AI is a Viable Alternative to High Throughput Screening

Published in:
Scientific Reports

DOI:
10.1038/s41598-024-54655-z

Publication date:
2024

Licence:
CC BY

Document Version
Publisher's PDF, also known as Version of record

scientific reports

Check for updates

OPEN

AI is a viable alternative to high throughput screening: a 318-target study

The Atomwise AIMS Program[1]✉*

High throughput screening (HTS) is routinely used to identify bioactive small molecules. This requires physical compounds, which limits coverage of accessible chemical space. Computational approaches combined with vast on-demand chemical libraries can access far greater chemical space, provided that the predictive accuracy is sufficient to identify useful molecules. Through the largest and most diverse virtual HTS campaign reported to date, comprising 318 individual projects, we demonstrate that our AtomNet® convolutional neural network successfully finds novel hits across every major therapeutic area and protein class. We address historical limitations of computational screening by demonstrating success for target proteins without known binders, high-quality X-ray crystal structures, or manual cherry-picking of compounds. We show that the molecules selected by the AtomNet® model are novel drug-like scaffolds rather than minor modifications to known bioactive compounds. Our empirical results suggest that computational methods can substantially replace HTS as the first step of small-molecule drug discovery.

Despite present interest in AI/ML and thirty years of case studies[1–4], computational screening techniques have achieved limited adoption within the pharmaceutical industry. A recent investigation into the origins of 156 clinical candidates[5] found that only 1% came from virtual screening; in contrast, over 90% of clinical candidates were derived from patent busting or high throughput screening (HTS). Unfortunately, these sources are increasingly challenged, given the pharmaceutical industry's shift to novel target classes, such as proximity-induced protein degradation[6], protein–protein interactions[7], and RNA targeting[8].

Currently, HTS is the critical tool in drug discovery, providing most novel scaffolds of recent clinical candidates[5,9,10]. These initial starting points crucially shape the course of downstream medicinal chemistry efforts, as most drugs preserve at least 80% of the scaffold of the initially identified lead[11]. Despite these foundational contributions, HTS suffers from practical limitations. Principally, HTS, like all physical experiments, requires that the compounds exist. However, with the advent of synthesis-on-demand libraries, most commercially-available molecules have yet to be synthesized. Still, they can be made and delivered for testing in a matter of weeks[12–14]. These libraries comprise trillions of molecules[14,15] that exemplify millions of otherwise-unavailable scaffolds[12], providing an opportunity to substantially expand the scope and diversity of available chemical space explored in the standard drug discovery process.

Computational approaches unlock this opportunity by reversing the requirement to make molecules before testing them. When computational experiments replace HTS as the primary screen, molecules are tested *before* they are made, and the results from these experiments can inform which molecules are worth synthesizing. Computational experiments further promise to improve upon HTS in terms of cost, speed, need to produce significant quantities of protein[16], effort of miniaturizing assay formats while maintaining experimental integrity[17–19], and reducing false-positive and false-negative rates[16,20–23] including artifacts from aggregation, covalent modification of the target, autofluorescence, or interactions with the reporter rather than the target[20,24,25]. Historical computational techniques such as ligand-based QSAR[26–28], structure-based docking[29,30], and machine learning[31,32] purport to address these limitations of physical screening methods. Unfortunately, these techniques have not replaced HTS; in fact, despite increasing interest in ML, the proportion of drugs discovered with computational techniques has remained steady over the past decades[5,10].

Because there will always be individual targets for which one screening technique can identify more hits than another, the key question governing if computation is ready to be the default hit discovery technique is whether computational screens can identify hits successfully across a broad range of diverse targets. Unfortunately, despite excellent benchmark accuracies[33–35], prospective discovery accuracy remains modest[33,36,37]. For

[1]San Fransisco, CA, USA. *A list of authors and their affiliations appears at the end of the paper.
✉email: izhar@atomwise.com

example, Cerón-Carrasco[38] reported over 700 virtual screens against the SARS-CoV-2 main protease. However, when the author sought to validate the computational predictions via physical experiments, the identified compounds were barely active (800uM). Computational approaches have also been limited by a need for extensive target-specific training data[31,39–41], a requirement for high-quality X-ray crystal structures[42,43], dependence on human adjudication (so-called 'cherry-picking')[12], or a limited domain of applicability[44–48]. Even recent systems have demonstrated utility only in identifying minor variants of known molecules for well-studied proteins with tens of thousands of known binders in their training data[49,50]. Figure 1 exemplifies the striking similarities between recently ML-developed compounds and their preceding published chemical matter. This is particularly concerning, as a myopic focus on well-studied proteins has been identified as a cause of low productivity in pharmaceutical discovery[51].

Nevertheless, we have observed that deep learning approaches are not as limited as these historical examples would imply. Using our AtomNet[52–54] screening system, we have previously reported success in finding novel scaffolds for targets without known ligands[55–57], X-ray crystal structures[56–60], or both[56,57], as well as challenging modulation via protein–protein interaction[59,61] or allosteric binding[60] (see Supplementary Table S1 for examples). However, individual examples do not demonstrate the overall success of such deep learning systems. We therefore report our internal discovery efforts against 22 targets of pharmaceutical interest. We then attempted to further assess the generalizability and robustness of deep learning predictive systems by identifying bioactive molecules for a diverse set of targets. We partnered with 482 academic labs and screening centers, from 257 different academic institutions across 30 countries, through our academic collaboration program, the Artificial Intelligence Molecular Screen (AIMS). This collaboration afforded an opportunity to prospectively evaluate the utility of the AtomNet model as a primary screen across a broad range of diverse, challenging, and realistic targets. In aggregate, we report successes and failures from 318 prospective experiments and evaluate our AtomNet machine-learning technology's ability to serve as a viable alternative to physical HTS campaigns.

Results

We investigated the ability of deep learning-based methods to identify novel bioactive chemotypes by applying the AtomNet model to identify hits for 22 internal targets of pharmaceutical interest. We also explored the breadth of applicability of this approach by attempting to identify drug-like hits in single-dose screens for 296 academic targets, of which 49 were followed up with dose–response experiments, and 21 were further validated by exploring analogs of the initial hits. The average hit rate for our internal projects (6.7%) was comparable to the hit rate for our academic collaborations (7.6%).

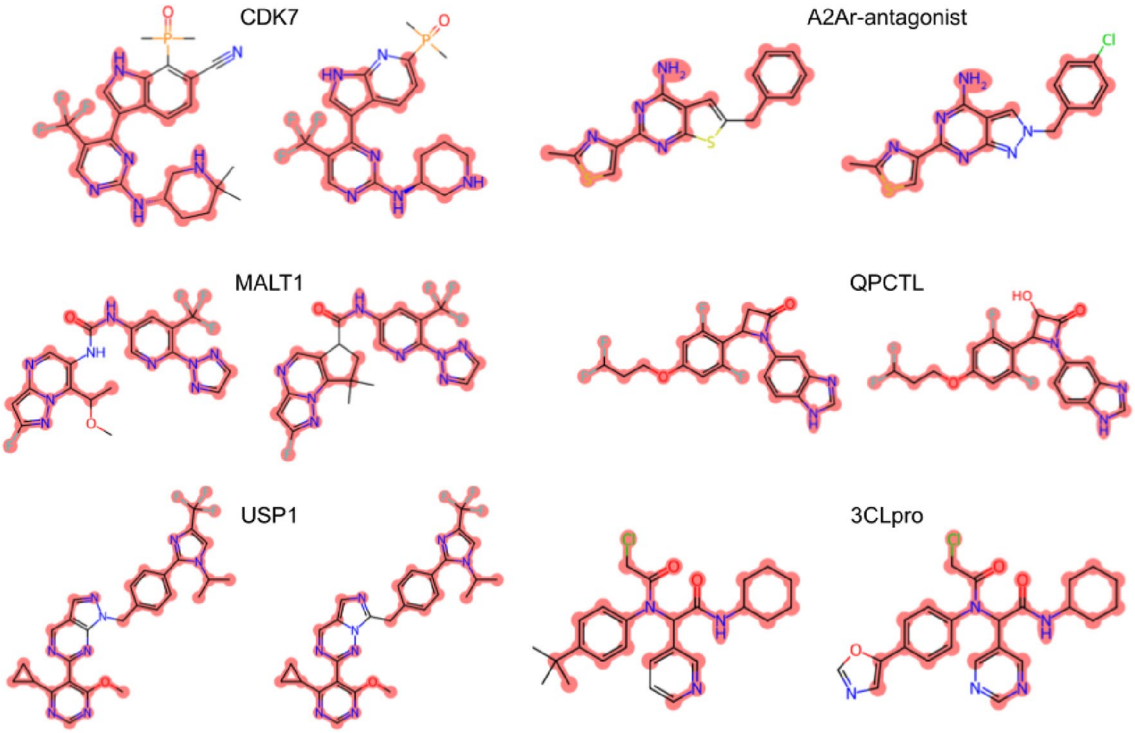

Figure 1. Pairs of representative compounds extracted from AI patents (right) and corresponding prior patents (left) for clinical-stage programs (CDK7[92,93], A2Ar-antagonist[94,95], MALT1[96,97], QPCTL[98,99], USP1[100,101], and 3CLpro[102,103]). The identical atoms between the chemical structures are highlighted in red.

Internal portfolio validation

As part of Atomwise's internal drug discovery efforts, we used the AtomNet model instead of high-throughput or DNA-encoded library (DEL) screening. We screened a 16-billion synthesis-on-demand chemical space[62], which is several thousand times larger than HTS libraries and even exceeds the size of most DELs without suffering limitations of DNA-compatible chemistry[16,23]. Each screen requires over 40,000 CPUs, 3,500 GPUs, 150 TB of main memory, and 55 TB of data transfers. We describe the protocol in detail in the Methods section; briefly, we computationally scored each catalog compound after removing molecules that were prone to interfere with the assays or were too similar to known binders of the target or its homologs. The neural network analyzes and scores the 3D coordinates of each generated protein–ligand co-complex, producing a list of ligands ranked by their predicted binding probability. Our workflow then clusters the top-ranked molecules to ensure diversity and algorithmically selects the highest-scoring exemplars from each cluster. At no point are compounds manually cherry-picked. The molecules were synthesized at Enamine (https://enamine.net) and quality controlled by LC–MS to purity > 90%, in agreement with HTS standards[63]. Hits were further validated using NMR. We then physically tested, on average, 440 compounds per target at reputable contract research organizations (CROs), while attempting to mitigate assay interferences such as aggregation and oxidation with standard additives (*e.g.*, Tween-20, Triton-X 100, and dithiothreitol (DTT)). We describe the assay protocols in detail in the Supplementary Data S1.

We describe the results of the 22 experiments in Table 1. In 91% of the experiments, we identified single-dose (SD) hits that were reconfirmed in dose–response (DR) experiments. The average target DR hit rate was 6.7% compared to 8.8% from the SD screens. Only 16 of the 22 projects were structurally enabled with X-ray crystallography; one used a cryo-EM structure, while five used homology models with an average sequence identity of 42% to their template protein. The DR hit rate for the cryo-EM project was 10.56%, while the average hit rate for the homology models was a similar 10.8%.

We then advanced 14 projects with at least one dose-responsive scaffold to a round of analog expansion. We found new bioactive analogs in the SD screen for all projects, with an average hit rate of 29.8%. Further validation with DR resulted in an average hit rate of 26% per project, which compares favorably with typical HTS hit rates ranging from 0.151 to 0.001%[64,65]. We note that the size and chemical diversity within and between physical[66] and virtual[14] HTS libraries prevent an explicit evaluation of the methods over the same chemical space. The most potent analogs ranged from single-digit nanomolar, against a kinase, to double-digit micromolar, against a transcription factor (Supplementary Table S2). Additionally, we present two internal studies in detail. For Large Tumor Suppressor Kinase 1 (LATS1), we identified potent compounds despite the lack of a crystal structure or known active compounds. For ATP-driven chaperone Valosin Containing Protein (VCP) we identified novel allosteric and orthosteric modulators.

