Published in final edited form as:

Nature. 2020 April ; 580(7805): 663–668. doi:10.1038/s41586-020-2117-z.

An open-source drug discovery platform enables ultra-large virtual screens

Christoph Gorgulla[1,2,3,*], **Andras Boeszoermenyi**[1,3,11], **Zi-Fu Wang**[1,11], **Patrick D. Fischer**[1,3,9], **Paul Coote**[1,3], **Krishna M. Padmanabha Das**[1,3], **Yehor S. Malets**[5,7], **Dmytro S. Radchenko**[5,7], **Yurii S. Moroz**[6,7], **David A. Scott**[1,3], **Konstantin Fackeldey**[8,10], **Moritz Hoffmann**[4], **Iryna Iavniuk**[5], **Gerhard Wagner**[1], **Haribabu Arthanari**[1,3,*]

[1]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Harvard University, Boston, USA.

[2]Department of Physics, Faculty of Arts and Sciences, Harvard University, Cambridge, USA.

[3]Department of Cancer Biology, Dana-Farber Cancer Institute, Boston, USA.

[4]Department of Mathematics and Computer Science, Freie Universität Berlin, Berlin, Germany.

[5]Enamine, Kiev, Ukraine.

[6]Chemspace, Riga, Latvia.

[7]National Taras Shevchenko University of Kiev, Kiev, Ukraine.

[8]Zuse Institute Berlin (ZIB), Berlin, Germany.

[9]Department of Pharmacy, Pharmaceutical and Medicinal Chemistry, Saarland University, Saarbrücken, Germany.

[10]Institute of Mathematics, Technical University Berlin, Berlin, Germany.

[11]These authors contributed equally to this work.

Abstract

* cgorgulla@g.harvard.edu, hari@hms.harvard.edu.

Author contributions C.G. conceived the project, and designed and implemented the drug discovery platform (VirtualFlow). H.A. and C.G designed the experimental workflow. A.B., H.A., Z.W., P.D.F. and K.M.P.D. designed and carried out the FP and NMR experiments. Z.W. designed and carried out the SPR and BLI experiments. K.M.P.D. and Z.W. carried out the DLS experiments. M.H. provided technical assistance regarding the code and homepage. C.G. designed the applications (screening of KEAP1, preparation of the REAL library). P.C. analysed the NMR data. C.G. carried out the computations done with VFLP (preparation of the REAL database) and VFVS (screening/rescoring of KEAP1). Y.Ma. created the web interface to the REAL database. C.G. prepared the VirtualFlow homepage. Y.Mo. prepared the REAL database in the initial SMILES format. C.G. and Y.Mo. designed the structure of the VirtualFlow-version of the REAL database. Y.Mo., I.I., and D.S.R. supervised and directed the synthesis and purification of the on-demand compounds from the REAL library. D.A.S. helped to evaluate the screening hits. C.G., H.A., A.B., K.F., Z.W., P.C. and G.W. prepared the manuscript. K.F., H.A. and G.W. supervised the project.

Competing interests The authors declare the following competing interests: I.I., D.S.R., and Ye.M. work for Enamine, a company that is involved in the synthesis and distribution of drug-like compounds. Yu.M. is a scientific advisor to Enamine.

Additional information

Supplementary Information is available for this paper at ...

On average, an approved drug today costs $2–3 billion and takes over ten years to develop[1]. In part, this is due to expensive and time-consuming wet-lab experiments, poor initial hit compounds, and the high attrition rates in the (pre-)clinical phases. Structure-based virtual screening (SBVS) has the potential to mitigate these problems. With SBVS, the quality of the hits improves with the number of compounds screened[2]. However, despite the fact that large compound databases exist, the ability to carry out large-scale SBVSs on computer clusters in an accessible, efficient, and flexible manner has remained elusive. Here we designed VirtualFlow, a highly automated and versatile open-source platform with perfect scaling behaviour that is able to prepare and efficiently screen ultra-large ligand libraries of compounds. VirtualFlow is able to use a variety of the most powerful docking programs. Using VirtualFlow, we have prepared the largest and freely available ready-to-dock ligand library available, with over 1.4 billion commercially available molecules. To demonstrate the power of VirtualFlow, we screened over 1 billion compounds and discovered a small molecule inhibitor (iKeap1) that engages KEAP1 with nanomolar affinity (K_d = 114 nM) and disrupts the interaction between KEAP1 and the transcription factor NRF2. We also identified a set of structurally diverse molecules that bind to KEAP1 with submicromolar affinity. This illustrates the potential of VirtualFlow to access vast regions of the chemical space and identify binders with high affinity for target proteins.

Repeated optimization of lead compounds and late-stage failure of drug candidates are the primary causes of longer development times and increased costs in drug development. Improving the quality of the initial lead compounds would minimize these lead optimization cycles and result in drug candidates entering (pre-)clinical phases with greater specificity and higher affinity. Virtual screening to identify molecules that bind to a specified site on a receptor protein has become an important part of the drug discovery pipeline[2–5]. Current virtual screening paradigms routinely sample only a tiny fraction, on the order of 10^6 −10^7 molecules, of the total chemical space of small organic compounds suitable for drug discovery, estimated to encompass more than 10^{60} molecules[6]. However, the scale of a virtual screen is of central importance because the more compounds that are screened, (a) the lower the rate of false positives, and (b) the more favourable the quality of the lead compounds (*e.g.*, higher-affinity binders). It was recently shown experimentally that ultra-large scale screening improves the rate of true positives[2]. Here we derived a probabilistic model of the true-positive rate as a function of the number of compounds screened, and analysis of our ultra-large screen confirms that the docking score of the highest-scoring compounds improve with the scale. Increasing the scale of a virtual screen can improve the quality of initial hits in two distinct ways: (1) by identifying hits with tighter binding affinity, which can result in lowered dosages and fewer off-target effects, and (2) by discovering compounds with more favourable pharmacokinetic and/or less inherent cytotoxic properties.

To increase the number of compounds evaluated in a virtual screen by orders of magnitude and make it accessible to any researcher, there is a dire need for a platform that can integrate all the tasks in the virtual screening process. Such a platform should ideally (1) scale linearly with the number of CPUs, (2) efficiently handle billions of files, (3) minimize input and output (I/O) load, (4) run robustly (*e.g.*, skip incorrectly encoded ligands, resist temporary I/O problems, and resume following unexpected termination), (5) run on any

type of computing cluster (including cloud platforms), and (6) be user-friendly and easy to use for non-computational scientists. Furthermore, to provide flexibility, a SBVS platform should be able to interface with a variety of docking programs, support both rigid and flexible receptor docking, test multiple docking scenarios in a single workflow, allow for consensus and ensemble docking, and carry out multiple replicas of the same docking scenario. Lastly, to democratize access, facilitate widespread usage, and catalyse further development, such a platform would need to be open source.

With these requirements in mind, we designed VirtualFlow, an open-source platform that is able to screen chemical space on an unprecedented scale. Screening one billion compounds on a single processor core, with an average docking time of 15 seconds per ligand, would take 475 years. By contrast, VirtualFlow can dock one billion compounds in approximately two weeks by leveraging 10,000 CPU cores simultaneously. Such high performance computing facilities are available to researchers through several potential sources, including local institute computer clusters, national supercomputing centres, or cloud computing platforms.

Targeting KEAP1 using VirtualFlow

To test the advantages of ultra-large-scale *in silico* screening and the performance of the VirtualFlow platform we decided to target the challenging and therapeutically relevant protein-protein interaction (PPI) between nuclear factor erythroid-derived 2-related factor 2 (NRF2) and Kelch-like ECH-associated protein 1 (KEAP1). NRF2 is a master regulator of cellular resistance to oxidative stress and cellular repair[7]. Under unstressed conditions, NRF2 is sequestered by KEAP1, an E3 ubiquitin ligase substrate adaptor, and targeted for degradation[8]. However, upon oxidative stress, reactive oxidants dissociate NRF2 from KEAP1 and NRF2 translocates to the nucleus to activate its transcriptional program of approximately 250 genes[9]. The NRF2-KEAP1 pathway is critical in protecting the cell under oxidative stress and inflammation and is implicated in a number of diseases[10]. There are ten drugs targeting KEAP1 that are in clinical trials and nine more that are at the preclinical stage[10]. Using VirtualFlow, we screened ~1.3 billion compounds (~1 billion compounds from the Enamine REAL Library and ~330 million compounds from the ZINC library) against the NRF2 interaction interface on KEAP1. First, we would like to describe the salient features of VirtualFlow and its scalability.

Characteristic features of VirtualFlow

One of the key features of VirtualFlow is its linear scaling behaviour ($\mathcal{O}(N)$) with respect to the number of CPUs and nodes utilized. VirtualFlow can run on computer clusters operated with any of the major resource managers (SLURM[11], Moab/TORQUE[12], PBS[13], LSF[14] and SGE[15]), and compatibility with additional job schedulers can be easily added. Thus VirtualFlow is also ideally configured for cloud computing platforms like Amazon's Web Services (AWS), Microsoft's Azure and Google's Cloud Platform (GCP). VirtualFlow is able to run autonomously from the first to the last ligand in the screening pipeline, a feature facilitated by automatic submission of new batch system jobs. The workflow can be monitored and controlled in real time. The VirtualFlow package consists of two applications

that work seamlessly together: The VFLP (VirtuaFlow for Ligand Preparation) module, which prepares small molecules for screening; and the VFVS (VirtualFlow for Virtual Screening) module, which executes the virtual screening procedures (Fig. 1). The separation of ligand preparation and virtual screening is desirable because the same ready-to-dock ligand library can be used in any number of VFVS virtual screens.

VirtualFlow for Ligand Preparation (VFLP)

VFLP (VirtualFlow for Ligand Preparation) prepares ligand databases by converting them from the SMILES format into any desired target format (*e.g.*, the PDBQT format, which is required by many of the AutoDock-based docking programs). VFLP uses ChemAxon's JChem package as well as Open Babel to desalt ligands, neutralize them, generate (even multiple) tautomeric states, compute protonation states at specific pH values, calculate 3D coordinates, and convert the molecules into desired target formats (Extended Data Fig. 2). The output file formats currently supported by VFLP are shown in Supplementary Table 7.

Preparation of the Enamine REAL library

Commercially available compounds constitute the most interesting subset of the chemical space, since these compounds can be readily purchased. The largest vendor library available today is the REAL library of Enamine, containing approximately 1.4 billion make-on-demand compounds (as of October 2019 the ZINC 15 database contained 1.46 billion compounds, but only provided 630 million molecules in a ready-to-dock format). We have used VFLP to convert the ~1.4 billion compounds of the REAL library into PDBQT format (see Methods), and have made it freely available on the VirtualFlow homepage, accessible via a graphical interface (Supplementary Fig. 5). The entire database has a six-dimensional lattice architecture, the general concept of which was modelled after the ZINC 15 database[16], where each dimension corresponds to a physico-chemical property of the compounds (molecular weight, partition coefficient, number of hydrogen bond donors and acceptors, number of rotatable bonds, and the topological polar surface area). The preparation of ligands using VFLP is a one-time effort.

VirtualFlow for Virtual Screening (VFVS)

To set up a virtual screen with VFVS, a set of docking scenarios is specified by the user. Docking scenarios are defined by the choice of the external docking program, the receptor structure, and the docking parameters (which include the predefined docking surface on the receptor, residues on the receptor that are allowed to be flexible during docking, and the rigor of the docking routine). VirtualFlow currently supports the following docking programs: AutoDock Vina[17], QuickVina 2[18], Smina (which includes the Vinardo and AutoDock 4 scoring functions)[19], AutoDockFR (ADFR)[20], QuickVina-W[5], VinaXB[21], and Vina-Carb[22]. By supporting an array of different docking programs, VFVS can be used in a variety of cases by leveraging the unique advantages of each program. VFVS allows the specification of multiple docking scenarios to be carried out for each ligand, enabling consensus docking procedures, as well as ensemble docking procedures[23,24]. VirtualFlow is

also amenable to the integration of other docking programs that are not currently a part of this platform.

Scaling behaviour of VFVS

In order to measure the scaling behaviour of VFVS, we measured the performance on two local clusters, LC1 and LC2. On LC1, we used 18,000 CPU cores of heterogeneous composition (different models of Intel Xeon and AMD Opteron processors), whereas on LC2 we employed up to 30,000 Intel Xeon 8268 CPUs. The scaling behaviour was effectively linear in both cases (*i.e.*, $\mathcal{O}(N)$, where N is the number of cores), see Extended Data Fig. 3a. These results meet theoretical expectations since there is no direct communication between the processes running in parallel, which is key to perfect scaling behaviour without bounds. The independence of its parallel processes means that VirtualFlow is expected to scale linearly even if millions of cores are used. We also tested the performance of the platform on cloud-based computing systems including GCP and AWS. On the GCP we carried out large-scale benchmarks with up to 160,000 CPUs, and despite this massive scaling in CPU volume, VirtualFlow still exhibited linear scaling behaviour (Extended Data Fig. 3a). A typical high-throughput screen, such as the one described in this study, of 1 billion compounds will take ~15 hours on the GCP with 160,000 CPUs, making VirtualFlow suitable for the highly anticipated exascale computing age.

Multistaging using VFVS

VFVS can also be used to organize virtual screens with multiple stages to substantially increase the quality of the results (Fig. 2a). In the multi-staging approach, several virtual screens are executed in succession. The number of top-scoring compounds that advance from one stage to the next is successively reduced, with concomitant increases in docking accuracy and computational cost.

Using VFVS to screen 1.3 billion ligands

In order to validate the performance of VFVS we screened a virtual library of 1.3 billion commercially available compounds (~330 million compounds from the ZINC 15 database[16], and ~1 billion compounds from the Enamine REAL library) against KEAP1. This effort was completed in around four weeks, using on average approximately 8,000 cores on a heterogeneous Linux cluster.

To illustrate the benefit of an ultra-large-scale screen, we chose a subsets of the ligands (0.1, 1, 10, and 100 million compounds) randomly from the 1 billion compound screen of the REAL library and considered the scores of the top 50 compounds.(Fig. 2b). As the scale of the screen increased, the average docking score increased thus improving the chances of identifying tighter binders. This in turn leads to higher true hit rates and tighter experimental binding affinities, as predicted by a probabilistic model which we derived (Supplementary Section D). and which was experimentally demonstrated previously[2].

