Journal of Medicinal Chemistry

pubs.acs.org/jmc

Perspective

Decision Making in Structure-Based Drug Discovery: Visual Inspection of Docking Results

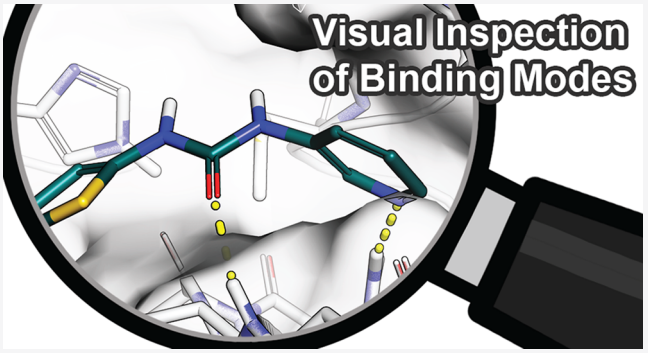
André Fischer, Martin Smieško, Manuel Sellner, and Markus A. Lill*

Cite This: J. Med. Chem. 2021, 64, 2489−2500

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Molecular docking is a computational method widely used in drug discovery. Due to the inherent inaccuracies of molecular docking, visual inspection of binding modes is a crucial routine in the decision making process of computational medicinal chemists. Despite its apparent importance for medicinal chemistry projects, guidelines for the visual docking pose assessment have been hardly discussed in the literature. Here, we review the medicinal chemistry literature with the aim of identifying consistent principles for visual inspection, highlighting cases of its successful application, and discussing its limitations. In this context, we conducted a survey reaching experts in both academia and the pharmaceutical industry, which also included a challenge to distinguish native from incorrect poses. We were able to collect 93 expert opinions that offer valuable insights into visually supported decision-making processes. This perspective shall motivate discussions among experienced computational medicinal chemists and guide young scientists new to the field to stratify their compounds.



INTRODUCTION

Computational procedures such as structure-based virtual screening of large digital compound libraries and the evaluation of ligand binding modes have become an essential component in modern drug discovery.[1] They may not only provide insights into ligands that potentially bind to a particular target but also assist in the interpretation of experimental observations such as structure−activity relationships.[2] Thus, structure-based computational methods can be applied in various stages of drug discovery and development. The popularity of virtual screening originates from its cost-effectiveness, its potential to identify novel chemotypes, and mechanistic insights into ligand−protein interactions.[3,4] One of the primary methods applied in structure-based drug discovery is molecular docking. In molecular docking, favorable ligand−protein configurations are identified yielding so-called ligand poses. Their interaction with the protein is quantified in a process termed scoring.[1,5−7] To achieve the necessary efficiency of molecular docking, the quantification of ligand−protein interactions is based on a simplistic scoring function. This simplified modeling of protein−ligand interactions is a substantial limitation of the methodology, which is manifested by inaccuracies in pose ranking and typically poor performance in the prediction of absolute or relative binding free energies. Furthermore, correct sampling of ligand poses can be limited by induced-fit effects as well as distinct conformational states of the protein, which is often treated as a rigid body, leading to imprecisions beyond those imposed by scoring functions.[8] In addition, intramolecular interactions are sometimes ill-modeled as highlighted by the occurrence of poses with, for example, twisted amide groups, intraligand clashes, or generally unrealistic three-dimensional (3D) structures.[3,4,9,10] Another factor essential for the performance of molecular docking is the quality of the input structures. Common pitfalls and caveats include the protonation state and tautomeric forms of the ligand, as well as peculiar protonation states of acidic and basic residues and flips of glutamine, asparagine, and histidine residues in the protein binding site. In virtual screening applications, the methodology frequently produces a large number of false-positive hits, limiting its applicability as a stand-alone platform.[3] Even though much effort has been spent to improve results of molecular docking, e.g., by running molecular dynamics (MD) simulations followed by molecular mechanics/generalized Born surface area (MM/GBSA), alchemical binding free energy calculations,[11] or modern machine learning methods, the inspection

Received: December 23, 2020
Published: February 22, 2021

https://dx.doi.org/10.1021/acs.jmedchem.0c02227
J. Med. Chem. 2021, 64, 2489−2500

Table 1. Criteria Used in Successful Applications of Visual Inspection[a]

considered criteria	targets
structural novelty, hydrogen bonding network, interactions with specific residues,[b] complementarity, ligand flexibility, water interactions, ligand strain, hydrophobic interactions	purine nucleoside phosphorylase[24]
binding mode consensus,[c] ligand protonation, tautomers, complementarity, interactions with specific residues[b]	aromatase[25]
unique interactions,[d] interactions with specific residues,[b] ligand flexibility, comparison to crystal structures, structural diversity	κ-opioid receptor[31]
interactions with specific residues,[b] hydrophobic interactions, ligand strain	RANKL receptor[30]
interactions with specific residues,[b] complementarity, hydrophobic interactions, π-stacking, hydrogen bonding network, water interactions, ligand strain	ecto-5′-nucleotidase[22]
distorted ligand geometry,[e] ligand strain, nature of solvent-exposed groups, electrostatic repulsion, unsatisfied ligand heteroatoms, hydrophobic interactions, π-stacking, cation−π interactions	enoyl-acyl-carrier protein reductase[21]
structural novelty, hydrogen bonding network, solubility	enoyl-acyl-carrier protein[61]
unsatisfied ligand heteroatoms, ligand flexibility, distorted ligand geometry, hydrogen bonding network, complementarity	EphB4 kinase[34]
complementarity, interactions with specific residues,[b] hydrogen bonding network, hydrophobic interactions, π-stacking, binding mode consensus[c]	γ-glutamylcysteine synthase[43]
complementarity, binding mode consensus,[c] improper atom types, hydrophilic−hydrophobic mismatch, nature of solvent-exposed groups	Bcl-xL[28]
complementarity, π-stacking, hydrogen bonding network interactions with specific residues[b]	thymidine monophosphate kinase[17]
structural diversity, hydrophobic interactions, hydrogen bonding network, complementarity	β2-adrenergic receptor[36]
complementarity, structural diversity, ligand flexibility	aldose reductase[44]
interactions with specific residues,[b] π-stacking, hydrophobic interactions, hydrogen bonding network	*N*-myristoyltransferase[62]

[a]Studies in which bioactive compounds were discovered and the criteria that were applied during visual inspection. [b]Interaction with specific amino acids or cofactors. [c]Consensus among binding modes of multiple docked ligands. [d]Interactions not present in homologous targets. [e]Distorted amides, esters, and other functional groups.

of the docking results by a specialist, especially regarding the predicted binding modes, is of major importance for a successful hit identification or lead optimization.[4,6,10,12−15] A recent review reported approximately 50% of 250 virtual screening publications to conduct visual inspection in compound prioritization.[16] It was suggested that any person with proper experience can visually evaluate the quality of a docking pose better than any available scoring function.[12] Still, several computational methods and tools have been developed to assist the decision making. They employ techniques such as interaction fingerprints,[17,18] scaffold docking,[19] or the comparison of binding modes obtained for multiple congeneric compounds[20] to assist in the decision making process. Some of these tools have been only recently developed and are available in the form of a user-friendly web server.

As we will describe below, visual inspection of docking results has been used in several successful discovery projects.[3,5,9,21−31] Additionally, visual inspection was also applied during the validation of docking protocols, in particular to evaluate pose prediction by redocking or cross-docking to supplement standard measures such as the root-mean-square deviation (RMSD) and native contacts.[32,33] Unfortunately, most of those studies did only vaguely specify applied criteria or did not mention them at all.[9,23,26−29,34] A recent review focused on decision making in medicinal chemistry in the evaluation of large data sets with respect to potency, selectivity, pharmacokinetics, structural biology, and computational results. With regard to compound prioritization based on such data sets, the authors reported a considerable disagreement among medicinal chemists, presumably arising from their different experience, background, and training.[13,35] This discrepancy among different researchers may also exist for visual evaluation of docking poses. Therefore, there is a need to comprehensively review applied criteria to identify consensus guidelines for the visual inspection of ligand binding modes.

Here, we reviewed medicinal and computational chemistry publications, in which visual inspection was applied to complement molecular docking and pose scoring in virtual screening campaigns. We report its successful applications, its limitations, and its direct impact on the accuracy of predictions in compound prioritization. In addition, we conducted a survey by addressing experts in computational medicinal chemistry to identify consensus and disagreement among criteria used for the visual inspection of docking poses. This perspective aims to provide a foundation for a broader discussion in the community for the utility and best strategies for visual inspection of docking results and to offer a guidance to new or young researchers in the field.

