Chapter 6

Considerations Around Structure-Based Drug Discovery for KRAS Using DOCK

Mayukh Chakrabarti, Y. Stanley Tan, and Trent E. Balius

Abstract

Molecular docking is a popular computational tool in drug discovery. Leveraging structural information, docking software predicts binding poses of small molecules to cavities on the surfaces of proteins. Virtual screening for ligand discovery is a useful application of docking software. In this chapter, using the enigmatic KRAS protein as an example system, we endeavor to teach the reader about best practices for performing molecular docking with UCSF DOCK. We discuss methods for virtual screening and docking molecules on KRAS. We present the following six points to optimize our docking setup for prosecuting a virtual screen: protein structure choice, pocket selection, optimization of the scoring function, modification of sampling spheres and sampling procedures, choosing an appropriate portion of chemical space to dock, and the choice of which top scoring molecules to pick for purchase.

Key words Molecular docking, DOCK, Virtual screening, Druggable pockets, Computational drug discovery

1 Introduction

The objective of molecular docking is to generate and score poses of small molecules to cavities on biological macromolecules, typically proteins, to enable drug discovery. We focus on UCSF DOCK, the first docking software [1]. UCSF DOCK deploys a receptor-based approach with physics-based scoring functions. Currently, there are two actively developed versions of DOCK: DOCK 6 and DOCK 3. These programs differ in their ligand sampling algorithms: *Anchor-and-grow* (DOCK 6), and *hierarchical database (HDB) search* (DOCK 3). We have implemented *HDB search* into DOCK 6 [2]. DOCK 3 can perform covalent docking by using a modified *HDB search* method [3]. We direct the reader

Mayukh Chakrabarti and Y. Stanley Tan contributed equally.

Andrew G. Stephen and Dominic Esposito (eds.), *KRAS: Methods and Protocols*, Methods in Molecular Biology, vol. 2797,
https://doi.org/10.1007/978-1-0716-3822-4_6,

to a protocol paper [4] for a discussion of large-scale docking with DOCK 3.

There are numerous molecular docking software programs, including Autodock, Autodock Vina, Glide, FRED, EnzyDock, Rosettaligand, Gold, Cdocker, Surflex, PLANTS, DockThor, FlexX, ICM, LigandFit, MCDock, MOE-Dock, LeDock, rDOCK, RxDock, DiffDOCK, DPL, GNINA, and others [5–8]. Recently, great excitement has emerged around machine learning, particularly deep neural networks, and seems to be impacting computational drug discovery and design [9, 10]. Docking programs deploy an array of sampling and scoring methods. Scoring functions classify into physics-based, knowledge-based, and empirical [11–16]. Recently, scoring functions have also been derived from machine learning, e.g., GNINA [17], MedusaNet [18], and DeepVS [19].

KRAS, a member of the superfamily of guanine nucleotide-binding proteins, is a GTPase, hydrolyzing GTP to GDP by releasing a phosphate group. The first 165 residues of RAS, referred to as the G-domain, adopts a topology consisting of 5 α-helices surrounding a 6-stranded β-sheet, arranged in a 3-layer α/β/α sandwich architecture known as a Rossmann fold. The secondary structural elements of KRAS are arranged as follows: β-sheet 1, phosphate-binding loop (p-loop), α-helix 1, Switch I, β-sheet 2, β-sheet 3, Switch II (which contains α-helix 2), β-sheet 4, α-helix 3, β-sheet 5, α-helix 4, β-sheet 6, α-helix 5, and the hypervariable region (HVR) (Fig. 1). The G-domain has two highly flexible regions: Switch I (residues 30–38) and Switch II (residues 60–76). The C-terminal HVR of RAS (residues 167–188/189) can be subdivided into a disordered linker region and a membrane interacting region that is the locus of posttranslational modifications (palmitoylation and prenylation), facilitating RAS association with the cell membrane. In all RAS isoforms, a farnesyl group present on the hypervariable region (residue 185/186) acts as a lipid membrane anchor. GTP and GDP bind under Switch I and interact with the p-loop, Switch I, α-helix 1, and Switch II.

In the Protein Data Bank (PDB), there are two main pockets on KRAS that bind small molecules (*see* Fig. 1b, c): the Switch II pocket formed by molecules wedged between Switch II and α-helix 3 (*see* Fig. 1d, e), and the Switch I/Switch II pocket (also called the SOS pocket, *see* Fig. 1f). Other pockets have also been observed, including a pocket under Switch I that is less well-represented in the PDB [20]. Switch I and Switch II are dynamic and play a role in effector binding [21–30].

Here, we focus on considerations for drug discovery against KRAS. Misregulation of RAS is implicated in several human disease conditions. RAS has historically been a difficult molecular target for several reasons, including the highly potent binding affinity of GTP to its orthosteric site (K_d of 10^{-11} M), precluding the development

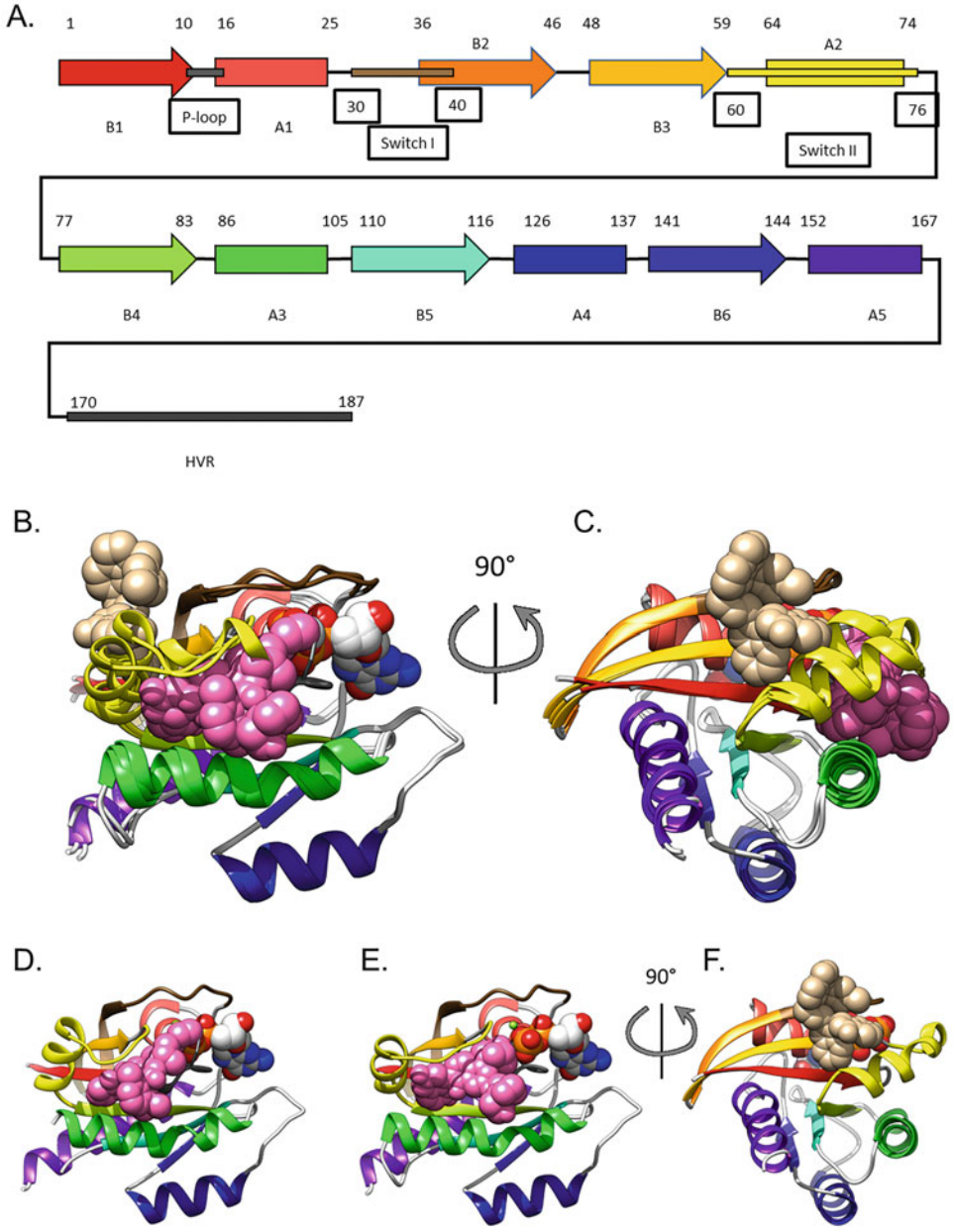


Fig. 1 KRAS secondary and tertiary structure, with two prominent binding pockets indicated in tertiary depictions. (**a**) Secondary elements are shown in a cartoon (not to scale). Rainbow colors applied where β-sheet 1, p-loop, α-helix 1, Switch I, β-sheet 2, β-sheet 3, Switch II (which includes α-helix 2), β-sheet 4, α-helix 3, β-sheet 5, α-helix 4, β-sheet 6, α-helix 5, and the hypervariable region are colored red, grey (p-loop), red-orange, brown (Switch I), orange, yellow-orange (amber), yellow (Switch II), yellow-green (lime), green, blue-green, blue, indigo, violet, and dark gray, respectively. (**b**) Overlay of three RAS structures with three small molecules in two distinct pockets. (**c**) Rotation of B by 90°. (**d**) PDB Code 6OIM with covalent ligand sotorasib (AMG 510) in the Switch II pocket. (**e**) PDB Code 7RPZ with ligand MRTX-1133 in the Switch II pocket. (**f**) PDB Code 6GJ8 SOS pocket (Switch I/Switch II pocket) shown, indicated by small molecule BI-2852. (**b**)–(**f**) are made with UCSF Chimera (orthographic mode)

