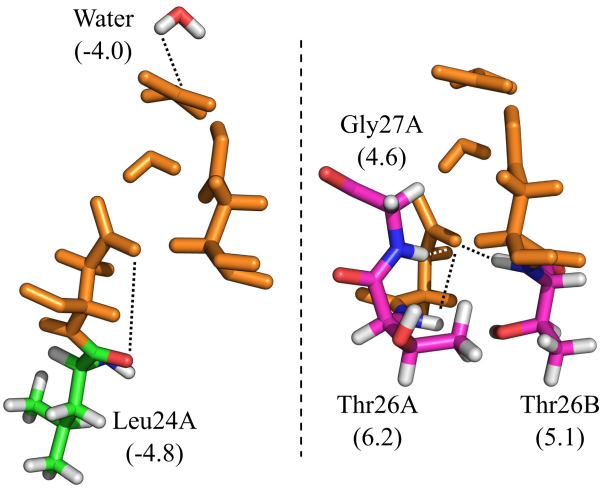


Figure 6.6: Neutral residues that have the most impact on the activation energies. Residues with a negative value of contribution lower the barrier, and are represented with green carbons at the left. Residues with a positive value of contribution increase the barrier and are represented with purple carbons at the right. The catalytic aspartates, the nucleophile and the substrate are colored in orange. Energies are in kcal mol^{-1} .

the barrier, while positive residues that are closer to Asp25A (light blue) increase the activation energy. For negative residues the opposite holds: residues near Asp25A (in red) are those that decrease the activation energy, while residues that increase it are near the peptide bond (in pink).

In figure 6.8, this same information is shown quantitatively. The contribution of the residues to the barrier is plotted against the difference between the distance of the residue to Asp25A and the distance of the residue to the peptide bond. Again, a positive energy value means that the residue increases the barrier by that amount, while a negative value means the residues stabilizes the barrier. Once more, the pattern is consistent with our previous results: positive residues populate the lower left and upper right quadrants; and negative residues populate the upper left and lower right quadrants. Furthermore, the greater the difference between the distances, the greater is the influence of the charged residues.

After establishing how and by how much the protein amino acids influence catalysis, we are ready to tackle the original question: Do fluctuations in the protein structure justify the dispersion observed in the catalytic barriers? Figure 6.9 is a plot of the standard deviation of the residues contribution to catalysis against the standard deviation of their distance to the active center. The figure tells us that certain residues move significantly from state to state, and that this movement affects the activation barrier significantly, especially if the residues are charged, or near the active center.

This result is more than enough to justify the dispersion of the activation energies. Considering that the contribution of each residue follows a normal distribution and the movement of the residues is not correlated, we can calculate the expected overall dispersion of results to be 10.2 kcal mol^{-1} , by the quadratic sum of the individual SDs. This value is the maximum deviation obtainable for the case where there is no correlation among residues (note that large-scale corre-

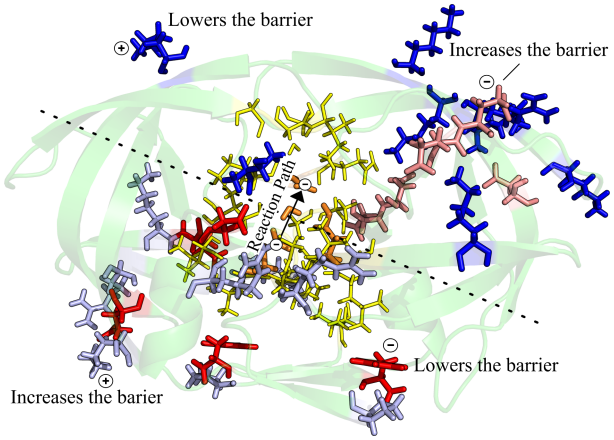


Figure 6.7: Representation of the protease enzyme with the residues that affect more the activation energy (> 0.5 kcal mol^{-1} . The catalytic aspartates, the nucleophile and the portion of the substrate in the high layer are colored in orange. Neutral residues are colored in yellow. Positively charged residues that decrease the activation energy are colored in dark blue, and positively charged residues that increase the barrier are colored in light blue. Negatively charged residues that decrease the activation energy are colored in red, and negatively charged residues that increase the barrier are colored in pink. The arrow represents the redistribution of negative charge from the reactants to the transition state.

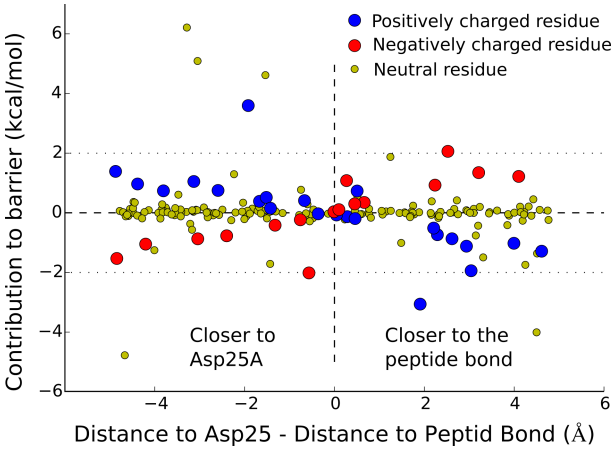


Figure 6.8: Averaged contribution to the barrier for each residue plotted against the difference of the distance between the residue and Asp25, and the distance between the residue and the peptide bond to be cleaved.

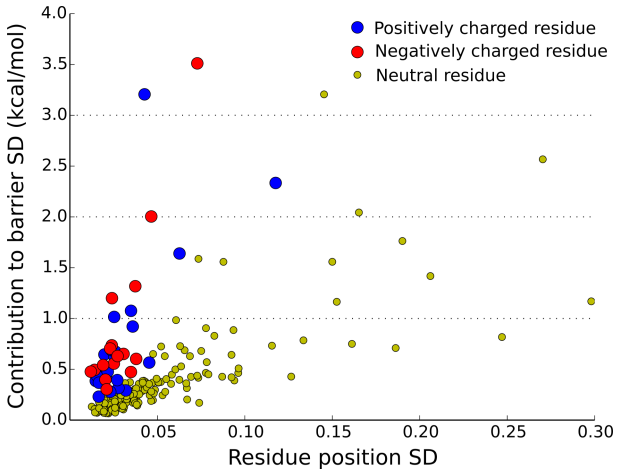


Figure 6.9: Standard deviation of the contribution of the residues to the barrier plotted against the standard deviation of the relative residue position. This last value is calculated as ((distance to Asp25A - distance to peptide bond)/average distance to active center).

lated movements can make these fluctuations wider, not smaller. Such backbone fluctuations and their effects over the turnover have been discussed before [217]. However, at the timescale studied here they are not present). The actual value for the protease system is smaller, 6.6 kcal mol^{-1} (for mechanism A.1), due to the existence of correlation between residues. Positive and negative residues interacting with their side chains, for example, will move in tandem, and will cancel each other in terms of contribution to the activation energy. This result gives us enough confidence to assert that enzyme structure fluctuations are responsible for the activation fluctuations.

6.5 Conclusions

In this study we aimed at a better understanding of the effect of conformational fluctuations in the activation free energy of enzymatic reactions by using a computational model of HIV-1 Protease. We took 40 equally time spaced snapshots from 80 ns of MD simulations, and used these structures as initial points for subsequent QM/MM studies of the reaction path first step. The results show that, along time, the enzyme goes through many reactant states that are associated with different activation energies (instantaneous disorder), from 14.5 kcal mol^{-1} up to 51.3 kcal mol^{-1}. The results point to a disordered energetic landscape where the barrier associated with each microstate varies several orders of magnitude in very short periods of time (ns). The overall apparent barrier calculated from 39 productive structures is 16.5 kcal mol^{-1} , which is in very good agreement with the experimental barrier of 15.9 kcal mol^{-1} for product release. In the second part of the work, we tried to understand the reasons behind such diversity in the activation energies of different microstates. We identified three main causes. The first is the existence of a different mechanism, where the role of Asp25A is diminished. The second is related to different conformation of the active center, in particular, the orientation of a single proton and two key distances along the reaction path. The proton in question is bonded to the nitrogen atom of the peptide bond to be broken.

In the main conformation it is orientated towards the nucleophile and helps in the stabilization of the transition state. On the other conformation, it is pointing to the opposite side. With respect to relation to the two interatomic distances, we found that when the distances at the reactants are closer to the transition state conformation, the associated barriers are smaller. The third reason for the large span in the activation barriers is explained by variations in the electrostatic environment of the active site, due to distinct structural conformations of the rest of the enzyme. The contribution of the residues to the instantaneous barrier is related to their influence on the movement of electronic charge from the reactants to the transition state. A negative charge near Asp25A, for example, pushes the negative charge away from Asp25A, destabilizing the reactants state and stabilizing the transition state. The contribution of each residue also fluctuates along time, according to its position relative to the active center. The fluctuations of all residues taken together are enough to justify the overall dispersion observed in the activation energies.

We think our results are of significance to a better understanding of enzymatic catalysis in general. We show very clearly that the conformational and catalytic landscape of enzymes is very heterogeneous, and that this heterogeneity can be traced back not only to different active site conformations but also to fluctuating interatomic interactions with the rest of the enzyme. Most importantly, we have shown that the fluctuations in the barriers are fundamental for the enzymatic rate constant, as most of the products will be formed by a very few transient enzyme conformations that provide very low barriers.

Chapter 7

Water controls reactivity in alpha-amylase on a sub-nanosecond timescale

Keywords: Enzymatic Catalysis, Near Attack Conformation, Transition State Stabilization, Glucosidase, Carbohydrates, Diabetes

7.1 Abstract

The subset of catalytically competent conformations can be significantly small in comparison with the full conformational landscape of enzyme-substrate complexes. In some enzymes, the probabilty of finding a reactive conformation can increase the activation barrier by at least 4 kcal/mol, even when the substrate remains tightly bound. In this study, we sampled conformations of alpha-amylase with bound substrate in a MD simulation of over 100 ns, and calculated energy profiles along the reaction coordinate. We found that reactive states require a hydrogen bond between a water molecule and E233, which is the general acid in the glycolysis mechanism. The effect of this single, non-reactive, intermolecular interaction is as much important as the correct positioning and orientation of the reacting residues to achieve a competent energy barrier. This hydrogen bond increases the acidity of E233, facilitating proton transfer to the glycosidic oxygen. In the MD simulation, this required hydrogen bond was observed in about half of the microstates, indicating that alpha-amylase is efficient at maintaining this important interaction in the reactants state. Importantly, this hydrogen bond formed and vanished on a sub-nanosecond time scale, suggesting that the instantaneous activation barrier oscillates at a much smaller timescale than turnover rate, from 11.2 kcal/mol to 31.2 kcal/mol. Interactions between the reacting groups, specifically the nucleophile D196 to the scissile carbon in the glycosidic bond, also changes at this timescale. Our results support the view of kinetics being determined by few low energy barriers.

7.2 Introduction

Chemical reactions can only occur if the reacting atoms are in close proximity and in a suitable orientation, allowing the necessary rearrangements of the electronic structure. Therefore, reaction rates are directly proportional to the fraction of reactant molecules in reactive conformations. The isomerization of chorismate to prephenate in water is a severe example of this effect because reactive conformers correspond to only 0.0001% of the conformational space [218, 219, 220], contributing 8.4 kcal/mol to the free energy of activation $\Delta G^{\ddagger}$ [221]. It is important to state that these 8.4 kcal/mol are not associated with changes in the electronic structure of chorismate, but instead with the probability of finding chorismate in a suitable geometry for the reaction to take place.

Based on extensive studies of the aforementioned reaction, the activation free energy $\Delta G^{\ddagger}$ was decomposed into chemical and nonchemical components [221, 219]:

$$\Delta G^{\ddagger} = \Delta G_{NAC} + \Delta G_{Chem} \tag{7.1}$$

where ΔG_{NAC} is the free energy difference between reactive substrate conformations (NAC stands for near attack conformation) and the full conformational space of the reactants (the ground state), and ΔG_{Chem} is the free energy of activation associated with the chemical transformation of a NAC to TS. The chemical component ΔG_{Chem} is primarily associated with the internal energy of the transition state and its stabilization by the surrounding environment — its value reflects the ammount of kinetic energy necessary to overcome the TS potential energy. On the other hand, the nonchemical component ΔG_{NAC} reflects the probability of finding the reagents in a reactive conformation, which is 8.4 kcal/mol for the isomerization of chorismate.

Please note that we do not refer to NAC as a catalytic effect, but simply as the subset of microstates that obey geometrical criteria for reactions to take place. Such geometrical criteria can be applied to enzyme-substrate complexes and to uncatalyzed reactions independently from each other. Since geometrical definitions are necessarily subjective, the accuracy of ΔG_{NAC} values must be interpreted with care.

In enzymes, ΔG_{NAC} may be surprisingly large in view of the conformational confinement of substrates inside active sites, which are expected to maintain required interactions for transition state stabilization (such as hydrogen bonds) consistently throughout the existence of a reactive enzyme-substrate complex. Molecular dynamics (MD) simulations estimate a ΔG_{NAC} up to 4.6 kcal/mol in HIV-1 protease [196] and a similar value of 4 kcal/mol was calculated for barnase [222]. By examining why an enzyme design failed, Ruscio et. al concluded that enzyme design needs to verify the NAC condition using a dynamical approach [223]. These studies indicate that enzyme-substrate complexes navigate a complex conformational landscape where critical interactions for catalysis have a remote chance of occurrence, at least in some enzymes. In computational studies this problem is easily solved as many authors who use protocols based in cluster models or QM/MM geometry optimizations at high theoretical levels already start from the NAC conformation, and easily get productive PES, providing accurate calculations of ΔG_{Chem}. However,

ΔG_{NAC} is mostly ignored, as these procedures miss the entropic contribution corresponding to the frequency of NACs among reactant microstates.