CRITERIA DURING VISUAL INSPECTION

One of the most frequently assessed criteria in visual inspection is the steric complementarity of a compound and its binding site (Table 1). Isothermal titration calorimetry experiments have revealed a correlation between the buried nonpolar surface and the binding free energy of a ligand.[2] Furthermore, in rational structure-based lead optimization, a partial lack of complementarity can guide the expansion or modification of the ligand. In some studies, shape complementarity was further differentiated into electrostatic and hydrophobic complementarity adding physicochemical properties of protein and ligand in the visual inspection, thus providing a more detailed picture of the interactions in a complex.[36−38]

Nearly all studies we assessed reported the inspection of hydrogen bonds (Table 1) between the ligand and protein. Due to their relative strength and clearly defined geometry, hydrogen bonds are easy to visually identify and are important factors for binding affinity and selectivity.[41,42] Frequently, hydrogen bonding networks with specific binding site residues, often in comparison to co-crystallized ligands or other ligands with a similar chemotype,[28,43,44] were considered during inspection. For example, in the design of enzyme inhibitors, specific interactions with catalytic residues are important for potent inhibition of the catalytic reaction and for resistance against potential mutations.[45,46] Generally, the strength, the stability, and thus the final contribution of hydrogen bonds to

https://dx.doi.org/10.1021/acs.jmedchem.0c02227
J. Med. Chem. 2021, 64, 2489−2500

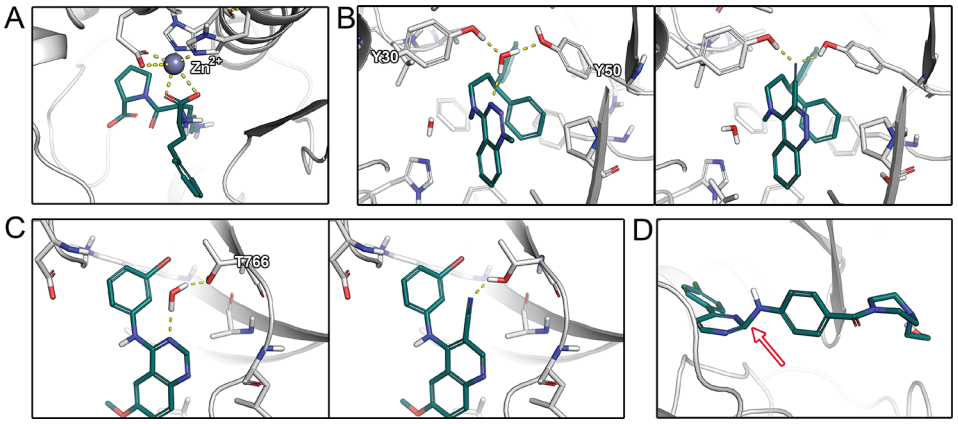


Figure 1. Examples of binding modes. (A) Ligand−metal interaction of lisinopril with angiotensin-converting enzyme [Protein Data Bank (PDB) entry 1O86]. (B) *N*-(3,3-Diphenylpropyl)-4-benz-1,2,3-triazine-amine (left) and 3-cyano-*N*-(3,3-diphenylpropyl)-4-cinnolin-amine (right) bound to scytalone dehydratase in the presence or absence of a water molecule. Displacement of a water molecule by the cyano group resulted in a 2.0 kcal/mol increase in binding free energy. The figure was adapted after a crystal structure (PDB entry 5STD) and a study on hydration site prediction.[39] (C) 4-Anilino-6,7-dialkoxyquinazoline (left) and 4-anilino-6,7-dialkoxyquinoline-3-carbonitrile (right) bound to EGFR kinase. Displacement of the water molecule resulted in an unfavorable binding free energy. Panel C was adapted after a crystal structure (PDB entry 1M17) and a study on hydration site prediction.[39] (D) Example for a strained ligand conformation in a crystal structure (PDB entry 3QAD)[40] with the 2-aminopyrimidine moiety out of plane as indicated by an arrow. The entry was since refined with a more reasonable conformation (PDB entry 3RZF).

binding affinity can vary significantly on the basis of the environment. Hydrogen bonds in a hydrophobic protein environment typically contribute more favorably to the free energy of binding than hydrogen bonds in solvent-accessible regions due to significant differences in desolvation free energy and attenuation of its electrostatic component. In particular, this is observed for charge-assisted hydrogen bonds, also termed salt bridges.[2,47] The contributions of salt bridges on the solvent-exposed surface of a binding site to the binding free energy, for example, are often negligible due to the high desolvation cost of the charged protein residue.[2,48] Hydrogen bonds to the protein backbone are generally stronger and often less affected by side chain mutations compared to hydrogen bonds formed with side chain atoms.[2,7] The loss of entropy of the protein is typically smaller for a backbone hydrogen bond because side chains are inherently more flexible.[49] The rigidification of the ligand−protein complex upon formation of a hydrogen bond is coupled to reduced ligand and protein mobility and, therefore, leads to a reduction in entropy. This can be exploited by establishing multiple synergistic hydrogen bonds by modifying a rigid ligand moiety to decrease the effects of entropy−enthalpy compensation.[2] Thus, the location of hydrogen bonds is essential to achieve potent ligand binding and the effect of protein environment on this type of interaction should be carefully considered during visual analysis. In addition to formed hydrogen bonds, all unsatisfied hydrogen bond donor and acceptor groups without interacting counterparts, in particular those buried in the binding site, need to be evaluated, as they generally result in reduced binding free energies due to an uncompensated desolvation free energy penalty.[2,21] In metalloproteins, distinct electrostatic interactions with metal cofactors can be of major importance for ligand potency (Figure 1A).[22,50,51] Interestingly, none of the studies reviewed in this perspective reported the consideration of weak hydrogen bonds, halogen bonds, orthogonal dipole interactions, or amide−π interactions as a criterion during visual inspection (see also survey results for additional details).

Solvation effects are a major contributor to the protein−ligand binding thermodynamics. During ligand binding, water molecules are displaced from the solvated cavities.[2] While the displacement of water molecules is favorable due to an increase in entropy and sometimes enthalpy, some may form enthalpically strong interactions with the protein and bridge ligand−protein interactions favorably contributing to the binding free energy. Thus, solvation effects such as water-mediated interactions should be considered in visual inspection (Figure 1B,C).[22,24] However, the proper assessment of the multifaceted water behavior is by far not trivial even for extremely advanced experts. Moreover, the solvation scenario in the binding site in the apo state (no ligand present) is not routinely included in the visual comparison. Additional software for estimating the enthalpic and entropic contributions for the displacement of water molecules, for example, based on MD simulations[39,52,53] or derived neural networks,[54] is available and can augment docking and scoring. Another related evaluation criterion is the nature of solvent-exposed ligand atoms, as hydrophobic groups protruding into the surrounding solvent are generally considered as unfavorable if they otherwise would have the chance to interact with the protein.[21,28]

In the formation of most complexes between ligand and protein, both entities have to adapt their configuration for optimizing their complementarity. It was estimated that 25% of all proteins undergo a certain degree of induced fit during ligand binding.[7] A subsequent study of 188 proteins in the Protein Data Bank, with both holo and apo structures available, reported a much larger fraction of 90% of the proteins undergoing rotamer changes upon binding.[55] This energetic cost upon deformation of either the ligand or the protein is termed strain, and its energetic evaluation is typically not optimally represented in the scoring function. Furthermore, due to the often rigid representation of the protein in molecular docking calculations employed for large-scale virtual screening, both binding partners cannot properly adapt to each other. Both shortcomings can result in unrealistic internal geometries of ligands (Figure 1D)[7,11,56] or overly close

https://dx.doi.org/10.1021/acs.jmedchem.0c02227
J. Med. Chem. 2021, 64, 2489−2500

contacts. For example, twisted amide or ester groups were observed, leading to the rejection of respective binding modes during visual inspection.[3,22,56,57] Thus, ligand and protein strain should be accounted for in a thorough visual inspection. Because protein strain can be considered only if the protein is treated flexibly, flexible docking campaigns and MD-based postprocessing should be used to take protein and ligand conformational adaptations into account.[11] However, this may still be computationally too expensive for a large number of ligand−protein complexes as is usually the case in virtual screening tasks. In lead optimization, however, the use of such methods often becomes accessible due to the smaller number of compounds.

Preexisting knowledge of a protein−ligand complex may assist in determining the quality of a binding pose. In the most direct way, available crystal structures of the investigated target protein in complex with known ligands could guide the detection of common or especially relevant ligand−protein interactions. It was shown that the training of docking protocols with the native ligand improved the accuracy of pose prediction. In addition, as homologous targets often share conserved ligands and binding modes,[58−60] crystal structures of related proteins with a similar co-crystallized ligand can be included in the decision making process.

CASE STUDIES FOR APPLICATIONS OF VISUAL INSPECTION IN COMPUTER-AIDED DRUG DESIGN

As mentioned above, visual inspection has been applied in a vast number of drug discovery projects. In most cases, the inspection was conducted as a final step of compound prioritization before experimental testing.[22,25,44,62] In Table 1, we summarized several studies that provided a detailed description of the criteria used for visual inspection and that led to the discovery of potent ligands. As mentioned in the previous section, the most commonly assessed feature in visual inspection was complementarity between the ligand and respective binding site, and the presence of ligand−protein interactions such as hydrophobic interactions and hydrogen bonds. Furthermore, interactions with specific binding site residues or cofactors were often accounted for. This was typically motivated by interactions observed between co-crystallized ligands and protein[44] or based on conserved interactions among various ligands docked to the same target protein.[28,43] Distorted ligand conformations were also frequently considered as an exclusion criterion for docking poses as they represent unrealistic ligand strain, for example, distortions of amide bonds, esters, or unfavorable *cis*−*trans* isomery.[2,10,21,24,30]

On the contrary, features such as interactions with conserved or crystallographic water molecules as well as interactions relevant for ligand selectivity were rarely mentioned in the considered studies.[22,24,31] Chemical diversity, novelty, and commercial availability of the best scored ligands were not a direct part of visual inspection, but rather criteria for postprocessing of docking results in general.[10,24,31,36,44,61]

Most drug design studies that mention visual inspection did not detail the criteria used in this process, despite being an essential component of the selection process. A more detailed description of criteria relevant for the decision process would be desirable as it allows for reproducibility of the results, a common requirement for publishing other elements of computer-aided drug design (CADD) studies.