To demonstrate VirtualFlow in a multi-staging context we subjected the top ~3 million ranking compounds from the primary virtual screen to a rescoring procedure (Fig. 2a). In stage-2, the 13 residues of KEAP1 at the NFR2 interaction interface were allowed to be flexible. This flexibility accounted for the movement/dynamics of the amino acids at the binding interface, not captured by a static structure. In the re-scoring procedure we utilized two different docking programs (Smina Vinardo and AutoDock Vina), and two replicas of each docking scenario were carried out to further increase the conformational space sampled during the docking runs. The necessity of multi-stage screening depends on the target of choice and the computational resources available, but this type of virtual screen is particularly useful in cases where dynamics at the docking interface is expected to play a significant role.

Experimental validation

From the *in silico* screen described above, we chose 590 hits for experimental validation. Of these, 492 compounds were from the top 0.03% of stage-2 screen and 98 compounds were from the top 0.0001% of stage-1. Hits from stage-1 were ordered to compare the true hit rate between stage-1 and stage-2 hits, in a multistage setting. In addition to the ranking by docking score, the choice of these compounds was based on factors like drug-likeness, availability for procurement, ligand efficiency and chemical diversity. We used four established biophysical methods: fluorescence polarization (FP), surface plasmon resonance (SPR), nuclear magnetic resonance (NMR), and bio-layer interferometry (BLI) to experimentally validate the binding of the VirtualFlow-derived hits to KEAP1. FP and SPR were initially used in a high-throughput fashion (Level-1) to detect binding and the compounds identified here were subsequently validated with more scrutiny in a detailed and low-throughput fashion (Level-2). We used a recombinantly expressed and purified Kelch domain of mouse KEAP1, henceforth referred to as KEAP1. An overview of the experimental verification workflow is graphically represented in Extended Data Fig. 6, and a detailed description of the experimental procedure is provided in the methods section. Of these four biophysical methods, FP and BLI detect the ability of the hits to displace the NRF2 peptide from KEAP1, identifying hits we refer to as displacers. SPR and NMR directly detect binding of hits to KEAP1, identifying hits referred to as binders. VirtualFlow identifies molecules that potentially bind to the NRF2-interacting interface on KEAP1, but the *in silico* screen is performed using KEAP1 alone, in the absence of NRF2. The NRF2 binding surface on KEAP1 is part of the deep pocket/tunnel of the KEAP1 β-barrel with NRF2 binding to the entrance of this tunnel. However, some compounds could bind more tightly by inserting deep into this central tunnel of KEAP1 rather than embracing the surface like the NRF2 peptide, and/or bind to parts of KEAP1 not engaged by NRF2. Such binders might not effectively disrupt the interaction with NRF2, while still engaging KEAP1 with high affinity (Extended Data Fig. 9). In our experimental validation we identified both displacers and binders.

Out of the cherry-picked 590 compounds, 69 were confirmed to bind to KEAP1 by Level-2 SPR. To assess the ability of the compounds to displace the NRF2 peptide we used the FP assay. Ten compounds were confirmed to be displacers with an $IC_{50} < 60$ μM by FP and all of them were identified as a binder by Level-2 SPR. Interference by autofluorescence

from the compounds themselves prevented the analysis of some of the compounds by FP. Thus, we used BLI as an orthogonal assay to assess the ability of the compounds to displace NRF2. The binding affinity of the NRF2 peptide to KEAP1 as measured by BLI was 1.86 nM which is similar to that measured by FP, 3.67 nM (Extended Data Fig. 4). 40 compounds of the 69 SPR Level-2 active compounds were able to disrupt the NRF2-KEAP1 interaction as observed by BLI. Of these 40 compounds, 16 were able to displace NRF2 from KEAP1 at a compound concentration of 20 μM, while all 40 compounds could do so at 100 μM. Using BLI, we were able to identify displacers that were missed by FP due to autofluorescence (an example is shown in Extended Data Fig. 8). We tested all the SPR Level-2 active compounds for potential aggregation by Dynamic Light Scattering (DLS). We identified seven compounds that aggregated in the DLS assay and hence were not considered for further evaluation (Supplementary Table 5). Based on the SPR Level-2 and the FP Level-2 binding data, we selected 23 compounds for SPR Level-3 experiments to determine the binding affinity. All 23 compounds had affinities in the low micromolar to nanomolar range, and 12 compounds had submicromolar K_d values. From these 23 compounds, we tested the binding of six compounds (iKeap1, 2, 7, 8, 9 and 22) to KEAP1 by a suite of NMR-based ligand-detected experiments. Out of these six compounds, five are displacers and one of them (iKeap9) is a binder. These six compounds were selected on the basis of the solubility constraints of the NMR experiments, the SPR K_d value, and/or their ability to displace the peptide. We used differential line broadening (DLB), saturation transfer difference (STD), Car-Purcell-Meiboom-Gill (CPMG)-based transverse relaxation time experiments, and protein-observed ^{1}H-^{13}C heteronuclear multiple-quantum correlation (HMQC) experiments to confirm binding of the compounds to KEAP1. The ligand-detected NMR experiments confirmed that all six of the tested SPR Level-3 active compounds bind to KEAP1 (Fig. 3, Extended Data Figs. 7, 8). Protein-detected ^{1}H-^{13}C HMQC experiments show that the compounds engage KEAP1 in a specific manner, at the targeted NRF2 binding site. In the absence of resonance assignments, we use the fact that the compounds perturb a subset of KEAP1 resonances affected by the addition of the NRF2 peptide as evidence for competitive binding. These compounds are shown in Supplementary Figs. 1, 2 and 3. Details about the other active compounds is provided in Supplementary Information Section B.

Two of our top hits, iKeap1 and iKeap2 are able to displace the NFR2 peptide from KEAP1. Both of the compounds are predicted to engage the NRF2 binding pocket on KEAP1, located at the entrance to the tunnel formed by the β-barrel. (Fig. 3a, b). In comparison to iKeap2, iKeap1 descends deeper into this central tunnel of KEAP1. SPR results showed that iKeap1 and iKeap2 bind to KEAP1 with a binding affinity of 114 nM and 158 nM, respectively (Fig. 3c, d). NMR-based ligand-detected experiments confirmed that both iKeap1 and iKeap2 directly bind to KEAP1 (Fig. 3e, f). FP assays showed that iKeap1 is able to displace NRF2 peptide with an IC_{50} of 258 nM and iKeap2 displaces the NRF2 peptide with an IC_{50} of 2.7 μM (Fig. 3g, h). BLI measurements additionally confirmed that both iKeap1 and iKeap2 are able to displace the NRF2 peptide from KEAP1. iKeap1 exhibits similarity to a previously reported naphthalene-based compound with a lower IC_{50} (IC_{50} = 2.7 μM; compound C17 in Supplementary Table 1 and Extended Data Fig. 5d)[25]. C17 was identified as the best hit in a high-throughput screen (HTS) of 270,000 compounds[25].

We would also like to highlight iKeap7, which has the highest affinity as assayed by SPR (K_d = 15 nM) and displaces the NRF2 peptide with an IC_{50} of 38.2 μM (Extended Data Fig. 8). It should be noted that of the 14 hits described in the manuscript, only two hits, namely iKeap2 and iKeap7, contain pan-assay interference compound (PAINS) substructures. However, we performed a series of orthogonal binding assays, which confirmed that iKeap2 and iKeap7 are not experimental false positives. For details and discussion on how we verified that our experimental results were not affected by PAINS see Supplementary Section B.

Typically PPIs have a larger interaction interface as compared to that of the active site of an enzyme. Hence the *in silico* screen can identify binders that either partially overlap with the binding site of the interacting protein, such as iKeap9 (see Extended Data Fig. 7), or those that bind in a manner which energetically favours the formation of the protein-protein complex. Examples of the latter, referred to as glues, have been previously described in the literature[26].

An open source platform

To allow VirtualFlow to be used widely and develop dynamically, it is set up as a free and open source (FOSS) project. GPU support is planned for the future and will be incorporated into VirtualFlow both natively and via external docking programs such as Gnina[27]. We encourage scientists to join the project and contribute to improving existing features, adding new features and functionality. The primary homepage of VirtualFlow, which provides additional resources, can be accessed at https://www.virtual-flow.org.

Outlook

VFVS can be used to search extremely large regions of the chemical space, which is the key to identifying promising small-molecule binders. VFVS is able to accomplish this by efficiently utilizing high-performance computing resources, which will continue to increase in availability and power in the years to come, and novel virtual screening databases such as the Chemical Universe Databases (GDBs), which contain billions to trillions of compounds, are still waiting to be explored[28].

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements and competing interests; and statements of data and code availability are available at https://doi.org.

Methods

Parallelization of the virtual screen with VirtualFlow

VirtualFlow employs four levels of parallelization in a hierarchical manner to permit it to run on batch system-managed Linux clusters of any configuration while allowing for perfect scaling behaviour. Each instance of VirtualFlow can submit multiple jobs, each job may use several job steps (currently only supported when using SLURM and Moab/TORQUE/PBS

as the resource manager, while for SGE and LSF only single job steps per job are possible), one job step is able to execute an arbitrary number of queues, and each queue executes the external programs which are processing the ligands (Extended Data Fig. 1). These programs may be additionally parallelized internally, for instance via multithreading. Details about the workflow within a single queue is provided in Supplementary Section G.

Workload balancing

When processing ligands in parallel, there needs to be a mechanism which makes sure that each ligand is treated only once. However, one main problem with parallelization is that most cluster file systems are too slow to work off a single simple task list. This is because when different processes access the file at the same time, clashes can occur, as it may take up to several seconds until one job sees the changes made to a file by another job. These latency problems also mean that file locking mechanisms do not prevent these clashes. The standard solution for solving this kind of problem is to let different processes communicate directly with each other or via a central master process. However, in most cases, this results in sub-linear scaling behaviour, which normally worsens as more and more parallel running processes become involved. Moreover, many advanced parallelization methods such as MPI or OpenMP, do not allow for inter-job communication, while in many cases multiple simultaneously running jobs are needed. Therefore, in order to maintain perfect scaling behaviour which allows multiple jobs and a virtually unrestricted number of CPUs, we have developed an advanced task-list mechanism. The key is to minimize the number of instances that the parallel processes need to access the task list. The mechanism we have implemented requires only a single access per batch system job, each of which can contain a large number of parallel running processes. For this purpose we have implemented a workload balancer, which distributes the tasks from the central task list at the beginning of each job to all the queues belonging to it. The central task list contains collections of ligands as elementary components (rather than individual ligands), and the workload balancer takes into account the length of each collection when distributing them among the queues. This approach dramatically reduces the number of times the central task list has to be accessed. For example, if the workflow employs 10 jobs in parallel, and each job runs on 100 nodes with a wall time (real run time) of one week and 24 CPU cores per node, and one ligand requires approximately 30 seconds to be docked, then the central task list needs to be accessed only 10 times per week to feed a total of 24,000 parallel running queues (assuming each queue runs on one CPU core). In this case, approximately 483,840,000 ligands are processed in one week, which means that the advanced task list approach reduced the number of accesses to the central task list by a factor of 48,384,000 in comparison to the number needed by a trivial task list approach (one access per ligand processed). This factor can be improved even further depending on the job size and the cluster wall time. In case two parallel processes want to access the central task list simultaneously, two backup mechanisms were implemented. The first mechanism is a time-dispersion mechanism, which spreads out simultaneously arriving jobs in time, and further stalls subsequent jobs until the workload balancer of the current job is finished. If this mechanism should fail to prevent a simultaneous access event, which could result in a damaged or empty task file, a second mechanism restores the task list using an automatically backed-up copy of a previous

version of the central task list. More information about the input and output file structures of VirtualFlow is provided in the Supplementary Information Section F.

Reduction of I/O load

One of the potential bottlenecks of computer clusters is the I/O load they can handle, even when they utilize shared cluster file systems with high bandwidth. The limit of the I/O capacities of a cluster can be easily reached if many small processes that individually handle their I/O and use the shared file system are running in parallel. This circumstance can pose a serious problem when running large-scale workflows with thousands of queues working in parallel, and can easily lead to crashing the cluster file system. To address this problem and dramatically minimize the load on the shared file system, VirtualFlow is able to perform most I/O operations on the local temporary file systems of the computing nodes, which are normally fast RAM-based (virtual) drives readily available on any Linux system (usually /dev/shm). The final output files are then stored in batches at large time intervals on the permanent cluster file system.

Preparation of the ligand databases

One of the ligand databases which was screened originates from the state of the ZINC 15 database in the autumn of 2016. Approximately 330 million compounds were downloaded in the SMILES format and converted into three-dimensional PDBQT files with VFLP because, at the time, the ZINC 15 database only provided a fraction of the compounds in a ready-to-dock format. During the conversion, the molecules were protonated with ChemAxon's cxcalc and the three-dimensional structure of the ligand was computed by ChemAxon's molconvert tool[29]. If protonation or the generation of the three-dimensional structure failed, Open Babel[30] was employed as a fallback option. Other preparation steps, such as desalting, were not carried out on these compounds as they had already undergone these basic preparation steps for the ZINC 15 database. We also prepared the compounds in the REAL database provided by Enamine[31]. Approximately 700 million partially-stereospecific SMILES were expanded into fully stereospecific SMILES, resulting in around 1.4 billion molecules. These were then prepared with VFLP into a ready-to-dock format. Specifically, the compounds were desalted and neutralized with ChemAxon's cxcalc, major tautomers were computed with cxcalc, and then protonated with cxcalc (using Open Babel as fallback), the 3D coordinates were computed with ChemAxon's molconvert[29] (using Open Babel as a fallback), and finally converted into the PDBQT format with Open Babel. This library has been made available via an interactive web interface (Supplementary Section C). The scaling behaviour of VFLP was measured on the GCP up to 20,000 CPU cores and this data is shown in Supplementary Fig. 7.