Besides active site organization, also the overall folding influences the activation free energy. In this regard, the strong dependence of reactivity on enzyme conformation is well documented by a large body of studies in which activation barriers were calculated for varying conformations of the same enzyme. These studies typically employ QM/MM methodologies to compute activation energies, reliably informing about the reactivity of each conformation, and avoiding geometrical criteria (NACs) which are unlikely to fully correlate with changes in electronic structure. A conformational change in ketosteroid isomerase raised the activation energy by around 20 kcal/mol [214]; variations up to 17 kcal/mol were found in P450 catalyzed reactions [215]; in fatty acid amide hydrolase the range of activation barriers was 11 kcal/mol [172]. Many other studies found significant variations of activation barriers for varying enzyme-substrate conformations [169, 168, 170, 171, 173, 174, 175, 176, 224, 150, 225, 226, 227, 228, 208, 229]. In our previous study on HIV-1 protease, [150] interactions of the nucleophilic water in the active site and the alignment of charge transfer within reactive groups with the electrostatic potential generated by the whole enzyme explained most of the observed fluctuations in activation barriers. Interestingly, such fluctuations occur on the nanosecond timescale, while the turnover rate is on the second time scale. This means that even enzymes without dynamic disorder, [230, 231] which appear to have a constant rate throughout many cycles, experience 'instantaneous disorder'. Instantaneous disorder are the fluctuations of activation energy on a timescale orders of magnitude faster than turnover rate. In view of the fact that chemical reactions occur on a fs timescale [232, 233], chemical reactions probably take place at specific conformations with low activation barriers. The rate constant depends on both the frequency of low activation barriers and their magnitude.

Here, we report MD and QM/MM studies on human pancreatic α-amylase (HPA), focusing on the relationship between activation energies $\Delta E^{\ddagger}$ and interactions occurring in the active site. This enzyme is a good case study because the glycosylation mechanism is well established [234] and is also relevant to a very large class of enzymes: glucosidases, implied in a variety of diseases [235]. Our results highlight the importance of solvent water in determining reactivity in HPA. The difference between this case and the ones previously discussed is that here we study the effect of a non-protein, very labile, hydrogen bond with a non-reactive water molecule, that turned out to be highly important for the observed reaction rate.

7.3 Methods

Our work consisted of two main stages: (i) conformational sampling of the reactants state of human alpha-amylase on over 100ns of molecular dynamics (MD) simulation and (ii) calculation of activation energies in selected microstates from the MD trajectory.

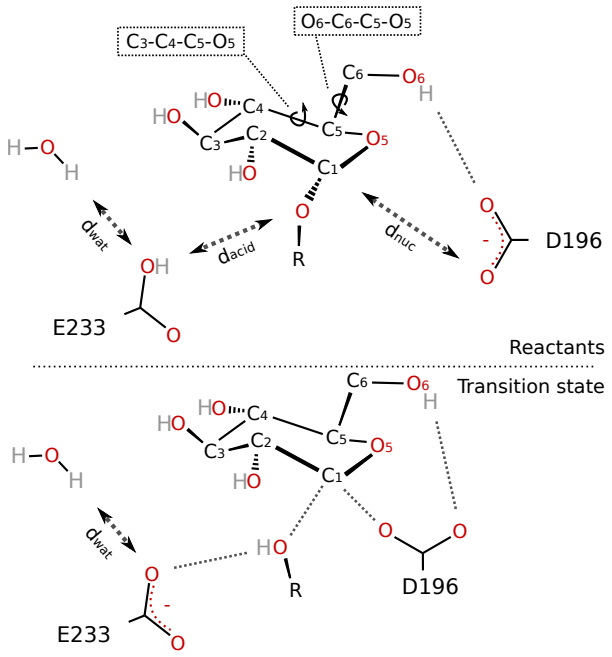


Figure 7.1: Reactants and transition state of glycolysis step. Important distances and dihedrals are defined: d_{wat} between a water hydrogen and the protonated oxygen of E233, d_{acid} between the acidic hydrogen of E233 and the glycosidic oxygen, d_{nuc} between the C_1 and a carboxylate oxygen of D196, and the dihedral angles $\theta_{C_3C_4C_5O_5}$ and $\theta_{O_6C_6C_5O_5}$.

7.3.1 Molecular Dynamics

7.3.1.1 System details

Human pancreatic α-amylase (HPA) was simulated by molecular dynamics (MD) with substrate maltopentaose (G5). Initial coordinates were retrieved from PDB[236] structure '1cpu', which was co-crystallized with an acarbose-like inhibitor with five monosaccharide rings [237]. Modeling G5 from this inhibitor involved two simple modifications: changing the N-glycosidic bond to a O-glycosidic bond and replacing a C=C double bond by a C-O single bond in the ring adjacent to N-glycosidic bond of the inhibitor. The binding mode of G5 was such that the catalytic machinery of HPA was positioned to cleave the glycosidic bond between the third and fourth glucose units, producing maltotriose (G3) and maltose (G2). In the X-ray structure, N461 was glycosylated — we removed the carbohydrate. All titratable groups were simulated in their standard protonation states (pH 7) except E233 which was simulated in the neutral state. E233 works as the general acid and donates a proton to the scissile glycosidic bond [234]. The hydrogen bond corresponding to d_{acid} in figure 7.1 was restrained at 2.0 Å with a harmonic potential with a force constant of 50 kcal mol^{-1} Å^{-2} to increase the sampling of catalytically competent conformations. The system was solvated in a periodic box filled with TIP3P water molecules such that at least 12 Å exist between the protein or G5. Sodium counterions were added to neutralize the charge of the system.

7.3.1.2 Simulation

Simulations were carried out using AMBER 12 [238]. The ff99SB forcefield [239] was used for HPA, the TIP3P model [60] for water molecules and GLYCAM_06h [240] parameters for G5. The *leap* program in Antechamber [241] was used to assign parameters. Particle Mesh Ewald (PME) [242, 198, 243] was used for electrostatics, with the real part truncated at 10 Å. The integration time step was 2fs, using the Shake algorithm [199]. Langevin dynamics were used with a collision frequency of 1ps, using the NPT ensemble at 1 bar and a pressure relaxation time of 2ps, at 300K. Microstates were recorded every 100 ps. Total simulation time for this production run was 109 ns.

Waters and counterions were equilibrated in prior by a 10 ns NVT ensemble run, maintaining all protein and G5 atoms fixed at their crystallographic coordinates. Reported simulation times are zeroed at the start of the production run, thus excluding the 10 ns equilibration time.

7.3.2 Snapshot selection

We selected snapshots from the MD simulation that met what we predicted to be adequate criteria for reactivity, based on interatomic distances that most probably would lead to catalytically relevant activation barriers. After calculating the barriers we concluded that one of our criteria (number 3) was irrelevant. The four criteria we used are: (1) nucleophilic aspartate (closest oxygen) to scissile carbon under 3.5 Å, (2) structural hydrogen bonds between D300 and the hydroxyl groups attached to C2 and C3 in figure 7.1 under 2.5 Å and (3) the existence of a water molecule with a proton within 2.5 Åfrom the glycosidic oxygen. If we had not restrained our MD, we would have to add a fourth condition, which would be the distance of the acidic proton in E233 to the glycosidic oxygen under a given value (e.g. 2.5 Å). This restraing biases the free energy, by increasing the frequency of NACs. Since the bias is constant for all the sampled conformations, it does not affect the relative barrier fluctuations, only their absolute values.

Out of the 42 selected snapshots, only 18 were used in our energetic analysis. The remaining 24 were excluded due to difficulties in characterizing the stationary points (GS or TS), or in guaranteeing that they lie in the same global minimum with respect to all degrees of freedom orthogonal to the reaction coordinate.

7.3.3 ONIOM

Structures from selected snapshots were studied using the ONIOM approach [205]. The high layer was studied using B3LYP [49, 202] with the 6-31G(d) basis-set [203]. Atoms in the low layer were described by ff99SB [239]. The high layer included two glucose monomers — before and after the scissile O-glycosidic bond, the neutral E233, the nucleophile D196 and structural D300, and the solvent water closest to the glycosidic oxygen. Several different water molecules occupied this position, so we used cpptraj [30] and custom scripts to find which water is closest at each recorded snapshot. The 1000 closest waters to any protein atom were kept. Residues were frozen at 15Å from the high layer, limiting the degrees of freedom during geometry optimizations. All water molecules except the one closest to the glycosidic oxygen were frozen. Frozen waters within 3 Å

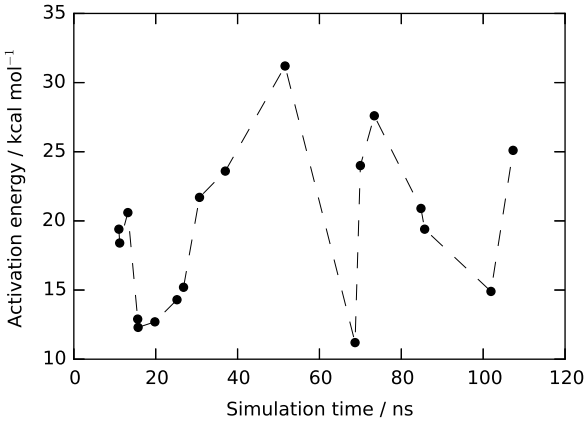


Figure 7.2: Activation energy for selected snapshots from MD simulation. The lowest activation barrier was found at the 68.7 ns mark and was 11.2 kcal/mol. The largest barrier of 31.2 kcal/mol corresponded to the snapshot recorded at 51.6 ns. There is a subtle tendency for structures closer in time to display similar activation barriers but large variations occurred at a nanosecond timescale. The dashed line provides visual guidance into the chronological order of snapshots.

from the high layer were removed in order to avoid artificial geometrical constraints, as we were not interested in studying the effect of local minima of the second solvation layer.

7.4 Results and Discussion

7.4.1 Energies and kinetics

We calculated the activation energy of the glycosylation step in human pancreatic α-amylase (HPA) using the adiabatic mapping approach, for 18 different conformations of the HPA-G5 complex sampled by MD simulation. We obtained activation energies $\Delta E^{\ddagger}$ ranging from 11.2 kcal/mol to 31.2 kcal/mol. Activation energies are reported in figure 7.2 and in table 7.1. Other studies on enzymatic catalysis have also found a wide range of activation barriers [228, 150, 208].

Importantly, we observe that potential energy barriers can change many orders of magnitude faster than the turnover rate; they change at the ns timescale. At 13.2 ns the barrier was 20.6 kcal/mol and 2.4 ns later, at the 15.6 mark, the barrier dropped to 12.9 kcal/mol. At the nanosecond timescale, where barrier fluctuations occur, conformational changes are mostly associated with thermal vibrations of bonds, angles or rotamers around a very well defined folding, or small movements of water molecules. There are no significant folding changes at this timescale. So, in most of the previously studied cases the enzyme responds to minimal thermal vibrations with enormous fluctuations in the reaction rate, due to its sheer number of interacting atoms. As a consequence, observed kinetics are a consequence of few activation barriers — occurring at specific conformations with critical interactions for catalysis — and the frequency at which such low barrier conformations are visited, i.e. the partition function of reactive conformers which can, in principle, be approximated by geometric definitions such as NACs.

Table 7.1: Activation energies $\Delta E^{\ddagger}$ and relevant distances and dihedrals.

time / ns	$\Delta E^{\ddagger}$ / kcal mol^{-1}	d^{R}_{wat} / Å	d^{R}_{acid} / Å	d^{R}_{nuc} / Å	d^{TS}_{wat} / Å	$\theta_{C_3C_4C_5O_5}$ / deg	$\theta_{O_6C_6C_5O_5}$ / deg
11.0	19.4	3.68	2.95	3.44	2.36	-22.9	-67.6
11.2	18.4	3.03	2.95	3.43	2.11	-26.3	-64.6
13.2	20.6	3.75	3.00	3.30	2.58	-17.6	-65.9
15.6	12.9	2.30	1.94	3.21	1.98	-16.3	-71.9
15.7	12.3	2.64	1.95	3.17	2.14	-23.6	-70.8
19.8	12.7	2.36	2.76	3.29	2.01	-22.4	-65.5
25.2	14.3	2.91	2.74	3.41	2.09	7.8	46.8
26.8	15.2	2.90	2.54	3.57	2.08	29.0	45.9
30.7	21.7	2.52	2.67	3.70	2.07	27.5	41.0
37.0	23.6	4.40	2.83	4.10	2.17	30.6	37.4
51.6	31.2	3.29	2.79	3.44	2.10	26.6	41.7
68.7	11.2	2.92	1.89	3.24	2.08	-22.6	-75.4
70.0	24.0	4.42	2.60	3.85	4.34	31.6	35.4
73.4	27.6	3.53	2.63	3.67	2.13	27.4	32.0
84.8	20.9	3.83	2.54	3.53	3.61	19.6	41.4
85.7	19.4	2.89	2.48	3.47	2.14	27.0	41.0
101.9	14.9	2.63	1.98	3.14	2.09	-27.6	-69.2
107.3	25.1	4.44	3.15	3.51	2.11	-33.2	-60.5

7.4.2 Structural Analysis

In order to understand the structural reasons underlying the fluctuations in the activation barrier, we superimposed the structures of reactants (fig. 7.3) and transition states (fig. 7.4). It is important to note that each R/TS pair lies on the same global minimum as could be verified by visual inspection of the structures and energy profiles along the reaction coordinate.

Transition state structures display better superimposition than reactant structures, implying that the conformational landscape of transition states is better defined than that of the reagents. This observation suggests a negative activation entropy $\Delta S^{\ddagger}$ for the glycosylation step in α-amylase. The fact that TS structures are confined around a well defined conformation may be of great utility for the design of transition state analogues.