In the following, we highlight some recent studies that demonstrate the usability of visual inspection of binding modes. For example, a large-scale virtual screening study discovered ligands binding to AmpC β-lactamase and D_4 dopamine receptor using a computational library consisting of 170 million compounds.[3] The authors not only used visual inspection for the discovery of highly potent ligands with novel chemotypes but also compared compounds prioritized with or without prior visualization. Interestingly, the latter analysis revealed that ligands selected on the basis of visual inspection generally presented superior binding affinity in the following experiments, confirming the usefulness of such procedures.

The Drug Design Data Resource (D3R) regularly invites experts in the field of computational chemistry and CADD to participate in a challenge to predict ligand poses, rank compounds, and predict relative binding free energies in a blinded fashion.[13,14,63,64] While the best-performing results in the second iteration of the challenge did not employ visual inspection, top submissions in the remaining challenges regularly relied on human intervention (Table 2). Similarly,

Table 2. Direct Evaluation of Visual Inspection

outcome	targets	challenge
two best submissions employed visual inspection	JNK3 kinase	SAMPL1[65]
best submission employed visual inspection	HIV integrase	SAMPL4[65]
four of 11 best and none of the 28 worst submissions employed visual inspection	HSP90, MAP4K4	D3R 2015[64]
none of the top submissions employed visual inspection	farnesoid X receptor	D3R #2[63]
all but two top submissions employed visual inspection, but only one of the worst submissions	cathepsin S, VEGFR2, JAK2, p38-α, TIE2, ABL1	D3R #3[14]
visual inspection augmented docking results in six of 10 top submissions	β-secretase 1, cathepsin S	D3R #4[13]
improved ligand affinity with visual inspection	D_4 dopamine receptor, AmpC β-lactamase	none[3]

the best performing submissions in the preceding SAMPL1 challenge reported visual inspection.[65] A study published by participants of the SAMPL4 challenge provided details on their decision process and retrospectively evaluated their predictions with and without human intervention. The best hit rate was achieved when visual inspection was coupled with rescoring.[66] Before D3R, the Community Structure−Activity Resource (CSAR) held similar challenges, but visual inspection was only briefly mentioned and could not be compared to submissions without human intervention.[67]

LIMITATIONS OF HUMAN INTERVENTION

Even though the studies mentioned above highlight the advantages of human intervention in structure-based drug discovery projects, this practice also suffers from several limitations. Most obviously, the expertise and intuition of the individuals or teams conducting visual inspection strongly influence its outcome,[13,16,35] as we further substantiate in the evaluation of our survey. Another limitation is the number of binding modes that can be visually inspected in a given amount of time.[68] While some studies report the inspection of <100 binding modes or compounds,[69,70] in the majority of the studies between 100 and 300 compounds were inspected.[37,44,61,71] Some studies reported the examination of ≤1500 compounds.[3,34,57,72−74] The latter likely presents the

https://dx.doi.org/10.1021/acs.jmedchem.0c02227
J. Med. Chem. 2021, 64, 2489−2500

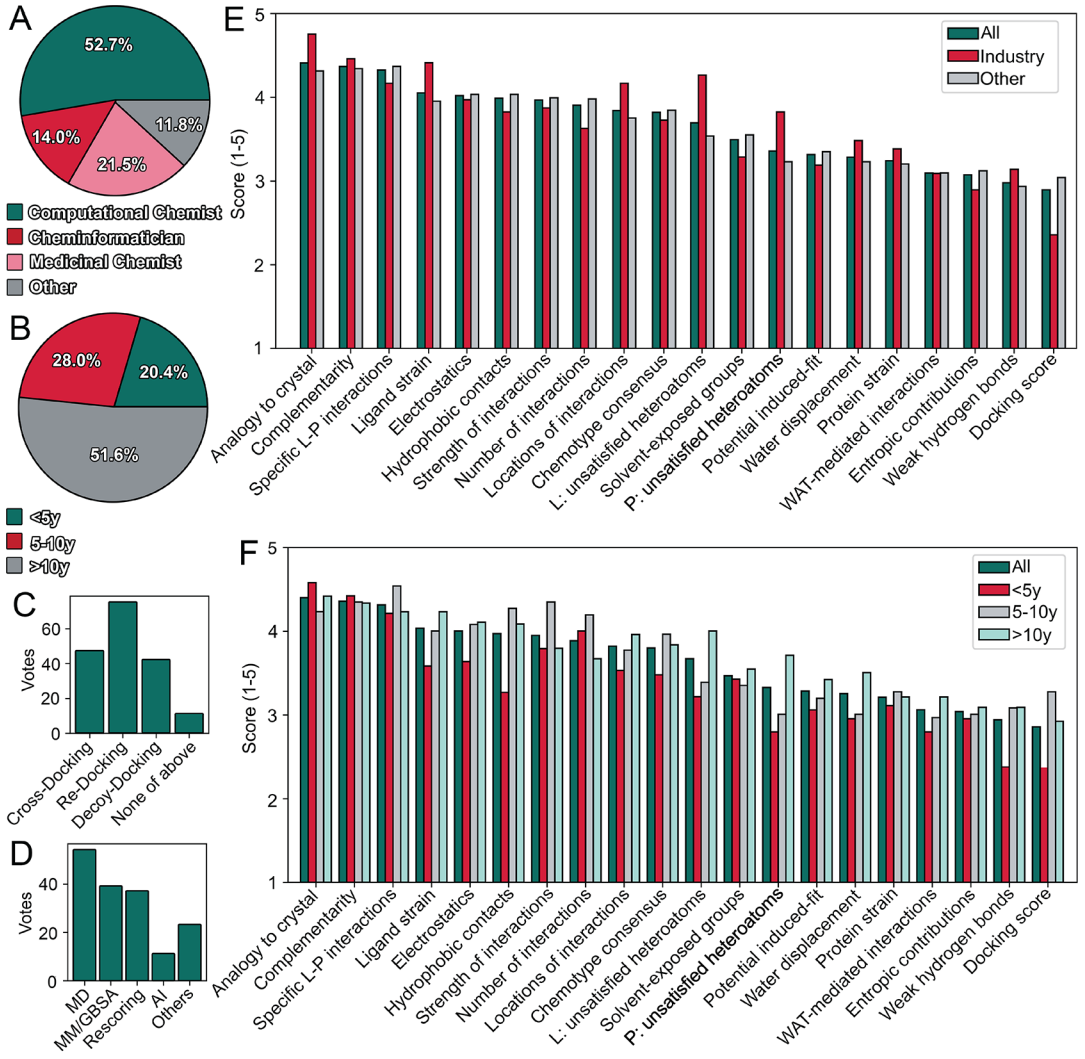


Figure 2. Results of the survey. (A) Participants of the survey divided according to their current scientific position. (B) Participants of the survey divided according to years of experience in computer-aided molecular design. (C) Procedures taken by all participants to validate the used docking protocols. (D) Methods of all participants used for postprocessing of docking results. (E) Importance of different criteria valued by all participants, participants from the pharmaceutical industry, and the remaining participants. (F) Importance of different criteria valued by all participants divided according to their years of experience compared to the overall result.

upper limit for the application of visual inspection as it becomes highly time-consuming and more importantly difficult to consistently evaluate such a large number of compounds. As the size of compound libraries that are screened in a single project has grown continuously with recent reports of multiple billion compounds,[75,76] the visual inspection might be unfeasible or the hit list has to be narrowed down with other techniques. However, the assessed studies did not state the time frame in which the visual inspection was conducted.

Furthermore, the underlying docking methodology harbors inherent limitations that cannot be compensated by visual inspection alone. For example, it is known that protein flexibility plays an important role in ligand binding. Despite progress in flexible docking engines, most large-scale docking campaigns are still performed without including protein flexibility and structural adaptations. Thus, visual inspection is also based on a static view of the formed complex analogous to a crystal structure.[5,7,77−79] Furthermore, loss of configurational entropy or desolvation free energy is difficult to evaluate by visual inspection.[2] Classical scoring functions, however, primarily evaluate enthalpic contributions using a simplified mathematical representation of interaction terms relevant to ligand−protein binding. Entropic effects, poorly represented in scoring functions, are mostly limited to the simplistic evaluation of ligand desolvation (e.g., hydrophobic contacts) and ligand configurational entropy (e.g., number of rotatable bonds).[5,80−83] This improper handling of binding thermodynamics presents an important limitation of docking procedures as, for example, compensatory entropic contributions can significantly reduce binding affinity.[2] As visual inspection alone cannot properly investigate those effects, it needs to be coupled with a more dynamic representation of ligand−protein binding, for example, MD simulations. MD simulations can intrinsically account for protein and ligand flexibility, and estimates based on trajectory postprocessing may incorporate entropic effects and thus provide a more precise estimation of free energies of binding compared to simple scoring functions. MD simulation can also be used to provide time-evolved insights into the stability of the ligand−protein complex.[79,84] Studies showed that MD simulations combined with MM/GBSA calculations or derivations thereof can improve the binding affinity prediction of poses obtained from docking

https://dx.doi.org/10.1021/acs.jmedchem.0c02227
J. Med. Chem. 2021, 64, 2489−2500

protocols, although the outcomes are highly case specific.[85−88] Currently, free energy perturbations (FEP) are considered the most accurate methodology for computing binding free energies with an accuracy of approximately 1 kcal/mol.[89,90] Despite recent developments in the improved treatment of solvation effects[39,91] and its integration in neural network postprocessing methods, standard docking programs can typically reproduce only ∼60% of native binding modes of a ligand within an RMSD threshold of 2 Å using the best scored pose as demonstrated, for example, in a large-scale evaluation of 10 docking programs.[83] This underlines the necessity of a visual inspection to discard non-native poses.