Computation time of VFVS

The total computation time (T) is directly proportional to the number of ligands screened (N) and the processing time per ligand (P), and inversely proportional to the number of CPUs used (C):

$$T \approx \frac{PN}{C}, \quad (1)$$

The processing time per ligand (P) depends mainly on the specific docking scenario (which includes the receptor and all the possible docking options/parameters) and the speed of CPUs used, and can be approximated by the equation

$$P \approx \frac{E\Theta + \zeta}{\eta}, \quad (2)$$

where η is a factor representing the CPU speed relative to a reference CPU, E is the docking exhaustiveness parameter (elaborated in the next paragraph below), Θ is the docking time per unit exhaustiveness on the reference CPU (i.e. the slope of the lines shown in Extended Data Fig. 3c), and ζ is the initial setup time required by the docking program on the reference CPU (i.e. the y-intercept of the lines in Extended Data Fig. 3c). For a typical case of a large-scale first-stage virtual screen on one of the newer Intel CPUs, the average processing time per ligand (P) is roughly 5 seconds using the fastest docking settings. It follows that when 5,000 CPUs are used, the total screening time for 100 million compounds will be roughly 30 hours. Extended Data Fig. 3b illustrates the relationship between the computation time and the number of CPUs for a given number of ligands (assuming an average processing time of 5 seconds per compound).

Relationship between the exhaustiveness parameter and the docking time

The time to dock a single molecule depends on the number of conformations that are sampled, and this number is largely independent of the size of the docking box or surface area. The number of conformations sampled can be controlled by the exhaustiveness parameter of the docking programs. Docking time has a linear dependency on the exhaustiveness parameter, as shown in Extended Data Fig. 3c. The inset in the graph shows the slope for each of the docking programs, providing an estimate of the degree of dependency between the computational time and the exhaustiveness parameter for individual docking programs.

Since most docking programs utilize a probabilistic search algorithm, the results of separate iterations with the same starting set-up can differ. This circumstance can be beneficial as it can be more efficient to carry out multiple less exhaustive docking iterations than to run one highly exhaustive iteration. The exhaustiveness here is a measure of the extent to which the conformational space of the ligand, and potentially the protein side chains, is explored by the search algorithm during the docking procedure. In light of this, VFVS can be configured to carry out multiple replicas per docking scenario, thus improving the overall efficiency.

Lead optimization using VFVS

The operational flexibility enables VFVS to also be used during lead optimization (*cf.* Fig. 1). In this context, a library of analogues of a chosen lead compound can be prepared with VFLP and screened by VFVS with high docking accuracy (e.g. setting the exhaustive parameter to a high value, allowing specific amino acids in the binding interface

to be flexible, using multiple docking programs, and/or multiple receptor (backbone) conformations), which can considerably accelerate the lead optimization process.

Parameters of the virtual screen against the KEAP1 target

For the virtual screening validation test, the crystal structure of the KEAP1 Kelch domain (PDB-ID 5FNQ[9]) was used. The protein was stripped of all small molecules present (including water), was protonated at physiological pH, and then converted into PDBQT format using AutoDockTools[32].

The NRF2 binding interface on KEAP1 was chosen as the target of the screening, and the exact location determined by previously published co-crystal structures of KEAP1 and the NRF2 peptide (such as PDB ID: 4IFL). The *in silico* screen was carried out as follows: VFVS used the docking program QuickVina 2 in an initial (primary) virtual screen with the mouse KEAP1 as a rigid receptor structure. In this primary virtual screening, the docking search space was a rectangular parallelepiped (i.e. a cuboid) of size 15.0 × 16.5 × 14.275 Å. The exhaustiveness parameter was set to 1, which favours fast computational times. The quality of individual docking results, and therefore the ranking, depends largely on the external docking program chosen (which is independent of VirtualFlow).

In the rescoring procedure, the following amino acid side chains at the binding interface were allowed to be flexible: Tyr334, Arg380, Asn382, Arg415, Cys434, His436, Ile461, Phe478, Arg483, Ser508, Tyr525, Tyr572 and Phe577. AutoDockTools was used to generate the rigid and flexible receptor structures in PDBQT format. The exhaustiveness was set to 1, and two replicas (iterations) were carried out of each docking scenario (with Smina Vinardo and AutoDock Vina as the docking programs). The size of the docking box was set to 27.0 × 27.0 × 24.0 Å.

Expression and purification of GST-KEAP1

A codon optimized vector of the mouse KEAP1 Kelch domain (residues 322–624) cloned into a pGEX-6P-3 vector with BamHI and XhoI cloning sites, and an NRF2 peptide (AFFAQLQLDEETGEFL) with an N-terminal tetramethylrhodamine (TAMRA) fluorophore were purchased from GenScript USA Inc. (NJ, USA). The pGEX-6P-3 vector contains an N-terminal glutathione S-transferase (GST) tag which is expressed as a fusion with the target sequence, resulting in a gene product that will henceforth be referred to as GST-KEAP1. The vector carrying GST-KEAP1 was transformed into BL21(DE3) *E. coli*. The transformed cells were grown at 37 °C to an optical density of 0.6 at a measurement wavelength of 600 nm and protein expression was induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). The cells were allowed to grow for 12 – 16 h at 18 °C and subsequently harvested by centrifugation at 4,200 rpm for 20 min at 4°C. To purify GST-KEAP1, cell pellets from 2 L of culture were resuspended in 40 mL of GST binding buffer (25 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1 mM EDTA) supplemented with 3.5 mM β-mercaptoethanol and protease inhibitors (Roche). Cells were lysed by sonication, the insoluble fraction was removed by centrifugation at 16,000 rpm and the soluble fraction was applied to 10 mL of GST slurry (GoldBio, MO). The suspension was nutated for 4 hours at 4 °C, and the unbound fraction was removed by gravity-flow chromatography. The slurry was washed twice with

GST binding buffer supplemented with 3.5 mM β-mercaptoethanol. The bound fraction was eluted from the slurry with 20 mM reduced glutathione in GST binding buffer. The resulting eluate was loaded on a Superdex 200 size exclusion column (SEC) pre-equilibrated in SEC buffer (20 mM Tris-HCl, pH 8.0, 50 mM NaCl, 10 mM dithiothreitol).

Fluorescence polarization (FP) assays

i. Dissociation constant of the NRF2-KEAP1 interaction: We prepared 2 nM TAMRA-NRF2 peptide in FP buffer (20 mM Tris-HCl pH: 8.0, 50 mM NaCl, 10 mM DTT, 2 mM 3-[(3-Cholamidopropyl)-dimethylammonio]-1-propanesulfonate (CHAPS), 0.005% BSA, 1% DMSO) in Corning 3575 384-well plates, to establish the dissociation constant (K_d) of the TAMRA-NRF2-GST-KEAP1 interaction. GST-KEAP1 was titrated into the TAMRA-NRF2 peptide starting at a concentration of 76 μM GST-KEAP1 followed by two-fold dilutions for a total of 24 points. A K_d of 3.67 ± 0.35 nM was determined for the interaction (Extended Data Fig. 4a).

ii. FP Level-1 (high-throughput screening of compounds): All 590 compounds ordered for testing were dissolved in DMSO-d_6 to a final concentration of 10 mM. Two AB1056 (Abgene, NH, USA) plates were prepared as source plates for screening. The first source plate contained 11 μL of each of the 10 mM compounds. The second source plate was filled with 9 μL DMSO and 1 μL from the first source plate was transferred into the second via pin transfer with a Vprep liquid handling pipetting station (Agilent, CA), resulting in a final concentration of 1 mM for each compound in the second source plate. 384-well Corning 3575 (Corning, NY, USA) assay plates were pre-loaded with 7 nM GST-KEAP1 in FP buffer (30 μL/well). 300 nL, 100 nL, and 33 nL volumes were transferred from each source plate (the 10 mM and 1 mM plates) to pre-loaded 384-well assay plates. The assay plates were incubated for 1 hour at room temperature before 2 nM TAMRA-NRF2 peptide was added to each well with an HP D300 (Hewlett-Packard, CA). After 3 hours of incubation at room temperature, fluorescence polarization (excitation 485 nm / emission 520 nm) was measured using an EnVision plate reader (PerkinElmer, MA, USA). This assay resulted in six-point titrations, which are not sufficient to calculate accurate IC_{50} values, but allow for the selection of top binders.

iii. FP Level-2 (screening of top hits): The 27 compounds which were active in the FP Level-1 assay were subjected to a second 24-point FP screen (Level-2), starting from 500 μM compound followed by 1.5-fold serial dilution. For the best compound, iKeap1, the starting concentration was lowered to 30 μM and the following concentrations were used in the titration: 30.00 μM, 21.60 μM, 15.50 μM, 11.10 μM, 8.00 μM, 5.76 μM, 4.14 μM, 2.98 μM, 2.15 μM, 1.54 μM, 1.11 μM, 0.80 μM, 0.576 μM, 0.414 μM, 0.298 μM, 0.215 μM, 0.154 μM, 0.111 μM, 0.080 μM, 0.0606 μM, 0.0459 μM, 0.0348 μM, 0.0264 μM, 0.02 μM. The measurements were carried out in triplicate, and the three data points for each concentration averaged. IC_{50} values were determined by fitting the averaged data points to a four parameter logistic curve using the non-linear least squares

method provided by the SciPy library for Python[33]. The standard error (see Supplementary Table 4) on the IC_{50} was computed by taking the square root of the diagonal of the parameter covariant matrix.

Bio-layer interferometry (BLI) assays

i. Bio-layer interferometry binding and displacement assays: NRF2-KEAP1 binding BLI experiments were performed on an Octet RED384 (ForteBio, Menlo Park, CA, USA) using streptavidin-coated Dip and Read Biosensors (ForteBio) and 384 well plates with 120 µL volume. The sensors were incubated for 5 minutes in 500 nM biotinylated NRF2 peptide in binding buffer (10 mM HEPES, pH 7.5, 50 mM NaCl, 0.1%(v/v) Tween20 with 0.5 mM TCEP and 1% DMSO). To test for nonspecific binding of GST-KEAP1 protein, reference tips were incubated in buffer alone. The tips were washed with buffer for 2 minutes to obtain a baseline reading and then transferred to wells in various concentrations of GST-KEAP1 protein (100 nM, 50 nM, 25 nM, 12.5 nM, 6.75 nM, 3.375 nM, 1.679 nM, 0.844 nM) for 10 minutes. After measuring association, tips were moved to wells containing buffer, and dissociation was measured for 5 minutes. The data were processed and analysed using the Octet data analysis software version 11.0 (ForteBio, Inc., Menlo Park, CA, USA). The association-dissociation curve for each concentration was fitted using a 1:1 model given by the equations

$$R_t^{\text{on}} = \frac{k_{\text{on}} * C}{k_{\text{on}} * C + k_{\text{on}}} R_{\text{max}} \left(1 - e^{-(k_{\text{on}} * C + k_{\text{off}}) * t}\right), \tag{3}$$

$$R_t^{\text{off}} = R_{\text{eq}} * e^{-k_{\text{off}} * t}, \tag{4}$$

where R_t^{on} and R_t^{off} are the BLI signals at time t, R_{eq} is the equilibrium response, k_{on} is the association rate constant, k_{off} is the dissociation rate constant, C is the analyte (protein) concentration, and R_{eq} is the signal level at the equilibrium of association which depends on the analyte (protein) concentration and the maximal capacity (R_{max}) of the sensor surface. By computing the ratio $k_{\text{off}}/k_{\text{on}}$ the apparent equilibrium constant K_{d} is obtained. The resulting apparent K_{d} values were averaged.

ii. Compound screening by BLI displacement assay: The BLI displacement assays were setup as described above. The biotinylated NRF2 peptide was used at a concentration of 500 nM and GST-KEAP1 protein was used at a concentration of 25 nM. The compounds were used at concentrations of 20 and 100 µM, and pre-incubated with GST-KEAP1 protein. The association phase was measured in the well containing compound with GST-KEAP1 protein for 10 minutes, and followed by a dissociation phase in buffer for 5 minutes. The inhibition percentage was the average BLI signal in the last 50 seconds of the dissociation phase, normalized against the condition of GST-KEAP1 protein in the absence of compound. The dose-dependent experiment with iKeap22 was carried out

at 10 μM, 20 μM, 40 μM, 80 μM and 100 μM compound concentration and pre-incubated with 25 nM GST-KEAP1 protein.

iii. To test for nonspecific binding of the compounds, the sensor was coupled with biotinylated NRF2 peptide and the compounds were used at 20 μM concentration without protein.