Low activation barriers only occurred if a set of geometric conditions were verified: short distances for d_{wat}, d_{acid} and d_{nuc} (see figure 7.1). Two of these distances were expected to be important as they are implicated in the reaction coordinate — d_{nuc} corresponds to the nucleophile attack of D196 to the carbon in the scissile glycosidic bond (C_1 in fig. 7.1) and d_{acid} corresponds to the proton transfer from E233 to the glycosidic oxygen. The importance of the third distance d_{wat} was surprising because it is not directly involved in the reaction coordinate, it corresponds to a hydrogen bond between a solvent water and E233. This hydrogen bond stabilizes the TS because E233 becomes negatively charged after donating its acidic proton to the glycosidic oxygen. In figure 7.3, the dependence of $\Delta E^{\ddagger}$ on these distances is displayed (superscripts ^{R} and ^{TS} indicate if the distances were calculated at reactants or transition states, respectively). $\Delta E^{\ddagger}$ depends heavily d^{TS}_{wat} as most TS structures displayed this required hydrogen bond, and the few that didn't had a $\Delta E^{\ddagger}$ of about 20 kcal/mol, well above the lower values around 12 kcal/mol (see panel B in figure 7.5).

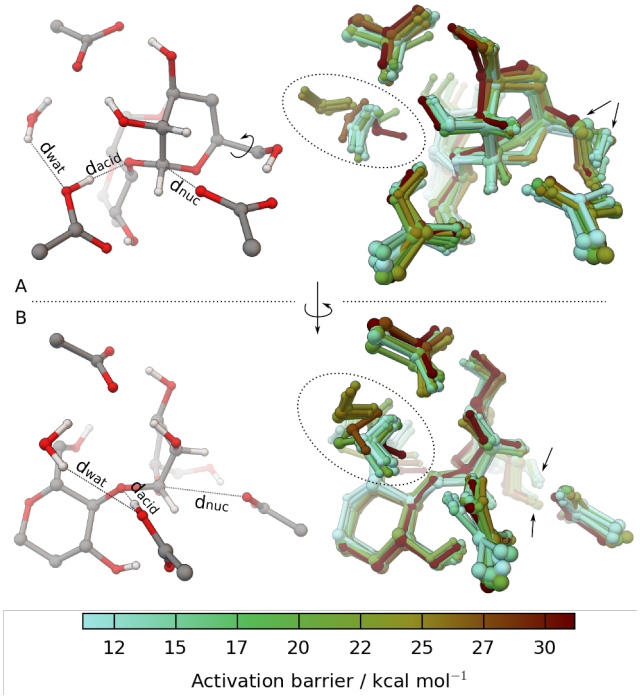


Figure 7.3: Reactant structures at the B3LYP/6-31g(d):ff99SB level of theory. Panels A and B represent the same structures rotated by about 60°. In each panel, a single structure is represented along with the superimposed structures (for visual guidance). The single water molecule can adopt a variety of interactions as is highlighted by the dashed ellipse. The dihedral angle of the hydrogen bond with D196 can also adopt one of two positions (highlighted with arrows) depending on whether the adjacent monosaccharide unit is in the boat or chair conformation. Important distances d_{wat}, d_{acid} and d_{nuc} are represented with dashes. Overall, reactant structures do not align as well as transition state structures (see figure 7.4).

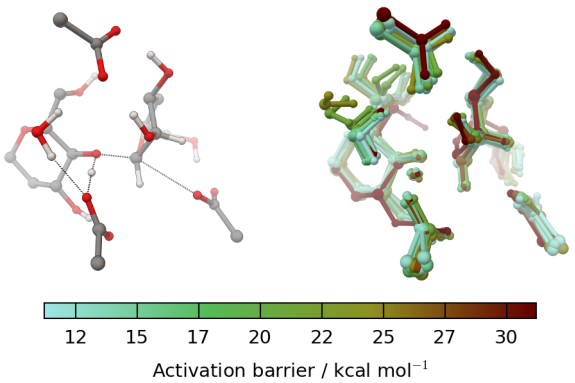


Figure 7.4: Transition state structures at the B3LYP/6-31g(d):ff99SB level of theory. A single structure is represented along with the superimposed structures for visual guidance. Important distances d_{wat}, d_{acid} and d_{nuc} are represented with dashes. Transition state structures display better alignment than reactant structures (see figure 7.3).

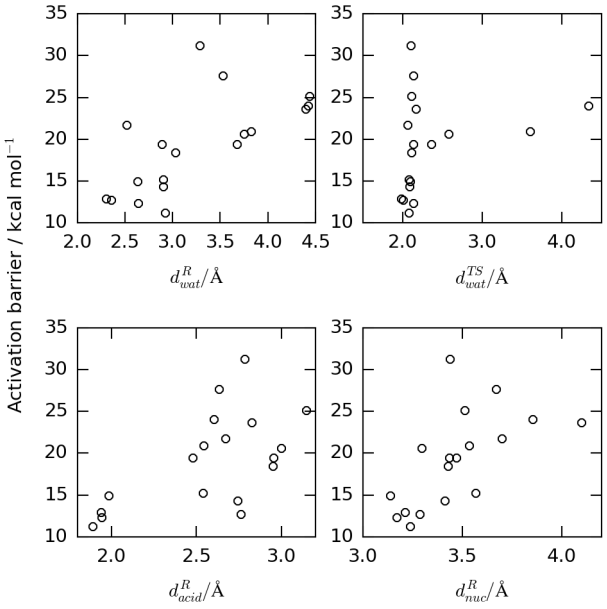


Figure 7.5: Correlation between distances and activation barriers.

Surprisingly, this hydrogen bond was equally critical in the reactants state (panel A in fig. 7.5). It appears that there are many hydrogen bond acceptors in the vicinity of E233, and many of them are stronger acceptors than the neutral E233. Thus, if the solvent water establishes a hydrogen bond with these better acceptors, the energy of the reagents will be lower (without affecting the TS energy), effectively increasing the energy difference between R and TS — that is $\Delta E^{\ddagger}$. The dependence of $\Delta E^{\ddagger}$ on d^R_{wat} is evident in figures 7.4 and 7.5, indicating that this interaction must be sustained most of the time for HPA to be efficient, and should be included in NAC definitions for this enzyme.

It is interesting to note that a hydrogen bond with a water molecule has an effect on $\Delta E^{\ddagger}$ that is of similar magnitude than the effect of the distance between the nucleophile D196 and C_1. The shorter the distance d^R_{nuc}, the lower the activation barrier (see figure 7.5).

The dependence of ONIOM activation barriers on d_{wat}, d_{nuc} and d_{acid} revealed that these distances must be below a certain threshold for the reaction to occur, otherwise the activation barrier is too large. Thus, we can significantly improve our definition of what constitutes a reactive conformation and estimate the frequency of NACs during the MD simulation with improved accuracy. Since the distance d_{acid} was restrained in our simulation, all sampled microstates will display this interaction. In figure 7.6 we represent each recorded microstate of the MD simulation (small circles) as a function of d_{nuc} and d_{wat}. There is a significantly high number of frames displaying short distances for both interactions. The two interactions appear to be independent from each other. The values of these distances for the reactant state of ONIOM calculations are also represented (the color indicates the activation barrier). It seems like HPA is efficient at sustaining these critical interactions.

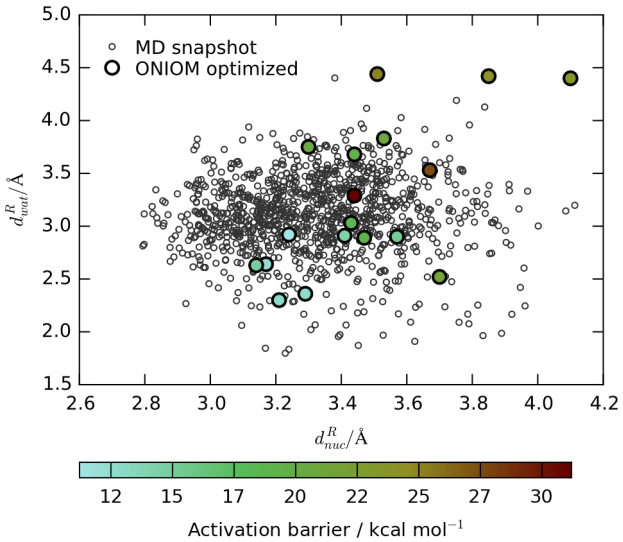


Figure 7.6: Overlay of MD distances and ONIOM barriers.

Retrospectively, the third criteria we used for selecting snapshots from the MD simulation significantly reduced the number of selected snapshots. The water molecule spends much more time interacting with the neutral E233 than with the glycosidic oxygen. This directly translates in a much higher number of reactive microstates, well beyond the 42 snapshots chosen due to incorrect criteria. Importantly, the crystallographic structure '1cpu' displays this interaction at 2.9 Å distance, supporting the results we found in this study about the importance of this hydrogen bond for catalysis. Accordingly, the ONIOM activation energy for the X-ray structure was close to our lower limit of 11.2 kcal/mol.

7.5 Conclusions

We identified important interactions in HPA that are associated with catalytic proficiency. Conformations of the HPA-G5 complex that display these important interactions display low activation barriers using the adiabatic mapping approach at the B3LYP/6-31g(D):ff99SB level of theory. This data may be useful to devise accurate definitions of NACs, which in turn can be used to estimate reactivity from MD simulations alone, provided that a reasonable value of ΔG_{Chem} (the energy from a reactive reactant state to the TS) has already been computed by a suitable level of theory.

The importance of a hydrogen bond between a solvent water molecule and E233 was surprising. This hydrogen lowers activation barriers by stabilizing the proton transfer from E233 to the glycosidic oxygen in the scissile bond. The volatility of this interaction (and also d^R_{nuc} as observed in our MD trajectory) suggest that activation barriers fluctuate rapidly on the nanosecond timescale or faster.

Chapter 8

Improving AutoDock4 for Glucosidases

8.1 Overview

This chapter summarizes the facts that suggest desolvation as a key AutoDock descriptor to model glucosidase inhibitors, and our main docking studies regarding a dataset of glucosidase inhibitors we assembled.

8.2 Dataset of glucosidase-inhibitor complexes

In order to build a dataset of glucosidase-ligand complexes we searched the Binding MOAD database [91] for the Enzyme Commission number EC 3.2.1.* (glucosidases). The resulting entries were crossed with the PDBBind dataset [244] to ensure the accuracy of the data (some discrepancies in inhibition constants were found during this step, and manual inspection of the original publications was required to retrieve the correct affinity values). The following requirements were applied: existence of experimentally determined inhibition (K_i) or dissociation (K_d) constants; no alternate conformation or missing atoms in the ligand; no alternate conformations in receptor residues within 5 Å from any ligand atom. We removed ligands that do not bind in an active site capable of performing glycosylation (due to missing carboxylate containing residues in appropriate orientation/positioning). We also removed cases in which ligands coordinated metal ions, or interacted extensively with the solvent. After applying the aforementioned filters our dataset had 105 complexes.

8.3 Performance of Autodock

8.3.1 Standard AutoDock4.2

We tested AutoDock4.2 using a standard approach. Receptors and ligands were prepared using openbabel [121] for atom-typing, charge assignment using Gasteiger-Marsili charges [97], and for defining protonation states. All crystallographic waters were removed from the receptors. The results, specially docking power, depended on the number of water molecules that were removed

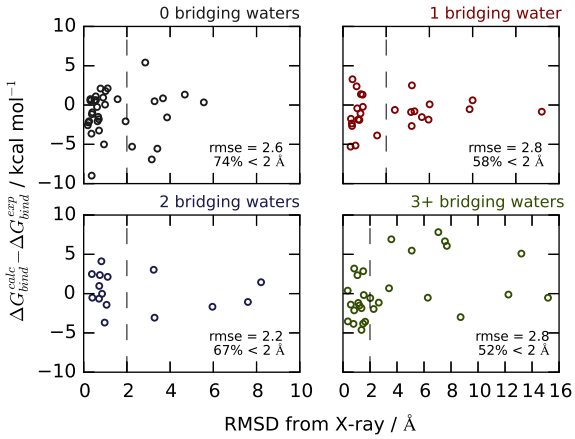


Figure 8.1: Docking power and scoring power of AutoDock4.2 in our glucosidase-ligand dataset. Each circle represents one protein-ligand complex as a function of RMSD from the corresponding X-ray structure (x-axis) and error in predicting ΔG_{bind} (y-axis). Results are organized in four plots according to the number of water molecules that mediate intermolecular interactions between the ligand and protein. Root mean squared errors (rmse) are reported for complexes that were docked within 2 Å, in kcal/mol.

from the receptor (figure 8.1). Docking power dropped from 74% of successful re-dockings (< 2 Å) when no bridging waters exist, to 52% with three or more waters.

Unlike in ref. [5] and in chapter 5, poses were not generated beforehand, as the search implemented in AutoDock4.2 (genetic algorithm) was used. Therefore, docking power also reflects the performance of the underlying search algorithm.

Based on the results shown in figure 8.1, we considered that water molecules in the protein-ligand interface may negatively impact the discovery of novel glucosidase inhibitors.

8.3.2 Re-calibrated AutoDock4.2

We intended to improve AutoDock using an approach similar to that reported in chapter 3. This approach consists in taking the ligand x-ray structure, performing a local geometry optimization and computing the contributions of individual AutoDock terms (van der Waals, hydrogen bonds, desolvation, electrostatics and torsional entropy of the ligand). This is performed by building a linear regression model, where the free energy of binding for each protein-ligand complex is calculated as the sum of each term multiplied by its weight. The least squares approach was used to minimize the difference between calculated and experimental ΔG_{bind}. We abandoned this approach because three out of five terms displayed poor statistics: desolvation, electrostatics and torsional entropy, suggesting a problem with the model, i.e. the scoring function should be improved.