Gene name	# of compounds tested	SD hit rate (%)	DR hit rate (%)	Potency range (IC50/Ki, uM)	# of analog tested	SD analog hit rate (%)	DR analog hit rate (%)	Analog potency range (IC50/Ki, uM)
ASAH1	376	10.64	7.71	0.3–102	–	–	–	–
AXL	597	12.06	8.21	0.181–71	3200	35.59	33.56	0.079–86
BCL2	422	3.08	0.00	–	–	–	–	–
CBLB	422	1.66	0.00	–	–	–	–	–
CDK5	786	10.69	10.43	0.049–79	587	47.53	43.61	0.43–76
CDK7	786	10.69	10.56	0.099–60	735	28.44	27.35	0.191–10
GFPT1	384	6.51	2.34	31–86	734	24.93	24.11	1–194
KCNT1	416	9.62	7.69	1.1–30	–	–	–	–
KDM6A	356	3.93	1.12	24–58	–	–	–	–
LATS1	418	18.18	17.94	0.077–82	841	51.72	45.78	0.034–98
MC2R	208	11.54	9.62	16–68	419	39.38	38.42	2.4–97
MDM4	422	2.37	0.47	5.9–29.8	192	18.23	18.23	4.4–90
NT5E	335	1.49	0.30	176	221	9.95	1.81	8.3–65
PARG	334	7.78	7.78	15–250	–	–	–	–
PARP14	576	5.38	2.95	3–96	616	26.46	26.30	0.2–95
POLQ	330	11.82	11.52	1.2–49	559	11.27	8.77	1.5–42
PPARA	422	4.03	0.24	131	211	14.22	3.79	59–95
PPM1D	530	11.89	6.98	4.5–98	–	–	–	–
PRMT5	422	4.03	0.95	7.2–79	415	7.95	5.54	19–114
PRODH2	542	2.77	1.11	15–84	–	–	–	–
TYK2	189	38.10	34.39	0.016–9	457	71.33	60.39	0.006–10
VCP	416	4.81	4.81	2.4–64	738	–	–	–

Table 1. Results from 22 Atomwise internal programs. SD and DR denote single-dose and dose–response, respectively.

Academic validation

In addition to our internal discovery efforts, we performed virtual screens for 296 targets, comprising more than 20 billion individual neural network scores of generated protein–ligand co-complexes. We purchased, on average, 85 off-the-shelf commercially available compounds, quality controlled by NMR and LC–MS to > 90% purity[63], and plated in a single 96-well plate. The compounds were then physically screened for activity against the target of interest in single-dose assays (see Supplemental Data S1 for assay protocols). As with HTS primary screens, additional characterization studies are required to validate the initially identified hits so, in 49 projects, we performed dose–response studies and analog expansion. We present a summary of our results in Supplementary Table S3.

Figure 2 illustrates the distributions of projects across therapeutic areas, protein families, and assay types. Every major therapeutic area is represented, with the most frequent area being oncology, comprising 35% of projects, followed by infectious diseases and neurology, comprising 27% and 9% of projects, respectively. Breaking down the projects by protein families reveals that all major enzyme classes are represented, with enzymes comprising 59% of the targets and membrane proteins such as GPCR, transporters, and ion channels, representing 12% of the targets. Working on a large and diverse set of therapeutic targets requires a heterogeneous collection of biological assays; 20% of the assays measured direct binding, whereas 56% and 20% were functional and phenotypic.

In 215 projects, we identified at least one bioactive compound for the target in a biochemical or cell-based assay. This 73% success rate substantially improves over the ~50% success rate for HTS[21,67]. On average, we screened 85 compounds per project and discovered 4.6 active hits, with an average hit rate of 5.5%. For the subset of targets where we found any hits, the average was 6.4 hits per project. Thus, we achieved an average hit rate of 7.6%, which again compares favorably with typical HTS hit rates. See Supplementary Material S1 for all assay definitions and conditions. Supplementary Table S4 shows a representative bioactive compound from each of the 215 successful projects, and Supplementary Fig. S2 shows that the physicochemical properties of the identified hits are largely druglike and Lipinski-compliant.

The AtomNet technology robustly identified active molecules, even for targets that lacked prior on-target bioactivity data. This ability to identify hits for previously undrugged targets is critical if machine learning-based approaches are to replace HTS as the default primary screening approach. For 207 out of the 296 targets (70%), the training data available for AtomNet models lacked a single active molecule for that target or any closely related protein (i.e., proteins with sequence identity greater than 70%). We interpret this as evidence of the ability of properly-architected machine learning systems to extrapolate to novel biological space. Figure 3A illustrates the hit rate versus the number of training examples available to our model. Although previous computational

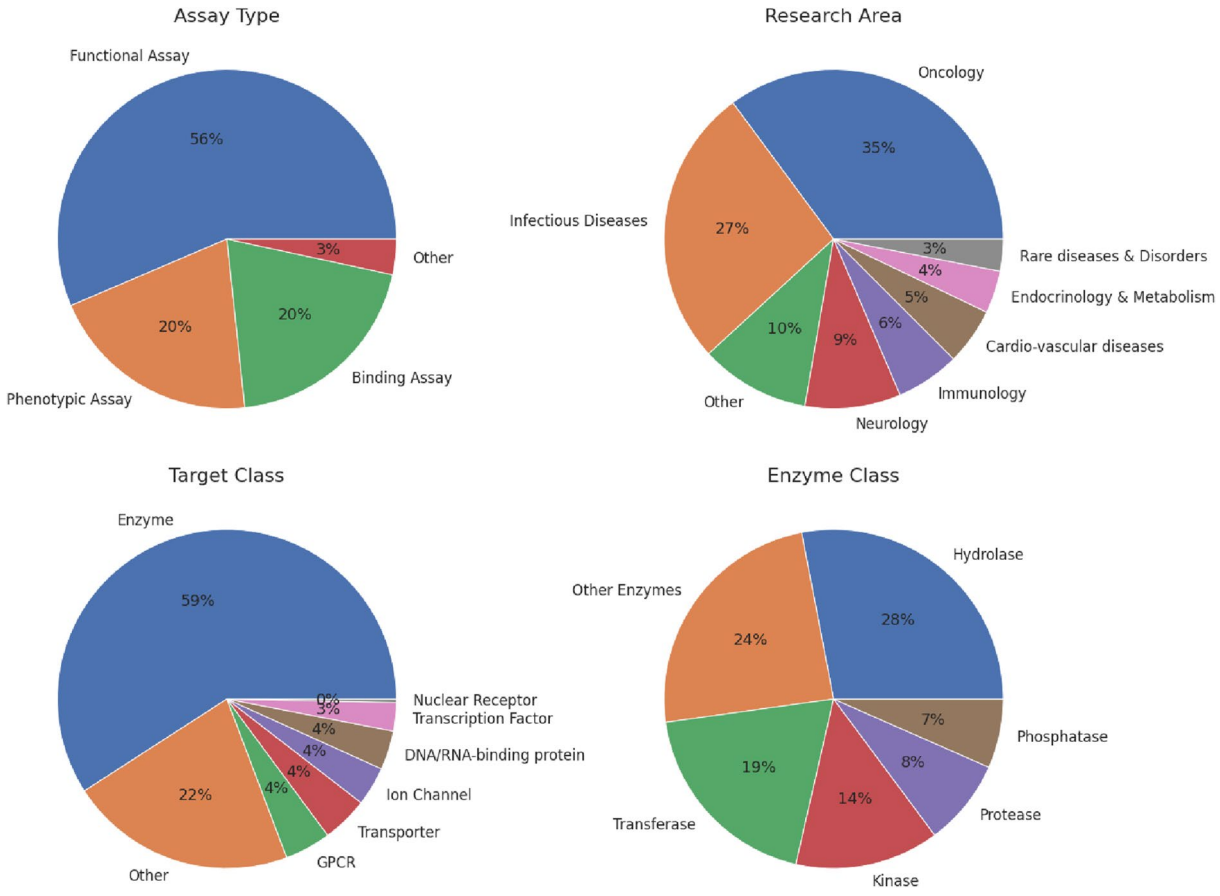

Figure 2. The distributions of 296 AIMS projects across assay types used in the primary screen, research areas, target classes, and further breakdown to enzyme classes when applicable.

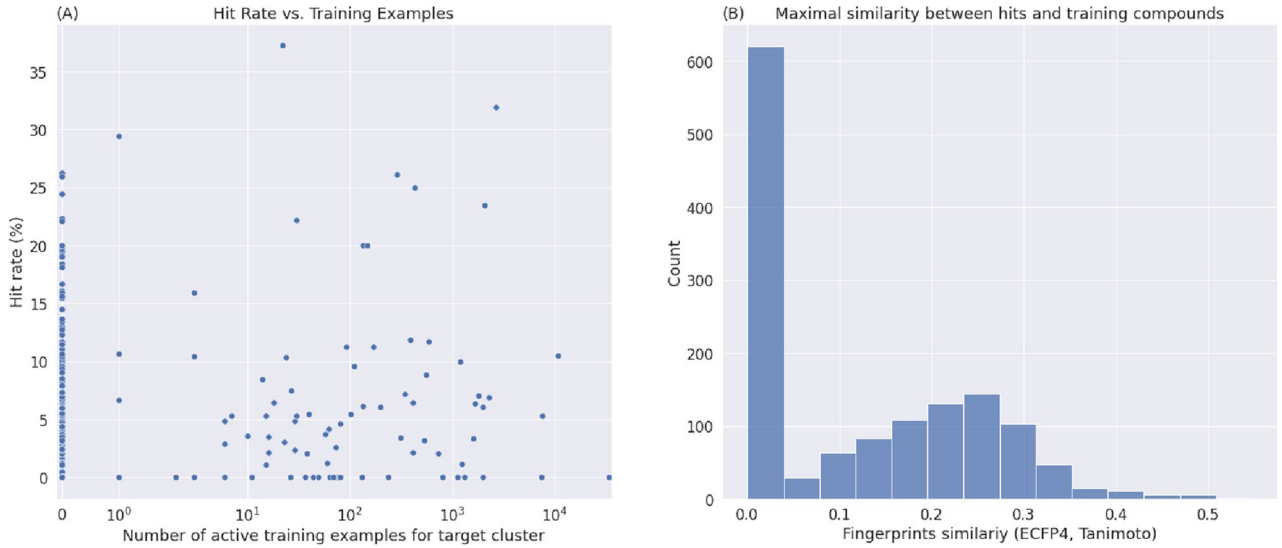

Figure 3. (**A**) An illustration of the hit rate versus the number of training examples available to our model. Each point represents a project, with the x-axis denoting the number of active molecules in our training for the target protein or homologs and the y-axis denoting the hit rate of the project (the percentage of molecules tested in the project that were active). The model shows no dependence on the availability of on-target training examples. For 70% of the targets, the AtomNet model training data lacked any active molecules for that target or any similar targets with greater than 70% sequence identity, yet the model achieved a hit rate of 5.3% compared to 6.1% when on-target data was available. (**B**) The distribution of similarities between hits and their most-similar bioactive compounds in our training data. Our screening protocol ensures that the compounds subjected to physical testing are not similar to known active compounds or close homologs (< 0.5 Tanimoto similarity using ECFP4, 1024 bits). Because 70% of the AIMS targets had no annotated bioactivities in our training dataset, hits identified in these projects have a similarity value of zero.

approaches typically require thousands of on-target training examples[31,39,42], the lack of correlation between training examples and hit rate ($R^2 = 0.0021$, p-value = 0.43) shows that our ML algorithm is agnostic to the availability of such data. We achieved an average success rate of 75% and hit rates of 5.3% when no training data was available, comparable to the 67% and 6.1% success and hit rates achieved when binding data was available in the training set. Interestingly, we also do not see a significant increase in hit rate attributable to the proportion of binding data available for a target ($R^2 = 0.008$, p-value = 0.39). This reflects the robustness of the screening protocol and the chemical dissimilarity of scaffolds identified by AtomNet models to previously known bioactive compounds.