Surface Plasmon Resonance (SPR) binding assays

All SPR binding experiments were performed on a BiacoreT200 (GE Healthcare, Sweden) instrument at 25°C in running buffer (10 mM HEPES pH 7.5, 50 mM NaCl, 0.1% (v/v) Tween20 with or without 0.5 mM TCEP, 1% DMSO). The running buffer was prepared freshly on each day of use, filtered and degassed prior to the SPR experiments. The target protein (GST-KEAP1) was anchored on a CM5 chip via a GST labelling kit (GE Healthcare, Sweden)[34], where a polyclonal goat anti-GST antibody was immobilized on a CM5 sensor chip by the amine coupling method 1.

i. SPR Level-1 (1-point HTS): The SPR Level-1 screening was carried out as previously reported[35]. First, we prepared 10 mM d6-DMSO stock solutions of the 590 compounds which were procured in powder form. 20 μM samples of the compounds were made by diluting the stock compounds in running buffer with 0.5 mM TCEP and 1% DMSO. The anti-GST immobilizing chip was saturated with GST at the reference channel and GST-KEAP1 at the target channel with resonance unit (RU) values of 750–800 for the GST and 2,000–3,000 for GST-KEAP1. Binding of compounds to the immobilized protein was monitored for 60 seconds in both the association and dissociation phase. Additional injection of the running buffer was performed after every compound binding. All binding signals (RU_{max} = 16–29 RU, 1:1 stoichiometry) were corrected for the signals from the reference channel and buffer blank. Compounds were classified as an SPR Level-1 hit if the condition RU > 4 was satisfied. This criterion was based on the positive control (iKeap1, RU = 4.65 ± 0.74).

ii. SPR Level-2 (5-point HTS of the SPR Level-1 hits): The hits from the SPR Level-1 assay were re-screened at five different compound concentrations (0.5, 1, 5, 10 and 20 μM), in running buffer with 0.5 mM TCEP and 1% DMSO, at a rate of 30 μL/min. The hits were classified as hits if they produced a concentration dependent SPR response and an RU value > 4 at a compound concentration of 20 μM.

iii. SPR Level-3 (SPR experiments of selected SPR Level-2 hits): 23 out of the 69 SPR Level-2 hits were chosen for Level-3 analysis. Given the low-throughput of the Level-3 SPR assay, we chose a subset of the SPR Level-2 hits, which included the displacers from the Level-3 FP assay, the compounds that were tested by NMR, and select SPR Level-2 hits. SPR experiments were carried out in which the target protein (GST-KEAP1) was captured and regenerated in each compound cycle. All SPR data processing and analyses were performed using the BiaEvaluation software (version 3.0). For steady-state binding, the R_{eq} signal was plotted against the analyte concentration and fitted to the one-site

or the biphasic binding model (see Supplementary Table 3) via the Levenberg-Marquardt algorithm used by the BiaEvaluation software. The one-site binding model is given by the equation

$$R_{\mathrm{eq}} = \frac{R_{\mathrm{max}} * C}{K_d + C} + b, \tag{5}$$

where R_{eq} is the SPR signal at equilibrium, R_{max} is the SPR signal at saturation of the binding mode, K_{d} is the dissociation constant of the compound, b is the offset, and C is the concentration of the compound. The biphasic binding model is given by the equation

$$R_{\mathrm{eq}} = \frac{R_{\mathrm{max}_1} * C}{K_{\mathrm{d}_1} + C} + \frac{R_{\mathrm{max}_2} * C}{K_{\mathrm{d}_2} + C} + b, \tag{6}$$

where R_{eq} the SPR signal at equilibrium, R_{max_1} and R_{max_2} are the SPR signals at saturation of the two binding modes, K_{d_1} and K_{d_2} are the dissociation constants of the compound corresponding to the two binding modes, b is the offset, and C is the concentration of the compound.

Standard errors of the estimated K_{d} values were computed with the BiaEvaluation software, which computes them via the diagonal elements of the covariance matrix and the residual. The software operates based on the equations found on page 378 in the book “Receptor-Ligand Interactions: A Practical Approach”[36].

Ligand-detected NMR experiments

The differential line broadening (DLB) experiments serve as simple one-dimensional experiments, where the proton signal of the ligand is monitored. The ligand concentration exceeds the receptor concentration (e.g. 10–20-fold) in this experiment and broadening of the resonance frequencies in presence of the receptor is a consequence of ligand molecules shuttling between free and bound states. DLB manifests as a broadening of the ligand resonance due to binding a protein. The ligand is in equilibrium between the free and protein-bound states dictated by the equilibrium constant. DLB is the result of the change in relaxation rate and the difference in chemical shift of the bound ligand. In the STD experiments, a region of the spectral space (−1 to 0.5 ppm) that has resonances from the receptor but not the ligand is selectively saturated. Resonances from methyl bearing amino acids (Ile, Leu, Val) often populate this region of the spectral space. This saturation is transferred to the rest of the protein and eventually to the bound ligand by spin diffusion (NOE). In the implementation of STD, two one-dimensional spectra are recorded in an interleaved fashion. In the first experiment neither the receptor nor the ligand is saturated (off-resonance) and in the second the receptor is selectively saturated (on-resonance). Spectra of free ligand are observed in both experiments. However, if the ligand transiently binds to the receptor then the saturated receptor will transfer magnetization to the ligand. This transfer will be reflected as reduced intensity in the on-resonance saturated spectrum compared to the off-resonance saturation. The results are often presented as a difference

spectrum between the on and off-resonance saturation experiments. The appearance of ligand resonances in the difference spectrum is indicative of ligand binding. Measurement of the transverse relaxation rate of the ligand is another complementary strategy to detect ligand binding to a receptor. The free ligand behaves like a small molecule and experiences slow transverse relaxation, however transient binding to the receptor enhances the transverse relaxation rate of the ligand. Thus, an increased transverse relaxation rate in presence of a receptor directly indicates binding to the receptor. In the experimental setup a series of 1D experiments where the coherences of ligand spends increasing amounts of time in the transverse plane is recorded. Ligands that engage the protein will relax faster than unbound ligand. We refer to these experiments as hear as CPMG-R2 or CPMG experiments. While any of these experiments are in principle sufficient to demonstrate ligand binding, false positives for either of these experiments have been reported. However, a detection of a false positive hit is highly unlikely if all three experiments indicate binding, which is the case with all the hits reported here.

All the ligand-detected experiments were performed with 50 μM compound alone or in presence of 5 μM KEAP1 (without the GST tag) in NMR buffer (phosphate saline buffer supplemented with 5% DMSO-d_6 and 4 mM deuterated DTT at pH 7.4) unless otherwise noted. For iKeap1 and iKeap2, the protein concentration was kept at 2.5 μM due to tight binding. ^{1}H 1D spectra of the compounds were recorded in the absence and in the presence of KEAP1 to assess line broadening. Saturation transfer difference spectra of the compounds in presence of KEAP1 were recorded with 3 second saturation time on (0 ppm) and off (−20 ppm) resonance, respectively. The relaxation rate of the compounds was measured in the absence and the presence of KEAP1 with a series of ^{1}H 1D experiments with CPMG-based transverse relaxation time filters of various lengths: 1 ms, 25 ms, 50 ms, 100 ms, 300 ms, 500 ms and 800 ms. Data were analysed and visualized in Matlab (MathWorks, MA).

Protein-detected NMR experiments

The cleaved mouse KEAP1 Kelch domain (residues 322–624) consists of 308 amino acids with close to 300 detectable amide resonances. Therefore, correlating chemical shift perturbations of small molecule inhibitors to perturbations introduced by NRF2 would have been prohibitively difficult in ^{1}H-^{15}N HSQC spectra without full backbone assignment. Our aim here was to rely on methodology that can be quickly and easily implemented even for very large proteins for which backbone assignment might not be feasible. Indeed, with a molecular weight of 33.7 kDa, the mouse KEAP1 Kelch domain (322–624) is already on the larger side for NMR backbone assignment. To overcome the spectral crowding in a ^{1}H-^{15}N HSQC spectrum and minimize problems due to low ligand solubility, we implemented a ^{1}H-^{13}C TROSY-HMQC experiment coupled with fast data acquisition. For the protein detected ^{1}H-^{13}C HMQC experiments a sample of KEAP1 which is selectively ^{1}H and ^{13}C labelled at the methyl groups of isoleucine, leucine, and valine residues, in an otherwise deuterated background was used. This labelling strategy is referred to as ILV labelling. ILV-labelled samples of KEAP1 were prepared by culturing BL21(DE3) cells containing a plasmid for GST-KEAP1, in perdeuterated M9 medium with 1 g ^{15}N-NH_4Cl and 2 g ^{2}H-^{12}C-glucose in $^{2}H_2O$. One hour before induction with IPTG, 330 mg/L 2-(^{13}C)methyl-4-($^{2}H_3$)-acetolactate (precursor for leucine and valine) was added. Prior to

that the acetolactate was activated as previously described[37]. 20 min before induction 75 mg/L $^{13}C/^{1}H$ methyl - and otherwise deuterated - ketobutyrate sodium (precursor for isoleucine) was added. The use of acetolactate resulted in stereospecific ^{1}H-^{13}C labelling of only one of the leucine($\lambda 2$) and valine($\gamma 2$) methyl groups as previously described[37]. The protein was purified as described above. The GST-tag was cleaved by preScission protease cleavage and the free mouse KEAP1 Kelch domain was eluted in NMR buffer from a size exclusion column. All NMR measurements for the ILV-labelled KEAP1 were performed at a protein concentration of 5 μM. The protein concentration was kept low to account for poor solubility (for NMR) of some of the compounds. The concentrations of the compounds were 50 μM, except for iKeap1 and iKeap2, where the concentrations were 25 μM, due to poor solubility.

Given the low concentration of the protein, we used the methyl SOFAST methyl TROSY with 46 ms and 18 ms acquisition times in the direct and indirect dimensions, respectively[38]. The spectral width was set to 14 ppm (^{1}H) and 20 ppm (^{13}C) in the direct and indirect dimensions, respectively and the spectrum was recorded at 298 K on an 800 MHz Bruker spectrometer equipped with an AVANCE III console and a cryogenically cooled probe. A 4.5 ms Pc9_4_90.1000 pulse was used for selective excitation of the methyl ^{1}H resonances and a 1.2 ms Rsnob.1000 pulse was used to selectively refocus proton chemical shift evolution and ^{1}H-^{13}C J-coupling during ^{13}C chemical shift evolution. Proper choice and calibration of the excitation and refocusing pulses is crucial to avoid perturbing the water signal, which can significantly lower the achievable signal to noise. Fast data acquisition was achieved with a 150 ms recycling delay, which allowed for the recording of experiments with 512 scans in 5 hours.

Detecting aggregation using Dynamic Light Scattering (DLS)

To test the potential aggregation of hits we used DLS experiments. The experiments were performed on a ZS90 Zetasizer instrument (Malvern Panalytical, UK). Measurements were done in triplicate with 10 scans per run (100 s). The compounds were used at 20 μM concentration in running buffer (10 mM HEPES pH 7.5, 50 mM NaCl, 0.1% (v/v) Tween20 with or without 0.5 mM TCEP, 2% DMSO) which was filtered before usage. The 20 μM working solution was made from a 1 mM stock of the compound in DMSO. The data was analysed by the built-in software. Compounds were classified as aggregated when the radius of the measured particles was above the minimum colloidal aggregate size (for small molecules) of 50 nm[39].

In addition, the solubility of iKeap1, our most potent displacer, was analysed with an NMR solubility assay based on a technique described previously[40]. We made individual samples of iKeap1 at various concentrations (in PBS buffer, pH 7.4) ranging from 5μM to 30μM and measured the 1D NMR spectrum of each sample with identical experimental conditions. The resonances of iKeap1 were then integrated and plotted as a function of concentration. The plot shows a linear trend ($R^2 = 0.996$) indicating that iKeap1 does not aggregate in this concentration range (see Supplementary Fig. 4).

Excluding interference from PAINS

PAINS comprise 480 markers initially identified as moieties postulated to cause interference in experimental high-throughput screens[41]. PAINS compounds are often found in the databases commonly used for *in silico* screens, and the user should be cognizant of the fact that a potential hit could harbour a PAINS sub-structure. However, it should also be noted that certain PAINS-like aspects can be mitigated by judicious use of medicinal chemistry, and some aspects of PAINS could have no effect, depending on the target of choice and/or the experimental assays used[42,43]. Attention should be paid to identifying and rigorously characterizing any PAINS compounds amongst the hits identified in an *in silico* screen.

Two of the hit compounds (iKeap2 and iKeap7) reported in this manuscript harbour PAINS substructures. We performed additional experiments to confirm that iKeap2 and iKeap7 are not false positives due to assay interference. Primarily, we used 1) DLS to confirm that all the compounds shown here do not aggregate at the concentrations used in the various experiments (Supplementary Table 5), 2) ligand-detected NMR experiments, STD and CPMG, performed with a 10-fold excess of the compound to show that iKeap2 and iKeap7 bind KEAP1 in a reversible manner (Fig. 3 and Extended Data Fig. 8), and 3) protein-observed ^{1}H-^{13}C HMQC experiments to show that both iKeap2 and iKeap7 engage KEAP1 in a specific manner at the NRF2 binding site and do not aggregate the protein (Supplementary Figs. 1, 3). In the event these compounds caused KEAP1 to aggregate, all the resonances will be broadened, which is not the case here.

Statistics and reproducibility

Characterization of the violin plots in Fig. 2b: Screening size 100K: minimum: −10.3; maximum: −11.6; median: −10.4; Q_1: −10.4, Q_3: −10.6. Screening size 1M: minimum: −10.9; maximum: −12; median: −11; Q_1: −11.1, Q_3: −11.3. Screening size 10M: minimum: −11.675; maximum: −12.3; median: −11.5; Q_1: −11.4, Q_3: −11.5. Screening size 100M: minimum: −11.8; maximum: −12.6; median: −11.9; Q_1: −11.8, Q_3: −12.1. Screening size 1B: minimum: −12.3; maximum: −13.4; median: −12.4; Q_1: −12.3, Q_3: −12.6.