We tested the influence of re-calibrating AutoDock weights using poses from independent docking runs (typically 10 to 100 runs are performed), where the free energy of binding ΔG_{bind} is calculated as an exponentially weighted average of the free energy of binding of each individual

docking run:

$$\Delta G_{bind} = -\frac{1}{A} ln \left(\frac{1}{N} \sum_{i=1}^{N} e^{-A\ \Delta G^{i}_{bind}} \right) \tag{8.1}$$

where A is a constant (we set A = 0.64) and ΔG^{i}_{bind} is the binding free energy of the pose from the i^{th} docking run. In total, we performed N = 100 docking runs for each protein-ligand complex. The number of degrees of freedom with this approach increases by 1 because of the A factor. The idea is to use of the number of times the search algorithm samples the same binding pose (or binding poses of similar energy) as a proxy to its entropy. This method resulted in a marginal gain in scoring power, around 0.1 kcal/mol improvement over a single pose re-calibration. We considered these results inconclusive.

8.3.3 'Wet' docking

Since re-calibration of the five AutoDock4.2 terms seemed like an inappropriate approach (due to poor statistics in the linear regression model), we turned our attentions to the problem of bridging water molecules. We used the method from Forli et. al [44] which is able to predict the presence of bridging water molecules during the docking search. It works by 'decorating' the ligand with pseudo-waters at 3 Å distance from any hydrogen bond acceptor/donor. For this reason it is called 'wet' docking. Waters can be kept if they interact favourably with the receptor, or displaced if they collide with the protein. We find this approach more attractive than keeping X-ray waters fixed in the receptor, because different ligands often induce different positions and orientations of bridging water molecules, preventing the discovery of new inhibitors that would require a different disposition of water molecules. Unfortunately this method did not increase docking power (see figure 8.2) as the number of successful re-dockings decreased slightly: 12 complexes displayed a RMSD under 2 Å with the standard approach and over 2 Å with the 'wet' approach. On the other hand, only 6 complexes were improved using the 'wet' approach.

8.4 Hydration of Carbohydrates

Carbohydrates have a large number of hydroxyl groups. Each hydroxyl group has one hydrogen bond donor site and two hydrogen bond acceptor sites. Consequently, carbohydrates can make a large number of hydrogen bonds with solvent water molecules (figure 8.3), and display more negative ΔG^{solv}_{water} than other molecules of similar size.

In order to estimate if our method for calculating free energies of solvation (chapter 4) is reliable for carbohydrates, we re-fitted the atomic solvation parameters without carbohydrates, and evaluated the performance of the method on carbohydrates (figure 8.4). The results are satisfactory.

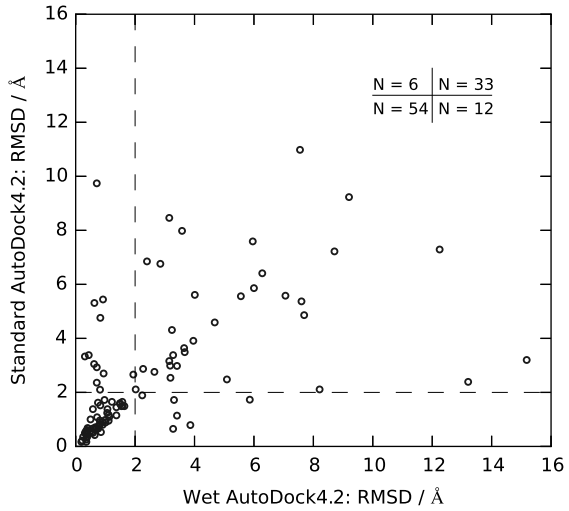


Figure 8.2: Docking power of wet and standard AutoDock4.2. Numbers in the top right corner indicate the number of complexes that fall within each quadrant delimited by dashed lines.

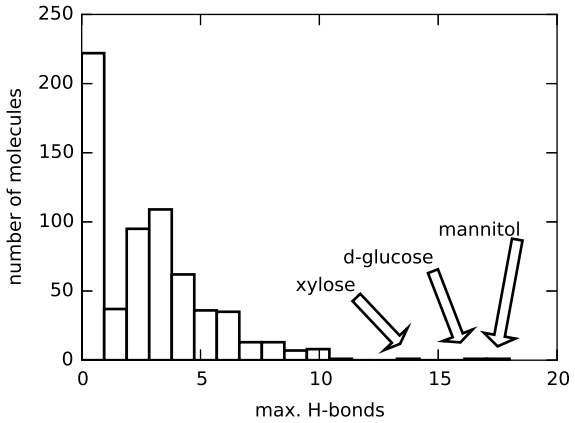


Figure 8.3: Number of hydrogen bond acceptors/donors in the FreeSolv-0.32 database.

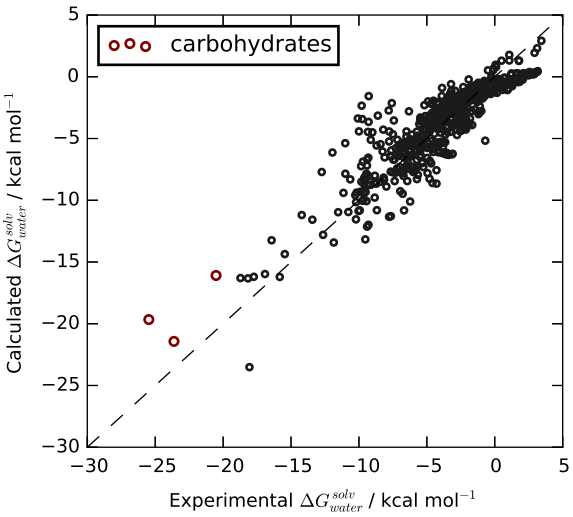


Figure 8.4: Validation of the solvation approach described in chapter 4 on carbohydrates. The three carbohydrates (xylose, d-glucose and mannitol) were excluded from the training set, i.e. atomic solvation parameters were determined exclusively on other molecules.

8.5 Further Evidence and Outlook

Finally, we'd like to refer to studies where AutoDock3 performs better than AutoDock4 for carbohydrate binding [245, 246]. One of the differences between AutoDock4 and the preceding version is a different desolvation model, supporting the development of an improved desolvation model to achieve a better description of glucosidase inhibitors.

We haven't yet integrated our new approach for solvation into molecular docking because the spatial confinement of water molecules inside binding pockets is extremely challenging from the energetic viewpoint (see section 1.1.1.2), and we have no reliable means to model such effect yet. Desolvation terms would benefit from a better description of water under confinement. This is certainly a hot topic for further developments of scoring functions.

References

[1] JianMin Jia, XiaoLi Xu, Fang Liu, XiaoKe Guo, MingYe Zhang, MengChen Lu, LiLi Xu, JinLian Wei, Jia Zhu, ShengLie Zhang, et al. Identification, design and bio-evaluation of novel hsp90 inhibitors by ligand-based virtual screening. *PLoS One*, 8(4):e59315, 2013.

[2] Hao-Peng Sun, Jian-Min Jia, Fen Jiang, Xiao-Li Xu, Fang Liu, Xiao-Ke Guo, Bahidja Cherfaoui, Hao-Ze Huang, Yang Pan, and Qi-Dong You. Identification and optimization of novel hsp90 inhibitors with tetrahydropyrido [4, 3-d] pyrimidines core through shape-based screening. *European Journal of Medicinal Chemistry*, 79:399–412, 2014.

[3] Joffrey Gabel, Jérémy Desaphy, and Didier Rognan. Beware of machine learning-based scoring functions: On the danger of developing black boxes. *Journal of Chemical Information and Modeling*, 54(10):2807–2815, 2014.

[4] Caitlin C Bannan, Kalistyn H Burley, Michael Chiu, Michael R Shirts, Michael K Gilson, and David L Mobley. Blind prediction of cyclohexane–water distribution coefficients from the sampl5 challenge. *Journal of Computer-Aided Molecular Design*, 30(11):927–944, 2016.

[5] Yan Li, Li Han, Zhihai Liu, and Renxiao Wang. Comparative assessment of scoring functions on an updated benchmark: 2. evaluation methods and general results. *Journal of Chemical Information and Modeling*, 54(6):1717–1736, 2014.

[6] Christopher J Fennell, Charlie Kehoe, and Ken A Dill. Oil/water transfer is partly driven by molecular shape, not just size. *Journal of the American Chemical Society*, 132(1):234–240, 2009.

[7] FV Grigoriev, MV Basilevsky, SN Gabin, AN Romanov, and VB Sulimov. Cavitation free energy for organic molecules having various sizes and shapes. *The Journal of Physical Chemistry B*, 111(49):13748–13755, 2007.

[8] David Chandler. Interfaces and the driving force of hydrophobic assembly. *Nature*, 437(7059):640–647, 2005.

[9] P Buchanan, N Aldiwan, AK Soper, JL Creek, and CA Koh. Decreased structure on dissolving methane in water. *Chemical Physics Letters*, 415(1):89–93, 2005.

[10] HJ Rezus, YLANand Bakker. Observation of immobilized water molecules around hydrophobic groups. *Physical Review Letters*, 99(14):148301, 2007.

[11] A Wallqvist and BJ Berne. Molecular dynamics study of the dependence of water solvation free energy on solute curvature and surface area. *The Journal of Physical Chemistry*, 99(9):2885–2892, 1995.

[12] David L Mobley, Christopher I Bayly, Matthew D Cooper, Michael R Shirts, and Ken A Dill. Small molecule hydration free energies in explicit solvent: an extensive test of fixed-charge atomistic simulations. *Journal of Chemical Theory and Computation*, 5(2):350–358, 2009.

[13] Duc D Nguyen and Guo-Wei Wei. The impact of surface area, volume, curvature, and lennard–jones potential to solvation modeling. *Journal of Computational Chemistry*, 38(1):24–36, 2017.

[14] Aleksandr V Marenich, Christopher J Cramer, and Donald G Truhlar. Universal solvation model based on solute electron density and on a continuum model of the solvent defined by the bulk dielectric constant and atomic surface tensions. *The Journal of Physical Chemistry B*, 113(18):6378–6396, 2009.

[15] FV Grigoriev, SN Gabin, AN Romanov, and VB Sulimov. Computation of the contribution from the cavity effect to protein- ligand binding free energy. *The Journal of Physical Chemistry B*, 112(48):15355–15360, 2008.

[16] Richard A Friesner, Robert B Murphy, Matthew P Repasky, Leah L Frye, Jeremy R Greenwood, Thomas A Halgren, Paul C Sanschagrin, and Daniel T Mainz. Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *Journal of Medicinal Chemistry*, 49(21):6177–6196, 2006.

[17] Thomas A Halgren, Robert B Murphy, Richard A Friesner, Hege S Beard, Leah L Frye, W Thomas Pollard, and Jay L Banks. Glide: a new approach for rapid, accurate docking and scoring. 2. enrichment factors in database screening. *Journal of Medicinal Chemistry*, 47(7):1750–1759, 2004.

[18] Richard A Friesner, Jay L Banks, Robert B Murphy, Thomas A Halgren, Jasna J Klicic, Daniel T Mainz, Matthew P Repasky, Eric H Knoll, Mee Shelley, Jason K Perry, et al. Glide: a new approach for rapid, accurate docking and scoring. 1. method and assessment of docking accuracy. *Journal of Medicinal Chemistry*, 47(7):1739–1749, 2004.

[19] Gareth Jones, Peter Willett, Robert C Glen, Andrew R Leach, and Robin Taylor. Development and validation of a genetic algorithm for flexible docking. *Journal of Molecular Biology*, 267(3):727–748, 1997.

[20] Oliver Korb, Thomas Stutzle, and Thomas E Exner. Empirical scoring functions for advanced protein- ligand docking with plants. *Journal of Chemical Information and Modeling*, 49(1):84–96, 2009.

[21] Matthew D Eldridge, Christopher W Murray, Timothy R Auton, Gaia V Paolini, and Roger 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(5):425–445, 1997.

[22] Carol A Baxter, Christopher W Murray, David E Clark, David R Westhead, and Matthew D Eldridge. Flexible docking using tabu search and an empirical estimate of binding affinity. *Proteins: Structure, Function, and Bioinformatics*, 33(3):367–382, 1998.

[23] Phillip W Snyder, Matthew R Lockett, Demetri T Moustakas, and George M Whitesides. Is it the shape of the cavity, or the shape of the water in the cavity? *The European Physical Journal Special Topics*, 223(5):853–891, 2014.

[24] Jayendran C Rasaiah, Shekhar Garde, and Gerhard Hummer. Water in nonpolar confinement: From nanotubes to proteins and beyond*. *Annual Review of Physical Chemistry*, 59:713–740, 2008.

[25] Michael Schauperl, Maren Podewitz, Birgit J Waldner, and Klaus R Liedl. Enthalpic and entropic contributions to hydrophobicity. *Journal of Chemical Theory and Computation*, 2016.

[26] Joachim Dzubiella. How interface geometry dictates water's thermodynamic signature in hydrophobic association. *Journal of Statistical Physics*, 145(2):227–239, 2011.

[27] Phillip W Snyder, Jasmin Mecinović, Demetri T Moustakas, Samuel W Thomas, Michael Harder, Eric T Mack, Matthew R Lockett, Annie Héroux, Woody Sherman, and George M Whitesides. Mechanism of the hydrophobic effect in the biomolecular recognition of arylsulfonamides by carbonic anhydrase. *Proceedings of the National Academy of Sciences*, 108(44):17889–17894, 2011.

[28] Tom Young, Robert Abel, Byungchan Kim, Bruce J Berne, and Richard A Friesner. Motifs for molecular recognition exploiting hydrophobic enclosure in protein–ligand binding. *Proceedings of the National Academy of Sciences*, 104(3):808–813, 2007.