Next, we assessed the ability of the AtomNet models to identify novel scaffolds. This is a critical capability for primary screens, as follow-up assays tend to work within the chemical space uncovered in the initial screen. The task of novel scaffold identification appears in two distinct scenarios: (1) when no scaffold is known for the target and we wish to identify the first scaffold, and (2) when some scaffolds are known but we wish to identify dissimilar scaffolds because novel chemical matter can yield improved selectivity, toxicity, pharmacokinetics, or patentability. Performance of AtomNet models for the first scenario, when no scaffolds for the target existed in the AtomNet model training data, was evaluated on 70% of the targets, where the training data contained no active molecules for the target or its homologs (vide supra). We achieved an average hit rate of 5.3% for targets with no training data. For the second scenario, we analyzed the similarity of the identified hits to known bioactive compounds in our training data (Fig. 3B). Our screening protocol ensures that the compounds subjected to physical testing are not similar to known active compounds or close homologs (< 0.5 Tanimoto similarity using ECFP4[68], 1024 bits). We interpret this as evidence of the ability of properly-architected machine learning systems to extrapolate to novel chemical space as well. For cases where training data was available (i.e., the Tanimoto similarity is above zero), the similarity distribution is close to the one expected by random compound pairs[69]. The novelty of the small-molecule structures is striking because target-specific machine-learning algorithms tend to uncover highly similar analogs for known bioactive molecules[50,70,71]. The superior performance of the AtomNet model is expected, considering the bias-variance tradeoff[72] in machine learning algorithms. Because the AtomNet convolutional neural network is a global model, concurrently trained on millions of bioactivities, hundreds of thousands of small molecules, and thousands of protein binding sites, it can reduce both bias and variance of the model compared to target-specific ones[33]. Specifically, our global model can benefit from multiple levels of information captured in the structures of the small molecules, the sequences of the target proteins, and the three-dimensional interactions between the two.

AtomNet also successfully identified active molecules when there was no X-ray crystal structure of the receptor. Figure 4A compares the hit rates obtained with 3-dimensional crystal structures, cryo-EM, and homology modeling. We did not attempt to select targets based on the similarity to the template but rather used the best template available. We observe no substantial difference in success rate between the three, in contrast to the common challenges in using homology models or low-precision structures for structure-based discovery[42,43,73]. We achieved average hit rates of 5.6%, 5.5%, and 5.1% for crystal structures, cryo-EM, and homology modeling. We

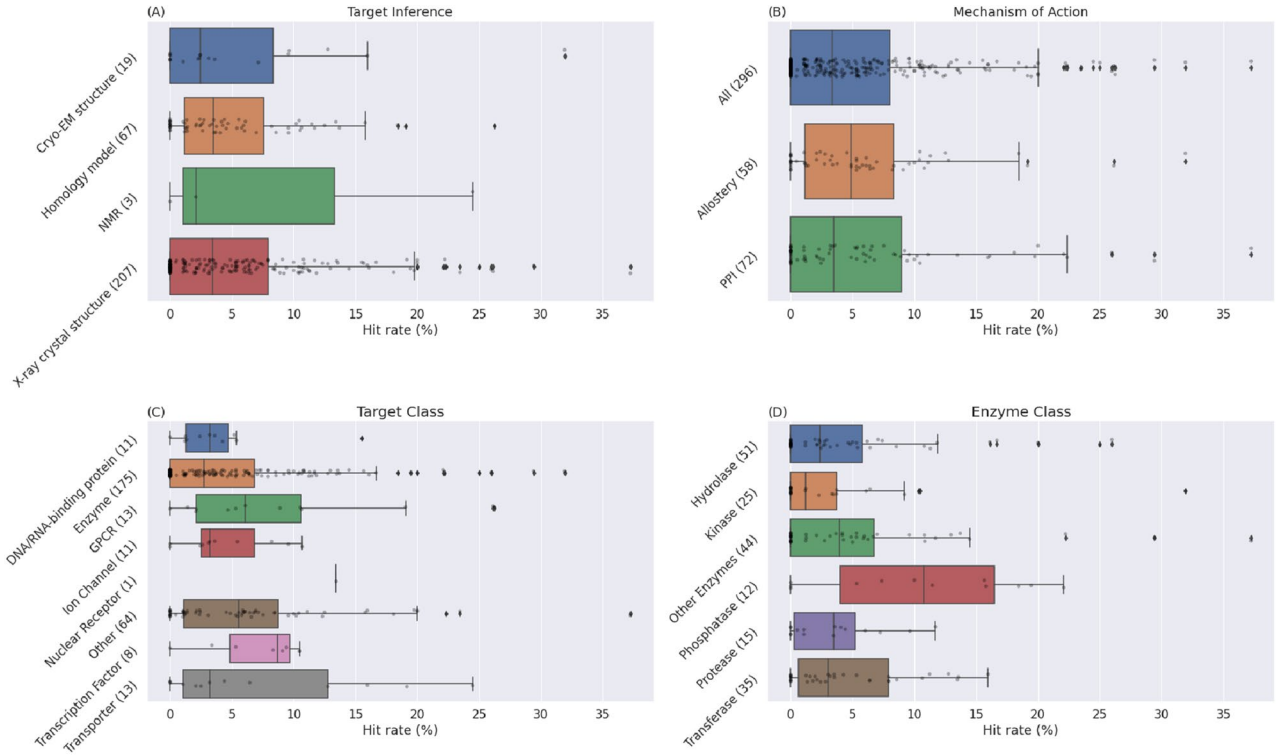

Figure 4. Hit rates obtained for the 296 AIMS projects. (**A**) A comparison of hit rates using X-ray crystallography, NMR, Cryo-EM, and homology for modeling the structure of the proteins. Each point represents a project with the x-axis denoting the hit rate of the project (the percentage of molecules tested in the project that were active). The number of projects of each type is given in parentheses. We observed no substantial difference in success rate between the physical and the computationally inferred models. We achieved average hit rates of 5.6%, 5.5%, and 5.1% for crystal structures, cryo-EM, and homology modeling, respectively. The number of projects using NMR structures is too small to make statistically-robust claims. (**B**) A comparison of hit rates observed for traditionally challenging target classes such as protein–protein interactions (PPI) and allosteric binding. Of the 296 projects, 72 targeted PPIs and 58 allosteric binding sites. The average hit rates were 6.4% and 5.8% for PPIs and allosteric binding, respectively. (**C**) Comparison of hit rates observed for different target classes and (**D**) enzyme classes. No protein or enzyme class falls outside the domain of applicability of the algorithm.

also successfully identified active compounds in projects with NMR structures, but the number of such targets is too small to make statistically-robust claims.

An interesting demonstration of the robustness of the AtomNet model to low data and poorly characterized protein structure is its ability to identify novel hits for traditionally challenging target classes such as protein–protein interaction (PPI) sites and allosteric binding sites (Fig. 3B). Of the 296 projects, 72 targeted PPIs and 58 allosteric binding sites. We identified hits for 53 (74%) PPI sites and 46 (79%) allosteric sites, with 13 projects representing allosteric sites at PPI interfaces. The average hit rate was 6.4% and 5.8% for PPIs and allosteric binding sites, respectively. The algorithm's success in these target classes, which often suffer from poorly characterized binding sites and a lack of bioactivity training data, is not surprising because Fig. 2A shows that our model is largely not dependent on the availability of on-target training data.

Finally, we investigated whether the algorithm exhibits domain of applicability limitations regarding different protein classes. Figures 4C and 3D illustrate the hit rate observed for each protein and enzyme class. No protein or enzyme class falls outside the domain of applicability of the algorithm, demonstrating that machine learning-based approaches are well-suited as a default technology for new scaffold identification. The hit rate for nuclear receptors is an outlier, with seemingly better accuracy than other classes, but a single data point is not statistically meaningful.

Dose–response validation studies

We performed additional validation studies for 49 AIMS projects with at least one reported hit. The objective of the validation studies was to establish dose–response (DR) relationships for the single-dose (SD) hits. We describe the protocol of the DR experiments in the Methods section. Briefly, we performed dose–response measurements for the reported hits from the single-dose primary screens. DR was determined using the same assay and screening protocol as the single-dose screens, at the same lab, and with the same personnel. Full dose response curves were obtained in most cases, however in some instances a full curve was not obtained, or concentration dependent activity was qualitatively determined by testing at concentrations other than that for the

primary screen. The distribution of assay types and target classes for the projects selected for DR validation also was similar to that of the AIMS projects (Supplementary Fig. S3).

We describe the results of the DR experiments in Supplementary Table S5. In 84% of the experiments, we validated at least one SD hit and got a DR readout. The median activity for the total of 144 DR measurements was 15.4 μM (which compares favorably with HTS[25,74]), of which 13% showed sub-μM potency. Overall, we achieved an average of 2.8 hits per validation study, resulting in a hit rate of 51%. The false positive rate of 49% observed in these experiments is favorably compared to HTS' which can be as high as 95%[20,75]. This difference in false positive rates may stem from the comparative ease and robustness of the low-throughput assay format we employed versus high-throughput assay. Representative dose–response curves for each of the 49 projects are shown in Supplementary Table S6.

Analog validation studies

For a subset of 21 projects, we further validated hits with DR activity by testing analogs of the active compounds. In those cases, we used the AtomNet platform to search a purchasable space for additional bioactive compounds chemically analogous to the SD hits. We selected up to 35 additional compounds for testing, including the active compounds from the SD screens.

We describe the results of the analoging experiments in Supplementary Table S7. We identified additional analogs with DR readouts for 16 projects (76%). The median DR activity of the 154 validated analogs was 7.4 μM compared to the median of 15.4 μM of the parent compound (Supplementary Fig. S4).

Methods

Screening protocols

AIMS screening protocol

We began by evaluating screening libraries of millions of catalog compounds from commercial vendors MCule (10 M)[76] and Enamine in-stock (2.5 M)[77]. We then selected a drug-like subset via algorithmic filtering by applying Eli Lilly medicinal chemistry filters[78] and removing likely false positives, such as aggregators, autofluorescers, and PAINS[79] (see Fig. 2 for the distributions of drug-like properties of the SD hits). The resulting library was virtually screened against the target of interest, removing any molecules with greater than 0.5 Tanimoto similarity in ECFP4 space to any known binders of the target and its homologs within 70% sequence identity. For kinase targets, we extend the exclusion to the whole kinome. The binding site was defined using co-complexes, mutagenesis studies, co-complexes of homologs, or by identifying potential sites using ICM Pocket Finder[80] or Fpocket[81]. Some were orthosteric, while others were allosteric, or as yet unestablished biological functions. In 64 cases, we built homology models using the closest sequence, with an average sequence similarity of 54%. We clustered the top 30,000 molecules using the Butina[82] algorithm with a Tanimoto similarity cutoff of 0.35 in ECFP4 space, selecting the highest-scoring exemplars. Additional computed physico-chemical property filters were applied as needed. At no point were compounds cherry-picked. We purchased, on average, 85 compounds, quality controlled by LC–MS to > 90% purity, generally dispensed as 10 mM DMSO stocks plated in a single 96-well plate. In addition, two vials of DMSO-only negative controls were included before scrambling the compound locations on the plate, by the supplier, for blinded experimental testing. To further control for potential artifacts, we removed compounds that showed measurable activity toward more than one target from the analysis.

Dose–response and analoging validation screening protocol

We considered advancing AIMS projects to additional validation studies based on the ability to reorder at least some of the initial SD hits, the availability of chemical analogs in the screening library to the initial hits, the capability to perform dose–response experiments, and the ability of the collaborators to perform additional screens and return results promptly.