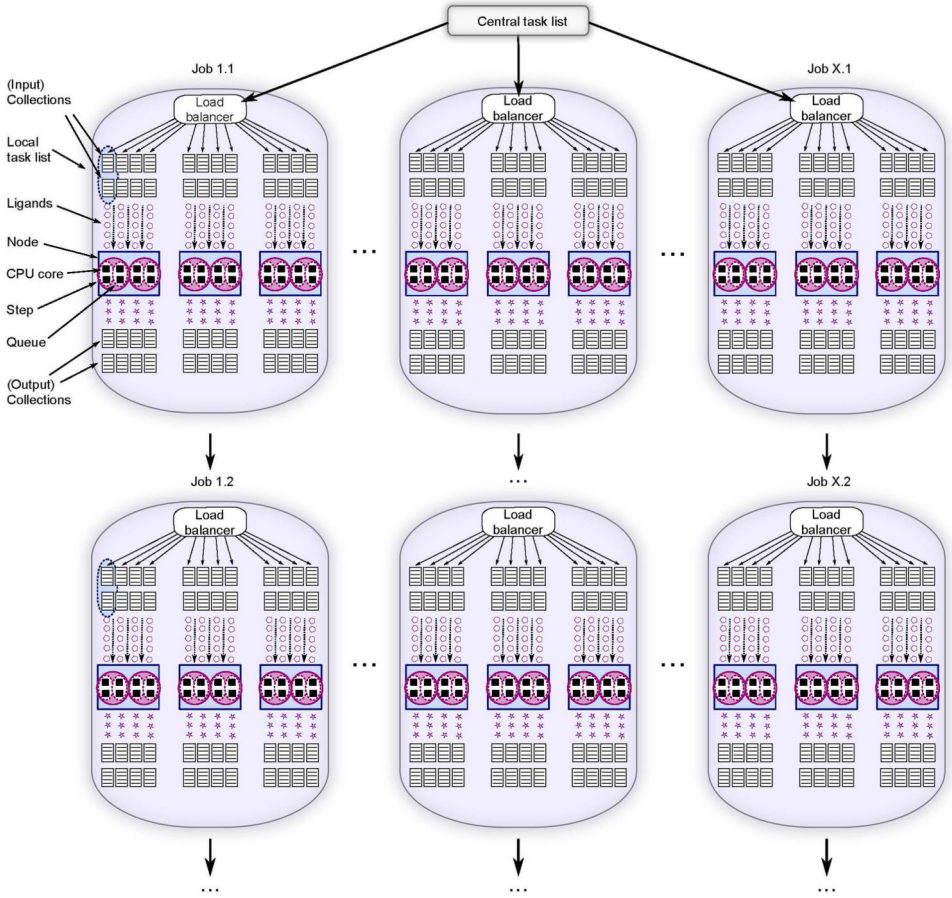
Code availability

VirtualFlow is mainly written in Bash (a Turing complete command language), which not only makes it simple for anyone to modify and extend the code, but also has essentially no computational overhead and is readily available in any major Linux distribution. The code for VirtualFlow is freely available on https://github.com/VirtualFlow, distributed under the GNU GPL open-source licence. The primary homepage for end users where additional resources including documentation, ligand libraries, tutorials and video demonstrations is available at https://www.virtual-flow.org. The external docking programs discussed here are available as follows: AutoDock Vina is available at http://vina.scripps.edu/, QuickVina 2 and QuickVina-W at https://qvina.github.io/, Vina-Carb at http://glycam.org/docs/othertoolsservice/download-docs/publication-materials/vina-carb/, Smina at https://sourceforge.net/projects/smina/, AutoDockFR at http://adfr.scripps.edu/ and VinaXB at https://github.com/ssirimulla/vinaXB.

Data availability

The ready-to-dock library from Enamine is freely available online on the homepage of VirtualFlow at http://virtual-flow.org/real-library. Source data of Figs. 2, 3 and Extended Data Figs. 7, 8 is available online at https://doi.org.

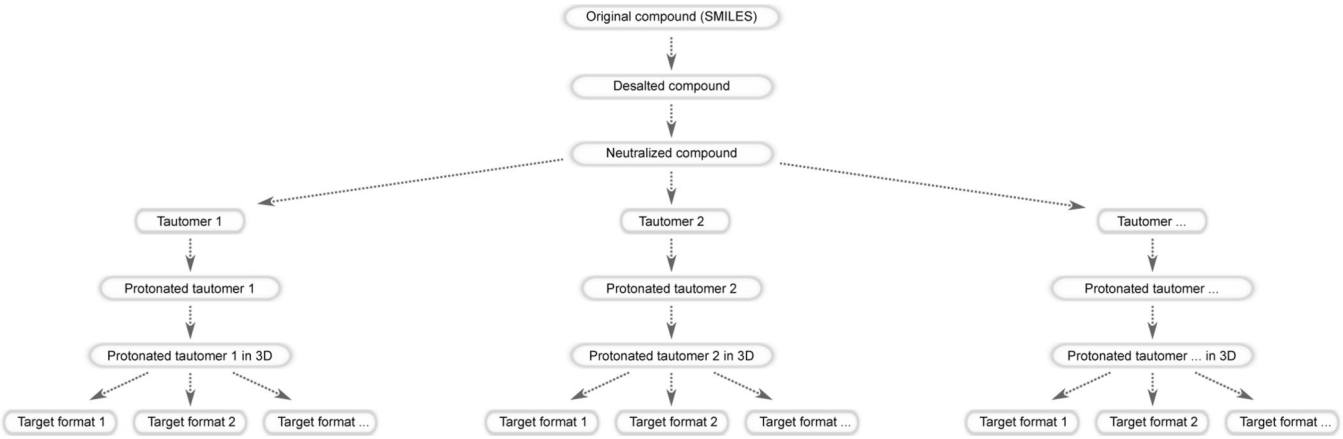
Extended Data



Extended Data Fig. 1: Schematic overview of the organization of the VirtualFlow workflow on computer clusters.

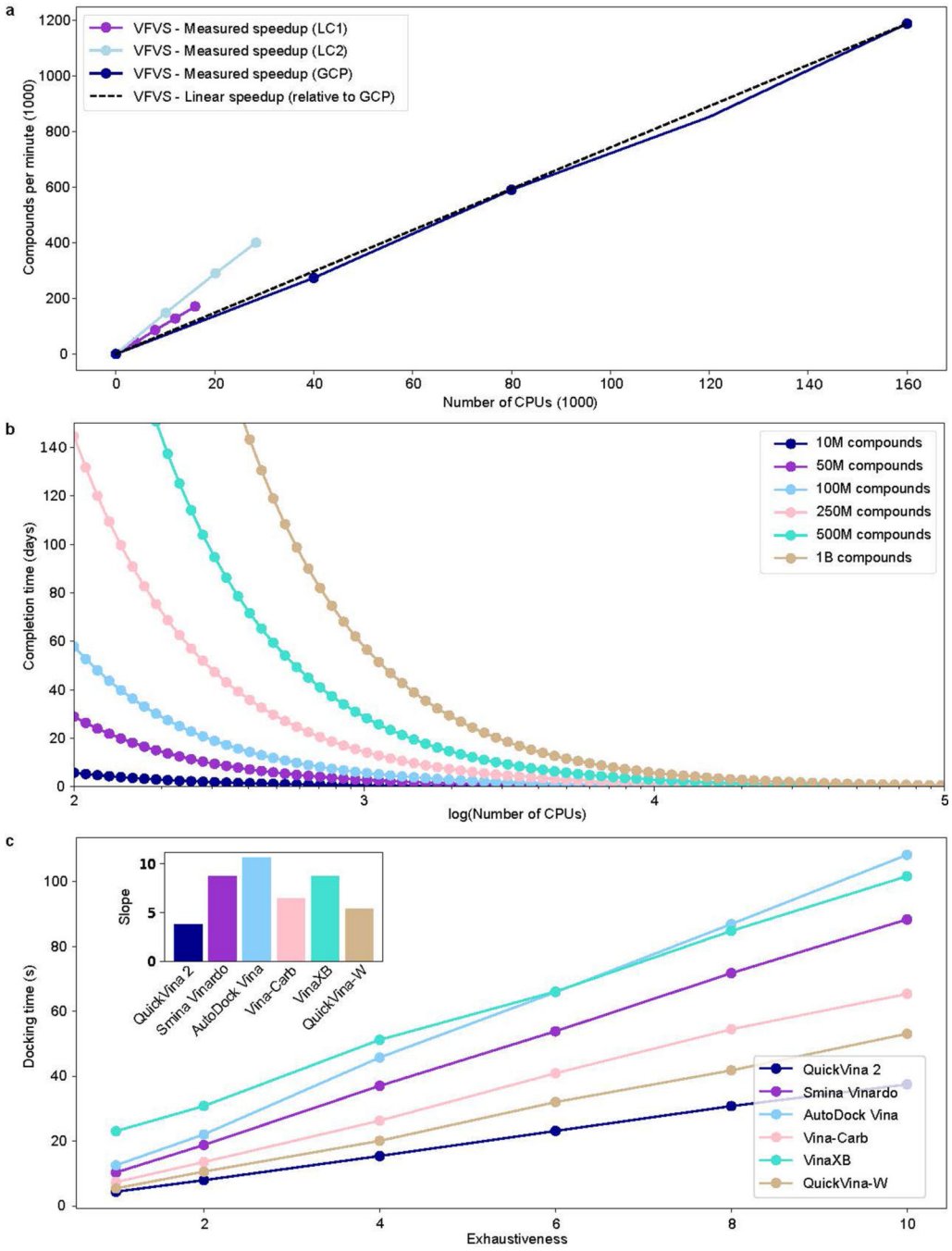
A computer cluster consists of compute nodes, i.e. single computers (blue boxes), which contain a certain number of CPU cores (black squares inside the blue boxes). The resource manager (batch system) of the cluster employs so-called jobs (large violet oval shapes), each of which uses a certain number of CPU cores and nodes. In the example shown above, each job uses three compute nodes, where each node has 8 CPU cores. Each job can employ multiple sub jobs, referred to as job steps (purple circles). With VirtualFlow, each job step employs multiple queues (white oval shapes within the purple circles). Often the workflow is setup such that on each CPU core one queue is running. The hierarchical multi-organization is required to allow VirtualFlow to run on any type of cluster, from the largest supercomputers (which often require that a single job employs multiple nodes) to very small clusters (which often allow a job to use single CPU cores). Each queue processes ligands, which are taken from the input collections in raw form and stored in the output

collection/database. The central task list contains all the ligand collections which should be processed by the workflow, and they are distributed among the queues (into local task lists) by a workload balancer at the beginning of each job. The user can choose any number of batch system jobs (first row with Job 1.1 through Job X.1), which will automatically start successive jobs (second row with Job 1.2 through X.2) after their completion.



Extended Data Fig. 2: Overview of possible processing steps during the ligand preparation with VFLP.

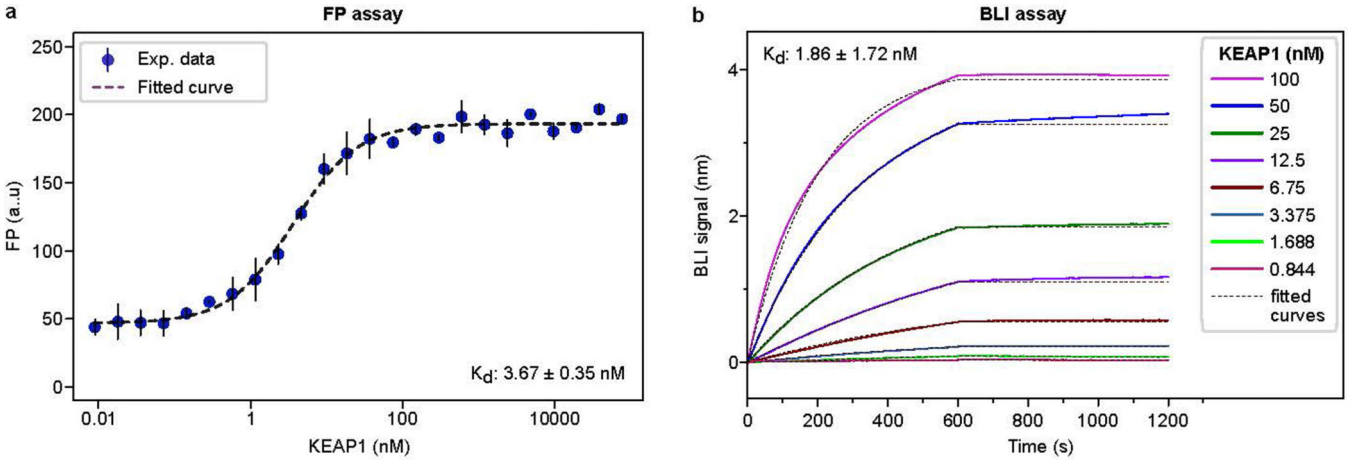
Ligands can be desalted, neutralized, (possibly multiple) tautomeric states can be generated, protonation states for each tautomer computed at specific pH values, 3D coordinates computed, and finally the molecules can be converted into (potentially multiple) desired target formats.



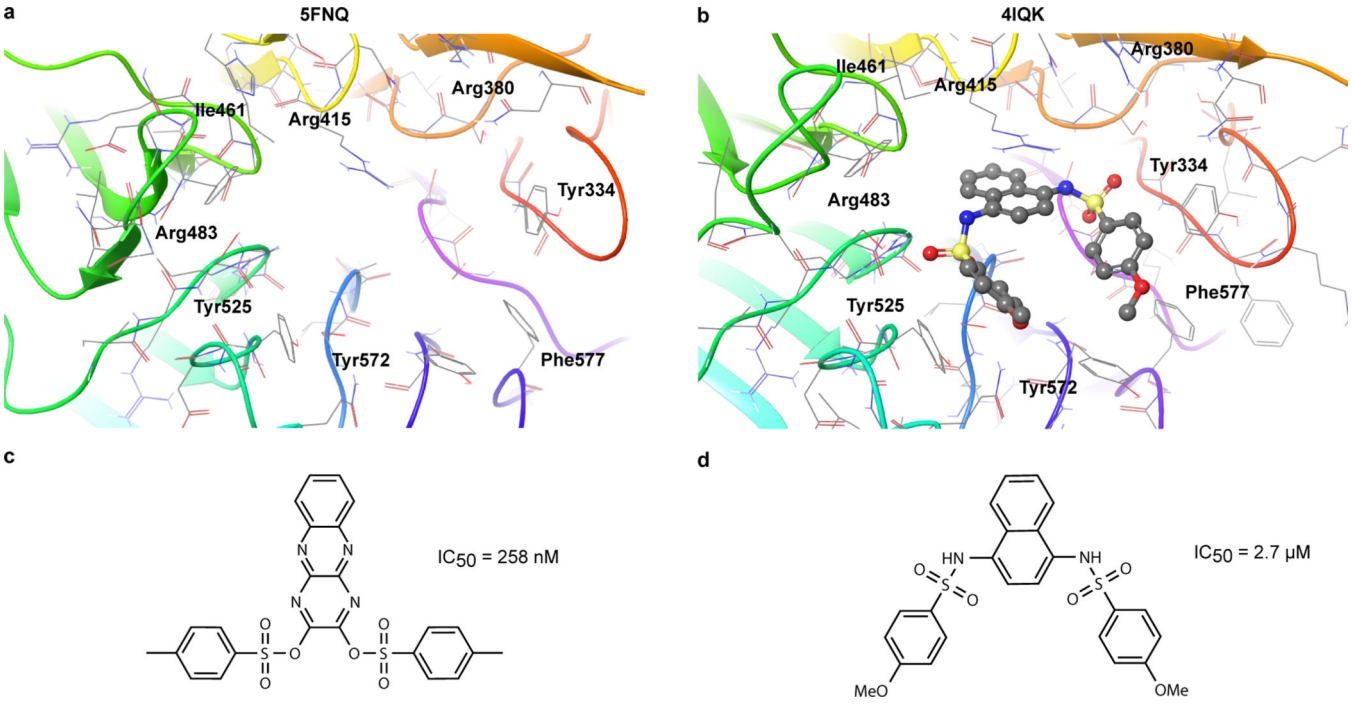
Extended Data Fig. 3: Docking and virtual screening metrics.

a, Scaling behaviour of VFVS using QuickVina 2 as the docking program. Tests with up to 30,000 cores on two local computer clusters (LC1, LC2) and up to 160,000 CPUs on the Google Cloud Platform were carried out. The measured speedup is linear. DOVIS 2.0, an alternative software for virtual screenings on Linux computer clusters using AutoDock, was shown to exhibit near-linear scaling only up to 256 cores as previously reported[4]. **b**,The computational time required (in days) for VFVS to complete virtual screens of different sizes, as a function of the number of CPUs being used in parallel. Each curve corresponds

to a different size input ligand library, and the average computation time per ligand was assumed to be 5 seconds per ligand. **c**, Docking time of an average-sized ligand on a modern Intel CPU (using only a single core) as a function of the exhaustiveness parameter for different docking programs supported by VFVS. The bar plot in the inset shows the slope of the curves, which corresponds to the docking time per exhaustiveness unit. The test ligand which was used for this purpose is given by the SMILES code CN1CCN(S(=O)(=O)N2CCN(C(=O)CCCNC(=O)C3CC3)CC2)CC1. More detailed benchmarks can be found in publications related to these docking programs[17,18,19,20,21,22,5].