of competitive inhibitors, and its relatively featureless surface topology, making it challenging to identify suitable pockets for small molecules. The major breakthrough in RAS drug discovery came in 2013, when Ostrem et al. targeted the KRAS G12C mutant in a covalent screen and discovered the galvanizing small molecule inhibitor termed Compound 12 [31], although three decades of effort to discover inhibitors of RAS proteins preceded it. This molecule eventually led to the development of the first two approved drugs: sotorasib (Amgen) and adagrasib (Mirati). The lead molecule, which binds to KRAS in the GDP-bound state, was discovered in a pocket that is located under Switch II (between Switch II and α-helix 3), termed the "Switch II pocket." The inhibitor was aided by two factors: the relatively high intrinsic hydrolysis rate of the G12C mutant relative to wild-type KRAS, allowing RAS to be trapped in the inactive state during the process of nucleotide exchange, and the catalytic interaction of residue K16 with the covalent warhead [32]. Other molecules have been identified that bind this site non-covalently [33, 34]. Other inhibitors that target multiple KRAS mutants are also currently being evaluated, including a pan-KRAS inhibitor by Boehringer Ingelheim, BI-2865 [35]. Several inhibitors that target KRAS in both the GDP-bound (inactive) and GTP-bound (active) states are presently in clinical trials [36–38], including MRTX-1133, a non-covalent inhibitor that binds to GDP-bound KRAS G12D with picomolar affinity. Docking to sotorasib and BI-2852 is explored below.

2 Theory and Methods

Before we can understand molecular docking, we first must understand the binding event (Fig. 2). This event can be computationally studied at different levels of theory, with a trade off between speed and accuracy. Usually, the more accurate the method, the slower it is. To calculate a ligand binding free energy to a receptor to form a complex, we can simulate each species: the complex, receptor, and ligand [39, 40]. Alternatively, we can run a simulation of the complex and split it into a receptor and a ligand piece, as is done in end-state free energy calculations, like MM-GBSA [41], with a binding energy for each snapshot. These snapshot energies can be averaged over the simulation, resulting in a binding energy comparable to simulating each species. In docking, the receptor is typically held rigid, while different orientations and conformations of the ligand are sampled, generating a large number of complexes [42]. Each complex is evaluated with a score that typically attempts to calculate the binding event. The binding event is an ensemble of conformations, while the docking score is generated by a single complex. Thus, docking is an approximation from the start, because of this one complex per score.

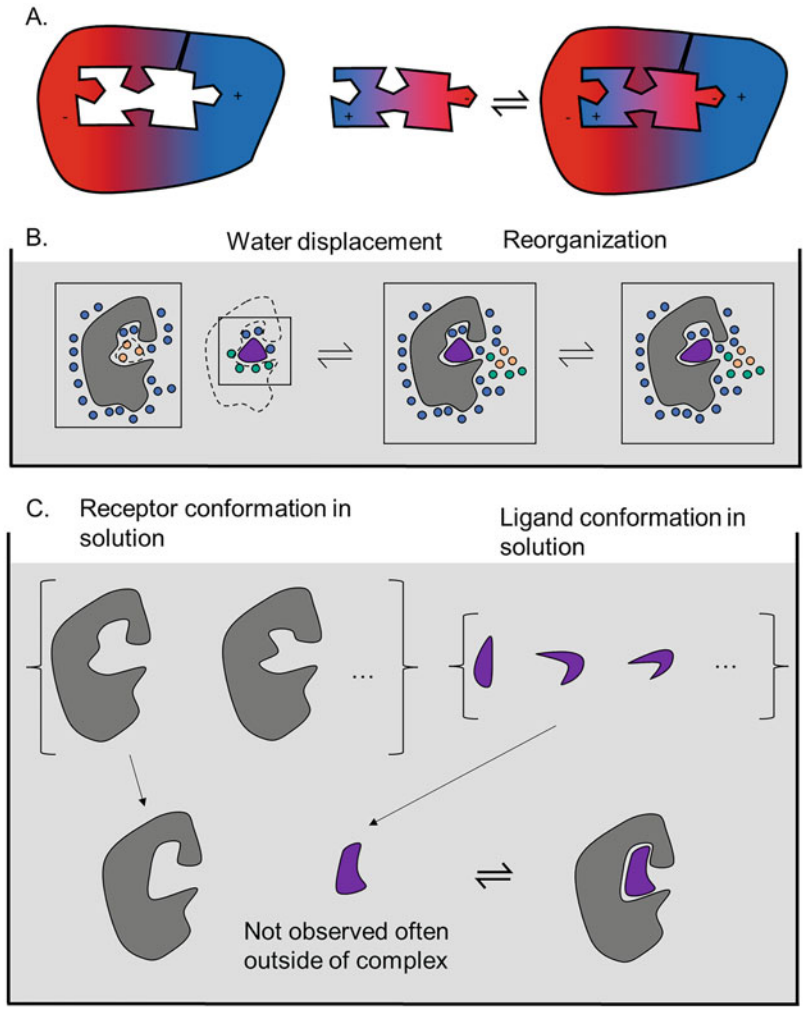


Fig. 2 Cartoons that describe the binding event. (**a**) Cartoon that shows the jigsaw analogy to the binding event. The receptor has a missing region that is perfectly filled by the ligand. The ligand and receptor have complementary charges (represented by the red and blue colors). This stresses the importance of shape (van der Waals (vdW)) and charge (electrostatics (ES)) complementarity. (**b**) Cartoon that illustrates the importance of water displacement and consideration of waters that remain. (**c**) Cartoon that shows the presence of multiple states of the receptor and ligand, and the possibility that the complex stabilizes a state that is not highly populated in the unbound ensemble. This figure illustrates how the receptor and ligand undergo conformational changes upon binding

Docking methods seek to approximate the binding event while remaining fast. Most molecular docking scoring functions crudely predict the small molecule binding to a protein. Because docking methods make approximations, we test the software to verify the speed advantage, and the deleterious impacts on success are manageable using retrospective tests before performing screens. Despite the approximations made, experimental approaches have validated docking predictions [43–45].

The two main scoring functions of DOCK 6 and DOCK 3 are Grid Score and ChemGrid Score, respectively. Grid Score is a two-component scoring function: $E_{\text{Grid}} = E_{\text{vdW}} + E_{\text{ES}}$. It consists of a van der Waals (vdW) component modeled using the Lennard-Jones equation, typically with 6–9 or 6–12 exponents pre-computed and stored on a grid, and an electrostatic component, calculated using Coulomb's law and typically with a distance-dependent dielectric, also stored on a grid. Grid Score typically uses an all-atom representation. ChemGrid Score is a three-component function: $E_{\text{ChemGrid}} = E_{\text{vdW}} + E_{\text{ES}} + E_{\text{lig, desolv}}$. It consists of a vdW component that uses the Lennard-Jones equation, with 6–12 exponents stored on a grid, and an electrostatic component, calculated by solving the Poisson-Boltzmann equation using *Qnifft* [46, 47]. In addition, ChemGrid Score has a ligand desolvation component [48, 49]. ChemGrid score uses a united atom representation, where only polar hydrogens are explicitly modeled. This DOCK 3.7 scoring function is cobbled together from different theories, and we often apply different modifications like *thin spheres* to improve docking performance on retrospective enrichment calculations. Additional terms exist and might be added to this scoring function, including GIST for receptor desolvation [50], ligand strain [51], and receptor strain when a flexible receptor is used [52]. For the two main tasks of docking (virtual screening and binding pose prediction), the balance of considerations is different. In virtual screening, we want to avoid testing molecules that do not bind (false positives). In binding mode prediction, we do not want to discard any molecules (false negatives).