SURVEY OF COMPUTATIONAL MEDICINAL CHEMISTS

In the Introduction, we recalled a significant disagreement among medicinal chemists regarding compound prioritization that is not surprising considering the high dimensionality of the problem.[35] To obtain more detailed insight into the current state of the art in visual pose inspection, we conducted a survey of experts in the field from both academia and the pharmaceutical industry. In total, we managed to collect 93 expert opinions. The survey consisted of three parts: (i) information about the participant such as profession and experience, (ii) general questions on visual inspection, used protocols, and decision criteria, and (iii) a challenge to recognize the native binding mode of three ligand−protein complexes. Potential participants for the survey were individually contacted by Email or LinkedIn. All participants have recently published in the area of structure-based drug discovery in peer-reviewed journals. The survey form was composed using a designated platform provided by Google.[92] In addition, participants were encouraged to share the survey with experienced colleagues.

A majority of the responders were from academia (71, 76.3%), followed by 20 (21.5%) experts from the pharmaceutical industry, while only two were from nonprofit organizations or government institutions. The largest group of participants consisted of computational chemists followed by medicinal chemists, and cheminformaticians (Figure 2A). More than half of the participants reported more than 10 years of experience in the field of computer-aided molecular design (Figure 2B). As a conservative summation, this amounted to >600 years of experience. The participants rated the importance of visual inspection during a drug design project with 4 (29.0%) or 5 points (64.5%) on a scale from 1 to 5 underlining the popularity of such procedures.

Prior to perspective docking and visual inspection, the validation of the used docking protocol to retrospectively assess pose prediction accuracy as well as the capability to distinguish actives from decoy compounds is generally recommended.[93] Whereas most participants (80.6%) reported that they performed redocking of co-crystallized ligands into the X-ray protein structure, only around half of them reported cross-docking of multiple ligands to the selected protein structures or using decoys to determine enrichment metrics (Figure 2C). Surprisingly, 11 participants (evenly spread between academia and industry) do not conduct any of the validation steps mentioned above.

Most participants do not rely on docking as a sole computational methodology. Typically, they perform post-processing of docking results, for example, by using MD simulations, likely to cope with the limited treatment of protein flexibility in standard docking protocols and evaluate the time-evolved stability of the obtained poses (Figure 2D).

With respect to visual inspection, a majority of participants conduct both individual assessments by themselves combined with an inspection in a team setting to discuss the results with colleagues. Around one-third of the participants perform visual inspection only by themselves, while 10 participants (10.8%) assess binding modes only with colleagues. Interestingly, when we subdivided the participants into the academic and private sector, a larger share of only 12.7% in the industry assessed binding modes with colleagues than in academia (4.6%). The integration of individual and collective opinions among a board of colleagues capitalizes on both the experience of single researchers as well as the excellence of a team, improving the quality of the inspection.

Whereas the options to exploit virtual reality (VR) in a CADD setting improved in recent years, which offers new possibilities for humans to interact with 3D data,[94] its use among experts is, to date, still highly limited (Figure 3A). More than 88% of the participants in our survey conduct visual inspection on a regular desktop computer screen. The use of 3D screens or 3D beamers at the interface of VR and regular screens was reported by 21.5% of participants. With regard to the time point at which visual inspection is conducted within the drug discovery pipeline, its popularity decreases from hit identification to the final drug candidate (Figure 3).

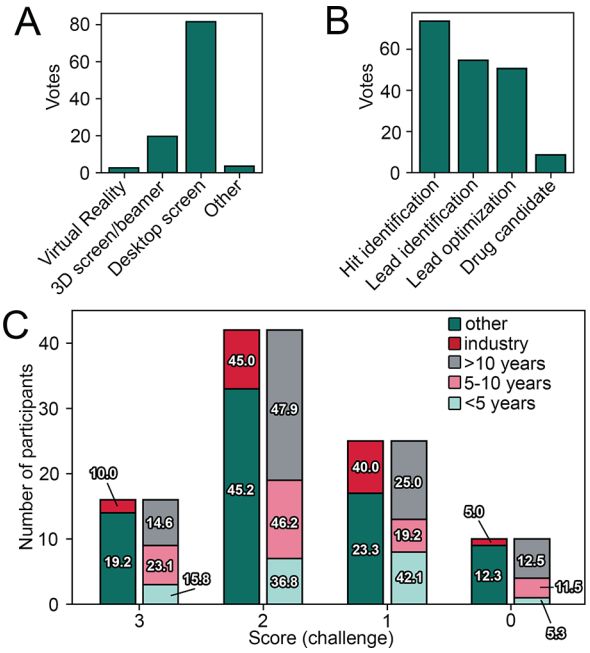


Figure 3. (A) Hardware used for visualization by the participants. (B) Tasks for which visual inspection is applied by the participants within a discovery pipeline. (C) Number of correct answers (score) of the participants divided according to affiliation and experience. The percentages indicate the fraction of the total number of participants categorized into a group (e.g., participants from the industry) who achieved a certain score.

In the survey, we evaluated the importance of different criteria used for the assessment of binding poses (Figure 2E,F). Consistent with the published studies previously discussed, shape complementarity was valued among the most relevant criteria independent of experience or affiliation, probably as it is simple to deduce at first sight. The analogy of predicted

https://dx.doi.org/10.1021/acs.jmedchem.0c02227
J. Med. Chem. 2021, 64, 2489−2500

binding poses to co-crystallized ligands in experimentally determined crystal structures was rated to be of similar importance. Interestingly, industrial scientists prioritized these assessments, potentially due to the better availability of such data in this setting. Ligand−protein interactions with specific binding site residues were valued highly in both our survey and the considered publications. Interestingly, ligand strain, the location of ligand−protein interactions with respect to the binding site, and unsatisfied polar ligand or protein heteroatoms were prioritized more highly by scientists in the pharmaceutical industry than in academia. Unsatisfied protein hydrogen bond-donating or -accepting groups were valued more highly by both the most experienced participants and those working in the pharmaceutical industry. However, our questionnaire did not distinguish between unsatisfied hydrogen bonding groups in different locations of the binding site such as deeply buried groups for which the desolvation penalty would have to be compensated by favorable interactions as opposed to solvent-exposed polar groups, which could favorably interact with the solvent. Solvation effects such as the displacement of water molecules during ligand−protein association were regarded more highly by the most experienced participants. Most remarkably, docking scores produced by the scoring functions were valued the least, especially by experts in the pharmaceutical industry. Seemingly, the known unreliability of scoring and ranking by docking algorithms affected its popularity among experts.

All statements regarding differences among the preference of certain criteria between different participant groups display statistical significance based on a two-sided t test at a $p = 0.05$ level. The scipy python package was used for those tests. Scientists involved in CADD projects, especially those new to the field, will likely profit from the results of our survey as it provides guidance and allows for a comparison of their own procedures and criteria with other scientists in the field.

BINDING MODE PREDICTION CHALLENGE

As mentioned above, our survey included three challenges focused on the identification of native binding modes, where "native" refers to the pose present in the crystal structure. The first example is the complex of (5-pyridin-3-ylfuran-2-yl)-methanamine bound to cytochrome P450 2A6 (CYP2A6), in which a terminal amine group is coordinated to the heme iron (Table 3). Due to the electrostatic interaction in particular with the metal ion, the pK_a value is significantly decreased compared to that of the unbound state of the ligand, forcing the amine to be uncharged in the complex. In addition, the protonation state of the amine might have been shifted to the neutral state due to a pH value of 8.5 in the crystallization medium as annotated in the Protein Data Bank. To obtain a non-native pose with a feasible interaction profile, we manually reversed the orientation of the ligand where the pyridine nitrogen is involved in metal coordination (Figure 4A). Surprisingly, the non-native orientation of the ligand was assumed as the correct pose by a slight majority of participants. In the survey, we offered the option to give a justification for the selection for all three examples. As expected, the rather unconventional protonation state of the amine was often referred to when the non-native complex was selected as the answer by the participants. Furthermore, heterocyclic nitrogen-containing ring moieties such as imidazoles and pyridines are classically regarded as coordination partners of the heme iron in CYPs.[95] The binding mode of the compound with the amine coordinated to the heme iron shows how the acido-basic character of the ligand—here the basicity, which is higher at the aliphatic amine group than at the aromatic ring nitrogen—can influence the interaction with metal ions and, thus, the binding mode of a compound.[96]

Table 3. Results of the Challenge and Machine-Based Ranking

	case 1	case 2	case 3
correct[a]	46.2%	63.4%	59.1%
wrong[a]	53.8%	36.6%	40.9%
Glide docking[b]	correct	wrong	correct
smina docking[b]	correct	wrong	correct
MM/GBSA[b]	correct	correct	inconclusive[c]

[a]Outcome of the survey regarding the respective case of the challenge. [b]Statement if ranking based on the computation of binding free energies using two different docking programs and MM/GBSA calculations did prioritize the correct pose. [c]No statistically significant difference between the calculated energies of the native and incorrect poses.

Thus, this case study represents an example in which the intuition and experience of most scientists deceived them from selecting the correct result. This was underlined by the fact that less experienced scientists (fewer than five years of experience) were selecting the correct pose more frequently than the more experienced scientist. Participants selecting the native binding mode mostly considered improved ligand−protein interactions, including the hydrogen bond of the pyridine ring as well as the unfavorable placement of the protonated amine in a hydrophobic region of the pocket. In contrast to the remaining examples, conventional scoring functions as well as MM/GBSA calculations outperformed the human prioritization in this case. The ligand protonation in the input structures for these scoring procedures was submitted as shown in Figure 4 with the native pose in the neutral form, while the terminal amine was protonated in the non-native complex. The largest differences among the individual components of the MM/GBSA term between the native and non-native poses were in the electrostatic and lipophilic contributions.