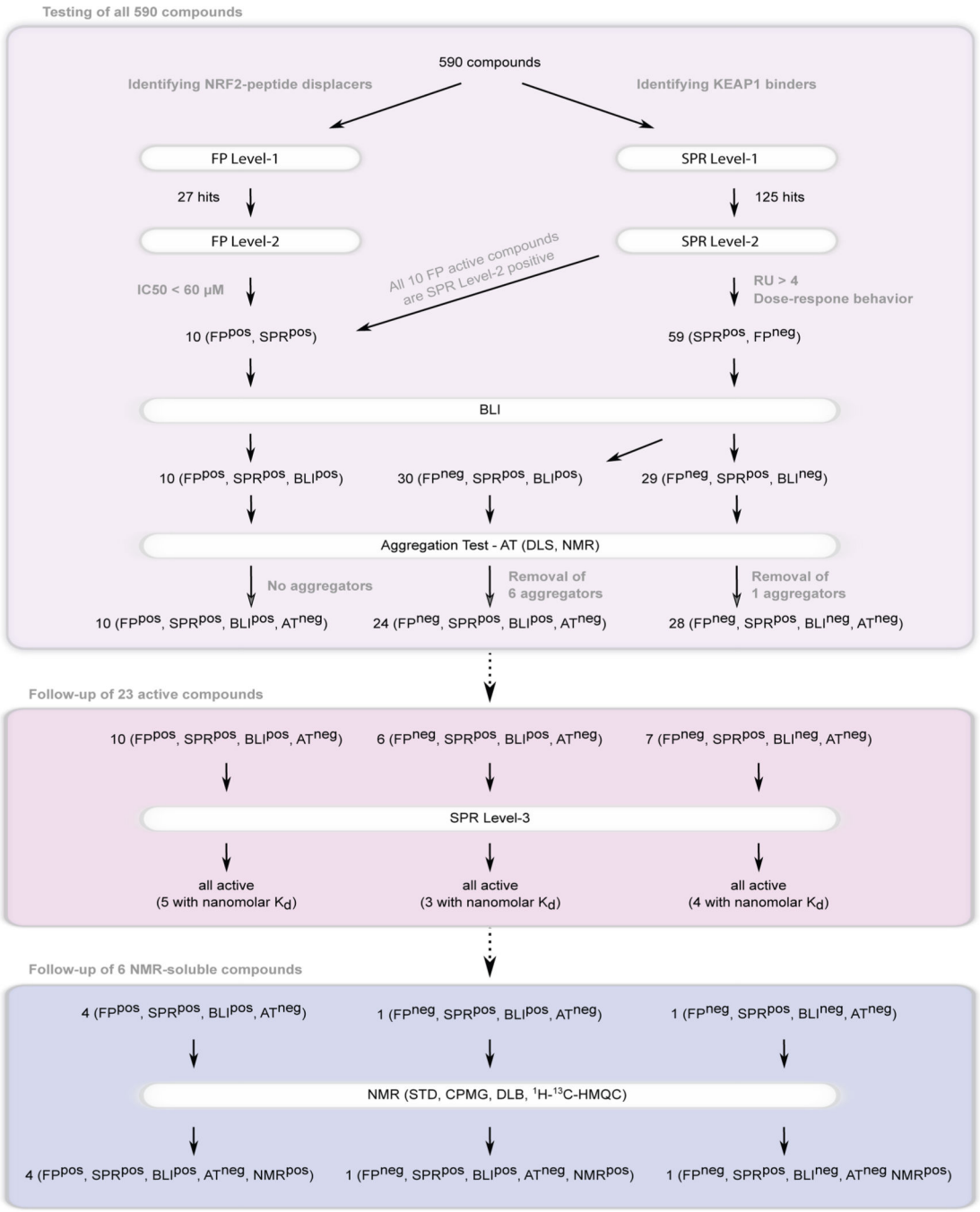


Extended Data Fig. 4: Binding of the NRF2 peptide to KEAP1 as assayed by FP (a) and BLI (b). **a**, For the FP assay a TAMRA-tagged NRF2 peptide and for the BLI assay a Biotin-tagged NRF2 peptide were used. The FP assay was performed with three technical replicates per point. The mean and standard deviation are shown for each titration point, along with the fitted curve. Two independent experiments (n=2) were performed, each with similar results and one representative result is shown here. **b**, For the BLI assay a biotin-tagged NRF2 peptide was used. The BLI experiment was repeated independently twice (n=2) with similar results and one representative result is shown here.



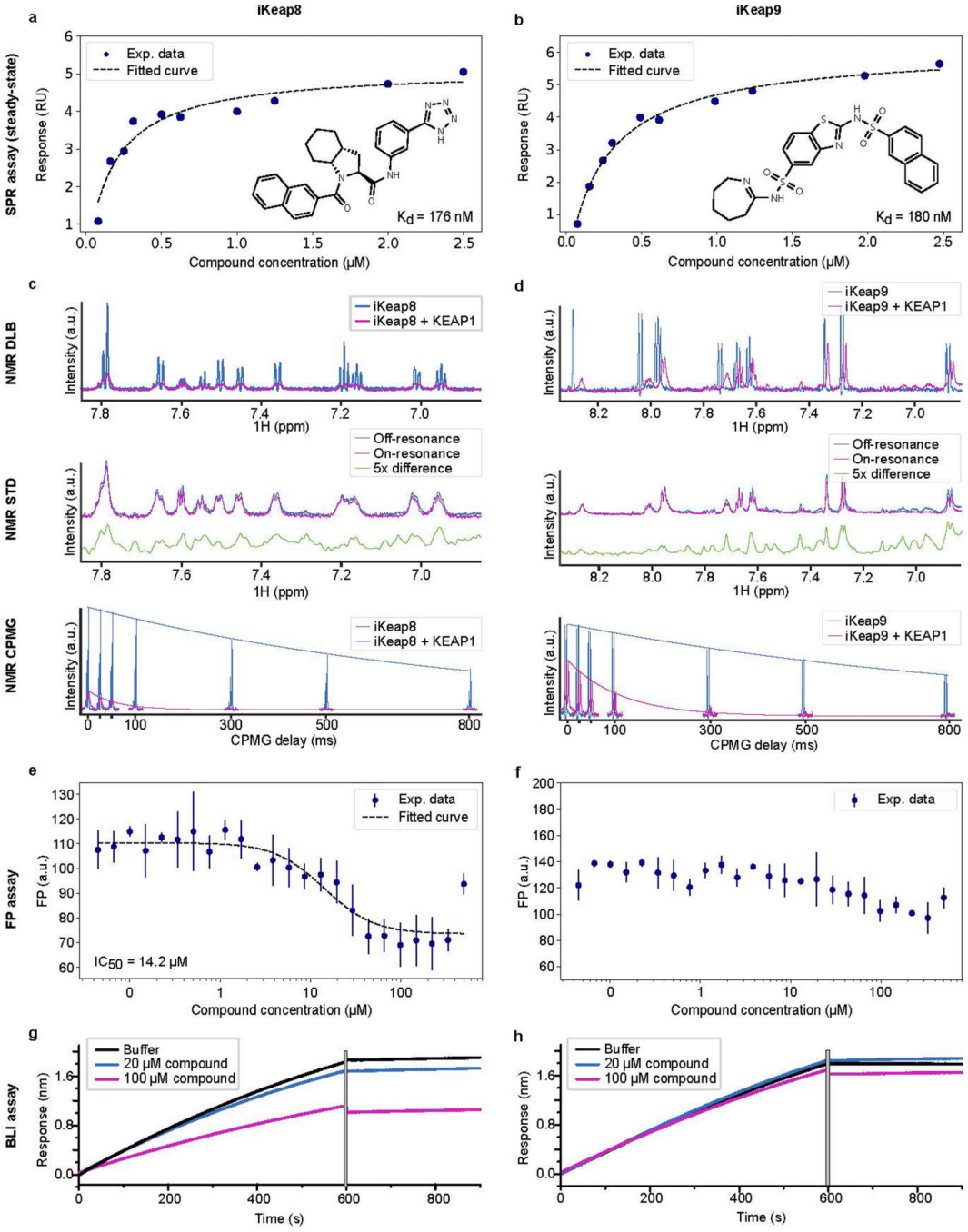
Extended Data Fig. 5: Comparison of iKeap1 with a previously identified displacer C17.

a, Crystal structure (PDB ID: 5FNQ[9]) of KEAP1 with its ligand removed, the structure used for the primary virtual screening procedure. **b**, Crystal structure of KEAP1 (PDB code 4IQK) with ligand C17 (Supplementary Table 1), which is also shown in panel **d. c**, iKeap1, the best binder as accessed by array of experimental validations, is similar to compound C17 previously identified by experimental methods (**d**). Though iKeap1 and C17 look similar they differ in a number of aspects in their core scaffold (therefore analogues of the two compounds cover distinct chemical spaces, assuming the analogues retain the core scaffold of the parent compound). This similarity, as well as the fact that the predicted docking positions (Fig. 3a) of both ligands (**b**) are nearly identical, is an additional evidence that iKeap1 is binding at the predicted site.



Extended Data Fig. 6: Overview of binding assays to determine the activity of the hits identified by VirtualFlow.

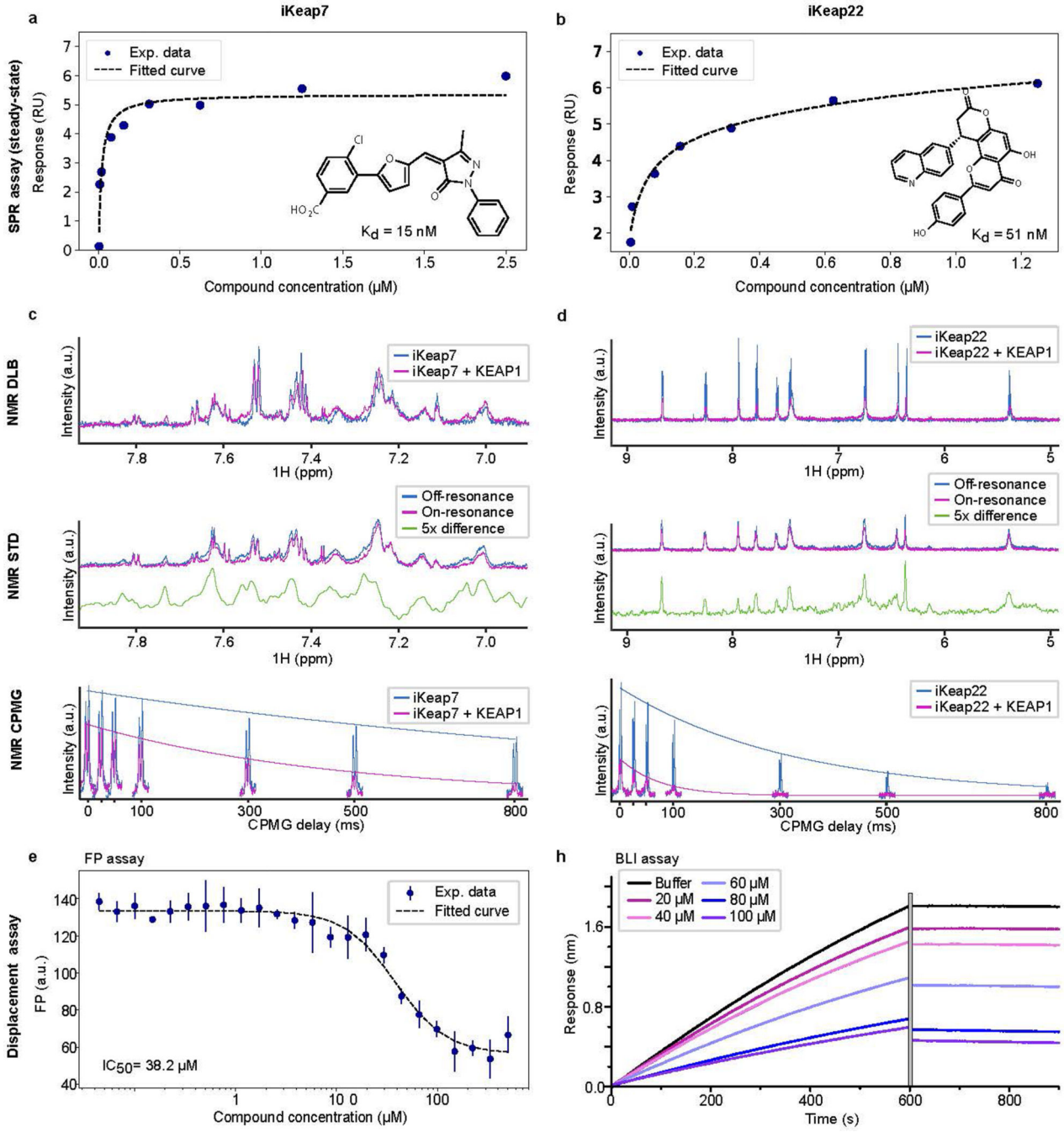
The binding experiments can be broadly classified into two categories i) assays that directly detect binding of the compounds to KEAP1 (SPR, NMR), and ii) assays that detect the displacement of the NRF2-peptide from KEAP1 (FP, BLI). Compounds in SPR Level-2 were classified as active if they exhibited dose-dependent activity (measured over a range of five concentrations) and had an RU value greater than 4 at a compound concentration of 20 µM.