A fast, accurate workflow can be thought of as an upside-down triangle. First, one runs fast calculations on a large molecular library and then performs more expensive calculations on the top scoring molecules, performed in layers. For example, (1) a fast docking can be performed with DOCK on hundreds of millions (or billons) of molecules. (2) The top 100,000 ranked molecules can be redocked with increased sampling in DOCK, and more poses can be saved. (3) These poses can then be minimized and rescored with other methods like MM-GBSA calculations or molecular footprint score [53]. (4) The top thousand or so could then be passed to AmberScore [54, 55] or Amber for molecular simulations, to conduct stability analysis and MM-GBSA binding energy calculations. (5) For these top scoring molecules, poses, properties, and energies are then visualized, and molecules are selected for purchase and experimental testing. (6) For experimentally verified hits, MM-GBSA, thermodynamic integration, or free energy perturbation calculations could be used for analog selection.

In docking, sampling approximations are made for speed and include the use of spheres for ligand placement, a rigid receptor, and an under-sampling of ligand internal degrees of freedom. Although there are many different sampling methods in other

docking software, here we discuss two methods implemented in DOCK. First, in *anchor-and-grow*, the ligand internal degrees of freedom are explored on-the-fly using an incremental construction procedure. The *anchor-and-grow* algorithm is a breadth-first search method. Second, in *Hierarchical database search*, the ligand internal degrees of freedom are sampled outside of docking. The *hierarchical database search* algorithm is a depth-first search method that traverses and scores pre-computed ligand conformations. This method is 16× faster than anchor-and-grow [2].

In standard docking, a ligand is placed into a pocket on a receptor by exploring the three translational and three rotational degrees of freedom, and the ligand's internal degrees of freedom. A pose can be thought of as the internal conformation and orientation (relative to the protein) of a molecule. *Anchor-and-grow* and *hierarchical database search* are both standard docking algorithms. Matching spheres are generated by converting crystallographic ligand atoms to spheres, or by filling a cavity on the protein with spheres, as is done with the auxiliary program *sphgen*. These *sphgen* spheres can be thought of as occupying the inverse image of the molecular surface of the protein. Spheres are used for orienting the ligand into the binding site.

Covalent molecules react with the protein (or biomolecule) to form an adduct. Most often, these molecules react with a cysteine residue, as is the case for sotorasib, which targets the KRAS C12 mutation. For covalent docking, we only explore orientations that are possible considering the covalent bond. Thus, the relative degrees of freedom between the ligand and receptor are greatly reduced compared to standard docking. The degrees that are explored are the dihedral angles, distances, and an angle around the bond. However, to be effective, the fineness of sampling is higher. In DOCK 3.6, 3.7, and 3.8, a covalent method has been implemented that attaches the ligand to the residue by placing a dummy atom onto the gamma sulfur and adjusting two dihedrals, the bond angle and a distance about the sulfur [3]. The ligand internal degrees of freedom are explored by traversing through the hierarchical database (*db2* file) [3].

We recommend, if structural and ligand data exists, that pose reproduction and enrichment control calculations be performed before any screen. In pose reproduction, our docking method is used to reproduce an experimentally determined binding pose. We define three outcomes of this type of calculation, guided by Mukherjee et al. [56]: (1) *Docking success* is when the experimental pose is reproduced by our docking method as the top scoring pose. (2) *Scoring failure* is when the experimental pose is reproduced, but not as the top scoring pose. (3) *Sampling failure* is when the experimental pose is not reproduced at all. To calculate these outcomes, we consider the number of poses kept during docking, and the method used to determine similarity to the experimental pose.

Here, we use a symmetry-corrected RMSD (Hungarian algorithm) [57] with a threshold of 2.0 Å. That is, any pose less than 2.0 Å RMSD is considered as reproducing the experimental pose. Alternatively, *docking success* can be defined by considering the top scoring pose, or the top five scoring poses. In enrichment calculations, a set of known-binding active molecules and a set of presumed nonbinding decoy molecules are docked to a protein, and the ability of our docking method to distinguish between the actives and decoys is measured. There are several metrics used to quantify enrichment. One metric is the area under the curve (AUC) of the receiver operator characteristic (ROC) curve [58]. The ROC curve AUC, plotted with a logarithmic scale on the *x*-axis (% of decoys found, or false positive rate), can be computed to weight early enrichment more (logAUC). Different docking parameter options can be examined in a semiautomated format to explore the enrichment space [59]. Other computational considerations include choice of ligands, decoy background, and receptor preparation. Focusing on decoy background, we can choose property-matched decoys, extrema decoys, a set of known non-binders, or other backgrounds. The background choice greatly impacts enrichment quantification.

Retrospective calculations are meant to be used as a control to ensure the setup is behaving as intended or that a code change is working as expected, before performing virtual screens. Despite their utility in virtual screening, there are still notable limitations. These include appropriately weighting early enrichment more than later enrichment [60], which can be addressed with log-adjusted AUCs; however, the lower bound then becomes important to consider [58]. In addition to retrospective tests, one can dock prospectively to simplified sites, i.e., host-guest systems and model cavities [50, 61–63].

The main application of molecular docking is virtual screening in pursuit of discovering small molecules that bind to a protein. These molecules can be developed further to become drug leads or biological tool compounds. We examine pitfalls in selecting molecules from a virtual screen (*see* **Note 1**) and rules of thumb for prioritizing molecules from a virtual screen (*see* **Note 2**). As a community, we aspire to dock all accessible chemical space. Libraries are growing ever larger. We examine the impact of chemical space growth on our calculations in a thought experiment (*see* **Note 3**) and discuss strategies for docking fewer molecules while maintaining the ability to accurately identify top-scoring hits (*see* **Note 4**). Before running a screen, we advise optimizing the docking setup with retrospective calculations and having a binding assay that is clear and high throughput, where 20–50 molecules and analogs can be tested without burden.

To prepare the receptor for ChemGrid Score, run the *blastermaster.py* script distributed with DOCK 3, which performs the

following: (1) Uses UCSF Chimera's Dock Prep function to remove solvent, and remove alternative conformations [64]. (2) Uses the program *reduce* [65] to add hydrogens to the protein. (3) Calculates hydrogens and partial charges for the cofactor as described below. These partial charges are stored, and vdW parameters are looked up based on the atom type and stored in text files for use in the next steps. (4) Uses *ChemGrid* to calculate the vdW grid. (5) Uses *Qnifft* to calculate the electrostatic grid. (6) Uses *Solvmap* to calculate the desolvation grid. (7) Creates matching spheres, using the experimental ligand atoms converted to spheres, supplemented with *sphgen* spheres. Run *blastermaster.py* once with the protein and the ion, but without the nucleotide cofactor (performing **steps 2**, **4**–7), and then a second time using a protonated protein and cofactor (obtained from **steps 2** and **3**, respectively), where *blastermaster.py* does not add hydrogens but uses existing ones (*see* **steps 4**–7). For both the ES grid and desolvation grid, thin spheres can be added to perturb the balance of energetic terms (use *blastermaster* defaults with desolvation spheres from *sphgen*). For more information about *blastermaster.py*, *see* refs. 4 and 47.

To prepare the receptor for *Grid Score* (the default DOCK 6 scoring method), perform the following: (1) Use UCSF Chimera's Dock Prep function to remove solvent, remove alternative conformations, and add charges to the protein and ion [64]. (2) Calculate hydrogens and partial charges for the cofactor, as described below. Chimera is used to combine the cofactor and protein. (3) Use the auxiliary program *sphgen* to calculate spheres that fill pockets on the protein and select spheres within 3 Å of the experimental ligand as matching spheres. (4) Use the program *Grid* to calculate the vdW and ES grids (see above). The grid is defined and bounded by a box placed about the spheres, padded with a margin of 10 Å.

Cofactors bound to the protein need to be included in docking calculations and should be prepared with the receptor. For RAS, both the nucleotide and a magnesium ion are treated as cofactors. There are four possibilities for the nucleotide cofactor: not present, GDP, GTP, or a GTP analog. There are three main analogs: GMPP[NH]P, GMPP[CH2]P, and GMPPP[γS], all of which differ from GTP by a single heavy atom. This atom can be mutated back to GTP, and either the GTP or GTP analog can be used. UCSF Chimera is used to protonate the cofactor and visually check the protonation state. Partial charges are computed using AM1-BCC charges calculated with Antechamber using the SQM program, as distributed with AmberTools [66]. Partial charges may also be obtained with Chimera's Dock Prep module, which for GDP and GTP can be looked up.

When starting from a chemical structure (SMILES), prepare a molecule for docking by performing the following steps (*see* Balius

et al. [2] for details): (1) tautomerization, (2) protonation, (3) conformation generation (going from a chemical structure to a 3D molecular conformation), (4) partial charge assignment, (5) desolvation decomposition, (6) conformational expansion (sampling the internal degrees of freedom of the molecule), and (7) *db2* file generation. To use *anchor-and-grow* and Grid Score, only **steps 1–4** are needed to generate *mol2* files with a single conformation. To use ChemGrid Score, **step 5** is needed, and desolvation per atom values need to be added to the mol2 file in a solvation section. For *HDB search*, **steps 1–7** should be performed. These steps are deployed by the ZINC database [67] to build molecules. If the molecule you are interested in docking exists in the ZINC database, simply download the molecule. There is no need to build the molecule yourself.