We performed two sets of experiments: DR validation of the SD hits from AIMS and analoging with DR readouts. We performed DR measurements using the same assays and protocols as SD.

We performed an analoging round by identifying, for each AIMS hit, its 1000 nearest neighbors from the Mcule library[76], using molecular fingerprints similarity[68]. We augmented the set with additional analogs using substructure[83] or FTrees[84] searches, if needed. We used an AtomNet regression model, trained to predict quantitative bioactivities (e.g., IC50 or Ki), to score and rank the analogs. A set of 20—35 compounds from the analogs space of an initial hit were then obtained based on similarity and top scores from the AtomNet model for testing.

Internal portfolio screening protocol

We followed a protocol similar to the AIMS screen with a few deviations. First, we used the Enamine REAL library of over 16 billion compounds[62]. Second, we used an ensemble of six AtomNet models for the screens. Last, on average, we selected a set of 440 compounds for testing.

The analoging protocol is similar to the AIMS validation studies, with the following deviations. First, we used the Enamine REAL library for analog search. Second, we selected an average of 676 analogs per project. Third, the analog search protocol was more complex, pulling nearest neighbors based on maximum common substructure and graph edit distance in addition to the ECFP4-based one.

AtomNet® model architecture

We previously published in detail[52,53,55,58,59,61,85,86] during the course of the AIMS program, and we described the most recent version of the AtomNet model architecture in detail elsewhere[53]. We provide a brief description below.

The AtomNet model is a Graph Convolution Network architecture with atoms represented as vertices and pair-wise, distance-dependent, edges representing atom proximities. The input is a graph network of features characterizing the atom types and topologies of an ensemble of protein–ligand complexes. Receptor atoms more than 7 Å away from any ligand atom are excluded from the complexes, and each node in the graph is associated with a feature vector representing the atom type using Sybyl typing[87].

The network has five graph convolutional blocks. In the first two graph convolution blocks, all ligand and receptor atoms 5 Å apart from each other are considered, and 64 filters per block are used. In the third block, the cutoff radius and filters are increased to 7 Å and 128, respectively. Only ligand features in the last two blocks are considered without changing the threshold cutoff or the number of filters. Finally, the sum-pool of the ligand-only layer creates a 3-task layer on top of the network. That multi-task layer predicts three endpoints: bioactivity, pose quality, and a physics-based docking score[88].

We trained an ensemble of 6 models, splitting the training data into sixfold cross-validation sets based on a protein sequence similarity cutoff of 70%. Then, each model in the ensemble was trained on a different fold for 10 epochs, using the ADAM optimizer[89] with a learning rate of 0.001, and targets were sampled with replacement, proportional to the number of active compounds associated with that target.

Data

All data generated or analyzed during this study are included in this published article (and its supplementary information S1 files). Boxplots illustrations show the quartiles (Q1 and Q3) of the dataset while the whiskers extend to show the rest of the distribution, except for points that are determined to be "outliers" (1.5 × of the inter-quartile range, as implemented in the Seaborn and Matplotlib toolboxes[90,91]).

Conclusion

HTS is the most widely-used tool for hit discovery for new targets. Unfortunately, all physical screening methods share the critical limitation that a molecule must exist to be screened. Computational methods enable a fundamental shift to a test-then-make paradigm. In this work, we report on 318 projects (22 internal projects and 296 collaborations) where we used the AtomNet platform as the primary screening tool coupled with low-throughput physical screens as validation. The AtomNet technology can identify bioactive scaffolds across a wide range of proteins, even without known binders, X-ray structures, or manual cherry-picking of compounds. Our empirical results suggest that machine learning approaches have reached a computational accuracy that can replace HTS as the first step of small-molecule drug discovery.

Data availability

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Received: 15 September 2023; Accepted: 15 February 2024
Published online: 02 April 2024

References

1. Kuntz, I. D. Structure-based strategies for drug design and discovery. *Science* **257**, 1078–1082 (1992).
2. Bajorath, J. Integration of virtual and high-throughput screening. *Nat. Rev. Drug Discov.* **1**, 882–894 (2002).
3. Walters, W. P., Stahl, M. T. & Murcko, M. A. Virtual screening—an overview. *Drug Discov. Today* **3**, 160–178 (1998).
4. Ring, C. S. *et al.* Structure-based inhibitor design by using protein models for the development of antiparasitic agents. *Proc. Natl. Acad. Sci. USA.* **90**, 3583–3587 (1993).
5. Brown, D. G. An analysis of successful hit-to-clinical candidate pairs. *J. Med. Chem.* https://doi.org/10.1021/acs.jmedchem.3c00521 (2023).
6. Békés, M., Langley, D. R. & Crews, C. M. PROTAC targeted protein degraders: The past is prologue. *Nat. Rev. Drug Discov.* **21**, 181–200 (2022).
7. Lu, H. *et al.* Recent advances in the development of protein–protein interactions modulators: Mechanisms and clinical trials. *Signal Transduct. Target. Ther.* **5**, 1–23 (2020).
8. Childs-Disney, J. L. *et al.* Targeting RNA structures with small molecules. *Nat. Rev. Drug Discov.* **21**, 736–762 (2022).
9. Brown, D. G. & Boström, J. Where do recent small molecule clinical development candidates come from?. *J. Med. Chem.* **61**, 9442–9468 (2018).
10. Dragovich, P. S., Haap, W., Mulvihill, M. M., Plancher, J.-M. & Stepan, A. F. Small-molecule lead-finding trends across the roche and genentech research organizations. *J. Med. Chem.* **65**, 3606–3615 (2022).
11. Perola, E. An analysis of the binding efficiencies of drugs and their leads in successful drug discovery programs. *J. Med. Chem.* **53**, 2986–2997 (2010).
12. Lyu, J. *et al.* Ultra-large library docking for discovering new chemotypes. *Nature* **566**, 224 (2019).
13. Sadybekov, A. A. *et al.* Synthon-based ligand discovery in virtual libraries of over 11 billion compounds. *Nature* **601**, 452–459 (2022).
14. Bellmann, L., Penner, P., Gastreich, M. & Rarey, M. Comparison of combinatorial fragment spaces and its application to ultralarge make-on-demand compound catalogs. *J. Chem. Inf. Model.* **62**, 553–566 (2022).
15. Neumann, A., Marrison, L. & Klein, R. Relevance of the trillion-sized chemical space "explore" as a source for drug discovery. *ACS Med. Chem. Lett.* **14**, 466–472 (2023).
16. Sunkari, Y. K., Siripuram, V. K., Nguyen, T.-L. & Flajolet, M. High-power screening (HPS) empowered by DNA-encoded libraries. *Trends Pharmacol. Sci.* **43**, 4–15 (2022).
17. Malo, N., Hanley, J. A., Cerquozzi, S., Pelletier, J. & Nadon, R. Statistical practice in high-throughput screening data analysis. *Nat. Biotechnol.* **24**, 167–175 (2006).
18. Iversen, P. W., Eastwood, B. J., Sittampalam, G. S. & Cox, K. L. A comparison of assay performance measures in screening assays: Signal window, Z' factor, and assay variability ratio. *J. Biomol. Screen.* **11**, 247–252 (2006).
19. Zhang, J.-H., Chung, T. D. Y. & Oldenburg, K. R. A simple statistical parameter for use in evaluation and validation of high throughput screening assays. *J. Biomol. Screen.* **4**, 67–73 (1999).

20. Jadhav, A. *et al.* Quantitative analyses of aggregation, autofluorescence, and reactivity artifacts in a screen for inhibitors of a thiol protease. *J. Med. Chem.* **53**, 37–51 (2010).
21. Fox, S. *et al.* High-throughput screening: Update on practices and success. *J. Biomol. Screen.* **11**, 864–869 (2006).
22. Owen, S. C., Doak, A. K., Wassam, P., Shoichet, M. S. & Shoichet, B. K. Colloidal aggregation affects the efficacy of anticancer drugs in cell culture. *ACS Chem. Biol.* **7**, 1429–1435 (2012).
23. Rössler, S. L., Grob, N. M., Buchwald, S. L. & Pentelute, B. L. Abiotic peptides as carriers of information for the encoding of small-molecule library synthesis. *Science* **379**, 939–945 (2023).
24. McGovern, S. L., Caselli, E., Grigorieff, N. & Shoichet, B. K. A Common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. *J. Med. Chem.* **45**, 1712–1722 (2002).
25. Feng, B. Y., Shelat, A., Doman, T. N., Guy, R. K. & Shoichet, B. K. High-throughput assays for promiscuous inhibitors. *Nat. Chem. Biol.* **1**, 146–148 (2005).
26. Martin, E. J., Polyakov, V. R., Tian, L. & Perez, R. C. Profile-QSAR 2.0: Kinase virtual screening accuracy comparable to four-concentration IC50s for realistically novel compounds. *J. Chem. Inf. Model.* **57**, 2077–2088 (2017).
27. Keiser, M. J. *et al.* Predicting new molecular targets for known drugs. *Nature* **462**, 175–181 (2009).
28. Svetnik, V. *et al.* Random forest: A classification and regression tool for compound classification and QSAR modeling. *J. Chem. Inf. Comput. Sci.* **43**, 1947–1958 (2003).
29. Kitchen, D. B., Decornez, H., Furr, J. R. & Bajorath, J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat. Rev. Drug Discov.* **3**, 935–949 (2004).
30. Shoichet, B. K. Virtual screening of chemical libraries. *Nature* **432**, 862–865 (2004).
31. Ma, J., Sheridan, R. P., Liaw, A., Dahl, G. E. & Svetnik, V. Deep neural nets as a method for quantitative structure-activity relationships. *J. Chem. Inf. Model.* **55**, 263–274 (2015).
32. Sheridan, R. P. *et al.* Machine Learning and Deep Learning Experimental error, kurtosis, activity cliffs, and methodology: What limits the predictivity of QSAR models?. *J. Chem. Inf. Model.* https://doi.org/10.1021/acs.jcim.9b01067 (2020).
33. Wallach, I. & Heifets, A. Most ligand-based classification benchmarks reward memorization rather than generalization. *J. Chem. Inf. Model.* **58**, 916–932 (2018).
34. Chen, L. *et al.* Hidden bias in the DUD-E dataset leads to misleading performance of deep learning in structure-based virtual screening. *PLOS ONE* **14**, e0220113 (2019).
35. Chuang, K. V. & Keiser, M. J. Comment on "Predicting reaction performance in C–N cross-coupling using machine learning". *Science* **362**, eaat8603 (2018).
36. Gaieb, Z. *et al.* D3R Grand Challenge 3: Blind prediction of protein–ligand poses and affinity rankings. *J. Comput. Aided Mol. Des.* **33**, 1–18 (2019).
37. Gabel, J., Desaphy, J. & Rognan, D. Beware of machine learning-based scoring functions on the danger of developing black boxes. *J. Chem. Inf. Model.* **54**, 2807–2815 (2014).
38. Cerón-Carrasco, J. P. When virtual screening yields inactive drugs: dealing with false theoretical friends. *ChemMedChem* **17**, e202200278 (2022).
39. McCloskey, K. *et al.* Machine learning on DNA-encoded libraries: A new paradigm for hit-finding. *J. Med. Chem.* **63**, 8857–8866 (2020).
40. Wenzel, J., Matter, H. & Schmidt, F. Predictive multitask deep neural network models for ADME-Tox properties: Learning from large data sets. *J. Chem. Inf. Model.* **59**, 1253–1268 (2019).
41. Feinberg, E. N. *et al.* PotentialNet for molecular property prediction. *ACS Cent. Sci.* **4**, 1520–1530 (2018).
42. Schindler, C. E. M. *et al.* Large-scale assessment of binding free energy calculations in active drug discovery projects. *J. Chem. Inf. Model.* **60**, 5457–5474 (2020).
43. Bordogna, A., Pandini, A. & Bonati, L. Predicting the accuracy of protein–ligand docking on homology models. *J. Comput. Chem.* **32**, 81–98 (2011).
44. Stokes, J. M. *et al.* A deep learning approach to antibiotic discovery. *Cell* **180**, 688-702.e13 (2020).
45. Melo, M. C. R., Maasch, J. R. M. A. & de la Fuente-Nunez, C. Accelerating antibiotic discovery through artificial intelligence. *Commun. Biol.* **4**, 1–13 (2021).
46. Skinnider, M. A. *et al.* A deep generative model enables automated structure elucidation of novel psychoactive substances. *Nat. Mach. Intell.* **3**, 973–984 (2021).
47. Muegge, I. & Oloff, S. Advances in virtual screening. *Drug Discov. Today Technol.* **3**, 405–411 (2006).
48. N. Muratov, E. *et al.* QSAR without borders. *Chem. Soc. Rev.* **49**, 3525–3564 (2020).
49. Zhavoronkov, A. *et al.* Deep learning enables rapid identification of potent DDR1 kinase inhibitors. *Nat. Biotechnol.* **37**, 1038–1040 (2019).
50. Walters, W. P. & Murcko, M. Assessing the impact of generative AI on medicinal chemistry. *Nat. Biotechnol.* **38**, 143–145 (2020).
51. Scannell, J. W., Blanckley, A., Boldon, H. & Warrington, B. Diagnosing the decline in pharmaceutical R&D efficiency. *Nat. Rev. Drug Discov.* **11**, 191 (2012).
52. Wallach, I., Dzamba, M. & Heifets, A. AtomNet: A Deep Convolutional Neural Network for Bioactivity Prediction in Structure-based Drug Discovery. *ArXiv Prepr. ArXiv151002855* 1–11 (2015).
53. Gniewek, P., Worley, B., Stafford, K., van den Bedem, H. & Anderson, B. *Learning physics confers pose-sensitivity in structure-based virtual screening.* https://doi.org/10.48550/arXiv.2110.15459 (2021).
54. Stafford, K. A., Anderson, B. M., Sorenson, J. & van den Bedem, H. AtomNet PoseRanker: Enriching ligand pose quality for dynamic proteins in virtual high-throughput screens. *J. Chem. Inf. Model.* **62**, 1178–1189 (2022).
55. Hsieh, C.-H. *et al.* Miro1 marks parkinson's disease subset and miro1 reducer rescues neuron loss in Parkinson's models. *Cell Metab.* **30**, 1131-1140.e7 (2019).
56. Reidenbach, A. G. *et al.* Multimodal small-molecule screening for human prion protein binders. *J. Biol. Chem.* **295**, 13516–13531 (2020).
57. Bon, C. *et al.* Discovery of novel trace amine-associated receptor 5 (TAAR5) antagonists using a deep convolutional neural network. *Int. J. Mol. Sci.* **23**, 3127 (2022).
58. Stecula, A., Hussain, M. S. & Viola, R. E. Discovery of novel inhibitors of a critical brain enzyme using a homology model and a deep convolutional neural network. *J. Med. Chem.* **63**, 8867–8875 (2020).
59. Su, S. *et al.* SPOP and OTUD7A Control EWS–FLI1 protein stability to govern ewing sarcoma growth. *Adv. Sci.* **8**, 2004846 (2021).
60. Pedicone, C. *et al.* Discovery of a novel SHIP1 agonist that promotes degradation of lipid-laden phagocytic cargo by microglia. *iScience* **25**, 104170 (2022).
61. Huang, C. *et al.* Small molecules block the interaction between porcine reproductive and respiratory syndrome virus and CD163 receptor and the infection of pig cells. *Virol. J.* **17**, 116 (2020).
62. Grygorenko, O. O. *et al.* Generating multibillion chemical space of readily accessible screening compounds. *iScience* **23**, 101681 (2020).
63. Dandapani, S., Rosse, G., Southall, N., Salvino, J. M. & Thomas, C. J. Selecting, acquiring, and using small molecule libraries for high-throughput screening. *Curr. Protoc. Chem. Biol.* **4**, 177–191 (2012).
64. Schuffenhauer, A. *et al.* Library design for fragment based screening. *Curr. Top. Med. Chem.* **5**, 751–762 (2005).