The second example consisted of two similar poses of 1-(5-methylthiophen-2-yl)-3-pyridin-3-ylurea bound to the main protease of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) in two slightly different orientations, of which one was obtained by using the Glide SP docking protocol (Figure 4B).[97] Both binding poses form two hydrogen bonds: the native pose interacts with the protein backbone as opposed to a flexible side chain in the non-native one. In this case, 63.4% of the participants correctly identified the native pose, the highest percentage among all three case studies. In contrast, both tested docking programs prioritized the non-native binding mode, supporting the previous assumption that expert knowledge can outperform conventional scoring functions.[12] Participants prioritizing the native pose frequently reported shape complementarity as well as the more stable hydrogen bond interaction with the protein backbone to be decisive for their choice. On the contrary, participants selecting the non-native pose mostly argued on the basis of the internal strain of the ligand and frequently referred to the conformation of the aminopyridine moiety. This example further underlined the benefit of MM/GBSA calculations, as they correctly

https://dx.doi.org/10.1021/acs.jmedchem.0c02227
J. Med. Chem. 2021, 64, 2489−2500

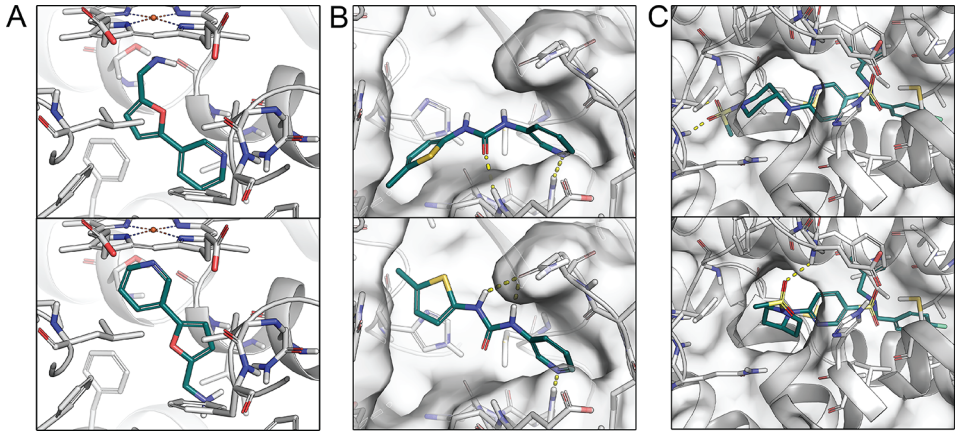


Figure 4. Structures that were included in the survey to distinguish native poses from artifacts. In the survey, the structures were presented in form of an animated GIF image, showing all relevant interactions. (A) Case 1: (5-pyridin-3-ylfuran-2-yl)methanamine bound to CYP2A6 (PDB entry 2FDW). The native pose is shown at the top, while a manually composed pose is shown at the bottom. (B) Case 2: 1-(5-methylthiophen-2-yl)-3-pyridin-3-ylurea bound to the main protease of SARS-CoV-2 (PDB entry 5RH0). The native pose is shown at the top, while a slightly different pose from docking is displayed at the bottom. (C) Case 3: *N*-[(4-fluorophenyl)methyl]-*N*-(2-methylpropyl)-6-[(1-methylsulfonylpiperidin-4-yl)amino]pyridine-3-sulfonamide bound to RAR-related orphan receptor γ (PDB entry 4WLB). The native pose is shown at the top, while a slightly different pose from docking is displayed at the bottom.

preferred the native pose as opposed to the used docking methods. This advantage originates from the use of time-evolved MD trajectories in combination with the more accurate treatment of solvation free energies.[89]

The third challenge consisted of an inverse agonist bound to RAR-related orphan receptor γ (RORγ) in two similar binding modes regarding the aromatic moieties of the compound. However, the sulfonamide moiety of the ligand is buried within the core of the receptor in the native binding mode, while it is partially solvent-exposed in the non-native docking pose we selected (Figure 4C). The native pose is preferred due to better hydrophobic complementarity and a more favorable hydrogen bonding network. Interestingly, even though the majority of participants prioritized the native pose, still more than 40% prioritized the non-native pose. As expected, the selection of the correct pose was frequently guided by the improved hydrogen bonding network and the increased extent of burial of the ligand and the two hydrogen bonds. The non-native pose was associated with unfavorable conformational strain and thus disregarded from being the native binding mode when the correct pose was selected. In contrast, the non-native pose was argued to be more favorable due to the solvent-exposed polar sulfonamide. While both docking programs were correctly predicting the native pose, MM/GBSA predicted similar free energies of binding for both poses. The latter might indicate multiple favorable binding modes, but the ligand atoms are well resolved according to the electron density in the accompanying crystal structure. The statistical significance of the MM/GBSA binding free energies, based on the quantification of 125 frames of the last quarter of a 20 ns MD simulation, was evaluated at a $p = 0.05$ level using the scipy python package. While the results of cases 1 and 3 were statistically significant, the average values in case 2 were only significant at a $p = 0.07$ level.

CONCLUSIONS

Due to the inherent drawbacks of molecular docking and pose scoring, visual inspection of predicted protein−ligand complexes was performed in a large number of drug discovery projects. Frequently, only limited insight into the actual decision making process was provided. We reviewed the available medicinal chemistry literature to deduce the applied criteria and summarize the studies that directly evaluated the gain or loss of prediction accuracy upon visual inspection of the docked binding modes. These comparison studies revealed that in most cases human intervention was improving the prediction performance, with only one study showing no improvement. The most frequently mentioned criteria included shape complementarity, hydrogen bonds, and hydrophobic contacts. Furthermore, we discussed the limitations of human interventions, which include the number of poses that can be accurately assessed, the different levels of experience, and the given limitations of the underlying methodology such as the frequent lack of protein flexibility or detailed consideration of entropy. To gain more insight into the decision making of experts in the field, we conducted a survey including a challenge to distinguish native from non-native poses in three case studies. While there was considerable disagreement among the participants, a majority detected the native pose in two of three cases. In the remaining case, in which an uncharged primary amine instead of a heterocyclic nitrogen of an inhibitor undergoes an interaction with the heme iron in CYP2A6, prior experience with CYP inhibitors negatively affected the detection of the native pose, demonstrating that human bias might exist, too. The individual inspection followed by a discussion with colleagues in a team setting, as the majority of participants reported it, can likely cope with such biases and improve the outcome. Similar to the literature review, the surveyed scientists specified shape complementarity, analogy to binding modes observed in co-crystallized complexes, and specific ligand−protein interactions as the most valuable criteria. Among scientists active in an industrial setting, ligand strain and the number of unsatisfied ligand or protein heteroatoms were valued comparatively highly. Docking scores, which are known to be frequently inaccurate, were valued least important among the available criteria. This suggests that the journey to reliable scoring functions is by far not over, as today's scoring functions are often no match for the complex knowledge and vast experience of computational medicinal chemists. Overall, many factors

https://dx.doi.org/10.1021/acs.jmedchem.0c02227
J. Med. Chem. 2021, 64, 2489−2500

have to be considered during visual inspection, and this perspective can support the decision making of experts as well as scientists new to the field. On the basis of the survey results, we would like to motivate scientists reviewing docking studies in peer-reviewed journals to routinely insist on docking protocol validation as well as on inclusion of visual analyses and other postprocessing methods.

AUTHOR INFORMATION

Corresponding Author

Markus A. Lill − *Computational Pharmacy, Department of Pharmaceutical Sciences, University of Basel, 4056 Basel, Switzerland;* orcid.org/0000-0003-3023-5188; Email: markus.lill@unibas.ch

Authors

André Fischer − *Computational Pharmacy, Department of Pharmaceutical Sciences, University of Basel, 4056 Basel, Switzerland*

Martin Smieško − *Computational Pharmacy, Department of Pharmaceutical Sciences, University of Basel, 4056 Basel, Switzerland;* orcid.org/0000-0003-2758-2680

Manuel Sellner − *Computational Pharmacy, Department of Pharmaceutical Sciences, University of Basel, 4056 Basel, Switzerland*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jmedchem.0c02227

Notes

The authors declare no competing financial interest.

Biographies

André Fischer obtained his Master's Degree in Drug Sciences from the University of Basel in 2018, completing it with his work on cytochrome P450 2D6 pharmacogenomics. In the same year, he started his Ph.D. studies in the research group of Computational Pharmacy at the University of Basel where he is currently active. His work includes contributions in the field of computational chemistry with regard to virtual screening, compound selectivity, and ligand-induced conformational changes.

Martin Smieško graduated in medicinal chemistry from the Faculty of Pharmacy of Comenius University (Bratislava, Slovakia) in 1999. In 2003, he completed his Ph.D. at the same university. In 2006, he joined the Molecular Modeling group of Prof. Angelo Vedani at the University of Basel as a research associate. In 2014, he was promoted to Senior Lecturer. He made various contributions to the field of computer-aided drug design and off-target binding prediction, including the development of a program for flexible docking.

Manuel Sellner obtained his Master's Degree in Drug Sciences following his Bachelor's Degree in Pharmaceutical Sciences at the University of Basel. He completed his studies with his work on the interactions between cytochrome P450 2D6 and its corresponding reductase. In early 2020, he initiated his Ph.D. studies in the research group of Computational Pharmacy at the University of Basel, where he works on the development of tools for predicting off-target binding in a structure-based approach.