[29] Steven Ramsey, Crystal Nguyen, Romelia Salomon-Ferrer, Ross C Walker, Michael K Gilson, and Tom Kurtzman. Solvation thermodynamic mapping of molecular surfaces in ambertools: Gist. *Journal of Computational Chemistry*, 37(21):2029–2037, 2016.

[30] Daniel R Roe and Thomas E Cheatham III. Ptraj and cpptraj: software for processing and analysis of molecular dynamics trajectory data. *Journal of Chemical Theory and Computation*, 9(7):3084–3095, 2013.

[31] Alexander S Bayden, Demetri T Moustakas, Diane Joseph-McCarthy, and Michelle L Lamb. Evaluating free energies of binding and conservation of crystallographic waters using szmap. *Journal of Chemical Information and Modeling*, 55(8):1552–1565, 2015.

[32] Daniel Robinson, Thomas Bertrand, Jean-Christophe Carry, Frank Halley, Andreas Karlsson, Magali Mathieu, Hervé Minoux, Marc-Antoine Perrin, Benoit Robert, Laurent Schio, et al. Differential water thermodynamics determine pi3k-beta/delta selectivity for solvent-exposed ligand modifications. *Journal of Chemical Information and Modeling*, 56(5):886–894, 2016.

[33] Matthew D Kelly and Ricardo L Mancera. A new method for estimating the importance of hydrophobic groups in the binding site of a protein. *Journal of Medicinal Chemistry*, 48(4):1069–1078, 2005.

[34] Yang Cao and Lei Li. Improved protein–ligand binding affinity prediction by using a curvature-dependent surface-area model. *Bioinformatics*, 30(12):1674–1680, 2014.

[35] Sergio Ruiz-Carmona, Daniel Alvarez-Garcia, Nicolas Foloppe, A Beatriz Garmendia-Doval, Szilveszter Juhos, Peter Schmidtke, Xavier Barril, Roderick E Hubbard, and S David Morley. rdock: a fast, versatile and open source program for docking ligands to proteins and nucleic acids. *PLoS Computational Biology*, 10(4):e1003571, 2014.

[36] Jie Liu and Renxiao Wang. Classification of current scoring functions. *Journal of Chemical Information and Modeling*, 55(3):475–482, 2015.

[37] Nuno MFSA Cerqueira, Diana Gesto, Eduardo F Oliveira, Diogo Santos-Martins, Natércia F Brás, Sérgio F Sousa, Pedro A Fernandes, and Maria J Ramos. Receptor-based virtual screening protocol for drug discovery. *Archives of Biochemistry and Biophysics*, 582:56–67, 2015.

[38] Hongjian Li, Kwong-Sak Leung, Man-Hon Wong, and Pedro J Ballester. Correcting the impact of docking pose generation error on binding affinity prediction. *BMC Bioinformatics*, 17(11):308, 2016.

[39] Diogo Santos-Martins. Interaction with specific hsp90 residues as a scoring function: validation in the d3r grand challenge 2015. *Journal of Computer-Aided Molecular Design*, 30(9):731–742, 2016.

[40] G Madhavi Sastry, Matvey Adzhigirey, Tyler Day, Ramakrishna Annabhimoju, and Woody Sherman. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *Journal of Computer-Aided Molecular Design*, 27(3):221–234, 2013.

[41] Symon Gathiaka, Shuai Liu, Michael Chiu, Huanwang Yang, Jeanne A Stuckey, You Na Kang, Jim Delproposto, Ginger Kubish, James B Dunbar, Heather A Carlson, et al. D3r grand challenge 2015: Evaluation of protein–ligand pose and affinity predictions. *Journal of Computer-Aided Molecular Design*, 30(9):651–668, 2016.

[42] Pradeep Anand Ravindranath, Stefano Forli, David S Goodsell, Arthur J Olson, and Michel F Sanner. Autodockfr: Advances in protein-ligand docking with explicitly specified binding site flexibility. *PLoS Computational Biology*, 11(12):e1004586, 2015.

[43] Yipin Lu, Renxiao Wang, Chao-Yie Yang, and Shaomeng Wang. Analysis of ligand-bound water molecules in high-resolution crystal structures of protein-ligand complexes. *Journal of Chemical Information and Modeling*, 47(2):668–675, 2007.

[44] Stefano Forli and Arthur J Olson. A force field with discrete displaceable waters and desolvation entropy for hydrated ligand docking. *Journal of Medicinal Chemistry*, 55(2):623–638, 2012.

[45] Robert B Murphy, Matthew P Repasky, Jeremy R Greenwood, Ivan Tubert-Brohman, Steven Jerome, Ramakrishna Annabhimoju, Nicholas A Boyles, Christopher D Schmitz, Robert Abel, Ramy Farid, et al. Wscore: A flexible and accurate treatment of explicit water molecules in ligand–receptor docking. *Journal of Medicinal Chemistry*, 59(9):4364–4384, 2016.

[46] Michael S Bodnarchuk, Russell Viner, Julien Michel, and Jonathan W Essex. Strategies to calculate water binding free energies in protein–ligand complexes. *Journal of Chemical Information and Modeling*, 54(6):1623–1633, 2014.

[47] Mette A Lie, René Thomsen, Christian NS Pedersen, Birgit Schiøtt, and Mikael H Christensen. Molecular docking with ligand attached water molecules. *Journal of Chemical Information and Modeling*, 51(4):909–917, 2011.

[48] Gianluca Rossato, Beat Ernst, Angelo Vedani, and Martin Smiesko. Acquaalta: a directional approach to the solvation of ligand–protein complexes. *Journal of Chemical Information and Modeling*, 51(8):1867–1881, 2011.

[49] Chengteh Lee, Weitao Yang, and Robert G Parr. Development of the colle-salvetti correlation-energy formula into a functional of the electron density. *Physical Review B*, 37(2):785, 1988.

[50] Sergio Filipe Sousa, Pedro Alexandrino Fernandes, and Maria Joao Ramos. General performance of density functionals. *The Journal of Physical Chemistry A*, 111(42):10439–10452, 2007.

[51] Wendy D Cornell, Piotr Cieplak, Christopher I Bayly, Ian R Gould, Kenneth M Merz, David M Ferguson, David C Spellmeyer, Thomas Fox, James W Caldwell, and Peter A Kollman. A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *Journal of the American Chemical Society*, 117(19):5179–5197, 1995.

[52] Lukas D Schuler, Xavier Daura, and Wilfred F Van Gunsteren. An improved gromos96 force field for aliphatic hydrocarbons in the condensed phase. *Journal of Computational Chemistry*, 22(11):1205–1218, 2001.

[53] William L Jorgensen and Julian Tirado-Rives. The opls [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society*, 110(6):1657–1666, 1988.

[54] Shina CL Kamerlin and Arieh Warshel. The empirical valence bond model: theory and applications. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 1(1):30–45, 2011.

[55] Junmei Wang, Romain M Wolf, James W Caldwell, Peter A Kollman, and David A Case. Development and testing of a general amber force field. *Journal of Computational Chemistry*, 25(9):1157–1174, 2004.

[56] William L Jorgensen, David S Maxwell, and Julian Tirado-Rives. Development and testing of the opls all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, 118(45):11225–11236, 1996.

[57] Christopher I Bayly, Piotr Cieplak, Wendy Cornell, and Peter A Kollman. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the resp model. *The Journal of Physical Chemistry*, 97(40):10269–10280, 1993.

[58] Araz Jakalian, Bruce L Bush, David B Jack, and Christopher I Bayly. Fast, efficient generation of high-quality atomic charges. am1-bcc model: I. method. *Journal of Computational Chemistry*, 21(2):132–146, 2000.

[59] Araz Jakalian, David B Jack, and Christopher I Bayly. Fast, efficient generation of high-quality atomic charges. am1-bcc model: Ii. parameterization and validation. *Journal of Computational Chemistry*, 23(16):1623–1641, 2002.

[60] William L Jorgensen, Jayaraman Chandrasekhar, Jeffry D Madura, Roger W Impey, and Michael L Klein. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2):926–935, 1983.

[61] Diogo Santos-Martins, Stefano Forli, Maria João Ramos, and Arthur J Olson. Autodock4zn: an improved autodock force field for small-molecule docking to zinc metalloproteins. *Journal of Chemical Information and Modeling*, 54(8):2371–2379, 2014.

[62] Bert L Vallee and David S Auld. Zinc coordination, function, and structure of zinc enzymes and other proteins. *Biochemistry*, 29(24):5647–5659, 1990.

[63] John H Laity, Brian M Lee, and Peter E Wright. Zinc finger proteins: new insights into structural and functional diversity. *Current Opinion in Structural Biology*, 11(1):39–46, 2001.

[64] Sérgio Filipe Sousa, Ana Branca Lopes, Pedro Alexandrino Fernandes, and Maria João Ramos. The zinc proteome: a tale of stability and functionality. *Dalton Transactions*, (38):7946–7956, 2009.

[65] J Nawarskas, V Rajan, and WH Frishman. Vasopeptidase inhibitors, neutral endopeptidase inhibitors, and dual inhibitors of angiotensin-converting enzyme and neutral endopeptidase. *Heart disease (Hagerstown, Md.)*, 3(6):378–385, 2000.

[66] Yuan-Ping Pang. Novel zinc protein molecular dynamics simulations: Steps toward antiangiogenesis for cancer treatment. *Molecular Modeling Annual*, 5(10):196–202, 1999.

[67] T Scott Reid, Stephen B Long, and Lorena S Beese. Crystallographic analysis reveals that anticancer clinical candidate l-778,123 inhibits protein farnesyltransferase and geranylgeranyltransferase-i by different binding modes. *Biochemistry*, 43(28):9000–9008, 2004.

[68] Jean-Denis Docquier, Manuela Benvenuti, Vito Calderone, Magdalena Stoczko, Nicola Menciassi, Gian Maria Rossolini, and Stefano Mangani. High-resolution crystal structure of the subclass b3 metallo-β-lactamase bjp-1: rational basis for substrate specificity and interaction with sulfonamides. *Antimicrobial agents and chemotherapy*, 54(10):4343–4351, 2010.

[69] Stephan Schilling, Ulrike Zeitschel, Torsten Hoffmann, Ulrich Heiser, Mike Francke, Astrid Kehlen, Max Holzer, Birgit Hutter-Paier, Manuela Prokesch, Manfred Windisch, Jagla others Wolfgang, Schlenzig Dagmar, Lindner Christiane, Rudolph Thomas, Reuter Gunter, Cynis Holger, Montag Dirk, Demuth Hans-Ulrich, and Rossner Steffen. Glutaminyl cyclase inhibition attenuates pyroglutamate aβ and alzheimer's disease–like pathology. *Nature Medicine*, 14(10):1106–1111, 2008.

[70] Chidambar Balbhim Jalkute, Sagar Hindurao Barage, Maruti Jayram Dhanavade, and Kailas Dasharath Sonawane. Molecular dynamics simulation and molecular docking studies of angiotensin converting enzyme with inhibitor lisinopril and amyloid beta peptide. *The Protein Journal*, 32(5):356–364, 2013.

[71] Faith E Jacobsen, Jana A Lewis, and Seth M Cohen. The design of inhibitors for medicinally relevant metalloproteins. *ChemMedChem*, 2(2):152–171, 2007.

[72] Roland H Stote and Martin Karplus. Zinc binding in proteins and solution: a simple but accurate nonbonded representation. *Proteins: Structure, Function, and Bioinformatics*, 23(1):12–31, 1995.

[73] Ruibo Wu, Zhenyu Lu, Zexing Cao, and Yingkai Zhang. A transferable nonbonded pairwise force field to model zinc interactions in metalloproteins. *Journal of Chemical Theory and Computation*, 7(2):433–443, 2011.

[74] Xin Hu and William H Shelver. Docking studies of matrix metalloproteinase inhibitors: zinc parameter optimization to improve the binding free energy prediction. *Journal of Molecular Graphics and Modelling*, 22(2):115–126, 2003.

[75] Claudia Andreini and Ivano Bertini. A bioinformatics view of zinc enzymes. *Journal of Inorganic Biochemistry*, 111:150–156, 2012.

[76] Kirti Patel, Anil Kumar, and Susheel Durani. Analysis of the structural consensus of the zinc coordination centers of metalloprotein structures. *Biochimica et Biophysica Acta (BBA)-Proteins & Proteomics*, 1774(10):1247–1253, 2007.

[77] Ian L Alberts, Katalin Nadassy, and Shoshana J Wodak. Analysis of zinc binding sites in protein crystal structures. *Protein science*, 7(8):1700–1716, 1998.

[78] Mikko Laitaoja, Jarkko Valjakka, and Janne Jänis. Zinc coordination spheres in protein structures. *Inorganic Chemistry*, 52(19):10983–10991, 2013.

[79] K Håkansson and A Liljas. The structure of a complex between carbonic anhydrase ii and a new inhibitor, trifluoromethane sulphonamide. *FEBS Letters*, 350(2):319–322, 1994.

[80] Joshua Pottel, Eric Therrien, James L. Gleason, and Nicolas Moitessier. Docking ligands into flexible and solvated macromolecules. 6. development and application to the docking of hdacs and other zinc metalloenzymes inhibitors. *Journal of Chemical Information and Modeling*, 54(1):254–265, 2014.

[81] Tong Zhu, Xudong Xiao, Changge Ji, and John ZH Zhang. A new quantum calibrated force field for zinc–protein complex. *Journal of Chemical Theory and Computation*, 9(3):1788–1798, 2013.

[82] Dmitri V Sakharov and Carmay Lim. Zn protein simulations including charge transfer and local polarization effects. *Journal of the American Chemical Society*, 127(13):4921–4929, 2005.

[83] Jiajing Zhang, Wei Yang, Jean-Philip Piquemal, and Pengyu Ren. Modeling structural coordination and ligand binding in zinc proteins with a polarizable potential. *Journal of Chemical Theory and Computation*, 8(4):1314–1324, 2012.