3 Examples

3.1 Preparing RAS for Molecular Docking

3.1.1 Protein Structure Choice

There are over 119 KRAS protein-ligand PDB structures. How can you pick the best system for docking? Perform retrospective calculations on each, including both cognate and cross-docking pose reproduction experiments, and enrichment calculations.

3.1.2 Pocket Selection

How can you pick the right pocket for docking? The procedure of blind docking, where the whole surface of the protein is explored instead of selecting a pocket, is not practical with UCSF DOCK. We recommend that the user pick a pocket. The most obvious way is to pick a pocket with a small molecule present. The nucleotide binding site on RAS is not considered targetable. When a small molecule is not present, it is less obvious how to pick a good pocket. Cryptic pockets are difficult to target unless experimental structures or long molecular simulations reveal them. You should use a holostructure with a bound ligand. If this is not possible, then you may consider other factors, e.g., pocket volume, how enclosed it is, how hydrophilic, how hydrophobic, or pocket shape. All can help in making the choice of what pocket to target. The *sphgen* module can be used to identify pockets. Deep, narrow sites that are mostly hydrophobic, but with polar patches, are often the most accessible to target with molecular docking.

3.1.3 Optimization of the Scoring Function

There are a few ways to adjust the scoring to change the energetics of the pocket. (1) Adjust the protein structure in the following ways: (i) Modify the protonation state on titratable residues: the most obvious residue is histidine, which can be charged, or protonated on the ε- or δ-nitrogen. (ii) Flip residues: glutamines and asparagines are good candidates because the amide nitrogen (H-bond donor) vs carbonyl oxygen (h-bond acceptor) offer up a

different environment. (iii) Rotate hydroxyls on serines and threonines, or flip the hydroxyl on tyrosines. (iv) Explore alternative sidechain conformations: sometimes there are alternative conformations in the PDB, or a rotamer library can be used. (2) Alter the balance between electrostatics and ligand desolvation terms. (i) In DOCK 3.7 and 3.8, this is primarily accomplished with *thin spheres*, which will adjust the molecular surface used in the electrostatics and desolvation grid calculations. The thicker the layer added to the surface, the stronger the electrostatics contribution; likewise, the thicker the layer of desolvation spheres, the more the poses are desolvated. Because the geometric placement of this layer is extending the surface, this strengthening of the contribution is dependent on the poses. The electrostatics *thin spheres* will encourage the placement of complementary charges close to the surface to be inside the boundary, so that it experiences a dielectric of 2 rather than 80. You can use the DOCK 3.7 and 3.8 scoring function in DOCK 6 and thus take advantage of these procedures. (ii) You can also increase the polar nature of groups by subtracting partial charge from one atom and adding it to another. (3) Add structural waters to the pocket. A structural water may be determined by looking at preserved water across multiple structures, or, alternatively, with inhomogeneous solvation theory. If a structural water is unfavorable to displace, then including the water as part of the receptor is advisable. It might be clear how the water hydrogens are oriented based on the receptor; if it is unclear, alternative water orientations might impact the calculation and could be considered.

3.1.4 Modification of Matching Spheres and Sampling Procedure

The matching spheres control where to place a small molecule. Only the anchor of the small molecule is placed on the spheres in anchor-and-grow, and likewise, its rigid segment is placed in *HDB search*. If the spheres are altered, the sampling will be perturbed. To alter the sampling behavior, you can perform the following: (1) Change the spheres. (i) Delete one or more spheres. For example, if the molecule that is being docked is being placed outside the pocket, then remove all spheres near where the molecule is being undesirably placed, and re-dock. (ii) Move one or more spheres. For example, if you notice that the binding mode is not exactly to your liking, you can move spheres to better position the ligand. There are automated scripts in *teb_scripts_programs* [68] that will randomly perturb the sphere coordinates. This is to help understand the sensitivity of the results on sphere locations. (2) In DOCK 6, there are parameters that can be applied, like *critical points*. This allows the user to specify that a specific sphere be used for all placements (orients). Thus, instead of deleting the spheres, as was discussed above, you can insist that DOCK match the ligand with specific spheres.

3.1.5 Chemical Space

The ZINC database is a publicly available resource that provides accessible chemical space, aspiring to store all vendor catalogs in standard file formats as well as DOCK-ready formats. The choice of the slice of the database to dock depends on your objectives. Docking *fragments* are useful for initial hits. They are often more soluble but are also lower in affinity. *Lead-like* is the most general portion to dock. *Drug-like* are large (and thus more flexible) and more expensive to dock (we often approach *drug-like* for selecting for analogs). In a typical workflow, you can dock *fragments* and *lead-like* to generate hits for experimental testing. *See* **Note 4** for strategies for docking fewer molecules without detection-accuracy loss for top-scoring hits. Library bias is a concern for screens on the order of 10^6 molecules. Perhaps, as libraries continue to grow, this will become less of an issue. However, there are still clear biases that exist in libraries [69]. For specific targets, it is possible to design bespoke libraries to bias toward the region of chemical space that likely is richer in binding molecules [70].

3.1.6 Hit Picking

Notes 2 and **1**, respectively, provide guidance on which top-scoring molecules to purchase and which to discard because of artifacts (i.e., broken molecules).

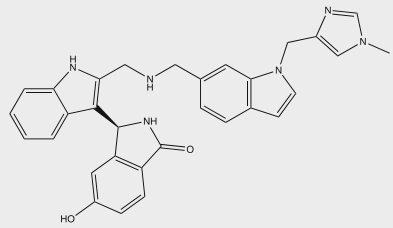
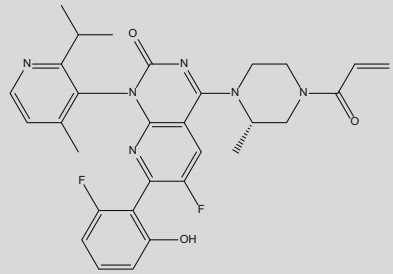
3.2 Two KRAS Examples

Here, we will walk through retrospective tests of docking to two KRAS systems (Table 1). (1) KRAS bound to inhibitor BI-2852, with GTP analog GMPPCP (PDB ID: 6GJ8). This ligand induces a dimer, and we dock to both the dimer and the monomer. (2) KRAS bound to covalent inhibitor sotorasib, the first approved drug that binds to KRAS, also called AMG 510, with cofactor GDP (PDB ID: 6OIM). We created online tutorials, available at https://github.com/tbalius/teb_docking_test_sets/wiki, where we go through these examples in step-by-step detail. In addition, we are developing a RAS docking test set, which includes these systems, that we plan to publish. These tutorials will be updated over time as DOCK develops and the RAS test set grows.

For both systems, start by preparing the receptor for docking by generating spheres and computing energetic grids (see above). Select the spheres within a certain distance from the experimentally determined ligand, and convert the experimental ligand atoms to spheres. To prepare the receptor for grid generation, complete missing side chains, add hydrogens to the protein, add hydrogens to the cofactor, and calculate the partial charges on the cofactors (*see* Subheading 2). Calculate the energetic grids. Next, prepare the crystallographic ligand for docking, and perform pose reproduction.

For BI-2852, dock to both the monomer and the dimer (*see* Table 2 and Fig. 3). When docking to the monomer, notice that DOCK correctly positions the portion of the ligand in the pocket to match the crystallographic ligand (*see* Fig. 3). However, DOCK

Table 1
Ligand and structures used for docking

Molecule	Chemical structure	SMILES	Notes
BI-2852		OC(C=C1)=CC([C@@H](C2=C(CNCC3=CC4=C(C=©CC5=CN(C)C=N5)C=C3)NC6=CC=CC=C62)N7)=C1C7=O	Binds to the SI/SII pocket. Appears to stabilize a nonfunctional dimer of RAS PDB ligand name: F0K PDB entries: 6GJ8, 6ZL5
Sotorasib		O=C(C=C)N1C[C@H]©N(C2=NC(N(C©(C)C=©C3C(C)C)C4=C2C=C(F)C(C5=C(O)C=CC=C5F)=N4)=O)CC1	First FDA-approved drug. Has two trade names: Lumakras and Lumykras Covalent molecule targeting G12C. Binds to the SII pocket PDB ligand name: MOV PDB entry: 6OIM

Underlined PDB entries used for docking in examples

Table 2
Pose reproduction results for non-covalent docking of BI-2852

Molecule	RMSD[a] (BE)[b]	Energy (BE)	RMSD (BR)[c]
BI-2852 monomer	3.43	−76.71	2.32
BI-2852 dimer	0.61	−116.33	Same

[a]RMSD calculated in DOCK using the Hungarian algorithm to correct for symmetry. Units are in Å
[b]BE stands for best energy pose
[c]BR stands for best RMSD pose from 25 poses kept

has difficulty orienting the methylimidazole ring. Despite a portion of the ligand matching the experimental pose closely, the RMSD is poor (RMSD = 3.43, *see* Table 2) because the pose diverges. If the ligand were to adopt the experimental pose in the context of the monomer, the methylimidazole ring would be solvent-exposed. Instead, a pose that lies along the surface is chosen. When docking to the dimer, additional interactions from the second protein allow DOCK to reproduce the experimental pose (RMSD = 0.61 Å).