Markus A. Lill received his Diploma in Physics from the University of Erlangen-Nürnberg and in 2002 his Ph.D. from the Max Planck Institute for Biophysics (Frankfurt, Germany). He completed four years of postdoctoral training at the Biographics Laboratory 3R and his Habilitation in 2006 at the University of Basel. He then became an Assistant Professor and 2012 Associate Professor for Computational Drug Design at Purdue University (West Lafayette, IN). Currently, he is leading the Computational Pharmacy group at the University of Basel, where his research is focused on the development and application of computational methods for drug discovery with a deep focus on physics-based artificial intelligence.

ACKNOWLEDGMENTS

The authors thank everyone who participated in our survey. In addition, the authors gratefully acknowledge the expert knowledge provided by Dr. Christian Kramer (Roche) and Dr. Andreas Bergner (Boehringer Ingelheim) supporting the design of the survey.

ABBREVIATIONS

D3R, Drug Design Data Resource; MM/GBSA, molecular mechanics/generalized Born surface area; RANKL, receptor activator of nuclear factor κ-B ligand; RORγ, RAR-related orphan receptor γ; SAMPL, Statistical Assessment of Modeling of Proteins and Ligands; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; VR, virtual reality

REFERENCES

(1) Seeliger, D.; de Groot, B. L. Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *J. Comput.-Aided Mol. Des.* **2010**, *24*, 417−422.

(2) Bissantz, C.; Kuhn, B.; Stahl, M. A medicinal chemist's guide to molecular interactions. *J. Med. Chem.* **2010**, *53*, 6241−6241.

(3) Lyu, J.; Wang, S.; Balius, T. E.; Singh, I.; Levit, A.; Moroz, Y. S.; O'Meara, M. J.; Che, T.; Algaa, E.; Tolmachova, K.; Tolmachev, A. A.; Shoichet, B. K.; Roth, B. L.; Irwin, J. J. Ultra-large library docking for discovering new chemotypes. *Nature* **2019**, *566*, 224−229.

(4) Waszkowycz, B. Towards improving compound selection in structure-based virtual screening. *Drug Discovery Today* **2008**, *13*, 219−226.

(5) Cappel, D.; Jerome, S.; Hessler, G.; Matter, H. Impact of different automated binding pose generation approaches on relative binding free energy simulations. *J. Chem. Inf. Model.* **2020**, *60*, 1432−1444.

(6) Kroemer, R. T. Structure-based drug design: Docking and scoring. *Curr. Protein Pept. Sci.* **2007**, *8*, 312−328.

(7) Mobley, D. L.; Dill, K. A. Binding of small-molecule ligands to proteins: "What you see" is not always "what you get. *Structure (Oxford, U. K.)* **2009**, *17*, 489−498.

(8) Lexa, K. W.; Carlson, H. A. Protein flexibility in docking and surface mapping. *Q. Rev. Biophys.* **2012**, *45*, 301−343.

(9) Damm-Ganamet, K. L.; Bembenek, S. D.; Venable, J. W.; Castro, G. G.; Mangelschots, L.; Peeters, D. C. G.; Mcallister, H. M.; Edwards, J. P.; Disepio, D.; Mirzadegan, T. A prospective virtual screening study: Enriching hit rates and designing focus libraries to find inhibitors of PI3Kδ and PI3Kγ. *J. Med. Chem.* **2016**, *59*, 4302−4313.

(10) Lionta, E.; Spyrou, G.; Vassilatis, D.; Cournia, Z. Structure-based virtual screening for drug discovery: Principles, applications and recent advances. *Curr. Top. Med. Chem.* **2014**, *14*, 1923−1938.

(11) Cournia, Z.; Allen, B. K.; Beuming, T.; Pearlman, D. A.; Radak, B. K.; Sherman, W. Rigorous free energy simulations in virtual screening. *J. Chem. Inf. Model.* **2020**, *60*, 4153.

(12) Schulz-Gasch, T.; Stahl, M. Scoring functions for protein-ligand interactions: A critical perspective. *Drug Discovery Today: Technol.* **2004**, *1*, 231−239.

(13) Parks, C. D.; Gaieb, Z.; Chiu, M.; Yang, H.; Shao, C.; Walters, W. P.; Jansen, J. M.; McGaughey, G.; Lewis, R. A.; Bembenek, S. D.; Ameriks, M. K.; Mirzadegan, T.; Burley, S. K.; Amaro, R. E.; Gilson, M. K. D3R grand challenge 4: Blind prediction of protein-ligand poses, affinity rankings, and relative binding free energies. *J. Comput.-Aided Mol. Des.* **2020**, *34*, 99−119.