[84] Martin B Peters, Yue Yang, Bing Wang, László Füsti-Molnár, Michael N Weaver, and Kenneth M Merz Jr. Structural survey of zinc-containing proteins and development of the zinc amber force field (zaff). *Journal of Chemical Theory and Computation*, 6(9):2935–2947, 2010.

[85] Birte Seebeck, Ingo Reulecke, Andreas Kämper, and Matthias Rarey. Modeling of metal interaction geometries for protein–ligand docking. *Proteins: Structure, Function, and Bioinformatics*, 71(3):1237–1254, 2008.

[86] Giovanni Cincilla, David Vidal, and Miquel Pons. An improved scoring function for suboptimal polar ligand complexes. *Journal of Computer-Aided Molecular Design*, 23(3):143–152, 2009.

[87] Oleksandr V Buzko, Anthony C Bishop, and Kevan M Shokat. Modified autodock for accurate docking of protein kinase inhibitors. *Journal of Computer-Aided Molecular Design*, 16(2):113–127, 2002.

[88] Fredrik Österberg, Garrett M Morris, Michel F Sanner, Arthur J Olson, and David S Goodsell. Automated docking to multiple target structures: incorporation of protein mobility and structural water heterogeneity in autodock. *Proteins: Structure, Function, and Bioinformatics*, 46(1):34–40, 2002.

[89] Sandro Cosconati, Luciana Marinelli, Francesco Saverio Di Leva, Valeria La Pietra, Angela De Simone, Francesca Mancini, Vincenza Andrisano, Ettore Novellino, David S Goodsell, and Arthur J Olson. Protein flexibility in virtual screening: the bace-1 case study. *Journal of Chemical Information and Modeling*, 52(10):2697–2704, 2012.

[90] Stefano Forli and Maurizio Botta. Lennard-jones potential and dummy atom settings to overcome the autodock limitation in treating flexible ring systems. *Journal of Chemical Information and Modeling*, 47(4):1481–1492, 2007.

[91] Liegi Hu, Mark L Benson, Richard D Smith, Michael G Lerner, and Heather A Carlson. Binding moad (mother of all databases). *Proteins*, 60(3):333–340, Aug 2005.

[92] Frances C Bernstein, Thomas F Koetzle, Graheme JB Williams, Edgar F Meyer Jr, Michael D Brice, John R Rodgers, Olga Kennard, Takehiko Shimanouchi, and Mitsuo Tasumi. The protein data bank: a computer-based archival file for macromolecular structures. *Archives of Biochemistry and Biophysics*, 185(2):584–591, 1978.

[93] Marjorie M Harding. Geometry of metal-ligand interactions in proteins. *Acta Crystallographica Section D: Biological Crystallography*, 57(3):401–411, 2001.

[94] Ulf Ryde. Carboxylate binding modes in zinc proteins: a theoretical study. *Biophysical Journal*, 77(5):2777–2787, 1999.

[95] Bradley A Katz and Christine Luong. Recruiting zn^{2+} to mediate potent, specific inhibition of serine proteases. *Journal of Molecular Biology*, 292(3):669–684, 1999.

[96] Ruth Huey, Garrett M Morris, Arthur J Olson, and David S Goodsell. A semiempirical free energy force field with charge-based desolvation. *Journal of Computational Chemistry*, 28(6):1145–1152, 2007.

[97] Johann Gasteiger and Mario Marsili. Iterative partial equalization of orbital electronegativity — a rapid access to atomic charges. *Tetrahedron*, 36(22):3219–3228, 1980.

[98] Ruibo Wu, Zhenyu Lu, Zexing Cao, and Yingkai Zhang. Zinc chelation with hydroxamate in histone deacetylases modulated by water access to the linker binding channel. *Journal of the American Chemical Society*, 133(16):6110–6113, 2011.

[99] Yu-Hung Chiu, Gregory J Gabriel, and James W Canary. Ternary ligand-zinc-hydroxamate complexes. *Inorganic Chemistry*, 44(1):40–44, 2005.

[100] Garrett M Morris, Ruth Huey, William Lindstrom, Michel F Sanner, Richard K Belew, David S Goodsell, and Arthur J Olson. Autodock4 and autodocktools4: Automated docking with selective receptor flexibility. *Journal of Computational Chemistry*, 30(16):2785–2791, 2009.

[101] Esther Kellenberger, Jordi Rodrigo, Pascal Muller, and Didier Rognan. Comparative evaluation of eight docking tools for docking and virtual screening accuracy. *Proteins: Structure, Function, and Bioinformatics*, 57(2):225–242, 2004.

[102] Christian Oefner, Bernard P Roques, M-C Fournie-Zaluski, and Glenn E Dale. Structural analysis of neprilysin with various specific and potent inhibitors. *Acta Crystallographica Section D: Biological Crystallography*, 60(2):392–396, 2004.

[103] Diogo Santos-Martins, Pedro Alexandrino Fernandes, and Maria João Ramos. Calculation of distribution coefficients in the sampl5 challenge from atomic solvation parameters and surface areas. *Journal of Computer-Aided Molecular Design*, 30(11):1079–1086, 2016.

[104] David L Mobley and J Peter Guthrie. Freesolv: a database of experimental and calculated hydration free energies, with input files. *Journal of Computer-Aided Molecular Design*, 28(7):711–720, 2014.

[105] David L. Mobley. Experimental and Calculated Small Molecule Hydration Free Energies. Retrieved from http://www.escholarship.org/uc/item/6sd403pz, 2013.

[106] David Eisenberg and Andrew D. McLachlan. Solvation energy in protein folding and binding. *Nature*, 319(6050):199–203, 1986.

[107] Robert D Boyer and Richard L Bryan. Fast estimation of solvation free energies for diverse chemical species. *The Journal of Physical Chemistry B*, 116(12):3772–3779, 2012.

[108] Junmei Wang, Wei Wang, Shuanghong Huo, Matthew Lee, and Peter A Kollman. Solvation model based on weighted solvent accessible surface area. *The Journal of Physical Chemistry B*, 105(21):5055–5067, 2001.

[109] Tingjun Hou, Xuebin Qiao, Wei Zhang, and Xiaojie Xu. Empirical aqueous solvation models based on accessible surface areas with implicit electrostatics. *The Journal of Physical Chemistry B*, 106(43):11295–11304, 2002.

[110] Tatsuo Ooi, Motohisa Oobatake, George Nemethy, and Harold A Scheraga. Accessible surface areas as a measure of the thermodynamic parameters of hydration of peptides. *Proceedings of the National Academy of Sciences*, 84(10):3086–3090, 1987.

[111] Jianfeng Pei, Qi Wang, Jiaju Zhou, and Luhua Lai. Estimating protein–ligand binding free energy: atomic solvation parameters for partition coefficient and solvation free energy calculation. *PROTEINS: Structure, Function, and Bioinformatics*, 57(4):651–664, 2004.

[112] Junmei Wang, George Krudy, Tingjun Hou, Wei Zhang, George Holland, and Xiaojie Xu. Development of reliable aqueous solubility models and their application in druglike analysis. *Journal of Chemical Information and Modeling*, 47(4):1395–1404, 2007.

[113] Jens Kleinjung, Walter RP Scott, Jane R Allison, Wilfred F van Gunsteren, and Franca Fraternali. Implicit solvation parameters derived from explicit water forces in large-scale molecular dynamics simulations. *Journal of Chemical Theory and Computation*, 8(7):2391–2403, 2012.

[114] Sheng-You Huang and Xiaoqin Zou. Inclusion of solvation and entropy in the knowledge-based scoring function for protein- ligand interactions. *Journal of Chemical Information and Modeling*, 50(2):262–273, 2010.

[115] Hwangseo Park. Extended solvent-contact model approach to SAMPL4 blind prediction challenge for hydration free energies. *Journal of Computer-Aided Molecular Design*, 28(3):175–186, feb 2014.

[116] David L. Mobley, Karisa L. Wymer, Nathan M. Lim, and J. Peter Guthrie. Blind prediction of solvation free energies from the SAMPL4 challenge. *Journal of Computer-Aided Molecular Design*, 28(3):135–150, mar 2014.

[117] Ariën S Rustenburg, Justin Dancer, Baiwei Lin, Jianwen A Feng, Daniel F Ortwine, David L Mobley, and John D Chodera. Measuring experimental cyclohexane-water distribution coefficients for the sampl5 challenge. *bioRxiv*, page 063081, 2016.

[118] Baiwei Lin and Joseph H Pease. A novel method for high throughput lipophilicity determination by microscale shake flask and liquid chromatography tandem mass spectrometry. *Combinatorial Chemistry & High Throughput Screening*, 16(10):817–825, 2013.

[119] Michel F Sanner, Arthur J Olson, and Jean-Claude Spehner. Reduced surface: an efficient way to compute molecular surfaces. *Biopolymers*, 38(3):305–320, 1996.

[120] J. Gasteiger and M. Marsili. A New Model for Calculating Atomic Charges in Molecules. *Tetrahedron Letters*, 34:3181–3184, 1978.

[121] Noel M O'Boyle, Michael Banck, Craig A James, Chris Morley, Tim Vandermeersch, and Geoffrey R Hutchison. Open babel: An open chemical toolbox. *Journal of Cheminformatics*, 3:33, 2011.

[122] Noel M O'Boyle, Chris Morley, and Geoffrey R Hutchison. Pybel: a python wrapper for the openbabel cheminformatics toolkit. *Chemistry Central Journal*, 2(5), 2008.

[123] R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2008.

[124] Alessandra Villa and Alan E. Mark. Calculation of the free energy of solvation for neutral analogs of amino acid side chains. *Journal of Computational Chemistry*, 23(5):548–553, 2002.

[125] SF Sousa, AJM Ribeiro, JTS Coimbra, RPP Neves, SA Martins, NSHN Moorthy, PA Fernandes, and MJ Ramos. Protein-ligand docking in the new millennium–a retrospective of 10 years in the field. *Current Medicinal Chemistry*, 20(18):2296–2314, 2013.

[126] Heather Ann Carlson, Richard Dayton Smith, Kelly L Damm-Ganamet, Jeanne A Stuckey, Aqeel Ahmed, Maire A Convery, Donald O Somers, Michael Kranz, Patricia A Elkins, Guanglei Cui, et al. Csar 2014: A benchmark exercise using unpublished data from pharma. *Journal of Chemical Information and Modeling*, 56(6):1063–1077, 2016.

[127] Heather A Carlson. Lessons learned over four benchmark exercises from the community structure–activity resource, 2016.

[128] Drug Design Data Resource (D3R). https://www.drugdesigndata.org/. Accessed: 2016-05-19.

[129] Tiejun Cheng, Xun Li, Yan Li, Zhihai Liu, and Renxiao Wang. Comparative assessment of scoring functions on a diverse test set. *Journal of Chemical Information and Modeling*, 49(4):1079–1093, 2009.

[130] Gregory L Warren, C Webster Andrews, Anna-Maria Capelli, Brian Clarke, Judith LaLonde, Millard H Lambert, Mika Lindvall, Neysa Nevins, Simon F Semus, Stefan Senger, et al. A critical assessment of docking programs and scoring functions. *Journal of Medicinal Chemistry*, 49(20):5912–5931, 2006.

[131] Regina Politi, Marino Convertino, Konstantin Popov, Nikolay V Dokholyan, and Alexander Tropsha. Docking and scoring with target-specific pose classifier succeeds in native-like pose identification but not binding affinity prediction in the csar 2014 benchmark exercise. *Journal of Chemical Information and Modeling*, 56(6):1032–1041, 2016.

[132] Andrew Anighoro and Jurgen Bajorath. Three-dimensional similarity in molecular docking: Prioritizing ligand poses on the basis of experimental binding modes. *Journal of Chemical Information and Modeling*, 56(3):580–587, 2016.

[133] Esben J Bjerrum. Machine learning optimization of cross docking accuracy. *Computational Biology and Chemistry*, 62:133–144, 2016.

[134] Oleg Trott and Arthur J Olson. Autodock vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of Computational Chemistry*, 31(2):455–461, 2010.

[135] Michael M Mysinger, Michael Carchia, John J Irwin, and Brian K Shoichet. Directory of useful decoys, enhanced (dud-e): better ligands and decoys for better benchmarking. *Journal of Medicinal Chemistry*, 55(14):6582–6594, 2012.

[136] http://dude.docking.org/targets/hs90a/P07900/pdb_codes.txt. Accessed: 2016-05-20.

[137] PDB IDs in the training set: 1uy6, 1uy7, 1uy8, 1uy9, 1uyc, 1uyd, 1uye, 1uyf, 1uyg, 1uyh, 1uyi, 1uyk, 1yc3, 1yc4, 2bsm, 2bt0, 2byh, 2byi, 2bz5, 2ccs, 2cct, 2ccu, 2fwy, 2fwz, 2h55, 2jjc, 2qf6, 2qfo, 2qg0, 2qg2, 2uwd, 2vci, 2vcj, 2wi1, 2wi4, 2wi5, 2wi6, 2wi7, 2xab, 2xdk, 2xdl, 2xds, 2xdx, 2xhr, 2xht, 2xhx, 2xjg, 2xjj, 2xjx, 2xk2, 3bm9, 3bmy, 3d0b, 3eko, 3ekr, 3ft5, 3ft8, 3hek, 3hyy, 3hyz, 3hz5, 3inw, 3inx, 3k97, 3k98, 3k99, 3mnr, 4ykr, 4yky.

[138] David Ryan Koes, Matthew P Baumgartner, and Carlos J Camacho. Lessons learned in empirical scoring with smina from the csar 2011 benchmarking exercise. *Journal of Chemical Information and Modeling*, 53(8):1893–1904, 2013.