For sotorasib, perform both non-covalent and covalent docking (*see* Table 3 and Fig. 4). Of the two, sotorasib is the most

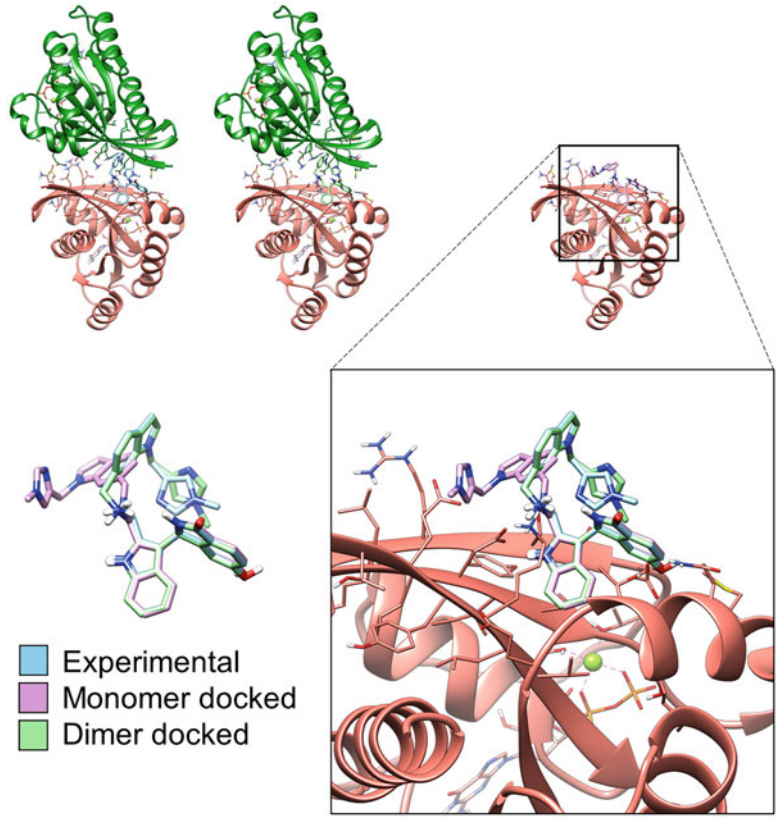


Fig. 3 Comparing poses of BI-2852 docked to the monomer or the dimer. Clockwise, starting from the top left: the dimer of KRAS is shown with BI-2852, with experimental pose in cyan, the pose docked to the dimer in light green, and the pose docked to the monomer in pink. In the expanded inset is the zoom-in of the pocket with all three poses overlaid. Finally, on the bottom left, the overlay of just the three poses is shown. (Images made with UCSF Chimera (orthographic mode))

challenging to successfully dock, due to considerations for how to deal with the covalent adduct, and issues arising from ligand strain of the experimental structure.

For non-covalent docking of sotorasib, alter the ligand by removing the atom that participates in the covalent bond, and computationally mutate the cysteine C12 to an alanine. The internal energy for the experimental pose is high. To reproduce the experimental pose, increase the internal energy parameter to 1,000,000.0 kcal/mol. This internal strain is not as large for the ligand built from SMILES, and the normal value of 100 can be used. This highlights that, for some ligands, particularly structurally congested molecules, the internal conformations may have large internal strain, and thus it is important to consider the baseline or minimum strain for that molecule (conformation). In pose

Table 3
Pose reproduction results for non-covalent docking and covalent docking of sotorasib

Molecule	RMSD[a] (BE)[b]	Energy (BE)	RMSD (BR)[c]	Notes
Sotorasib non-covalent from experimental structure	5.27	−57.17	4.29	Because of ligand strain of the experimental structure, you will need to adjust the parameter *internal_energy_cutoff* (1,000,000.0)
Sotorasib non-covalent from experimental structure (LAS subset[d])	0.40	−79.52	Same	DOCK reproduces the experimental pose
Sotorasib non-covalent from SMILES	4.56	−59.24	1.19	Conformation has less ligand strain energy (internal energy)
Sotorasib non-covalent from SMILES (LAS subset[d])	1.40	−62.25	Same	The pose looks good, but ring pucker is different and the 3-chlorophenol ring is flipped
Sotorasib covalent (SMILES)[e]	9.03	−11.97[f]	8.69	DOCK fails to reproduce the experimental pose
Sotorasib covalent (experimental structure)[e]	0.91	−43.93[f]	0.86	DOCK mostly reproduces the pose (but the pyridine ring is flipped)

[a]RMSD calculated in DOCK using the Hungarian algorithm to correct for symmetry. Units are in Å
[b]BE stands for best energy pose
[c]BR stands for best RMSD pose from 25 poses kept
[d]The *l*igand *a*toms are converted to *s*pheres. A solvent accessible group was removed to focus sampling in the cavity
[e]Result generated with DOCK 3.7
[f]ChemGrid Score (all others are Grid Score)

reproduction for sotorasib, the spheres make a difference. If you use spheres generated with *sphgen*, you will not find a good pose, with the best RMSD of 4.29 Å, and 1.19 Å when using SMILES (*see* Table 3). However, if you use the ligand atoms as spheres (*see* Fig. 4a, b), you will get better results: the RMSD is 0.40 Å for the ligand built from mol2 and 1.40 Å from SMILES (*see* Table 3). When you prepare the ligand database from SMILES, the pucker for the piperazine ring will not match the experimental structure unless additional, higher-energy ring conformations are generated.

For covalent docking, use DOCK 3.7, the DOCKovalent method [3]. DOCK 3 uses ChemGrid Score. Before generating energy grids, remove the gamma sulfur atom and its attached hydrogen atom from covalent residue CYS12. If you use the molecule generated from SMILES, you will not be able to reproduce the experimental pose (the best RMSD is 8.69 Å, *see* Table 3, Fig. 4c). However, when you use the molecule generated from the experimental structure (bond length and angles are preserved, but dihedral information is forgotten), you will be able to reproduce the

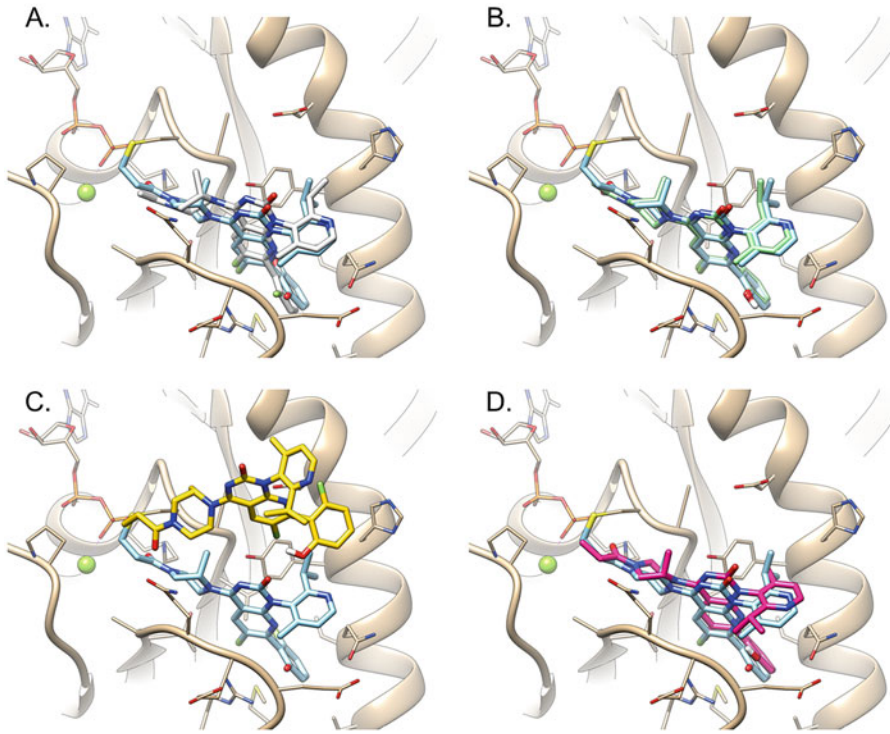


Fig. 4 Sotorasib docking pose reproduction experiments. (**a**) Non-covalent docking pose (white) generated from SMILES docked with DOCK6 using the anchor-and-grow algorithm and compared to experimental pose (sky blue). (**b**) Non-covalent docking pose (light green) generated from the experimental structure (i.e., bonds and angles are preserved from the experimental structure) docked with DOCK6 using the anchor-and-grow algorithm. (**c**) Covalent docking pose (yellow) from SMILES docked with DOCK3 using the DOCKovalent algorithm. (**d**) Covalent docking pose (magenta) from experimental structure docked with DOCK3 using DOCKovalent algorithm

pose of the experimental structure: 0.91 Å (best score) and 0.86 Å (best RMSD), as shown in Table 3 and Fig. 4d.