Extended Data Fig. 7: Binder versus displacer.

Here we highlight two new scaffolds, iKeap8 and iKeap9, to illustrate the difference between binders and displacers. SPR confirms that both iKeap8 and iKeap9 bind KEAP1 (**a** and **b**) with similar K_d values. Shown are representative results from the SPR assay for iKeap8 and iKeap9. For each compound, three independent SPR experiments were performed, each with similar results and one representative result is shown here. Ligand-detected NMR experiments shows that both iKeap8 and iKeap9 bind to KEAP1 (**c** and **d**). However, FP (**e** and **f**) and BLI (**g** and **h**) assays show that iKeap8 is able to displace

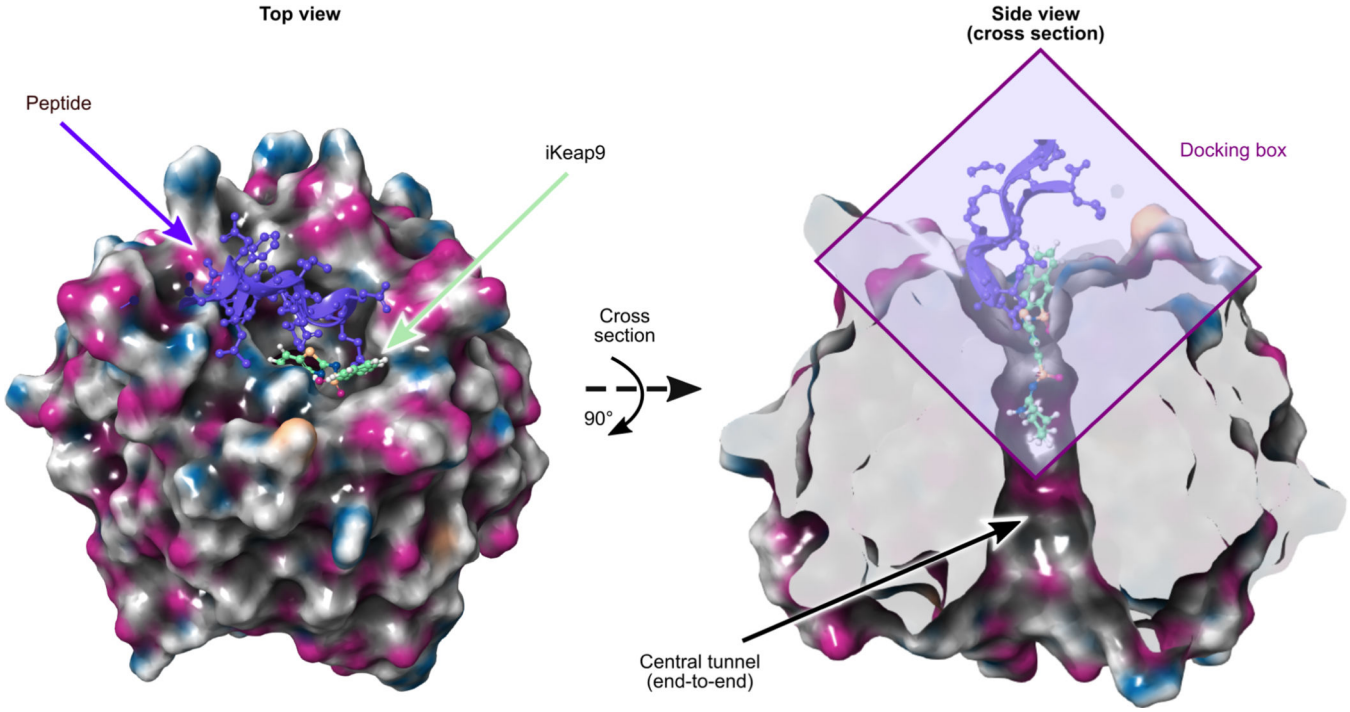
the NRF2 peptide while iKeap9 is not able to effectively displace the NRF2 peptide. The fluorescence polarization (FP) assay was performed with three technical replicates per concentration measured. The mean and standard deviation are shown for each titration point, along with the fitted curve.



Extended Data Fig. 8: Displacers validated by FP versus BLI.

Here we take the opportunity to present two more displacers, iKeap7 and iKeap22, both of which were confirmed as binders by SPR (top panels). Ligand-detected NMR experiments

shows that both iKeap7 and iKeap22 bind to KEAP1 (**c** and **d**). iKeap7 is confirmed to be a displacer of the NRF2 peptide by both FP (bottom left panel) and BLI (not shown). Since the FP experiments on iKeap22 were affected by autofluorescence, BLI (bottom right panel) was needed to confirm that this compounds is a displacer. The FP assay was performed with three technical replicates per concentration measured. The mean and standard deviation are shown for each titration point, along with the fitted curve. Two independent BLI experiment were performed with similar results and one representative result shown here.



Extended Data Fig. 9: NRF2 peptide and ligand binding sites, rationale for binder versus displacer.

Here we show the docking pose of one of the hit compounds (iKeap9, green ball-and-stick representation) bound to KEAP1, together with the NRF2 peptide (PDB ID: 4IFL; peptide in violet). iKeap9 is a tight binder (180 nM by steady-state SPR) but cannot displace NRF2. The left figure shows the top view, while the right figure shows the side view of the cross-section of KEAP1 along the central plane. The violet box in right figure indicates the docking region (where the ligands were allowed to bind) which was used in the virtual screening. The site of interest includes a part of the deep pocket/tunnel of the β-barrel-shaped KEAP1, since it can allow ligands to bind more tightly by insertion into the channel than on a shallow surface. However, the deep tunnel is largely non-overlapping with the peptide binding site (which binds to the entrance site of the tunnel). Thus, binding molecules might only partially interfere with the peptide binding, which might reduce or eliminate the ability of small molecule binders to displace the peptide. The ability of a small molecule to displace the peptide is hard to predict, and was not attempted in this study. In some cases, small molecules can also act as molecular glues and strengthen the interaction between NRF2 and KEAP1.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors would like to thank M. Zhang for help with the binding assays. We thank the RC teams of the Faculty of Arts and Sciences at Harvard University (esp. S. Yockel, J. Cuff, F. Pontiggia, and P. Edmon), the Jülich Supercomputing Centre, the Freie Universität (esp. Jens Dreger), the Harvard Medical School (HMS), the HLRN, and the IT support of HMS (esp. K. Bayer, G. Sekmokas and D. Morgan) for their support. We are grateful to K.E. Leigh, N. Gray, M. Kostic, A. Dubey, B. Klein, S. Schwaninger, and S. Wu for helpful discussions and manuscript preparation. We acknowledge the ICCB-Longwood Screening and East Quad NMR Facilities at HMS for assistance with the ligand screen. We thank K. Arnett and the Center for Macromolecular Interactions at the HMS for advice on the SPR and BLI experiments, and ChemAxon for a license for JChemSuite. We would like to thank the teams from the Google Cloud Platform (esp. S. Fang, R. Goldenbroit, and D. Payne) and Amazon Web Services, and Fluid Numerics for their support. This work was partially funded by a scholarship to C.G. from the Max Planck Institute for Molecular Genetics in Berlin, and a scholarship from the Einstein Center for Mathematics Berlin. C.G. and K.F. would like to thank the ECMath and MATHEON. M.H acknowledges funding from Deutsche Forschungsgemeinschaft (CRC 958/Project A04, CRC 1114/Project A04). A.B. was supported by the Austrian Science Fund's Schrödinger Fellowship J3872-B21 and the American Heart Association's fellowship 19POST34380800. We would like thank A. Jaffe for his support. This research was supported in part by the grant TRT 0159 from the Templeton Religion Trust and by the ARO Grant W911NF1910302 to A. Jaffe. K.M.P.D. was supported by a fellowship from the Max Kade Foundation and the Austrian Academy of Sciences. H.A. acknowledges funding from the Claudia Adams Barr Program for Innovative Cancer Research. G.W. acknowledges support from NIH grant CA200913.

References

1. DiMasi JA, Grabowski HG & Hansen RW Innovation in the pharmaceutical industry: New estimates of R&D costs. Journal of Health Economics 47, 20–33 (2016). [PubMed: 26928437]
2. Lyu J, Wang S, Balius TE, et al.Ultra-large library docking for discovering new chemotypes. Nature (2019).
3. Zhang S, Kumar K, Jiang X, et al.DOVIS: an implementation for high-throughput virtual screening using AutoDock. BMC bioinformatics9, 126 (2008). [PubMed: 18304355]
4. Jiang X, Kumar K, Hu X, et al.DOVIS 2.0: an efficient and easy to use parallel virtual screening tool based on AutoDock 4.0. Chemistry Central Journal2, 18 (2008). [PubMed: 18778471]
5. Hassan NM, Alhossary AA, Mu Y, et al.Protein-Ligand Blind Docking Using QuickVina-W With Inter-Process Spatio-Temporal Integration. Scientific Reports7, 15451 (2017). [PubMed: 29133831]
6. Bohacek RS, McMartin C & Guida WC The art and practice of structure-based drug design: A molecular modeling perspective. Medicinal Research Reviews 16, 3–50 (1996). [PubMed: 8788213]
7. Yonchuk JG, Foley JP, Bolognese BJ, et al.Characterization of the Potent, Selective Nrf2 Activator, 3-(Pyridin-3-Ylsulfonyl)-5-(Trifluoromethyl)-2 H -Chromen-2-One, in Cellular and In Vivo Models of Pulmonary Oxidative Stress. Journal of Pharmacology and Experimental Therapeutics363, 114–125 (2017).
8. Pallesen JS, Tran KT & Bach A Non-covalent Small-Molecule Kelch-like ECH-Associated Protein 1–Nuclear Factor Erythroid 2-Related Factor 2 (Keap1–Nrf2) Inhibitors and Their Potential for Targeting Central Nervous System Diseases. Journal of Medicinal Chemistry 61, 8088–8103 (2018). [PubMed: 29750408]
9. Davies TG, Wixted WE, Coyle JE, et al.Monoacidic Inhibitors of the Kelch-like ECH-Associated Protein 1: Nuclear Factor Erythroid 2-Related Factor 2 (KEAP1:NRF2) Protein–Protein Interaction with High Cell Potency Identified by Fragment-Based Discovery. Journal of Medicinal Chemistry59, 3991–4006 (2016). [PubMed: 27031670]
10. Cuadrado Aet al.Therapeutic targeting of the NRF2 and KEAP1 partnership in chronic diseases. Nature Reviews Drug Discovery18, 295–317 (2019). [PubMed: 30610225]
11. SLURMhttps://slurm.schedmd.com.Accessed: 2018-12-22.
12. Moab/TORQUEhttp://www.adaptivecomputing.com.Accessed: 2018-12-22.
13. PBS Professionalhttp://www.pbspro.org/.Accessed: 2018-12-22.

14. LFShttps://www.ibm.com/us-en/marketplace/hpc-workload-managementAccessed: 2018-12-22.
15. SGEhttp://gridscheduler.sourceforge.net.Accessed: 2018-12-22.
16. Sterling T & Irwin JJ ZINC 15 – Ligand Discovery for Everyone. Journal of Chemical Information and Modeling 55. 2324–2337. eprint: 10.1021/acs.jcim.5b00559 (2015). [PubMed: 26479676]
17. Trott O & Olson AJ AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. Journal of computational chemistry 31, 455–61 (2010). [PubMed: 19499576]
18. Alhossary A, Handoko SD, Mu Y & Kwoh C-K Fast, accurate, and reliable molecular docking with QuickVina 2. Bioinformatics 31, 2214–2216 (2015). [PubMed: 25717194]
19. Koes DR, Baumgartner MP & Camacho CJ Lessons Learned in Empirical Scoring with smina from the CSAR 2011 Benchmarking Exercise. Journal of Chemical Information and Modeling 53, 1893–1904. arXiv: NIHMS150003 (2013). [PubMed: 23379370]
20. Ravindranath PA, Forli S, Goodsell DS, et al.AutoDockFR: Advances in Protein-Ligand Docking with Explicitly Specified Binding Site Flexibility. PLOS Computational Biology11 (ed Fetrow JS) e1004586 (2015). [PubMed: 26629955]
21. Koebel MR, Schmadeke G, Posner RG, et al.AutoDock VinaXB: implementation of XBSF, new empirical halogen bond scoring function, into AutoDock Vina. Journal of Cheminformatics8, 27 (2016). [PubMed: 27195023]
22. Nivedha AK, Thieker DF, Makeneni S, et al.Vina-Carb: Improving Glycosidic Angles during Carbohydrate Docking. Journal of Chemical Theory and Computation12, 892–901 (2016). [PubMed: 26744922]
23. Amaro RE, Baudry J, Chodera J, et al.Ensemble Docking in Drug Discovery. Biophysical Journal114, 2271–2278 (2018). [PubMed: 29606412]
24. Houston DR & Walkinshaw MD Consensus Docking: Improving the Reliability of Docking in a Virtual Screening Context. Journal of Chemical Information and Modeling 53, 384–390 (2013). [PubMed: 23351099]
25. Marcotte D, Zeng W, Hus J-C, et al.Small molecules inhibit the interaction of Nrf2 and the Keap1 Kelch domain through a non-covalent mechanism. Bioorganic & Medicinal Chemistry21, 4011–4019 (2013). [PubMed: 23647822]
26. Andrei SAet al.Stabilization of protein-protein interactions in drug discovery. Expert Opinion on Drug Discovery12, 925–940 (2017). [PubMed: 28695752]
27. Ragoza M, Hochuli J, Idrobo E, Sunseri J & Koes DR Protein–Ligand Scoring with Convolutional Neural Networks. Journal of Chemical Information and Modeling 57. 942–957. eprint: 10.1021/acs.jcim.6b00740 (2017). [PubMed: 28368587]
28. Reymond J.-l.The Chemical Space Project. Accounts of Chemical Research48, 722–730 (2015). [PubMed: 25687211]
29. ChemAxon. JChem Suite 18.20.0https://chemaxon.com/products/jchem-engines.Accessed: 2020-01-10.
30. O'Boyle NM, Banck M, James CA, et al.Open Babel: An open chemical toolbox. Journal of Cheminformatics3, 1–14 (2011). [PubMed: 21214931]
31. Enamine. REAL Libraryhttps://enamine.net/library-synthesis/real-compounds/.Accessed: 2020-01-10.
32. Morris GM, Huey R, Lindstrom W, Sanner MF, et al.AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. Journal of Computational Chemistry30, 2785–2791 (2009). [PubMed: 19399780]
33. SciPy developers. SciPyhttps://www.scipy.org/.2019.
34. Surface Plasmon Resonance (eds Mol NJ & Fischer MJE) (Humana Press, Totowa, NJ, 2010).
35. Hämäläinen MDet al.Label-free primary screening and affinity ranking of fragment libraries using parallel analysis of protein panels. Journal of Biomolecular Screening13, 202–209 (2008). [PubMed: 18270366]
36. Receptor-Ligand Interactions A Practical Approach (ed Hulme EC) (Oxford University Press, 1992).