65. Jacoby, E. *et al.* Key aspects of the novartis compound collection enhancement project for the compilation of a comprehensive Chemogenomics drug discovery screening collection. *Curr. Top. Med. Chem.* **5**, 397–411 (2005).
66. Petrova, T., Chuprina, A., Parkesh, R. & Pushechnikov, A. Structural enrichment of HTS compounds from available commercial libraries. *MedChemComm* **3**, 571–579 (2012).
67. Macarron, R. *et al.* Impact of high-throughput screening in biomedical research. *Nat. Rev. Drug Discov.* **10**, 188–195 (2011).
68. Rogers, D. & Hahn, M. Extended-connectivity fingerprints. *J. Chem. Inf. Model.* **50**, 742–754 (2010).
69. Riniker, S. & Landrum, G. A. Open-source platform to benchmark fingerprints for ligand-based virtual screening. *J. Cheminformatics* **5**, 26 (2013).
70. Ren, F. *et al.* AlphaFold accelerates artificial intelligence powered drug discovery: Efficient discovery of a novel cyclin-dependent kinase 20 (CDK20) Small Molecule Inhibitor (2022).
71. Assessing structural novelty of the first AI-designed drug candidates to go into human clinical trials. *CAS* https://www.cas.org/resources/blog/ai-drug-candidates.
72. Kohavi, R. & Wolpert, D. Bias plus variance decomposition for zero-one loss functions. in *Proceedings of the Thirteenth International Conference on International Conference on Machine Learning* 275–283 (Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1996).
73. Ferrara, P. & Jacoby, E. Evaluation of the utility of homology models in high throughput docking. *J. Mol. Model.* **13**, 897–905 (2007).
74. Walters, W. P. & Namchuk, M. Designing screens: How to make your hits a hit. *Nat. Rev. Drug Discov.* **2**, 259–266 (2003).
75. Inglese, J. *et al.* High-throughput screening assays for the identification of chemical probes. *Nat. Chem. Biol.* **3**, 466–479 (2007).
76. mcule database. https://mcule.com/database/.
77. Screening Collections - Enamine. https://enamine.net/compound-collections/screening-collection.
78. Bruns, R. F. & Watson, I. A. Rules for identifying potentially reactive or promiscuous compounds. *J. Med. Chem.* **55**, 9763–9772 (2012).
79. Baell, J. B. & Holloway, G. A. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J. Med. Chem.* **53**, 2719–2740 (2010).
80. Abagyan, R. & Kufareva, I. The flexible pocketome engine for structural chemogenomics. *Methods Mol. Biol. Clifton NJ* **575**, 249–279 (2009).
81. Le Guilloux, V., Schmidtke, P. & Tuffery, P. Fpocket: An open source platform for ligand pocket detection. *BMC Bioinformatics* **10**, 168 (2009).
82. Butina, D. Unsupervised data base clustering based on daylight's fingerprint and tanimoto similarity: A fast and automated way to cluster small and large data sets. *J. Chem. Inf. Comput. Sci.* **39**, 747–750 (1999).
83. *RDKit: Open-Source Cheminformatics*.
84. Rarey, M. & Dixon, J. S. Feature trees: A new molecular similarity measure based on tree matching. *J. Comput. Aided Mol. Des.* **12**, 471–490 (1998).
85. Stafford, K., Anderson, B. M., Sorenson, J. & van den Bedem, H. *AtomNet PoseRanker: Enriching Ligand Pose Quality for Dynamic Proteins in Virtual High Throughput Screens.* https://doi.org/10.26434/chemrxiv-2021-t6xkj (2021).
86. Schroedl, S. Current methods and challenges for deep learning in drug discovery. *Drug Discov. Today Technol.* **32–33**, 9–17 (2019).
87. Bender, A., Mussa, H. Y., Glen, R. C. & Reiling, S. Molecular similarity searching using atom environments, information-based feature selection, and a Naïve Bayesian classifier. *J. Chem. Inf. Comput. Sci.* **44**, 170–178 (2004).
88. Trott, O. & Olson, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461 (2010).
89. Kingma, D. P. & Ba, J. Adam: A Method for Stochastic Optimization. *ArXiv14126980 Cs* (2017).
90. Waskom, M. L. seaborn: Statistical data visualization. *J. Open Source Softw.* **6**, 3021 (2021).
91. Hunter, J. D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).
92. Marineau, J. J. *et al.* Discovery of SY-5609: A selective, noncovalent inhibitor of CDK7. *J. Med. Chem.* **65**, 1458–1480 (2022).
93. Gu, X., BAI, H., Barbeau, O. R. & Besnard, J. Aromatic heterocyclic compound, and pharmaceutical composition and application thereof. (2022).
94. Barbay, J. K., Chakravarty, D., Leonard, K., Shook, B. C. & Wang, A. Phenyl and heteroaryl substituted thieno[2,3-d]Pyrimidines and their use as adenosine A2a receptor antagonists (2010).
95. Bell, A. S., Schreyer, A. M. & Versluys, S. Pyrazolopyrimidine compounds as adenosine receptor antagonists (2019).
96. Soldermann, C. P. *et al.* Pyrazolo pyrimidine derivatives and their use as MALT1 inhbitors (2019).
97. Feng, S. *et al.* Tricyclic compounds useful in the treatment of cancer, autoimmune and inflammatory disorders (2023).
98. Heiser, U. & Sommer, R. Inhibitors of glutaminyl cyclase (2020).
99. Cheng, X., Liu, Y., Qin, L., Ren, F. & Wu, J. Beta-lactam derivatives for the treatment of diseases (2023).
100. Wylie, A. A. *et al.* Therapeutic combinations comprising ubiquitin-specific-processing protease 1 (usp1) inhibitors and poly (adp-ribose) polymerase (parp) inhibitors (2021).
101. Wu, J., Qin, L. & Liu, J. Small molecule inhibitors of ubiquitin specific protease 1 (usp1) and uses thereof 2023).
102. Stille, J. *et al.* Design, Synthesis and Biological Evaluation of Novel SARS-CoV-2 3CLpro Covalent Inhibitors. https://doi.org/10.26434/chemrxiv.13087742.v1 (2020).
103. Zavoronkovs, A., Ivanenkov, Y. A. & Zagribelnyy, B. Sars-cov-2 inhibitors having covalent modifications for treating coronavirus infections. (2021).

Acknowledgements

See Supplementary section S1.

Author contributions

All authors have contributed to the publication, being variously involved in technology development, experimental protocol designs, experimental performance, data acquisition, statistical analysis, and manuscript writing.

Competing interests

The authors affiliated with Atomwise declare the existence of a financial competing interest.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-54655-z.