To perform enrichment, begin by preparing a database of ligand and property-matched decoys for docking. The database should include five ligands from the PDB that bind to the Switch I/Switch II (SOS) pocket, including the BI-2852 molecule (see tutorial instructions). Property-matched decoys are property-matched to the ligands by formal charge, molecular weight, calculated water-octanol partition coefficient, rotatable bonds, number of hydrogen-bond donors, and number of hydrogen-bond acceptors. Include potential decoys for each ligand protomer, discarding decoys too similar to ligands, and clustering decoys that are similar to each other. Aim to assign 50 decoys to each ligand protomer. Online resources for generating property-matched decoys from a set of ligand SMILES are available at https://dude.docking.org and https://tldr.docking.org [71, 72]. DUDE decoys can be used for this calculation (see the tutorial for more information). Using the

Table 4
Enrichment for the SOS pocket

	AUC	logAUC
Grid Score	77.60	19.45
Grid Score, keep 10	77.60	17.37
Grid Score, keep 10, torsion min	75.93	18.13
ChemGrid Score	63.60	17.01
ChemGrid Score, keep 10	63.67	19.99
ChemGrid Score, keep 10, torsion min	64.33	18.64

DOCK 3.8 pipeline, prepare ligand and decoy molecules for docking in the *db2* file format, and then perform docking with DOCK 6's *HDB search*.

Upon performing the calculation, you will see that Grid Score seems to outperform ChemGrid Score for AUC (77.6 compared to 63.60), but that they perform more similarly for logAUC (19.45 compared to 17.01) (*see* Table 4). For ChemGrid Score, if you conduct additional sampling, it seems to help (17.01 compared to 19.99), but minimization does not seem to help (*see* Table 4).

3.3 Illustrative Virtual Screening Results

We performed a large-scale virtual screen of ~158.7 million molecules with DOCK 3.8 on the SOS pocket defined by the BI-2852 molecule (PDB: 6GJ8) on RAS as a monomer. A similar docking setup was used to that described above, but the GTP analog was kept as GMPPCP (not converted to GTP). We visualized the top 1000 poses. Although many of the top-scoring molecules look promising, here we showcase four molecules, which we would *discard* based on the docked pose, to illustrate docking *issues*. We discard molecules that (1) have multidentate interactions (an atom of one residue should not accept many interactions from the small molecule pose) (*see* Fig. 5a); (2) look like known binders (*see* Fig. 5a); (3) are occupying a tautomeric or protonation state that is not well occupied at the relevant experimental pH, most often pH 7 (*see* Fig. 5b); (4) float above the pocket and do not fill the pocket or make good contact with the pocket (*see* Fig. 5c); and (5) have stranded interactions (*see* Fig. 5d). *See* **Note 2** for more details of rules of thumb. To go into more detail on the example from point 3 (Fig. 5b), the molecule is incorrectly protonated. The 2-aminoquinoline ring has a predicted pKa of 6.16, according to ChemAxon, and will be 91.64% in the neutral state at pH 7.2. When chlorine is in the ring, the predicted pKa is lower (5.41), resulting in 98.41% of the species being in a neutral state at pH 7.2. The conformation is strained because the 2-aminoquinoline ring group should be planar.

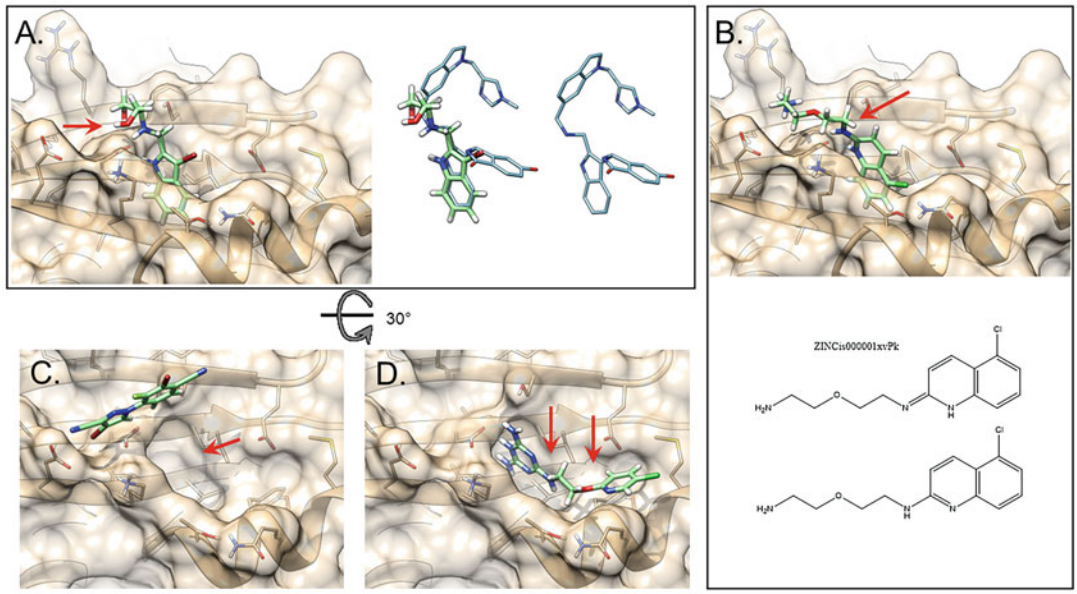


Fig. 5 Molecules that exhibit unwanted behavior or properties from an example screen. (**a**) A molecule that resembles the experimental molecule and that has a congested tri- or quadridentate interaction. Red arrow points to congested interaction. (**b**) A molecule that is an incorrect protomer and tautomer and is in a strained conformation. (**c**) A molecule that does not fill the site. (**d**) A molecule that has stranded h-bond donors and stranded acceptors

4 Concluding Thoughts

Here, we hope that you have learned about the following: (1) The molecular binding event, and how docking uses approximations to model binding at scale. (2) Molecular docking of covalent and non-covalent ligands using DOCK. (3) KRAS, and issues with docking to KRAS. Potential issues in these docking calculations associated with the small molecules, including ring puckers and internal strain, have been explored. (4) Rules of thumb for picking molecules from virtual screens. (5) How to perform retrospective calculations, including pose reproduction and enrichment calculations to evaluate a docking setup. (6) How to choose a database of small molecules for virtual screening. (7) Potential pitfalls using molecular docking for ligand discovery, and how to avoid them.

5 Notes

1. Pitfalls in molecular docking for virtual screening.

   Broken molecules habitually score well: These molecules exploit issues with the scoring function. Examples of broken molecules include (1) incorrect protonation states,

(2) unphysical tautomerization, (3) strained ligand conformation, (4) incorrect partial charge assignment, and (5) incorrect desolvation per atom assignment.

2. Rules of thumb for prioritizing molecules from a virtual screen.

   Discard the docked poses of molecules if they possess the following: (1) Unsatisfied h-bond acceptors and donors. Pay attention to internal h-bonds. One strained donor and two strained acceptors are permitted. (2) Poor surface contacts, large gaps between receptor and ligand surfaces. (3) Internal issues that indicate that it is a broken molecule (*see* **Note 1**). (4) Many internal degrees of freedom (entropic cost). (5) Extreme hydrophobicity. If it is too polar, it will pay a cost to come out of water. If it is too hydrophobic, it might not be able to be tested because it falls out of solution. (6) Extreme size. If it is too big, it will not fit in the pocket. Conversely, if it is too small, it might not fill the site, or it might be nonspecific. (7) Reactive groups. We might permit weak reactive groups, but will discard very reactive molecules. (8) Pan-assay interference compounds (PAINS). We will often permit molecules with these vulnerabilities, since these molecules are meant to be leads. (9) Similarity to known aggregators. One of the biggest nonspecific binding artifacts is colloidal aggregation, where a compound will form colloids and will absorb the proteins on their surface and interfere with assays. (10) Molecules that resemble undesirable molecules, for example, known binders (we are looking for the new, not the old), or bad actors (points 7–9). Molecules selected for experimental testing should satisfy critical interactions: e.g., they should make a key salt-bridge or a key hydrogen bond. These rules of thumb are not hard and uncompromising. If you believe in a molecule, test it.