37. Gans Pet al.Stereospecific Isotopic Labeling of Methyl Groups for NMR Spectroscopic Studies of High-Molecular-Weight Proteins. Angewandte Chemie International Edition49, 1958–1962 (2010). [PubMed: 20157899]

38. Lu Met al.Discovery of a Keap1-dependent peptide PROTAC to knockdown Tau by ubiquitination-proteasome degradation pathway. European Journal of Medicinal Chemistry146, 251–259 (2018). [PubMed: 29407955]

39. Irwin JJet al.An Aggregation Advisor for Ligand Discovery. Journal of Medicinal Chemistry58, 7076–7087 (2015). [PubMed: 26295373]

40. LaPlante SRet al.Compound Aggregation in Drug Discovery: Implementing a Practical NMR Assay for Medicinal Chemists. Journal of Medicinal Chemistry56, 5142–5150 (2013). [PubMed: 23730910]

41. Baell JB & Holloway GA New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. Journal of Medicinal Chemistry 53, 2719–2740 (2010). [PubMed: 20131845]

42. Baell JB & Nissink JWM Seven Year Itch: Pan-Assay Interference Compounds (PAINS) in 2017—Utility and Limitations. ACS Chemical Biology 13, 36–44 (2018). [PubMed: 29202222]

43. Capuzzi SJ, Muratov EN & Tropsha A Phantom PAINS: Problems with the Utility of Alerts for Pan- Assay IN terference Compound S. Journal of Chemical Information and Modeling 57, 417–427 (2017). [PubMed: 28165734]

44. Jnoff E, Albrecht C, et al.Binding Mode and Structure-Activity Relationships around Direct Inhibitors of the Nrf2-Keap1 Complex. ChemMedChem9, 699–705 (2014). [PubMed: 24504667]

45. Zhuang C, Narayanapillai S, Zhang W, et al.Rapid Identification of Keap1–Nrf2 Small-Molecule Inhibitors through Structure-Based Virtual Screening and Hit-Based Substructure Search. Journal of Medicinal Chemistry57, 1121–1126 (2014). [PubMed: 24417449]

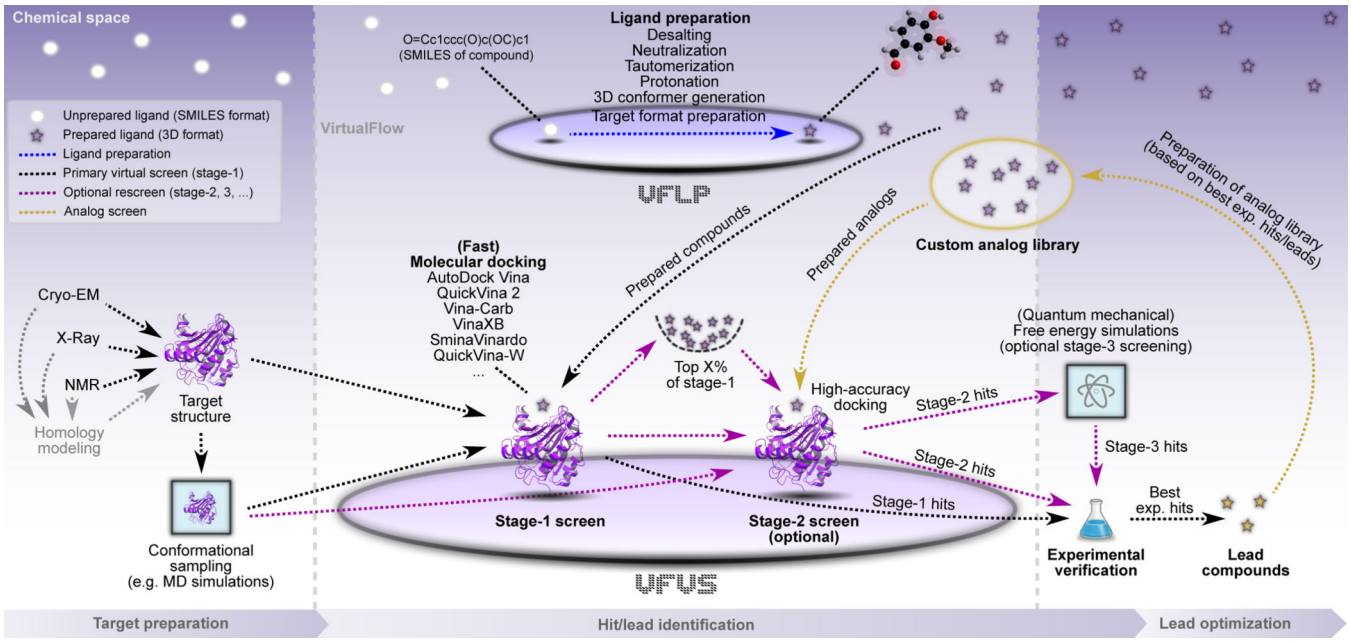


Figure 1: Application of VirtualFlow to the drug discovery process.

Before the screening can begin, the target structure, which is generally obtained from X-ray, NMR, cryo-EM, or homology modelling studies, needs to be prepared. The preparation step can include molecular dynamics (MD) simulations to obtain one or more relevant conformations of the target protein. Once the structure is prepared, it can be used to identify novel hit compounds by virtual screening-based approaches. The two independent modules of VirtualFlow, VFLP (VirtualFlow for Ligand Preparation) and VFVS (VirtualFlow for Virtual Screening), are designed to aid virtual screening-based needs. VFLP prepares (blue arrow) the desired chemical space into ready-to-dock ligand libraries, which can subsequently be used by VFVS during the virtual screen. The virtual screen consists normally of a primary virtual screening (stage-1), but can also be implemented in a multi-stage setting (violet arrows), which can include protein-side chain flexibility or inclusion of multiple protein conformations (*e.g.*, from the MD simulations or NMR structures). In addition, complimentary software packages can be leveraged during multi-staged screening (for example to carry out quantum mechanics (QM)-based free energy simulations as a final step) to improve the true hit rate and estimated binding affinities. After the virtual screening procedure, experimental verification can be carried out to identify true binders. Promising binders (lead compounds) can be further optimized by creating custom analogue libraries and screening them again with VirtualFlow (golden arrows), followed again by experimental verification of the hits.

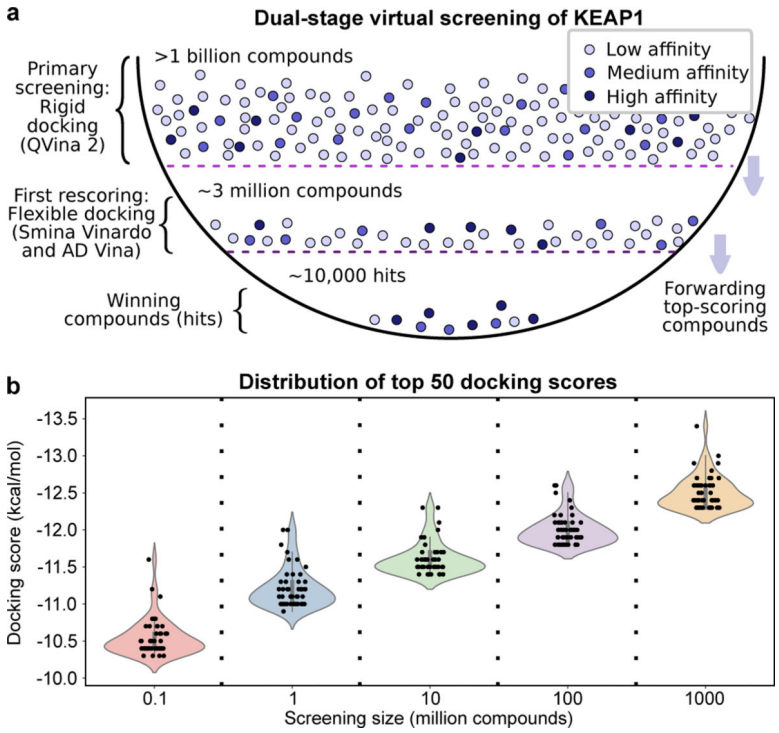


Figure 2: Schematic overview of the multi stage screen and benefits of ultra-large scale screens.

a, In stage-1, ~1.3 billion compounds were screened with the fast docking program QuickVina 2 at the lowest accuracy level. In stage-2, 13 residues of the receptor were allowed to be flexible and the top ~3 million scoring compounds were rescored with higher accuracy using AutoDock Vina and Smina Vinardo. **b**, Violin plots of the docking scores of the top 50 molecules from virtual screens targeting KEAP1 with different starting library sizes (0.1, 1, 10, 100, 1000 million ligands). In order to mimic virtual screens of smaller sizes we randomly chose a subset of the ligands from the ~1 billion compound screen of the REAL library and considered the scores of the top 50 compounds. The docking score is an estimation of the free energy of binding (in kcal/mol), and therefore the more negative the value the tighter the binding of the hit to the target. The given distributions show that the docking score of the top 50 compounds improves with the scale of the screen. The procedure was repeated independently five times with similar results.

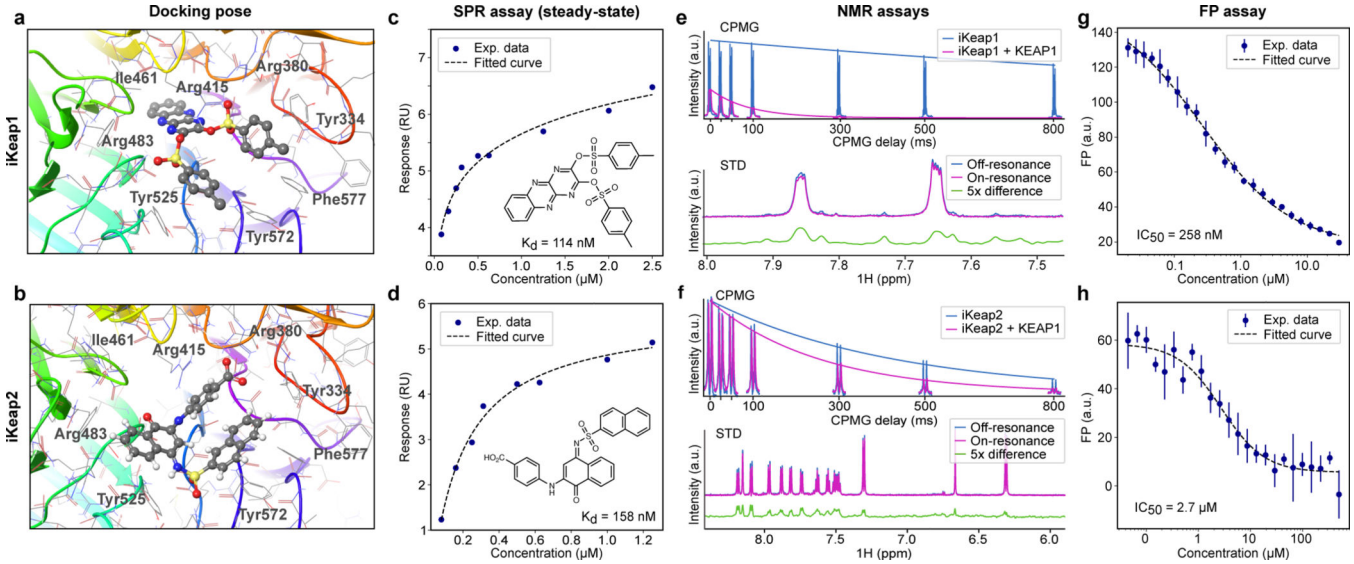


Figure 3: Docking poses (a-b) and experimental verification (c-h) of two hit compounds (iKeap1 and iKeap2).

The docking poses (**a** and **b**) were obtained from stage-2 of the virtual screening. SPR steady-state binding curves are shown for iKeap1 (**c**) and iKeap2 (**d**), showing clear binding with nanomolar K_d. We show one representative data from three independent experiments (n=3) with similar results. Ligand-detected NMR experiments,CPMG-R2 and STD-NMR (**e** and **f**) confirm the binding of the two compounds. The two hits were also functional in the FP assay (**g** and **h**) confirming that the compounds displace the peptide. The FP data shown here is from three technical replicates and the curve was fitted to the average value of the three the technical replicates. The mean and the standard deviation for the individual data points are shown. The FP was repeated independently twice with similar results and one representative result is shown here.