Correspondence and requests for materials should be addressed to

The Atomwise AIMS Program

Izhar Wallach[2], Denzil Bernard[2], Kong Nguyen[2], Gregory Ho[2], Adrian Morrison[2], Adrian Stecula[2], Andreana Rosnik[2], Ann Marie O'Sullivan[2], Aram Davtyan[2], Ben Samudio[2], Bill Thomas[2], Brad Worley[2], Brittany Butler[2], Christian Laggner[2], Desiree Thayer[2], Ehsan Moharreri[2], Greg Friedland[2], Ha Truong[2], Henry van den Bedem[2], Ho Leung Ng[2], Kate Stafford[2], Krishna Sarangapani[2], Kyle Giesler[2], Lien Ngo[2], Michael Mysinger[2], Mostafa Ahmed[2], Nicholas J. Anthis[2], Niel Henriksen[2], Pawel Gniewek[2], Sam Eckert[2], Saulo de Oliveira[2], Shabbir Suterwala[2], Srimukh Veccham Krishna PrasadPrasad[2], Stefani Shek[2], Stephanie Contreras[2], Stephanie Hare[2], Teresa Palazzo[2], Terrence E. O'Brien[2], Tessa Van Grack[2], Tiffany Williams[2], Ting-Rong Chern[2], Victor Kenyon[2], Andreia H. Lee[3], Andrew B. Cann[4], Bastiaan Bergman[5], Brandon M. Anderson[6], Bryan D. Cox[7], Jeffrey M. Warrington[8], Jon M. Sorenson[9], Joshua M. Goldenberg[10], Matthew A. Young[11], Nicholas DeHaan[12], Ryan P. Pemberton[13], Stefan Schroedl[14], Tigran M. Abramyan[11,15], Tushita Gupta[16], Venkatesh Mysore[17], Adam G. Presser[18], Adolfo A. Ferrando[19], Adriano D. Andricopulo[20], Agnidipta Ghosh[21], Aicha Gharbi Ayachi[22], Aisha Mushtaq[23], Ala M. Shaqra[24], Alan Kie Leong Toh[25], Alan V. Smrcka[26], Alberto Ciccia[27], Aldo Sena de Oliveira[28], Aleksandr Sverzhinsky[29], Alessandra Mara de Sousa[30], Alexander I. Agoulnik[31], Alexander Kushnir[32], Alexander N. Freiberg[33], Alexander V. Statsyuk[34], Alexandre R. Gingras[35], Alexei Degterev[36], Alexey Tomilov[37], Alice Vrielink[38], Alisa A. Garaeva[39], Amanda Bryant-Friedrich[40], Amedeo Caflisch[41], Amit K. Patel[35], Amith Vikram Rangarajan[42], An Matheeussen[43], Andrea Battistoni[44], Andrea Caporali[45], Andrea Chini[46], Andrea Ilari[47], Andrea Mattevi[48], Andrea Talbot Foote[49], Andrea Trabocchi[50], Andreas Stahl[51], Andrew B. Herr[52], Andrew Berti[40], Andrew Freywald[53], Andrew G. Reidenbach[54], Andrew Lam[55], Andrew R. Cuddihy[56], Andrew White[57], Angelo Taglialatela[19], Anil K. Ojha[58], Ann M. Cathcart[59], Anna A. L. Motyl[45], Anna Borowska[39], Anna D'Antuono[60], Anna K. H. Hirsch[61], Anna Maria Porcelli[62], Anna Minakova[48], Anna Montanaro[60], Anna Müller[41], Annarita Fiorillo[63], Anniina Virtanen[64], Anthony J. O'Donoghue[35], Antonio Del Rio Flores[51], Antonio E. Garmendia[65], Antonio Pineda-Lucena[66], Antonito T. Panganiban[67], Ariela Samantha[38], Arnab K. Chatterjee[68], Arthur L. Haas[69], Ashleigh S. Paparella[21], Ashley L. St. John[70], Ashutosh Prince[71], Assmaa ElSheikh[72], Athena Marie Apfel[57], Audrey Colomba[73], Austin O'Dea[74], Bakary N'tji Diallo[75], Beatriz Murta Rezende Moraes Ribeiro[76], Ben A. Bailey-Elkin[77], Benjamin L. Edelman[78], Benjamin Liou[52], Benjamin Perry[79], Benjamin Soon Kai Chua[80], Benjámin Kováts[81], Bernhard Englinger[59], Bijina Balakrishnan[82], Bin Gong[33], Bogos Agianian[21], Brandon Pressly[37], Brenda P. Medellin Salas[83], Brendan M. Duggan[35], Brian V. Geisbrecht[84], Brian W. Dymock[85], Brianna C. Morten[85], Bruce D. Hammock[37], Bruno Eduardo Fernandes Mota[76], Bryan C. Dickinson[86], Cameron Fraser[87], Camille Lempicki[88], Carl D. Novina[89], Carles Torner[90], Carlo Ballatore[35], Carlotta Bon[91], Carly J. Chapman[92], Carrie L. Partch[93], Catherine T. Chaton[94], Chang Huang[65], Chao-Yie Yang[95], Charlene M. Kahler[38], Charles Karan[27], Charles Keller[96], Chelsea L. Dieck[97], Chen Huimei[70], Chen Liu[98], Cheryl Peltier[77], Chinmay Kumar Mantri[70], Chinyere Maat Kemet[55], Christa E. Müller[99], Christian Weber[100], Christina M. Zeina[59],

Christine S. Muli[101], Christophe Morisseau[37], Cigdem Alkan[33], Clara Reglero[19], Cody A. Loy[101], Cornelia M. Wilson[102], Courtney Myhr[31], Cristina Arrigoni[48], Cristina Paulino[39], César Santiago[103], Dahai Luo[22], Damon J. Tumes[104], Daniel A. Keedy[105], Daniel A. Lawrence[57], Daniel Chen[106], Danny Manor[71], Darci J. Trader[101], David A. Hildeman[52], David H. Drewry[107], David J. Dowling[108], David J. Hosfield[86], David M. Smith[109], David Moreira[110], David P. Siderovski[111], David Shum[112], David T. Krist[113], David W. H. Riches[78], Davide Maria Ferraris[114], Deborah H. Anderson[115], Deirdre R. Coombe[116], Derek S. Welsbie[35], Di Hu[71], Diana Ortiz[117], Dina Alramadhani[118], Dingqiang Zhang[119], Dipayan Chaudhuri[82], Dirk J. Slotboom[39], Donald R. Ronning[120], Donghan Lee[121], Dorian Dirksen[122], Douglas A. Shoue[123], Douglas William Zochodne[124], Durga Krishnamurthy[125], Dustin Duncan[126], Dylan M. Glubb[92], Edoardo Luigi Maria Gelardi[127], Edward C. Hsiao[128], Edward G. Lynn[129], Elany Barbosa Silva[130], Elena Aguilera[131], Elena Lenci[50], Elena Theres Abraham[132], Eleonora Lama[62], Eleonora Mameli[45], Elisa Leung[126], Emily M. Christensen[133], Emily R. Mason[134], Enrico Petretto[70], Ephraim F. Trakhtenberg[135], Eric J. Rubin[18], Erick Strauss[136], Erik W. Thompson[25], Erika Cione[137], Erika Mathes Lisabeth[138], Erkang Fan[139], Erna Geessien Kroon[76], Eunji Jo[112], Eva M. García-Cuesta[103], Evgenia Glukhov[35], Evripidis Gavathiotis[21], Fang Yu[140], Fei Xiang[141], Fenfei Leng[142], Feng Wang[143], Filippo Ingoglia[82], Focco van den Akker[71], Francesco Borriello[144], Franco J. Vizeacoumar[145], Frank Luh[146], Frederick S. Buckner[139], Frederick S. Vizeacoumar[53], Fredj Ben Bdira[147], Fredrik Svensson[73], G. Marcela Rodriguez[148], Gabriella Bognár[81], Gaia Lembo[149], Gang Zhang[150], Garrett Dempsey[51], Gary Eitzen[151], Gaétan Mayer[152], Geoffrey L. Greene[86], George A. Garcia[57], Gergely L. Lukacs[153], Gergely Prikler[81], Gian Carlo G. Parico[93], Gianni Colotti[47], Gilles De Keulenaer[154], Gino Cortopassi[37], Giovanni Roti[60], Giulia Girolimetti[62], Giuseppe Fiermonte[155], Giuseppe Gasparre[156], Giuseppe Leuzzi[19], Gopal Dahal[157], Gracjan Michlewski[158,159], Graeme L. Conn[160], Grant David Stuchbury[85], Gregory R. Bowman[161], Grzegorz Maria Popowicz[162], Guido Veit[153], Guilherme Eduardo de Souza[20], Gustav Akk[163], Guy Caljon[43], Guzmán Alvarez[164], Gwennan Rucinski[165], Gyeongeun Lee[112], Gökhan Cildir[166], Hai Li[27], Hairol E. Breton[167], Hamed Jafar-Nejad[168], Han Zhou[169], Hannah P. Moore[170], Hannah Tilford[165], Haynes Yuan[171], Heesung Shim[37], Heike Wulff[37], Heinrich Hoppe[75], Helena Chaytow[45], Heng-Keat Tam[172], Holly Van Remmen[173], Hongyang Xu[174], Hosana Maria Debonsi[175], Howard B. Lieberman[27], Hoyoung Jung[176], Hua-Ying Fan[177], Hui Feng[55], Hui Zhou[19], Hyeong Jun Kim[178], Iain R. Greig[179], Ileana Caliandro[180], Ileana Corvo[181], Imanol Arozarena[182], Imran N. Mungrue[183], Ingrid M. Verhamme[184], Insaf Ahmed Qureshi[185], Irina Lotsaris[186], Isin Cakir[57], J. Jefferson P. Perry[195], Jacek Kwiatkowski[85], Jacob Boorman[71], Jacob Ferreira[188], Jacob Fries[189], Jadel Müller Kratz[79], Jaden Miner[82], Jair L. Siqueira-Neto[35], James G. Granneman[190], James Ng[165], James Shorter[161], Jan Hendrik Voss[99], Jan M. Gebauer[132], Janelle Chuah[109], Jarrod J. Mousa[191], Jason T. Maynes[192], Jay D. Evans[193], Jeffrey Dickhout[194], Jeffrey P. MacKeigan[138], Jennifer N. Jossart[195], Jia Zhou[33], Jiabei Lin[161], Jiake Xu[196], Jianghai Wang[146], Jiaqi Zhu[197], Jiayu Liao[195], Jingyi Xu[195], Jinshi Zhao[198], Jiusheng Lin[199], Jiyoun Lee[200], Joana Reis[48], Joerg Stetefeld[77], John B. Bruning[201], John Burt Bruning[80], John G. Coles[202], John J. Tanner[167], John M. Pascal[29], Jonathan So[59], Jordan L. Pederick[80], Jose A. Costoya[110], Joseph B. Rayman[19], Joseph J. Maciag[52], Joshua Alexander Nasburg[37], Joshua J. Gruber[203], Joshua M. Finkelstein[55], Joshua Watkins[165], José Miguel Rodríguez-Frade[204], Juan Antonio Sanchez Arias[205], Juan José Lasarte[206], Julen Oyarzabal[205], Julian Milosavljevic[88], Julie Cools[154], Julien Lescar[22], Julijus Bogomolovas[35], Jun Wang[148], Jung-Min Kee[176], Jung-Min Kee[178], Junzhuo Liao[207], Jyothi C. Sistla[118], Jônatas Santos Abrahão[76], Kamakshi Sishtla[208], Karol R. Francisco[35], Kasper B. Hansen[209], Kathleen A. Molyneaux[71], Kathryn A. Cunningham[33], Katie R. Martin[138], Kavita Gadar[210], Kayode K. Ojo[139], Keith S. Wong[126], Kelly L. Wentworth[128], Kent Lai[82], Kevin A. Lobb[75], Kevin M. Hopkins[27], Keykavous Parang[211], Khaled Machaca[212], Kien Pham[98], Kim Ghilarducci[213], Kim S. Sugamori[126], Kirk James McManus[77], Kirsikka Musta[64], Kiterie M. E. Faller[45], Kiyo Nagamori[96], Konrad J. Mostert[136], Konstantin V. Korotkov[94], Koting Liu[214], Kristiana S. Smith[215], Kristopher Sarosiek[216], Kyle H. Rohde[217], Kyu Kwang Kim[218], Kyung Hyeon Lee[219], Lajos Pusztai[98], Lari Lehtiö[220], Larisa M. Haupt[25], Leah E. Cowen[126], Lee J. Byrne[102], Leila Su[146], Leon Wert-Lamas[89], Leonor Puchades-Carrasco[221], Lifeng Chen[86], Linda H. Malkas[187], Ling Zhuo[222],