(14) Gaieb, Z.; Parks, C. D.; Chiu, M.; Yang, H.; Shao, C.; Walters, W. P.; Lambert, M. H.; Nevins, N.; Bembenek, S. D.; Ameriks, M. K.; Mirzadegan, T.; Burley, S. K.; Amaro, R. E.; Gilson, M. K. D3R grand challenge 3: Blind prediction of protein-ligand poses and affinity rankings. *J. Comput.-Aided Mol. Des.* **2019**, *33*, 1−18.
(15) Davis, A. M.; Teague, S. J.; Kleywegt, G. J. Application and limitations of X-ray crystallographic data in structure-guided ligand and drug design. 2007. http://europepmc.org/abstract/CTX/C7868.
(16) Ripphausen, P.; Stumpfe, D.; Bajorath, J. Analysis of structure-based virtual screening studies and characterization of identified active compounds. *Future Med. Chem.* **2012**, *4*, 603−613.
(17) Kumar, A.; Chaturvedi, V.; Bhatnagar, S.; Sinha, S.; Siddiqi, M. I. Knowledge based identification of potent antitubercular compounds using structure based virtual screening and structure interaction fingerprints. *J. Chem. Inf. Model.* **2009**, *49*, 35−42.
(18) Rodríguez-Pérez, R.; Miljković, F.; Bajorath, J. Assessing the information content of structural and protein-ligand interaction representations for the classification of kinase inhibitor binding modes via machine learning and active learning. *J. Cheminf.* **2020**, *12*, 36.
(19) Varela-Rial, A.; Majewski, M.; Cuzzolin, A.; Martínez-Rosell, G.; de Fabritiis, G. SkeleDock: A web application for scaffold docking in PlayMolecule. *J. Chem. Inf. Model.* **2020**, *60*, 2673−2677.
(20) Velázquez-Libera, J. L.; Durán-Verdugo, F.; Valdés-Jiménez, A.; Núñez-Vivanco, G.; Caballero, J. LigRMSD: A web server for automatic structure matching and RMSD calculations among identical and similar compounds in protein-ligand docking. *Bioinformatics* **2020**, *36*, 2912−2914.
(21) Perryman, A. L.; Yu, W.; Wang, X.; Ekins, S.; Forli, S.; Li, S.-G.; Freundlich, J. S.; Tonge, P. J.; Olson, A. J. A virtual screen discovers novel, fragment-sized inhibitors of Mycobacterium tuberculosis InhA. *J. Chem. Inf. Model.* **2015**, *55*, 645−659.
(22) Viviani, L. G.; Piccirillo, E.; Ulrich, H.; Amaral, A. T.-d. Virtual screening approach for the identification of hydroxamic acids as novel human ecto-5-nucleotidase inhibitors. *J. Chem. Inf. Model.* **2020**, *60*, 621−630.
(23) Toledo-Sherman, L.; Deretey, E.; Slon-Usakiewicz, J. J.; Ng, W.; Dai, J.-R.; Foster, J. E.; Redden, P. R.; Uger, M. D.; Liao, L. C.; Pasternak, A.; Reid, N. Frontal affinity chromatography with MS detection of EphB2 tyrosine kinase receptor. 2. identification of small-molecule inhibitors via coupling with virtual screening. *J. Med. Chem.* **2005**, *48*, 3221−3230.
(24) Postigo, M. P.; Guido, R. V. C.; Oliva, G.; Castilho, M. S.; da R. Pitta, I.; de Albuquerque, J. F. C.; Andricopulo, A. D. Discovery of new inhibitors of Schistosoma mansoni PNP by pharmacophore-based virtual screening. *J. Chem. Inf. Model.* **2010**, *50*, 1693−1705.
(25) Caporuscio, F.; Rastelli, G.; Imbriano, C.; del Rio, A. Structure-based design of potent aromatase inhibitors by high-throughput docking. *J. Med. Chem.* **2011**, *54*, 4006−4017.
(26) Li, H.; Xiao, H.; Lin, L.; Jou, D.; Kumari, V.; Lin, J.; Li, C. Drug design targeting protein-protein interactions (PPIs) Using multiple ligand simultaneous docking (MLSD) and drug repositioning: Discovery of raloxifene and bazedoxifene as novel inhibitors of IL-6/GP130 interface. *J. Med. Chem.* **2014**, *57*, 632−641.
(27) Hou, X.; Li, K.; Yu, X.; Sun, J.-p.; Fang, H. Protein flexibility in docking-based virtual screening: Discovery of novel lymphoid-specific tyrosine phosphatase inhibitors using multiple crystal structures. *J. Chem. Inf. Model.* **2015**, *55*, 1973−1983.
(28) Mukherjee, P.; Desai, P.; Zhou, Y.-D.; Avery, M. Targeting the BH3 domain mediated protein-protein interaction of Bcl-xL through virtual screening. *J. Chem. Inf. Model.* **2010**, *50*, 906−923.
(29) Ye, F.; Zhang, W.; Ye, X.; Jin, J.; Lv, Z.; Luo, C. Identification of selective, cell active inhibitors of protein arginine methyltransferase 5 through structure-based virtual screening and biological assays. *J. Chem. Inf. Model.* **2018**, *58*, 1066−1073.
(30) Jiang, M.; Peng, L.; Yang, K.; Wang, T.; Yan, X.; Jiang, T.; Xu, J.; Qi, J.; Zhou, H.; Qian, N.; Zhou, Q.; Chen, B.; Xu, X.; Deng, L.; Yang, C. Development of small-molecules targeting receptor activator of nuclear factor-κB ligand (RANKL)-receptor activator of nuclear factor-κB (RANK) protein-protein interaction by structure-based virtual screening and hit optimization. *J. Med. Chem.* **2019**, *62*, 5370−5381.
(31) Negri, A.; Rives, M.-L.; Caspers, M. J.; Prisinzano, T. E.; Javitch, J. A.; Filizola, M. Discovery of a novelselective kappa-opioid receptor agonist using crystal structure-based virtual screening. *J. Chem. Inf. Model.* **2013**, *53*, 521−526.
(32) Cole, J. C.; Murray, C. W.; Nissink, J. W. M.; Taylor, R. D.; Taylor, R. Comparing protein-ligand docking programs is difficult. *Proteins: Struct., Funct., Genet.* **2005**, *60*, 325−332.
(33) Hartshorn, M. J. AstexViewerTM †: A visualisation aid for structure-based drug design. *J. Comput.-Aided Mol. Des.* **2002**, *16*, 871−881.
(34) Zhou, T.; Caflisch, A. High-throughput virtual screening using quantum mechanical probes: Discovery of selective kinase inhibitors. *ChemMedChem* **2010**, *5*, 1007−1014.
(35) Gomez, L. Decision making in medicinal chemistry: The power of our intuition. *ACS Med. Chem. Lett.* **2018**, *9*, 956−958.
(36) Kolb, P.; Rosenbaum, D. M.; Irwin, J. J.; Fung, J. J.; Kobilka, B. K.; Shoichet, B. K. Structure-based discovery of beta2-adrenergic receptor ligands. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 6843−6848.
(37) Kumar, A.; Ito, A.; Hirohama, M.; Yoshida, M.; Zhang, K. Y. J. Identification of sumoylation activating enzyme 1 inhibitors by structure-based virtual screening. *J. Chem. Inf. Model.* **2013**, *53*, 809−820.
(38) Bauer, M. R.; Mackey, M. D. Electrostatic complementarity as a fast and effective tool to optimize binding and selectivity of protein-ligand complexes. *J. Med. Chem.* **2019**, *62*, 3036−3050.
(39) Yang, Y.; Hu, B.; Lill, M. A. In *WATsite2.0 with PyMOL plugin: Hydration site prediction and visualization BT - protein function prediction: Methods and protocols*; Kihara, D., Ed.; Springer: New York, 2017; pp 123−134.
(40) Liebeschuetz, J.; Hennemann, J.; Olsson, T.; Groom, C. R. The good, the bad and the twisted: A survey of ligand geometry in protein crystal structures. *J. Comput.-Aided Mol. Des.* **2012**, *26*, 169−183.
(41) Blundell, C. D.; Nowak, T.; Watson, M. J. In *Chapter 2: Measurement interpretation and use of free ligand solution conformations in drug discovery*; Lawton, G., et al., Eds.; Elsevier, 2016; Vol. *55*, pp 45−147.
(42) Kubinyi, H. Hydrogen bonding: The last mystery in drug design? *Pharmacokinetic Optimization in Drug Research* **2007**, 513−524.
(43) Agnihotri, P.; Mishra, A. K.; Mishra, S.; Sirohi, V. K.; Sahasrabuddhe, A. A.; Pratap, J. V. Identification of novel inhibitors of Leishmania donovani γ-glutamylcysteine synthetase using structure-based virtual screening, docking, molecular dynamics simulation, and in vitro studies. *J. Chem. Inf. Model.* **2017**, *57*, 815−825.
(44) Iwata, Y.; Arisawa, M.; Hamada, R.; Kita, Y.; Mizutani, M. Y.; Tomioka, N.; Itai, A.; Miyamoto, S. Discovery of novel aldose reductase inhibitors using a protein structure-based approach: 3D-database search followed by design and synthesis. *J. Med. Chem.* **2001**, *44*, 1718−1728.
(45) Da'adoosh, B.; Kaito, K.; Miyashita, K.; Sakaguchi, M.; Goldblum, A. Computational design of substrate selective inhibition. *PLoS Comput. Biol.* **2020**, *16*, e1007713.
(46) Ribeiro, A. J. M.; Tyzack, J. D.; Borkakoti, N.; Holliday, G. L.; Thornton, J. M. A global analysis of function and conservation of catalytic residues in enzymes. *J. Biol. Chem.* **2020**, *295*, 314−324.
(47) Luo, R.; David, L.; Hung, H.; Devaney, J.; Gilson, M. K. Strength of solvent-exposed salt-bridges. *J. Phys. Chem. B* **1999**, *103*, 727−736.
(48) Ferreira de Freitas, R.; Schapira, M. A systematic analysis of atomic protein-ligand interactions in the PDB. *MedChemComm* **2017**, *8*, 1970−1981.
(49) Fuchs, J. E.; Waldner, B. J.; Huber, R. G.; von Grafenstein, S.; Kramer, C.; Liedl, K. R. Independent metrics for protein backbone and side-chain flexibility: Time scales and effects of ligand binding. *J. Chem. Theory Comput.* **2015**, *11*, 851−860.

https://dx.doi.org/10.1021/acs.jmedchem.0c02227
J. Med. Chem. 2021, 64, 2489−2500