3. A thought experiment: Chemical space tractability.

   Assume a docking rate of 1 molecule per second, the average time for DOCK [73]. Docking 0.5 billion molecules takes 500 million seconds on one CPU core. If we run on 1000 CPU cores, it will take 500,000 s, which is ~5.8 days. If we dock 5 billion molecules on the same number of machines, it will take about 10 times longer: ~58 days, or ~2 months. This is too long, but there are ways to make it faster: (1) Dock with more CPUs: one way to do this is to use the cloud [74]; (2) make docking faster through algorithm improvements; and (3) focus on the productive portion of chemical space.

   Another consideration is disk space. 0.5 billion molecules use about 50 TB of disk space. Five billion would use ten times that, about 500 TB or 0.5 PB. This is used to store the precomputed conformations of the ligands. At scale, the overhead incurred by filesystem operations begins to play a role in

the throughput of molecular docking algorithms [75]. Parallelization of docking necessitates consideration of the costs incurred by file writing, reading, and memory usage, and algorithms developed for docking billions of molecules must consider their use of filesystem resources to ensure reasonable throughput and stability. Do we have any real advantage to docking more molecules? How big is too big? In Lyu et al. [76], the authors state that docking larger and larger databases in the foreseeable future is worth the additional investment in computational time and resources.

4. Approaches to more efficiently sample rapidly expanding chemical space.

   To address expanding chemical libraries, ideas for improving the efficiency and throughput of docking have been developed. For example, chemical space can be reduced. One can dock *synthons*, which are fragments based on reactants used to generate a chemical library, like Enamine Real. The molecules from the chemical library associated with the top scoring synthons can then be docked [77, 78]. These *synthons* scale on the order of the second or third root of the library size. Another method is to train a machine learning model, using chemical structures and docking screens, to predict docking scores to prefilter libraries to help prioritize molecules for docking. This has an effective speed up of 10×, with the potential for speedups of 100–1000× given improvements in model accuracy, and does not significantly impact the quality of hits [79]. This type of method would be trained on a specific docking setup (protein, pocket, setup parameters used), and this training needs to happen before each unique docking campaign. Related to this are active learning approaches, which have also been used effectively to increase throughput and reduce computing costs by more than an order of magnitude. In this process, a small subset of molecules is randomly selected from a library, docked, and used to train a machine learning model to predict the docking scores of the remaining library, followed by iterative rounds of selecting the top-ranked machine learning predictions for further docking to refine the model [80–82]. A similar method is Adaptive Target-Guided Virtual Screens [83]. Here, a sparse docking is performed, and then the algorithm focuses on the perceived richest part of chemical space. Alternatively, the user could pick the space based on their own knowledge of the pocket. Other approaches have been hardware-based, e.g., adapting docking algorithms to run on GPUs [84], and in combination with software that permits the usage of high-performance computing or cloud-based resources [74, 85].

Acknowledgments

We thank Alyssa Klein, Muyan Zhou, Dillon Chu, Maryam Abdur-Rahman, Yilin Liang, and Yunpeng Wang for contributions to the tutorials. We thank colleagues at the NCI RAS Initiative, including Dwight Nissley and Frank McCormick, and the Frederick Research Computing Environment (FRCE). This project was funded in whole or in part with federal funds from the National Cancer Institute, NIH Contract 75N91019D00024. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, and the mention of trade names, commercial products, or organizations does not imply endorsement by the US government.

References

1. Kuntz ID, Blaney JM, Oatley SJ et al (1982) A geometric approach to macromolecule-ligand interactions. J Mol Biol 161:269–288
2. Balius TE, Tan YS, Chakrabarti M (2024) DOCK 6: incorporating hierarchical traversal through precomputed ligand conformations to enable large-scale docking. J Comput Chem 45:47–63. https://doi.org/10.1002/jcc.27218
3. London N, Miller RM, Krishnan S et al (2014) Covalent docking of large libraries for the discovery of chemical probes. Nat Chem Biol 10:1066–1072
4. Bender BJ, Gahbauer S, Luttens A et al (2021) A practical guide to large-scale docking. Nat Protoc 16:4799–4832
5. Kamal IM, Chakrabarti S (2023) MetaDOCK: a combinatorial molecular docking approach. ACS Omega 8:5850–5860
6. Pagadala NS, Syed K, Tuszynski J (2017) Software for molecular docking: a review. Biophys Rev 9:91–102
7. Zev S, Raz K, Schwartz R et al (2021) Benchmarking the ability of common docking programs to correctly reproduce and score binding modes in SARS-CoV-2 protease Mpro. J Chem Inf Model 61:2957–2966
8. Shuya N, Yoshiharu M, Shigenori T (2022) End-to-end protein-ligand complex structure generation with diffusion-based generative models. bioRxiv. https://doi.org/10.1101/2022.12.20.521309
9. Dara S, Dhamercherla S, Jadav SS et al (2022) Machine learning in drug discovery: a review. Artif Intell Rev 55:1947–1999
10. Vamathevan J, Clark D, Czodrowski P et al (2019) Applications of machine learning in drug discovery and development. Nat Rev Drug Discov 18:463–477
11. Kitchen DB, Decornez H, Furr JR et al (2004) Docking and scoring in virtual screening for drug discovery: methods and applications. Nat Rev Drug Discov 3:935–949
12. Chen P, Ke Y, Lu Y et al (2019) DLIGAND2: an improved knowledge-based energy function for protein-ligand interactions using the distance-scaled, finite, ideal-gas reference state. J Cheminform 11:52
13. Guedes IA, Pereira FSS, Dardenne LE (2018) Empirical scoring functions for structure-based virtual screening: applications, critical aspects, and challenges. Front Pharmacol 9:1089
14. Liu J, Wang R (2015) Classification of current scoring functions. J Chem Inf Model 55:475–482
15. Meli R, Morris GM, Biggin PC (2022) Scoring functions for protein-ligand binding affinity prediction using structure-based deep learning: a review. Front Bioinform 2:885983
16. Torres PHM, Sodero ACR, Jofily P et al (2019) Key topics in molecular docking for drug design. Int J Mol Sci 20:4574
17. McNutt AT, Francoeur P, Aggarwal R et al (2021) GNINA 1.0: molecular docking with deep learning. J Cheminform 13:43
18. Jiang H, Fan M, Wang J et al (2020) Guiding conventional protein-ligand docking software with convolutional neural networks. J Chem Inf Model 60:4594–4602
19. Pereira JC, Caffarena ER, Dos Santos CN (2016) Boosting docking-based virtual screening with deep learning. J Chem Inf Model 56:2495–2506

20. Mathieu et al., KRAS G12C fragment screening renders new binding pockets, Small GTPases, 13:1, 225–238, https://doi.org/10.1080/21541248.2021.1979360
21. Liao J, Shima F, Araki M et al (2008) Two conformational states of Ras GTPase exhibit differential GTP-binding kinetics. Biochem Biophys Res Commun 369:327–332
22. Spoerner M, Hozsa C, Poetzl JA et al (2010) Conformational states of human rat sarcoma (Ras) protein complexed with its natural ligand GTP and their role for effector interaction and GTP hydrolysis. J Biol Chem 285:39768–39778
23. Ye M, Shima F, Muraoka S et al (2005) Crystal structure of M-Ras reveals a GTP-bound "off" state conformation of Ras family small GTPases. J Biol Chem 280:31267–31275
24. Shima F, Ijiri Y, Muraoka S et al (2010) Structural basis for conformational dynamics of GTP-bound Ras protein. J Biol Chem 285:22696–22705
25. Kalbitzer HR, Spoerner M, Ganser P et al (2009) Fundamental link between folding states and functional states of proteins. J Am Chem Soc 131:16714–16719
26. Chao F-A, Chan AH, Dharmaiah S et al (2023) Reduced dynamic complexity allows structure elucidation of an excited state of KRASG13D. Commun Biol 6:594
27. Shima F, Yoshikawa Y, Ye M et al (2013) In silico discovery of small-molecule Ras inhibitors that display antitumor activity by blocking the Ras-effector interaction. Proc Natl Acad Sci USA 110:8182–8187
28. Rosnizeck IC, Filchtinski D, Lopes RP et al (2014) Elucidating the mode of action of a typical Ras state 1(T) inhibitor. Biochemistry 53:3867–3878
29. Rosnizeck IC, Graf T, Spoerner M et al (2010) Stabilizing a weak binding state for effectors in the human ras protein by cyclen complexes. Angew Chem Int Ed Engl 49:3830–3833
30. Rosnizeck IC, Spoerner M, Harsch T et al (2012) Metal-bis(2-picolyl)amine complexes as state 1(T) inhibitors of activated Ras protein. Angew Chem Int Ed Engl 51:10647–10651
31. Ostrem JM, Peters U, Sos ML et al (2013) K-Ras(G12C) inhibitors allosterically control GTP affinity and effector interactions. Nature 503:548–551
32. McCormick F (2020) Sticking it to KRAS: covalent inhibitors enter the clinic. Cancer Cell 37:3–4
33. Broker J, Waterson AG, Smethurst C et al (2022) Fragment optimization of reversible binding to the switch II pocket on KRAS leads to a potent, in vivo active KRAS(G12C) inhibitor. J Med Chem 65:14614–14629
34. Wang X, Allen S, Blake JF et al (2022) Identification of MRTX1133, a noncovalent, potent, and selective KRAS(G12D) inhibitor. J Med Chem 65:3123–3133
35. Kim D, Herdeis L, Rudolph D et al (2023) Pan-KRAS inhibitor disables oncogenic signalling and tumour growth. Nature 619:160
36. Kwan AK, Piazza GA, Keeton AB et al (2022) The path to the clinic: a comprehensive review on direct KRAS(G12C) inhibitors. J Exp Clin Cancer Res 41:27
37. Punekar SR, Velcheti V, Neel BG et al (2022) The current state of the art and future trends in RAS-targeted cancer therapies. Nat Rev Clin Oncol 19:637–655
38. Rosen JC, Sacher A, Tsao MS (2023) Direct GDP-KRAS(G12C) inhibitors and mechanisms of resistance: the tip of the iceberg. Ther Adv Med Oncol 15:17588359231160141
39. Dong L, Qu X, Zhao Y et al (2021) Prediction of binding free energy of protein-ligand complexes with a hybrid molecular mechanics/generalized born surface area and machine learning method. ACS Omega 6:32938–32947
40. Kollman PA, Massova I, Reyes C et al (2000) Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc Chem Res 33:889–897
41. Genheden S, Ryde U (2015) The MM/PBSA and MM/GBSA methods to estimate ligand-binding affinities. Expert Opin Drug Discov 10:449–461
42. Lorber DM, Shoichet BK (1998) Flexible ligand docking using conformational ensembles. Protein Sci 7:938–950
43. Wang Z, Sun H, Yao X et al (2016) 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 18:12964–12975
44. Warren GL, Andrews CW, Capelli AM et al (2006) A critical assessment of docking programs and scoring functions. J Med Chem 49:5912–5931
45. Xu M, Shen C, Yang J et al (2022) Systematic investigation of docking failures in large-scale structure-based virtual screening. ACS Omega 7:39417–39428
46. Meng EC, Shoichet BK, Kuntz ID (1992) Automated docking with grid-based energy evaluation. J Comput Chem 13:505–524