Lizbeth Hedstrom[223], Lizbeth Hedstrom[223], Loren D. Walensky[59], Lorenzo Antonelli[63], Luisa Iommarini[62], Luke Whitesell[126], Lía M. Randall[224], M. Dahmani Fathallah[225], Maira Harume Nagai[198], Mairi Louise Kilkenny[226], Manu Ben-Johny[19], Marc P. Lussier[213], Marc P. Windisch[112], Marco Lolicato[48], Marco Lucio Lolli[180], Margot Vleminckx[43], Maria Cristina Caroleo[227], Maria J. Macias[90], Marilia Valli[20], Marim M. Barghash[126], Mario Mellado[204], Mark A. Tye[228], Mark A. Wilson[199], Mark Hannink[229], Mark R. Ashton[85], Mark Vincent C.dela Cerna[121], Marta Giorgis[179], Martin K. Safo[118], Martin St. Maurice[230], Mary Ann McDowell[123], Marzia Pasquali[82], Masfique Mehedi[231], Mateus Sá Magalhães Serafim[76], Matthew B. Soellner[57], Matthew G. Alteen[232], Matthew M. Champion[123], Maxim Skorodinsky[233], Megan L. O'Mara[234], Mel Bedi[40], Menico Rizzi[114], Michael Levin[119], Michael Mowat[235], Michael R. Jackson[236], Mikell Paige[219], Minnatallah Al-Yozbaki[102], Miriam A. Giardini[130], Mirko M. Maksimainen[220], Monica De Luise[62], Muhammad Saddam Hussain[208], Myron Christodoulides[165], Natalia Stec[158], Natalia Zelinskaya[160], Natascha Van Pelt[43], Nathan M. Merrill[57], Nathanael Singh[105], Neeltje A. Kootstra[237], Neeraj Singh[238], Neha S. Gandhi[25], Nei-Li Chan[214], Nguyen Mai Trinh[22], Nicholas O. Schneider[230], Nick Matovic[85], Nicola Horstmann[239], Nicola Longo[82], Nikhil Bharambe[22], Nirvan Rouzbeh[209], Niusha Mahmoodi[21], Njabulo Joyfull Gumede[240], Noelle C. Anastasio[33], Noureddine Ben Khalaf[225], Obdulia Rabal[205], Olga Kandror[216], Olivier Escaffre[33], Olli Silvennoinen[64], Ozlem Tastan Bishop[75], Pablo Iglesias[110], Pablo Sobrado[241], Patrick Chuong[242], Patrick O'Connell[138], Pau Martin-Malpartida[90], Paul Mellor[53], Paul V. Fish[73], Paulo Otávio Lourenço Moreira[30], Pei Zhou[198], Pengda Liu[107], Pengda Liu[107], Pengpeng Wu[243], Percy Agogo-Mawuli[111], Peter L. Jones[244], Peter Ngoi[93], Peter Toogood[57], Philbert Ip[126], Philipp von Hundelshausen[100], Pil H. Lee[57], Rachael B. Rowswell-Turner[218], Rafael Balaña-Fouce[245], Rafael Eduardo Oliveira Rocha[76], Rafael V. C. Guido[20], Rafaela Salgado Ferreira[76], Rajendra K. Agrawal[58], Rajesh K. Harijan[21], Rajesh Ramachandran[246], Rajkumar Verma[247], Rakesh K. Singh[248], Rakesh Kumar Tiwari[249], Ralph Mazitschek[228], Rama K. Koppisetti[167], Remus T. Dame[147], Renée N. Douville[250], Richard C. Austin[194], Richard E. Taylor[123], Richard G. Moore[218], Richard H. Ebright[148], Richard M. Angell[73], Riqiang Yan[238], Rishabh Kejriwal[65], Robert A. Batey[126], Robert Blelloch[128], Robert J. Vandenberg[186], Robert J. Hickey[187], Robert J. Kelm Jr.[49], Robert J. Lake[177], Robert K. Bradley[251], Robert M. Blumenthal[106], Roberto Solano[46], Robin Matthias Gierse[252], Ronald E. Viola[157], Ronan R. McCarthy[210], Rosa Maria Reguera[245], Ruben Vazquez Uribe[253], Rubens Lima do Monte-Neto[30], Ruggiero Gorgoglione[155], Ryan T. Cullinane[223], Sachin Katyal[171], Sakib Hossain[105], Sameer Phadke[57], Samuel A. Shelburne[239], Sandra E. Geden[217], Sandra Johannsen[61], Sarah Wazir[220], Scott Legare[77], Scott M. Landfear[117], Senthil K. Radhakrishnan[118], Serena Ammendola[44], Sergei Dzhumaev[254], Seung-Yong Seo[141], Shan Li[143], Shan Zhou[168], Shaoyou Chu[134], Shefali Chauhan[255], Shinsaku Maruta[256,257], Shireen R. Ashkar[57], Show-Ling Shyng[117], Silvestro G. Conticello[149,257], Silvia Buroni[48], Silvia Garavaglia[114], Simon J. White[65], Siran Zhu[158,159], Sofiya Tsimbalyuk[258], Somaia Haque Chadni[142], Soo Young Byun[112], Soonju Park[112], Sophia Q. Xu[259], Sourav Banerjee[260], Stefan Zahler[222], Stefano Espinoza[91], Stefano Gustincich[91], Stefano Sainas[180], Stephanie L. Celano[138], Stephen J. Capuzzi[107], Stephen N. Waggoner[261], Steve Poirier[262], Steven H. Olson[236], Steven O. Marx[263], Steven R. Van Doren[167], Suryakala Sarilla[184], Susann M. Brady-Kalnay[71], Sydney Dallman[231], Syeda Maryam Azeem[105], Tadahisa Teramoto[264], Tamar Mehlman[105], Tarryn Swart[75], Tatjana Abaffy[265], Tatos Akopian[216], Teemu Haikarainen[64], Teresa Lozano Moreda[266], Tetsuro Ikegami[33], Thaiz Rodrigues Teixeira[175], Thilina D. Jayasinghe[120], Thomas H. Gillingwater[45], Thomas Kampourakis[267], Timothy I. Richardson[208], Timothy J. Herdendorf[84], Timothy J. Kotzé[136], Timothy R. O'Meara[268], Timothy W. Corson[208], Tobias Hermle[88], Tomisin Happy Ogunwa[256], Tong Lan[86], Tong Su[229], Toshihiro Banjo[269], Tracy A. O'Mara[92], Tristan Chou[42], Tsui-Fen Chou[143], Ulrich Baumann[132], Umesh R. Desai[118], Vaibhav P. Pai[119], Van Chi Thai[38], Vasudha Tandon[260], Versha Banerji[77], Victoria L. Robinson[65], Vignesh Gunasekharan[169], Vigneshwaran Namasivayam[99], Vincent F. M. Segers[43], Vincent Maranda[53], Vincenza Dolce[137], Vinícius Gonçalves Maltarollo[76], Viola Camilla Scoffone[48], Virgil A. Woods[105], Virginia Paola Ronchi[270], Vuong Van Hung Le[271], W. Brent Clayton[101], W. Todd Lowther[272], Walid A. Houry[126], Wei Li[273], Weiping Tang[207],

Wenjun Zhang[51], Wesley C. Van Voorhis[139], William A. Donaldson[230], William C. Hahn[59], William G. Kerr[274], William H. Gerwick[130], William J. Bradshaw[275], Wuen Ee Foong[276], Xavier Blanchet[277], Xiaoyang Wu[86], Xin Lu[123], Xin Qi[246], Xin Xu[84], Xinfang Yu[168], Xingping Qin[278], Xingyou Wang[223], Xinrui Yuan[95], Xu Zhang[279], Yan Jessie Zhang[83], Yanmei Hu[148], Yasser Ali Aldhamen[138], Yicheng Chen[71], Yihe Li[71], Ying Sun[52], Yini Zhu[123], Yogesh K. Gupta[280], Yolanda Pérez-Pertejo[245], Yong Li[168], Young Tang[65], Yuan He[40], Yuk-Ching Tse-Dinh[142], Yulia A. Sidorova[281], Yun Yen[146], Yunlong Li[282], Zachary J. Frangos[283], Zara Chung[22], Zhengchen Su[33], Zhenghe Wang[71], Zhiguo Zhang[27], Zhongle Liu[126], Zintis Inde[216], Zoraima Artía[164] & Abraham Heifets[2]