[139] Aixia Yan, Guy H Grant, and W Graham Richards. Dynamics of conserved waters in human hsp90: implications for drug design. *Journal of The Royal Society Interface*, 5(3):199–205, 2008.

[140] Sayan Dutta Gupta, D Snigdha, Gisela I Mazaira, Mario D Galigniana, CVS Subrahmanyam, NL Gowrishankar, and NM Raghavendra. Molecular docking study, synthesis and biological evaluation of schiff bases as hsp90 inhibitors. *Biomedicine & Pharmacotherapy*, 68(3):369–376, 2014.

[141] Pei-Pei Kung, Piet-Jan Sinnema, Paul Richardson, Michael J Hickey, Ketan S Gajiwala, Fen Wang, Buwen Huang, Guy McClellan, Jeff Wang, Karen Maegley, et al. Design strategies to target crystallographic waters applied to the hsp90 molecular chaperone. *Bioorganic & Medicinal Chemistry Letters*, 21(12):3557–3562, 2011.

[142] Lisa Wright, Xavier Barril, Brian Dymock, Louisa Sheridan, Allan Surgenor, Mandy Beswick, Martin Drysdale, Adam Collier, Andy Massey, Nick Davies, et al. Structure-activity relationships in purine-based inhibitor binding to hsp90 isoforms. *Chemistry & Biology*, 11(6):775–785, 2004.

[143] http://dude.docking.org/targets/hs90a/pdb_analyze.txt. Accessed: 2016-05-20.

[144] Hongjian Li, Kwong-Sak Leung, Man-Hon Wong, and Pedro J Ballester. Improving autodock vina using random forest: the growing accuracy of binding affinity prediction by the effective exploitation of larger data sets. *Molecular Informatics*, 34(2-3):115–126, 2015.

[145] Xavier Barril and S David Morley. Unveiling the full potential of flexible receptor docking using multiple crystallographic structures. *Journal of Medicinal Chemistry*, 48(13):4432–4443, 2005.

[146] Eric Jones, Travis Oliphant, Pearu Peterson, et al. SciPy: Open source scientific tools for Python. http://www.scipy.org/", 2001–. Accessed 2016-05-13.

[147] Wei Zhao, Kirk E Hevener, Stephen W White, Richard E Lee, and James M Boyett. A statistical framework to evaluate virtual screening. *BMC Bioinformatics*, 10(1):1, 2009.

[148] Michael M Mysinger and Brian K Shoichet. Rapid context-dependent ligand desolvation in molecular docking. *Journal of Chemical Information and Modeling*, 50(9):1561–1573, 2010.

[149] Fedor N Novikov, Viktor S Stroylov, Oleg V Stroganov, and Ghermes G Chilov. Improving performance of docking-based virtual screening by structural filtration. *Journal of Molecular Modeling*, 16(7):1223–1230, 2010.

[150] António JM Ribeiro, Diogo Santos-Martins, Nino Russo, Maria J Ramos, and Pedro A Fernandes. Enzymatic flexibility and reaction rate: A qm/mm study of hiv-1 protease. *ACS Catalysis*, 5(9):5617–5626, 2015.

[151] David D Boehr, Dan McElheny, H Jane Dyson, and Peter E Wright. The dynamic energy landscape of dihydrofolate reductase catalysis. *Science*, 313(5793):1638–1642, 2006.

[152] Katherine Henzler-Wildman and Dorothee Kern. Dynamic personalities of proteins. *Nature*, 450(7172):964–972, 2007.

[153] Robert Callender and R Brian Dyer. The dynamical nature of enzymatic catalysis. *Accounts of Chemical Research*, 48(2):407–413, 2015.

[154] Amnon Kohen. Role of dynamics in enzyme catalysis: substantial versus semantic controversies. *Accounts of Chemical Research*, 48(2):466–473, 2014.

[155] R Derike Smiley and Gordon G Hammes. Single molecule studies of enzyme mechanisms. *Chemical Reviews*, 106(8):3080–3094, 2006.

[156] Qifeng Xue and Edward S Yeung. Differences in the chemical reactivity of individual molecules of an enzyme. *Nature*, 373(6516):681–683, 1995.

[157] H Peter Lu, Luying Xun, and X Sunney Xie. Single-molecule enzymatic dynamics. *Science*, 282(5395):1877–1882, 1998.

[158] Brian P English, Wei Min, Antoine M Van Oijen, Kang Taek Lee, Guobin Luo, Hongye Sun, Binny J Cherayil, SC Kou, and X Sunney Xie. Ever-fluctuating single enzyme molecules: Michaelis-menten equation revisited. *Nature Chemical Biology*, 2(2):87–94, 2006.

[159] Douglas B Craig, Edgar A Arriaga, Jerome CY Wong, Hui Lu, and Norman J Dovichi. Studies on single alkaline phosphatase molecules: Reaction rate and activation energy of a reaction catalyzed by a single molecule and the effect of thermal denaturation the death of an enzyme. *Journal of the American Chemical Society*, 118(22):5245–5253, 1996.

[160] Jue Shi, Bruce A Palfey, Joe Dertouzos, Kaj Frank Jensen, Ari Gafni, and Duncan Steel. Multiple states of the tyr318leu mutant of dihydroorotate dehydrogenase revealed by single-molecule kinetics. *Journal of the American Chemical Society*, 126(22):6914–6922, 2004.

[161] Taekjip Ha, Alice Y Ting, Joy Liang, W Brett Caldwell, Ashok A Deniz, Daniel S Chemla, Peter G Schultz, and Shimon Weiss. Single-molecule fluorescence spectroscopy of enzyme conformational dynamics and cleavage mechanism. *Proceedings of the National Academy of Sciences*, 96(3):893–898, 1999.

[162] Tatyana G Terentyeva, Hans Engelkamp, Alan E Rowan, Tamiki Komatsuzaki, Johan Hofkens, Chun-Biu Li, and Kerstin Blank. Dynamic disorder in single-enzyme experiments: facts and artifacts. *ACS Nano*, 6(1):346–354, 2012.

[163] Richard A Friesner and Victor Guallar. Ab initio quantum chemical and mixed quantum mechanics/molecular mechanics (qm/mm) methods for studying enzymatic catalysis. *Annual Review of Physical Chemistry*, 56:389–427, 2005.

[164] Hans Martin Senn and Walter Thiel. Qm/mm methods for biological systems. In *Atomistic approaches in modern biology*, pages 173–290. Springer, 2007.

[165] Hans Martin Senn and Walter Thiel. Qm/mm methods for biomolecular systems. *Angewandte Chemie International Edition*, 48(7):1198–1229, 2009.

[166] Adrian J Mulholland. Introduction. biomolecular simulation. *Journal of The Royal Society Interface*, 5(3):169–172, 2008.

[167] Sérgio Filipe Sousa, Pedro Alexandrino Fernandes, and Maria João Ramos. Computational enzymatic catalysis–clarifying enzymatic mechanisms with the help of computers. *Physical Chemistry Chemical Physics*, 14(36):12431–12441, 2012.

[168] Yingkai Zhang, Jeremy Kua, and J Andrew McCammon. Influence of structural fluctuation on enzyme reaction energy barriers in combined quantum mechanical/molecular mechanical studies. *The Journal of Physical Chemistry B*, 107(18):4459–4463, 2003.

[169] Silvia Ferrer, Iñaki Tuñón, Sergio Martí, Vicente Moliner, Mireia Garcia-Viloca, Àngels González-Lafont, and José M Lluch. A theoretical analysis of rate constants and kinetic isotope effects corresponding to different reactant valleys in lactate dehydrogenase. *Journal of the American Chemical Society*, 128(51):16851–16863, 2006.

[170] Po Hu and Yingkai Zhang. Catalytic mechanism and product specificity of the histone lysine methyltransferase set7/9: An ab initio qm/mm-fe study with multiple initial structures. *Journal of the American Chemical Society*, 128(4):1272–1278, 2006.

[171] Alessio Lodola, Marco Mor, Jolanta Zurek, Giorgio Tarzia, Daniele Piomelli, Jeremy N Harvey, and Adrian J Mulholland. Conformational effects in enzyme catalysis: reaction via a high energy conformation in fatty acid amide hydrolase. *Biophysical Journal*, 92(2):L20–L22, 2007.

[172] Alessio Lodola, Jitnapa Sirirak, Natalie Fey, Silvia Rivara, Marco Mor, and Adrian J Mulholland. Structural fluctuations in enzyme-catalyzed reactions: determinants of reactivity in fatty acid amide hydrolase from multivariate statistical analysis of quantum mechanics/molecular mechanics paths. *Journal of Chemical Theory and Computation*, 6(9):2948–2960, 2010.

[173] Stephen J Benkovic, Gordon G Hammes, and Sharon Hammes-Schiffer. Free-energy landscape of enzyme catalysis†. *Biochemistry*, 47(11):3317–3321, 2008.

[174] Maite Roca, Benjamin Messer, Donald Hilvert, and Arieh Warshel. On the relationship between folding and chemical landscapes in enzyme catalysis. *Proceedings of the National Academy of Sciences*, 105(37):13877–13882, 2008.

[175] António JM Ribeiro, Maria J Ramos, and Pedro A Fernandes. The catalytic mechanism of hiv-1 integrase for dna 3'-end processing established by qm/mm calculations. *Journal of the American Chemical Society*, 134(32):13436–13447, 2012.

[176] Melchor Sanchez-Martinez, Enrique Marcos, Romà Tauler, Martin Field, and Ramon Crehuet. Conformational compression and barrier height heterogeneity in the n-acetylglutamate kinase. *The Journal of Physical Chemistry B*, 117(46):14261–14272, 2013.

[177] PMD Fitzgerald and JP Springer. Structure and function of retroviral proteases. *Annual Review of Biophysics and Biophysical Chemistry*, 20(1):299–320, 1991.

[178] John M Louis, Rieko Ishima, Dennis A Torchia, and Irene T Weber. Hiv-1 protease: Structure, dynamics, and inhibition. *Advances in Pharmacology*, 55:261–298, 2007.

[179] K Suguna, Eduardo A Padlan, Clark W Smith, William D Carlson, and David R Davies. Binding of a reduced peptide inhibitor to the aspartic proteinase from rhizopus chinensis: implications for a mechanism of action. *Proceedings of the National Academy of Sciences*, 84(20):7009–7013, 1987.

[180] M Miller, J Schneider, BK Sathyanarayana, MV Toth, GR Marshall, L Clawson, L Selk, SB Kent, and A Wlodawer. Structure of complex of synthetic hiv-1 protease with a substrate-based inhibitor at 2.3 a resolution. *Science*, 246(4934):1149–1152, 1989.

[181] Lawrence J Hyland, Thaddeus A Tomaszek Jr, and Thomas D Meek. Human immunodeficiency virus-1 protease. 2. use of ph rate studies and solvent kinetic isotope effects to elucidate details of chemical mechanism. *Biochemistry*, 30(34):8454–8463, 1991.

[182] Evelyn J Rodriguez, Thelma S Angeles, and Thomas D Meek. Use of nitrogen-15 kinetic isotope effects to elucidate details of the chemical mechanism of human immunodeficiency virus 1 protease. *Biochemistry*, 32(46):12380–12385, 1993.

[183] Laszlo Polgar, Zoltan Szeltner, and Imre Boros. Substrate-dependent mechanisms in the catalysis of human immunodeficiency virus protease. *Biochemistry*, 33(31):9351–9357, 1994.

[184] Moses Prabu-Jeyabalan, Ellen Nalivaika, and Celia A Schiffer. How does a symmetric dimer recognize an asymmetric substrate? a substrate complex of hiv-1 protease. *Journal of Molecular Biology*, 301(5):1207–1220, 2000.

[185] Dexter B Northrop. Follow the protons: a low-barrier hydrogen bond unifies the mechanisms of the aspartic proteases. *Accounts of Chemical Research*, 34(10):790–797, 2001.

[186] Andrey Y Kovalevsky, Alexander A Chumanevich, Fengling Liu, John M Louis, and Irene T Weber. Caught in the act: the 1.5 å resolution crystal structures of the hiv-1 protease and the i54v mutant reveal a tetrahedral reaction intermediate. *Biochemistry*, 46(51):14854–14864, 2007.

[187] Amit Das, Smita Mahale, Vishal Prashar, Subhash Bihani, J-L Ferrer, and MV Hosur. X-ray snapshot of hiv-1 protease in action: observation of tetrahedral intermediate and short ionic hydrogen bond sihb with catalytic aspartate. *Journal of the American Chemical Society*, 132(18):6366–6373, 2010.

[188] Stefano Piana, Denis Bucher, Paolo Carloni, and Ursula Rothlisberger. Reaction mechanism of hiv-1 protease by hybrid car-parrinello/classical md simulations. *The Journal of Physical Chemistry B*, 108(30):11139–11149, 2004.

[189] V Carnevale, Simone Raugei, Simone Piana, and Paolo Carloni. On the nature of the reaction intermediate in the hiv-1 protease: a quantum chemical study. *Computer Physics Communications*, 179(1):120–123, 2008.

[190] Julian Garrec, P Sautet, and P Fleurat-Lessard. Understanding the hiv-1 protease reactivity with dft: what do we gain from recent functionals? *The Journal of Physical Chemistry B*, 115(26):8545–8558, 2011.

[191] Sinisa Bjelic and Johan Åqvist. Catalysis and linear free energy relationships in aspartic proteases. *Biochemistry*, 45(25):7709–7723, 2006.

[192] AJ Beveridge. An ab initio study of the first stage of catalysis in the monomeric aspartic proteinases. *Journal of Molecular Structure: THEOCHEM*, 900(1):1–8, 2009.

[193] Rajiv Singh, Arghya Barman, and Rajeev Prabhakar. Computational insights into aspartyl protease activity of presenilin 1 (ps1) generating alzheimer amyloid β-peptides (aβ40 and aβ42). *The Journal of Physical Chemistry B*, 113(10):2990–2999, 2009.