(50) Caballero, J. Considerations for Docking of Selective Angiotensin-Converting Enzyme Inhibitors. *Molecules* **2020**, *25*, 295.
(51) Kannan, S.; Melesina, J.; Hauser, A.-T.; Chakrabarti, A.; Heimburg, T.; Schmidtkunz, K.; Walter, A.; Marek, M.; Pierce, R. J.; Romier, C.; Jung, M.; Sippl, W. Discovery of inhibitors of Schistosoma mansoni HDAC8 by combining homology modeling, virtual screening, and in vitro validation. *J. Chem. Inf. Model.* **2014**, *54*, 3005−3019.
(52) Masters, M. R.; Mahmoud, A. H.; Yang, Y.; Lill, M. A. Efficient and accurate hydration site profiling for enclosed binding sites. *J. Chem. Inf. Model.* **2018**, *58*, 2183−2188.
(53) Abel, R.; Young, T.; Farid, R.; Berne, B. J.; Friesner, R. A. Role of the active-site solvent in the thermodynamics of factor Xa ligand binding. *J. Am. Chem. Soc.* **2008**, *130*, 2817−2831.
(54) Mahmoud, A. H.; Masters, M. R.; Yang, Y.; Lill, M. A. Elucidating the multiple roles of hydration for accurate protein-ligand binding prediction via deep learning. *Commun. Chem.* **2020**, *3*, 19.
(55) Gaudreault, F.; Chartier, M.; Najmanovich, R. Side-chain rotamer changes upon ligand binding: Common, crucial, correlate with entropy and rearrange hydrogen bonding. *Bioinformatics* **2012**, *28*, i423−i430.
(56) Ravindranathan, K. P.; Mandiyan, V.; Ekkati, A. R.; Bae, J. H.; Schlessinger, J.; Jorgensen, W. L. Discovery of novel fibroblast growth factor receptor 1 kinase inhibitors by structure-based virtual screening. *J. Med. Chem.* **2010**, *53*, 1662−1672.
(57) Cournia, Z.; Leng, L.; Gandavadi, S.; Du, X.; Bucala, R.; Jorgensen, W. L. Discovery of human macrophage migration inhibitory factor (MIF)-CD74 antagonists via virtual screening. *J. Med. Chem.* **2009**, *52*, 416−424.
(58) Dayhoff, J. E.; Shoemaker, B. A.; Bryant, S. H.; Panchenko, A. R. Evolution of protein binding modes in homooligomers. *J. Mol. Biol.* **2010**, *395*, 860−870.
(59) Kasahara, K.; Kinoshita, K. Landscape of protein-small ligand binding modes. *Protein science* **2016**, *25*, 1659−1671.
(60) Damm-Ganamet, K. L.; Smith, R. D.; Dunbar, J. B.; Stuckey, J. A.; Carlson, H. A. CSAR benchmark exercise 2011−2012: Evaluation of results from docking and relative ranking of blinded congeneric series. *J. Chem. Inf. Model.* **2013**, *53*, 1853−1870.
(61) Hevener, K. E.; Mehboob, S.; Su, P.-C.; Truong, K.; Boci, T.; Deng, J.; Ghassemi, M.; Cook, J. L.; Johnson, M. E. Discovery of a novel and potent class of F. tularensis enoyl-reductase (FabI) inhibitors by molecular shape and electrostatic matching. *J. Med. Chem.* **2012**, *55*, 268−279.
(62) Singh, N.; Shah, P.; Dwivedi, H.; Mishra, S.; Tripathi, R.; Sahasrabuddhe, A. A.; Siddiqi, M. I. Integrated machine learning, molecular docking and 3D-QSAR based approach for identification of potential inhibitors of trypanosomal N-myristoyltransferase. *Mol. BioSyst.* **2016**, *12*, 3711−3723.
(63) Gaieb, Z.; Liu, S.; Gathiaka, S.; Chiu, M.; Yang, H.; Shao, C.; Feher, V. A.; Walters, W. P.; Kuhn, B.; Rudolph, M. G.; Burley, S. K.; Gilson, M. K.; Amaro, R. E. D3R grand challenge 2: Blind prediction of protein-ligand poses, affinity rankings, and relative binding free energies. *J. Comput.-Aided Mol. Des.* **2018**, *32*, 1−20.
(64) Gathiaka, S.; Liu, S.; Chiu, M.; Yang, H.; Stuckey, J. A.; Kang, Y. N.; Delproposto, J.; Kubish, G.; Dunbar, J. B. J.; Carlson, H. A.; Burley, S. K.; Walters, W. P.; Amaro, R. E.; Feher, V. A.; Gilson, M. K. D3R grand challenge 2015: Evaluation of protein-ligand pose and affinity predictions. *J. Comput.-Aided Mol. Des.* **2016**, *30*, 651−668.
(65) Mobley, D. L.; Liu, S.; Lim, N. M.; Wymer, K. L.; Perryman, A. L.; Forli, S.; Deng, N.; Su, J.; Branson, K.; Olson, A. J. Blind prediction of HIV integrase binding from the SAMPL4 challenge. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 327−345.
(66) Perryman, A. L.; Santiago, D. N.; Forli, S.; Santos-Martins, D.; Olson, A. J. Virtual screening with AutoDock Vina and the common pharmacophore engine of a low diversity library of fragments and hits against the three allosteric sites of HIV integrase: Participation in the SAMPL4 protein-ligand binding challenge. *J. Comput.-Aided Mol. Des.* **2014**, *28*, 429−441.
(67) Smith, R. D.; Damm-Ganamet, K. L.; Dunbar, J. B.; Ahmed, A.; Chinnaswamy, K.; Delproposto, J. E.; Kubish, G. M.; Tinberg, C. E.; Khare, S. D.; Dou, J.; Doyle, L.; Stuckey, J. A.; Baker, D.; Carlson, H. A. CSAR benchmark exercise 2013: Evaluation of results from a combined computational protein design, docking, and scoring/ranking challenge. *J. Chem. Inf. Model.* **2016**, *56*, 1022−1031.
(68) Xing, L.; McDonald, J. J.; Kolodziej, S. A.; Kurumbail, R. G.; Williams, J. M.; Warren, C. J.; O'Neal, J. M.; Skepner, J. E.; Roberds, S. L. Discovery of potent inhibitors of soluble epoxide hydrolase by combinatorial library design and structure-based virtual screening. *J. Med. Chem.* **2011**, *54*, 1211−1222.
(69) Fischer, A.; Sellner, M.; Neranjan, S.; Smieško, M.; Lill, M. A. Potential inhibitors for novel coronavirus protease identified by virtual screening of 606 million compounds. *Int. J. Mol. Sci.* **2020**, *21*, 3626.
(70) Broccatelli, F.; Brown, N. Best of both worlds: On the complementarity of ligand-based and structure-based virtual screening. *J. Chem. Inf. Model.* **2014**, *54*, 1634−1641.
(71) Xu, W.; Lim, J.; Goh, C.-Y.; Suen, J. Y.; Jiang, Y.; Yau, M.-K.; Wu, K.-C.; Liu, L.; Fairlie, D. P. Repurposing registered drugs as antagonists for protease-activated receptor 2. *J. Chem. Inf. Model.* **2015**, *55*, 2079−2084.
(72) Xu, L.; Zhou, S.; Yu, K.; Gao, B.; Jiang, H.; Zhen, X.; Fu, W. Molecular modeling of the 3D structure of 5-HT1AR: Discovery of novel 5-HT1AR agonists via dynamic pharmacophore-based virtual screening. *J. Chem. Inf. Model.* **2013**, *53*, 3202−3211.
(73) Rudling, A.; Gustafsson, R.; Almlöf, I.; Homan, E.; Scobie, M.; Warpman Berglund, U.; Helleday, T.; Stenmark, P.; Carlsson, J. Fragment-based discovery and optimization of enzyme inhibitors by docking of commercial chemical space. *J. Med. Chem.* **2017**, *60*, 8160−8169.
(74) di Pizio, A.; Laghezza, A.; Tortorella, P.; Agamennone, M. Probing the S1' site for the identification of non-zinc-binding MMP-2 inhibitors. *ChemMedChem* **2013**, *8*, 1475−1482.
(75) Grebner, C.; Malmerberg, E.; Shewmaker, A.; Batista, J.; Nicholls, A.; Sadowski, J. Virtual screening in the cloud: How big is big enough? *J. Chem. Inf. Model.* **2020**, *60*, 4274−4282.
(76) Walters, W. P.; Wang, R. New trends in virtual screening. *J. Chem. Inf. Model.* **2020**, *60*, 4109−4111.
(77) Smieško, M. DOLINA - docking based on a local induced-fit algorithm: Application toward small-molecule binding to nuclear receptors. *J. Chem. Inf. Model.* **2013**, *53*, 1415−1423.
(78) Lill, M. A. Efficient incorporation of protein flexibility and dynamics into molecular docking simulations. *Biochemistry* **2011**, *50*, 6157−6169.
(79) de Vivo, M.; Masetti, M.; Bottegoni, G.; Cavalli, A. Role of molecular dynamics and related methods in drug discovery. *J. Med. Chem.* **2016**, *59*, 4035−4061.
(80) Guedes, I. A.; Pereira, F. S. S.; Dardenne, L. E. Empirical scoring functions for structure-based virtual screening: Applications, critical aspects, and challenges. *Front. Pharmacol.* **2018**, *9*, 1089.
(81) Huang, S.-Y.; Grinter, S. Z.; Zou, X. Scoring functions and their evaluation methods for protein-ligand docking: Recent advances and future directions. *Phys. Chem. Chem. Phys.* **2010**, *12*, 12899−12908.
(82) Li, J.; Fu, A.; Zhang, L. An overview of scoring functions used for protein-ligand interactions in molecular docking. *Interdiscip. Sci.: Comput. Life Sci.* **2019**, *11*, 320−328.
(83) Wang, Z.; Sun, H.; Yao, X.; Li, D.; Xu, L.; Li, Y.; Tian, S.; Hou, T. Comprehensive evaluation of ten docking programs on a diverse set of protein-ligand complexes: The prediction accuracy of sampling power and scoring power. *Phys. Chem. Chem. Phys.* **2016**, *18*, 12964−12975.
(84) Salmaso, V.; Moro, S. Bridging molecular docking to molecular dynamics in exploring ligand-protein recognition process: An overview. *Front. Pharmacol.* **2018**, *9*, 923.
(85) Maffucci, I.; Hu, X.; Fumagalli, V.; Contini, A. An efficient implementation of the Nwat-MMGBSA method to rescore docking results in medium-throughput virtual screenings. *Front. Chem.* **2018**, *6*, 43.

https://dx.doi.org/10.1021/acs.jmedchem.0c02227
J. Med. Chem. 2021, 64, 2489−2500

(86) Rastelli, G.; Pinzi, L. Refinement and rescoring of virtual screening results. *Front. Chem.* **2019**, *7*, 498.
(87) Zhang, X.; Wong, S. E.; Lightstone, F. C. Toward fully automated high performance computing drug discovery: A massively parallel virtual screening pipeline for docking and molecular mechanics/generalized born surface area rescoring to improve enrichment. *J. Chem. Inf. Model.* **2014**, *54*, 324−337.
(88) Virtanen, S. I.; Niinivehmas, S. P.; Pentikäinen, O. T. Case-specific performance of MM-PBSA, MM-GBSA, and SIE in virtual screening. *J. Mol. Graphics Modell.* **2015**, *62*, 303−318.
(89) Wang, E.; Sun, H.; Wang, J.; Wang, Z.; Liu, H.; Zhang, J. Z.; Hou, T. End-Point binding free energy calculation with MM/PBSA and MM/GBSA: Strategies and applications in drug design. *Chem. Rev.* **2019**, *119*, 9478−9508.
(90) Albanese, S. K.; Chodera, J. D.; Volkamer, A.; Keng, S.; Abel, R.; Wang, L. Is structure-based drug design ready for selectivity optimization? *J. Chem. Inf. Model.* **2020**, *60*, 6211−6227.
(91) Ghanbarpour, A.; Mahmoud, A. H.; Lill, M. A. On-the-fly prediction of protein hydration densities and free energies using deep learning. *arXiv* **2020**.
(92) Djenno, M.; Insua, G.; Pho, A. From paper to pixels: Using Google Forms for collaboration and assessment. *Library hi tech news* **2015**, *32*, 9−13.
(93) Abagyan, R.; Rueda, M.; Bottegoni, G. Recipes for the selection of experimental protein conformations for virtual screening. *J. Chem. Inf. Model.* **2010**, *50*, 186−193.
(94) Dalkas, G. A.; Vlachakis, D.; Tsagkrasoulis, D.; Kastania, A.; Kossida, S. State-of-the-art technology in modern computer-aided drug design. *Briefings Bioinf.* **2013**, *14*, 745−752.
(95) Sevrioukova, I. F.; Poulos, T. L. Pyridine-substituted desoxyritonavir is a more potent inhibitor of cytochrome P450 3A4 than ritonavir. *J. Med. Chem.* **2013**, *56*, 3733−3741.
(96) Yano, J. K.; Denton, T. T.; Cerny, M. A.; Zhang, X.; Johnson, E. F.; Cashman, J. R. Synthetic inhibitors of cytochrome P-450 2A6: Inhibitory activity, difference spectra, mechanism of inhibition, and protein cocrystallization. *J. Med. Chem.* **2006**, *49*, 6987−7001.
(97) Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: A new approach for rapid, accurate docking and scoring. 2. enrichment factors in database screening. *J. Med. Chem.* **2004**, *47*, 1750−1759.

https://dx.doi.org/10.1021/acs.jmedchem.0c02227
J. Med. Chem. 2021, 64, 2489−2500