47. Coleman RG, Carchia M, Sterling T et al (2013) Ligand pose and orientational sampling in molecular docking. PLoS One 8:e75992
48. Mysinger MM, Shoichet BK (2010) Rapid context-dependent ligand desolvation in molecular docking. J Chem Inf Model 50: 1561–1573
49. Shoichet BK, Leach AR, Kuntz ID (1999) Ligand solvation in molecular docking. Proteins 34:4–16
50. Balius TE, Fischer M, Stein RM et al (2017) Testing inhomogeneous solvation theory in structure-based ligand discovery. Proc Natl Acad Sci USA 114:E6839–E6846
51. Gu S, Smith MS, Yang Y et al (2021) Ligand strain energy in large library docking. J Chem Inf Model 61:4331–4341
52. Fischer M, Coleman RG, Fraser JS et al (2014) Incorporation of protein flexibility and conformational energy penalties in docking screens to improve ligand discovery. Nat Chem 6:575–583
53. Balius TE, Mukherjee S, Rizzo RC (2011) Implementation and evaluation of a docking-rescoring method using molecular footprint comparisons. J Comput Chem 32:2273–2289
54. Lang PT, Brozell SR, Mukherjee S et al (2009) DOCK 6: combining techniques to model RNA-small molecule complexes. RNA 15: 1219–1230
55. Allen WJ, Balius T, Bickel J et al (2023) DOCK 6.10 users manual. [cited May 3, 2023]. Available from https://dock.compbio.ucsf.edu/DOCK_6/dock6_manual.htm
56. Mukherjee S, Balius TE, Rizzo RC (2010) Docking validation resources: protein family and ligand flexibility experiments. J Chem Inf Model 50:1986–2000
57. Allen WJ, Rizzo RC (2014) Implementation of the Hungarian algorithm to account for ligand symmetry and similarity in structure-based design. J Chem Inf Model 54:518–529
58. Knight IS, Naprienko S, Irwin JJ (2022) Enrichment Score: a better quantitative metric for evaluating the enrichment capacity of molecular docking models. arXiv:2210.10905. https://doi.org/10.48550/arXiv.2210.10905
59. Knight IS, Mailhot O, Tang KG, Irwin JJ, (2024) DockOpt: A tool for automatic optimization of docking models. J Chem Inf Model Article ASAP. https://doi.org/10.1021/acs.jcim.3c01406
60. Hawkins PC, Warren GL, Skillman AG et al (2008) How to do an evaluation: pitfalls and traps. J Comput Aided Mol Des 22:179–190
61. Amezcua M, Setiadi J, Ge Y et al (2022) An overview of the SAMPL8 host-guest binding challenge. J Comput Aided Mol Des 36:707–734
62. Graves AP, Shivakumar DM, Boyce SE et al (2008) Rescoring docking hit lists for model cavity sites: predictions and experimental testing. J Mol Biol 377:914–934
63. Kamenik AS, Singh I, Lak P et al (2021) Energy penalties enhance flexible receptor docking in a model cavity. Proc Natl Acad Sci USA 118:e2106195118
64. Pettersen EF, Goddard TD, Huang CC et al (2004) UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 25:1605–1612
65. Word JM, Lovell SC, Richardson JS et al (1999) Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. J Mol Biol 285:1735–1747
66. Case DA, Ben-Shalom IY, Brozell SR et al (2019) AMBERTools 2019. University of California, San Francisco
67. Tingle BI, Tang KG, Castanon M et al (2023) ZINC-22 – a free multi-billion-scale database of tangible compounds for ligand discovery. J Chem Inf Model 63:1166–1176
68. Balius TE (2023) teb_scripts_programs. Available from https://github.com/tbalius/teb_scripts_programs/tree/master/zzz.scripts/
69. Hert J, Irwin JJ, Laggner C et al (2009) Quantifying biogenic bias in screening libraries. Nat Chem Biol 5:479–483
70. Kaplan AL, Confair DN, Kim K et al (2022) Bespoke library docking for 5-HT(2A) receptor agonists with antidepressant activity. Nature 610:582–591
71. Irwin JJ, Shoichet BK, Mysinger MM et al (2009) Automated docking screens: a feasibility study. J Med Chem 52:5712–5720
72. Mysinger MM, Carchia M, Irwin JJ et al (2012) Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. J Med Chem 55:6582–6594
73. Beroza P, Crawford JJ, Ganichkin O et al (2022) Chemical space docking enables large-scale structure-based virtual screening to discover ROCK1 kinase inhibitors. Nat Commun 13:6447
74. Tingle BI, Irwin JJ (2023) Large-scale docking in the cloud. J Chem Inf Model 63:2735–2741
75. Lu H, Wei Z, Wang C et al (2021) Redesigning Vina@QNLM for ultra-large-scale molecular docking and screening on a Sunway supercomputer. Front Chem 9:750325

76. Lyu J, Irwin JJ, Shoichet BK (2023) Modeling the expansion of virtual screening libraries. Nat Chem Biol 19:712–718
77. Sadybekov AA, Sadybekov AV, Liu Y et al (2022) Synthon-based ligand discovery in virtual libraries of over 11 billion compounds. Nature 601:452–459
78. Sadybekov AV, Katritch V (2023) Computational approaches streamlining drug discovery. Nature 616:673–685
79. Clyde A, Liu X, Brettin T et al (2023) AI-accelerated protein-ligand docking for SARS-CoV-2 is 100-fold faster with no significant change in detection. Sci Rep 13:2105
80. Gentile F, Yaacoub JC, Gleave J et al (2022) Artificial intelligence-enabled virtual screening of ultra-large chemical libraries with deep docking. Nat Protoc 17:672–697
81. Graff DE, Shakhnovich EI, Coley CW (2021) Accelerating high-throughput virtual screening through molecular pool-based active learning. Chem Sci 12:7866–7881
82. Yang Y, Yao K, Repasky MP et al (2021) Efficient exploration of chemical space with docking and deep learning. J Chem Theory Comput 17:7106–7119
83. Christoph G, AkshatKumar N, Matt K et al (2023) VirtualFlow 2.0 – the next generation drug discovery platform enabling adaptive screens of 69 billion molecules. bioRxiv. https://doi.org/10.1101/2023.04.25.537981
84. Yu Y, Cai C, Wang J et al (2023) Uni-Dock: GPU-accelerated docking enables ultralarge virtual screening. J Chem Theory Comput 19:3336–3345
85. Zhang B, Li H, Yu K et al (2022) Molecular docking-based computational platform for high-throughput virtual screening. CCF Trans High Perform Comput 4:63–74