[194] Ram Prasad Bora, Arghya Barman, Xiaoxia Zhu, Mehmet Ozbil, and Rajeev Prabhakar. Which one among aspartyl protease, metallopeptidase, and artificial metallopeptidase is the most efficient catalyst in peptide hydrolysis? *The Journal of Physical Chemistry B*, 114(33):10860–10875, 2010.

[195] Natércia F Brás, Maria J Ramos, and Pedro A Fernandes. The catalytic mechanism of mouse renin studied with qm/mm calculations. *Physical Chemistry Chemical Physics*, 14(36):12605–12613, 2012.

[196] S Kashif Sadiq and Peter V Coveney. Computing the role of near attack conformations in an enzyme-catalyzed nucleophilic bimolecular reaction. *Journal of Chemical Theory and Computation*, 11(1):316–324, 2014.

[197] Roy Dennington, Todd Keith, John Millam, et al. Gaussview, version 5. *Semichem Inc., Shawnee Mission, KS*, 2009.

[198] Ulrich Essmann, Lalith Perera, Max L Berkowitz, Tom Darden, Hsing Lee, and Lee G Pedersen. A smooth particle mesh ewald method. *The Journal of Chemical Physics*, 103(19):8577–8593, 1995.

[199] Jean-Paul Ryckaert, Giovanni Ciccotti, and Herman JC Berendsen. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics*, 23(3):327–341, 1977.

[200] Jay W Ponder and David A Case. Force fields for protein simulations. *Advances in Protein Chemistry*, 66:27–85, 2003.

[201] DA Case, TA Darden, TE Cheatham, CL Simmerling, J Wang, RE Duke, R Luo, M Crowley, RC Walker, W Zhang, et al. Amber 10; university of california: San francisco, 2010.

[202] Axel D Becke. Density-functional thermochemistry. iii. the role of exact exchange. *The Journal of Chemical Physics*, 98(7):5648–5652, 1993.

[203] DITCHFIE. R, WJ HEHRE, and JA Pople. Self-consistent molecular-orbital methods. 9. extended gaussian-type basis for molecular-orbital studies of organic molecules. *Journal of Chemical Physics*, 54(2):724, 1971.

[204] Stefan Grimme, Jens Antony, Stephan Ehrlich, and Helge Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (dft-d) for the 94 elements h-pu. *The Journal of Chemical Physics*, 132(15):154104, 2010.

[205] Thom Vreven, K Suzie Byun, István Komáromi, Stefan Dapprich, John A Montgomery Jr, Keiji Morokuma, and Michael J Frisch. Combining quantum mechanics methods with molecular mechanics methods in oniom. *Journal of Chemical Theory and Computation*, 2(3):815–826, 2006.

[206] MJEA Frisch, GW Trucks, Hs B Schlegel, GE Scuseria, MA Robb, JR Cheeseman, G Scalmani, V Barone, B Mennucci, GAea Petersson, et al. Gaussian 09, revision a. 02, gaussian. *Inc., Wallingford, CT*, 200, 2009.

[207] Barbara Maschera, Graham Darby, Giorgio Palú, Lois L Wright, Margaret Tisdale, Richard Myers, Edward D Blair, and Eric S Furfine. Human immunodeficiency virus mutations in the viral protease that confer resistance to saquinavir increase the dissociation rate constant of the protease-saquinavir complex. *Journal of Biological Chemistry*, 271(52):33231–33235, 1996.

[208] Stefano Piana, Paolo Carloni, and Michele Parrinello. Role of conformational fluctuations in the enzymatic reaction of hiv-1 protease. *Journal of Molecular Biology*, 319(2):567–583, 2002.

[209] Alexander L Perryman, Jung-Hsin Lin, and J Andrew McCammon. Hiv-1 protease molecular dynamics of a wild-type and of the v82f/i84v mutant: Possible contributions to drug resistance and a potential new target site for drugs. *Protein Science*, 13(4):1108–1123, 2004.

[210] Donald Hamelberg and J Andrew McCammon. Fast peptidyl cis-trans isomerization within the flexible gly-rich flaps of hiv-1 protease. *Journal of the American Chemical Society*, 127(40):13778–13779, 2005.

[211] Chia-En A Chang, Joanna Trylska, Valentina Tozzini, and J Andrew Mccammon. Binding pathways of ligands to hiv-1 protease: Coarse-grained and atomistic simulations. *Chemical Biology & Drug Design*, 69(1):5–13, 2007.

[212] Valentina Tozzini, Joanna Trylska, Chia-en Chang, and J Andrew McCammon. Flap opening dynamics in hiv-1 protease explored with a coarse-grained model. *Journal of Structural Biology*, 157(3):606–615, 2007.

[213] Joanna Trylska, Valentina Tozzini, A Chang Chia-en, and J Andrew McCammon. Hiv-1 protease substrate binding and product release pathways explored with coarse-grained molecular dynamics. *Biophysical Journal*, 92(12):4179–4187, 2007.

[214] Marc W Kamp, Robin Chaudret, and Adrian J Mulholland. Qm/mm modelling of ketosteroid isomerase reactivity indicates that active site closure is integral to catalysis. *FEBS Journal*, 280(13):3120–3131, 2013.

[215] Richard Lonsdale, Jeremy N Harvey, and Adrian J Mulholland. Compound i reactivity defines alkene oxidation selectivity in cytochrome p450cam. *The Journal of Physical Chemistry B*, 114(2):1156–1162, 2010.

[216] Shina CL Kamerlin and Arieh Warshel. At the dawn of the 21st century: Is dynamics the missing link for understanding enzyme catalysis? *Proteins: Structure, Function, and Bioinformatics*, 78(6):1339–1375, 2010.

[217] Vincenzo Carnevale, Simone Raugei, Cristian Micheletti, and Paolo Carloni. Convergent dynamics in the protease enzymatic superfamily. *Journal of the American Chemical Society*, 128(30):9766–9772, 2006.

[218] Sun Hur and Thomas C Bruice. Comparison of formation of reactive conformers (nacs) for the claisen rearrangement of chorismate to prephenate in water and in the e. c oli mutase: The efficiency of the enzyme catalysis. *Journal of the American Chemical Society*, 125(19):5964–5972, 2003.

[219] Jesús Giraldo, David Roche, Xavier Rovira, and Juan Serra. The catalytic power of enzymes: Conformational selection or transition state stabilization? *FEBS Letters*, 580(9):2170–2177, 2006.

[220] Hong Guo, Qiang Cui, William N Lipscomb, and Martin Karplus. Substrate conformational transitions in the active site of chorismate mutase: Their role in the catalytic mechanism. *Proceedings of the National Academy of Sciences*, 98(16):9032–9037, 2001.

[221] Sun Hur and Thomas C Bruice. The near attack conformation approach to the study of the chorismate to prephenate reaction. *Proceedings of the National Academy of Sciences*, 100(21):12015–12020, 2003.

[222] Maite Roca, Leonardo De Maria, Shoshana J Wodak, Vicente Moliner, Iñaki Tuñón, and Jesus Giraldo. Coupling of the guanosine glycosidic bond conformation and the ribonucleotide cleavage reaction: Implications for barnase catalysis. *Proteins: Structure, Function, and Bioinformatics*, 70(2):415–428, 2008.

[223] Jory Z Ruscio, Jonathan E Kohn, K Aurelia Ball, and Teresa Head-Gordon. The influence of protein dynamics on the success of computational enzyme design. *Journal of the American Chemical Society*, 131(39):14111–14115, 2009.

[224] Rui P Sousa, Pedro A Fernandes, Maria J Ramos, and Natércia F Brás. Insights into the reaction mechanism of 3-o-sulfotransferase through qm/mm calculations. *Physical Chemistry Chemical Physics*, 18(16):11488–11496, 2016.

[225] Eduardo F Oliveira, Nuno MFSA Cerqueira, Maria J Ramos, and Pedro A Fernandes. Qm/mm study of the mechanism of reduction of 3-hydroxy-3-methylglutaryl coenzyme a catalyzed by human hmg-coa reductase. *Catalysis Science & Technology*, 6(19):7172–7185, 2016.

[226] Yanwei Li, Ruiming Zhang, Likai Du, Qingzhu Zhang, and Wenxing Wang. How many conformations of enzymes should be sampled for dft/mm calculations? a case study of fluoroacetate dehalogenase. *International Journal of Molecular Sciences*, 17(8):1372, 2016.

[227] Claudia Loerbroks, Andreas Heimermann, and Walter Thiel. Solvents effects on the mechanism of cellulose hydrolysis: a qm/mm study. *Journal of Computational Chemistry*, 36(15):1114–1123, 2015.

[228] Tatiana Vasilevskaya, Maria G Khrenova, Alexander V Nemukhin, and Walter Thiel. Methodological aspects of qm/mm calculations: A case study on matrix metalloproteinase-2. *Journal of Computational Chemistry*, 2016.

[229] April M Cooper and Johannes Kästner. Averaging techniques for reaction barriers in qm/mm simulations. *ChemPhysChem*, 15(15):3264–3269, 2014.

[230] Nily Dan. Understanding dynamic disorder fluctuations in single-molecule enzymatic reactions. *Current Opinion in Colloid & Interface Science*, 12(6):314–321, 2007.

[231] Tatyana G Terentyeva, Hans Engelkamp, Alan E Rowan, Tamiki Komatsuzaki, Johan Hofkens, Chun-Biu Li, and Kerstin Blank. Dynamic disorder in single-enzyme experiments: facts and artifacts. *ACS Nano*, 6(1):346–354, 2011.

[232] Iñaki Tuñón, Damien Laage, and James T Hynes. Are there dynamical effects in enzyme catalysis? some thoughts concerning the enzymatic chemical step. *Archives of Biochemistry and Biophysics*, 582:42–55, 2015.

[233] J Javier Ruiz-Pernia, Louis YP Luk, Rafael García-Meseguer, Sergio Martí, E Joel Loveridge, Iñaki Tuñón, Vicent Moliner, and Rudolf K Allemann. Increased dynamic effects in a catalytically compromised variant of escherichia coli dihydrofolate reductase. *Journal of the American Chemical Society*, 135(49):18689–18696, 2013.

[234] Gaspar P Pinto, Natercia F Bras, Marta AS Perez, Pedro A Fernandes, Nino Russo, Maria J Ramos, and Marirosa Toscano. Establishing the catalytic mechanism of human pancreatic α-amylase with qm/mm methods. *Journal of Chemical Theory and Computation*, 11(6):2508–2516, 2015.

[235] Natércia F Brás, Nuno MFSA Cerqueira, Maria J Ramos, and Pedro A Fernandes. Glycosidase inhibitors: a patent review (2008–2013). *Expert Opinion on Therapeutic Patents*, 24(8):857–874, 2014.

[236] Helen M Berman, John Westbrook, Zukang Feng, Gary Gilliland, Talapady N Bhat, Helge Weissig, Ilya N Shindyalov, and Philip E Bourne. The protein data bank. *Nucleic acids research*, 28(1):235–242, 2000.

[237] Gary D Brayer, Gary Sidhu, Robert Maurus, Edwin H Rydberg, Curtis Braun, Yili Wang, Nham T Nguyen, Christopher M Overall, and Stephen G Withers. Subsite mapping of the human pancreatic α-amylase active site through structural, kinetic, and mutagenesis techniques. *Biochemistry*, 39(16):4778–4791, 2000.

[238] DA Case, TA Darden, TE Cheatham III, CL Simmerling, J Wang, RE Duke, R Luo, RC Walker, W Zhang, KM Merz, et al. Amber 12; university of california: San francisco, 2012.

[239] Viktor Hornak, Robert Abel, Asim Okur, Bentley Strockbine, Adrian Roitberg, and Carlos Simmerling. Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins: Structure, Function, and Bioinformatics*, 65(3):712–725, 2006.

[240] Karl N Kirschner, Austin B Yongye, Sarah M Tschampel, Jorge González-Outeiriño, Charlisa R Daniels, B Lachele Foley, and Robert J Woods. Glycam06: a generalizable biomolecular force field. carbohydrates. *Journal of Computational Chemistry*, 29(4):622–655, 2008.

[241] Junmei Wang, Wei Wang, Peter A Kollman, and David A Case. Automatic atom type and bond type perception in molecular mechanical calculations. *Journal of Molecular Graphics and Modelling*, 25(2):247–260, 2006.

[242] Tom Darden, Darrin York, and Lee Pedersen. Particle mesh ewald: An n log (n) method for ewald sums in large systems. *The Journal of Chemical Physics*, 98(12):10089–10092, 1993.

[243] Michael Crowley, Tom Darden, Thomas Cheatham III, and David Deerfield II. Adventures in improving the scaling and accuracy of a parallel molecular dynamics program. *The Journal of Supercomputing*, 11(3):255–278, 1997.

[244] Yan Li, Zhihai Liu, Jie Li, Li Han, Jie Liu, Zhixiong Zhao, and Renxiao Wang. Comparative assessment of scoring functions on an updated benchmark: 1. compilation of the test set. *Journal of Chemical Information and Modeling*, 54(6):1700–1716, 2014.

[245] Sushil Kumar Mishra, Jan Adam, Michaela Wimmerová, and Jaroslav Koča. In silico mutagenesis and docking study of ralstonia solanacearum rsl lectin: performance of docking software to predict saccharide binding. *Journal of Chemical Information and Modeling*, 52(5):1250–1261, 2012.

[246] Sushil K Mishra, Gaetano Calabró, Hannes H Loeffler, Julien Michel, and Jaroslav Koča. Evaluation of selected classical force fields for alchemical binding free energy calculations of protein-carbohydrate complexes. *Journal of Chemical Theory and Computation*, 11(7):3333–3345, 2015.