[2]Atomwise Inc., San Fransico, USA. [3]Amgen, Thousand Oaks, USA. [4]OpenAI, San Francisco, USA. [5]Model Medicines, La Jolla, USA. [6]Atomic.AI, San Francisco, USA. [7]Edifice Health, Inc., San Mateo, USA. [8]METiS Therapeutics, Cambridge, USA. [9]Genentech, San Mateo, USA. [10]US Navy Medical Service Corps Officer (2300/1810D), San Mateo, USA. [11]Totus Medicines, Inc., Emeryville, USA. [12]Cytokinetics, Inc., South San Francisco, USA. [13]Nurix Therapeutics, San Francisco, USA. [14]Amazon Alexa, Suite, USA. [15]The University of North Carolina at Chapel Hill Eshelman School of Pharmacy, Chapel Hill, USA. [16]Refibered Inc., Cupertino, USA. [17]NVIDIA, Santa Clara, USA. [18]Harvard TH Chan School of Public Health, Boston, USA. [19]Columbia University, New York, USA. [20]University of São Paulo, São Paulo, Brazil. [21]Albert Einstein College of Medicine, Bronx, USA. [22]Nanyang Technological University, Singapore, Singapore. [23]University of Washington, Seattle, USA. [24]Chan Medical School, University of Massachusetts, Worcester, USA. [25]Queensland University of Technology, Brisbane, USA. [26]University of Michigan Medical School, Ann Arbor, USA. [27]Columbia University Irving Medical Center, New York, USA. [28]Universidade Federal de Santa Catarina, Florianópolis, Brazil. [29]Université de Montréal, Montreal, Canada. [30]Instituto René Rachou-Fundação Oswaldo Cruz/Fiocruz Minas, Belo Horizonte, Brazil. [31]Herbert Wertheim College of Medicine, Biomolecular Science Institute, Florida International University, Miami, USA. [32]NYU Langone Health, New York, USA. [33]The University of Texas Medical Branch at Galveston, Galveston, USA. [34]University of Houston, Galveston, USA. [35]University of California, San Diego, USA. [36]School of Medicine, Tufts University, Medford, USA. [37]University of California, Davis, Davis, USA. [38]University of Western Australia, Crawley, Australia. [39]University of Groningen, Groningen, The Netherlands. [40]Wayne State University, Detroit, USA. [41]University of Zurich, Zürich, Switzerland. [42]Stanford University, Stanford, USA. [43]University of Antwerp, Antwerp, Belgium. [44]University of Rome Tor Vergata, Rome, Italy. [45]University of Edinburgh, Edinburgh, UK. [46]Department of Plant Molecular Genetics, Centro Nacional de Biotecnología, Consejo Superior de Investigaciones Científicas (CNB-CSIC), Madrid, Spain. [47]CNR (Italian National Research Council), Rome, Italy. [48]University of Pavia, Pavia, Italy. [49]University of Vermont, Burlington, USA. [50]University of Florence, Florence, Italy. [51]University of California, Berkeley, Berkeley, USA. [52]Cincinnati Children's Hospital Medical Center, Cincinnati, USA. [53]University of Saskatchewan, Saskatoon, Canada. [54]Broad Institute of MIT and Harvard, Cambridge, USA. [55]Boston University, Boston, USA. [56]CancerCare Manitoba Research Institute, Winnipeg, Canada. [57]University of Michigan, Ann Arbor, USA. [58]Wadsworth Center, New York State Department of Health and University at Albany, Albany, USA. [59]Dana-Farber Cancer Institute, Boston, USA. [60]University of Parma, Parma, Italy. [61]Helmholtz Institute for Pharmaceutical Research Saarland, Saarbrücken, Germany. [62]University of Bologna, Bologna, Italy. [63]Sapienza University of Rome, Rome, Italy. [64]Tampere University, Tampere, Finland. [65]University of Connecticut, Storrs, USA. [66]Centro de Investigación Médica Aplicada, Universidad de Navarra, Pamplona, Spain. [67]Tulane National Primate Research Center, Tulane University, Covington, USA. [68]Scripps Research, San Diego, USA. [69]Louisiana State University School of Medicine, New Orleans, USA. [70]Duke-NUS Medical School, Singapore, Singapore. [71]Case Western Reserve University, Cleveland, USA. [72]Oregon Health and Science University and Tanta University in Tanta, Tanta, Egypt. [73]University College London, London, UK. [74]Saint Louis University, St. Louis, USA. [75]Rhodes University, Makhanda, South Africa. [76]Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Brazil. [77]University of Manitoba, Winnipeg, Canada. [78]National Jewish Health, Denver, USA. [79]Drugs for Neglected Diseases Initiative (DNDi), Geneva, Switzerland. [80]The University of Adelaide, Adelaide, Australia. [81]Mcule, Budapest, Hungary. [82]University of Utah, Salt Lake City, USA. [83]The University of Texas at Austin, Austin, USA. [84]Kansas State University, Manhattan, USA. [85]UniQuest Pty Ltd, St Lucia, Australia. [86]University of Chicago, Chicago, USA. [87]Harvard University, Cambridge, USA. [88]University of Freiburg, Freiburg Im Breisgau, Germany. [89]Dana-Farber Cancer Institute and Harvard Medical School, Boston, USA. [90]IRB Barcelona, Barcelona, Spain. [91]Istituto Italiano Di Tecnologia, Genoa, Italy. [92]QIMR Berghofer Medical Research Institute, Herston, Australia. [93]University of California, Santa Cruz, Santa Cruz, USA. [94]University of Kentucky, Lexington, USA. [95]University of Tennessee Health Science Center, Memphis, USA. [96]Children's Cancer Therapy Development Institute, Beaverton, USA. [97]Columbia University Medical Center, New York, USA. [98]Yale School of Medicine, New Haven, USA. [99]University of Bonn, Bonn, Germany. [100]Ludwig-Maximilians-Universität München, Munich, Germany. [101]Purdue University, West Lafayette, USA. [102]Canterbury Christ Church University, Canterbury, UK. [103]National Centre for Biotechnology (CNB-CSIC), Madrid, Spain. [104]University of South Australia and SA Pathology, Adelaide, Australia. [105]CUNY Advanced Science Research Center, New York, USA. [106]The University of Toledo, Toledo, USA. [107]University of North Carolina at Chapel Hill, Chapel Hill, USA. [108]Boston Children's Hospital and Harvard Medical School, Boston, USA. [109]West Virginia University, Morgantown, USA. [110]Universidade de Santiago de Compostela, Santiago, Spain. [111]University of North Texas Health Science Center at Fort Worth, Fort Worth, USA. [112]Institut Pasteur Korea, Seongnam, South Korea. [113]Carle Illinois College of Medicine, Urbana, USA. [114]Università del Piemonte Orientale, Vercelli, Italy. [115]Saskatchewan Cancer Agency, Saskatoon, Canada. [116]Curtin University, Bentley, Australia. [117]Oregon Health and Science University, Portland, USA. [118]Virginia Commonwealth University, Richmond, USA. [119]Tufts University, Medford, USA. [120]University of Nebraska Medical Center, Omaha, USA. [121]University of

Louisville, Louisville, USA. [122]Dana Farber Cancer Institute, Boston, USA. [123]University of Notre Dame, Notre Dame, USA. [124]University of Alberta, Edmonton, Canada. [125]Cincinnati Childrens Hospital Medical Center, Cincinnati, USA. [126]University of Toronto, Toronto, Canada. [127]University of Piemonte Orientale, Vercelli, Italy. [128]University of California, San Francisco, San Francisco, USA. [129]St. Joseph's Healthcare Hamilton, and Hamilton Center for Kidney Research, McMaster University, Hamilton, Canada. [130]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, San Diego, USA. [131]Universidad de La República, Montevideo, Uruguay. [132]University of Cologne, Cologne, Germany. [133]Johnson University, Knoxville, USA. [134]Indiana University, Bloomington, USA. [135]School of Medicine, University of Connecticut, Farmington, USA. [136]Stellenbosch University, Stellenbosch, South Africa. [137]University of Calabria, Arcavacata, Italy. [138]Michigan State University, East Lansing, USA. [139]University of Washington, Washington, USA. [140]Weill Cornell Medicine-Qatar, Ar-Rayyan, Qatar. [141]Gachon University, Seongnam, South Korea. [142]Florida International University, Miami, USA. [143]California Institute of Technology, Pasadena, USA. [144]Boston Children's Hospital, Boston, USA. [145]Saskatchewan Cancer Agency and University of Saskatchewan, Saskatchewan, Canada. [146]Sino-American Cancer Foundation, Covina, USA. [147]Leiden University, Leiden, The Netherlands. [148]Rutgers University, Newark, USA. [149]Core Research Laboratory, ISPRO, Florence, Italy. [150]Caltech, Pasadena, USA. [151]University of Alberta, Edmonton, USA. [152]Montreal Heart Institute and Université de Montréal, Montreal, Canada. [153]McGill University, Montreal, Canada. [154]Antwerp University, Antwerp, Belgium. [155]University of Bari Aldo Moro, Bari, Italy. [156]Alma Mater Studiorum-University of Bologna, Bologna, Italy. [157]University of Toledo, Toledo, USA. [158]International Institute of Molecular and Cell Biology in Warsaw, Warsaw, Poland. [159]Infection Medicine, University of Edinburgh The Chancellor's Building, Edinburgh, UK. [160]Emory University, Atlanta, USA. [161]University of Pennsylvania, Philadelphia, USA. [162]Helmholtz Zentrum München, Munich, Germany. [163]Washington University School of Medicine, St. Louis, USA. [164]CENUR Litoral Norte, Universidad de La República, Montevideo, Uruguay. [165]University of Southampton, Southampton, UK. [166]Centre for Cancer Biology, University of South Australia, Adelaide, Australia. [167]University of Missouri, Columbia, USA. [168]Baylor College of Medicine, Houston, USA. [169]Yale University, New Haven, USA. [170]Reno School of Medicine, University of Nevada, Reno, USA. [171]University of Manitoba and CancerCare Manitoba, Winnipeg, Canada. [172]Goethe University Frankfurt, Frankfurt, Germany. [173]Oklahoma Medical Research Foundation/Oklahoma City VA Medical Center, Oklahoma City, USA. [174]Oklahoma Medical Research Foundation, Oklahoma City, USA. [175]Department of Biomolecular Sciences, School of Pharmaceutical Sciences of Ribeirão Preto, University of São Paulo, Ribeirão Preto, SP, Brazil. [176]Ulsan National Institute of Science and Technology, Ulsan, South Korea. [177]University of New Mexico Comprehensive Cancer Center, Albuquerque, USA. [178]Ulsan National Institute of Science and Technology (UNIST), Ulsan, South Korea. [179]University of Aberdeen, Aberdeen, UK. [180]University of Turin, Turin, Italy. [181]Universidad de La República, CenUR LN, Montevideo, Uruguay. [182]Navarrabiomed-IdiSNA, Pamplona, Spain. [183]Independent, Los Angeles, USA. [184]Vanderbilt University Medical Center, Nashville, USA. [185]University of Hyderabad, Hyderabad, India. [186]University of Sydney, Sydney, Australia. [187]City of Hope Medical Center, Duarte, USA. [188]Weill Cornell Medicine, New York, NY 10065, USA. [189]University of Toledo College of Medicine and Life Sciences, Toledo, USA. [190]School of Medicine, Wayne State University, Detroit, USA. [191]University of Georgia, Athens, USA. [192]The Hospital for Sick Children, Toronto, Canada. [193]United States Department of Agriculture, Agricultural Research Service (USDA-ARS), Washington, DC, USA. [194]McMaster University, Hamilton, Canada. [195]University of California, Riverside, Riverside, USA. [196]The University of Western Australia, Perth, Australia. [197]The University of Connecticut, Storrs, USA. [198]Duke University School of Medicine, Durham, USA. [199]University of Nebraska-Lincoln, Lincoln, USA. [200]Sungshin University, Seoul, South Korea. [201]University of Adelaide, Adelaide, Australia. [202]University Toronto, Toronto, Canada. [203]University of Texas Southwestern Medical Center, Dallas, USA. [204]Centro Nacional de Biotecnologia/CSIC, Madrid, Spain. [205]Centro de Investigación Médica Aplicada, Pamplona, Spain. [206]Centro de Investigación Médica Aplicada, Universidad de Navarra, Pamplona, Spain. [207]University of Wisconsin-Madison, Madison, USA. [208]Indiana University School of Medicine, Indianapolis, USA. [209]University of Montana, Missoula, USA. [210]Brunel University London, London, UK. [211]Chapman University, Orange, USA. [212]Weill Cornell Medicine Qatar, Ar-Rayyan, Qatar. [213]Université du Québec À Montréal, Montreal, Canada. [214]National Taiwan University, Taipei, Taiwan. [215]Rhodes College, Memphis, USA. [216]Harvard School of Public Health, Boston, USA. [217]University of Central Florida, Orlando, USA. [218]University of Rochester, Rochester, USA. [219]George Mason University, Fairfax, USA. [220]University of Oulu, Oulu, Finland. [221]Instituto Investigación Sanitaria La Fe, Valencia, Spain. [222]Ludwig-Maximilians-University, Munich, Germany. [223]Brandeis University, Waltham, USA. [224]Universidad de La República, CENUR Litoral Norte, Montevideo, Uruguay. [225]Arabian Gulf University, Manama, Bahrain. [226]University of Cambridge, Cambridge, UK. [227]University of Magna Graecia, Catanzaro, Italy. [228]Massachusetts General Hospital, Boston, USA. [229]University of Missouri-Columbia, Columbia, USA. [230]Marquette University, Milwaukee, USA. [231]University of North Dakota, Grand Forks, USA. [232]Simon Fraser University, Burnaby, Canada. [233]CancerCare Manitoba Research Institute (CCMR), Winnipeg, Canada. [234]The University of Queensland, Brisbane, Australia. [235]University of Manitoba and CancerCare Manitoba Research Institute, Winnipeg, Canada. [236]Sanford Burnham Prebys, La Jolla, USA. [237]University of Amsterdam, Amsterdam, The Netherlands. [238]UConn Health, Farmington, USA. [239]The University of Texas MD Anderson Cancer Center, Houston, USA. [240]Walter Sisulu University, Mthatha, South Africa. [241]Virginia Tech, Blacksburg, USA. [242]University of Houston, Houston, USA. [243]Rutgers University, New Brunswick, USA. [244]University of Nevada, Reno, USA. [245]Universidad de León, León, Spain. [246]School of Medicine, Case Western Reserve University, Cleveland, USA. [247]School of Medicine, UConn Health, Farmington, USA. [248]University of Rochester Medical Center, Rochester, USA. [249]Chapman University School of Pharmacy, Irvine, USA. [250]University of Winnipeg/St. Boniface Research Centre, Winnipeg, Canada. [251]Fred Hutchinson Cancer Center, Seattle, USA. [252]Helmholtz Institute for Pharmaceutical Research Saarland (HIPS), Saarbrücken, Germany. [253]Technical University of Denmark, Kongens Lyngby, Denmark. [254]The City College of New York, New York, USA. [255]Children's Cancer, Therapy Development Institute (Cc-TDI), Beaverton, USA. [256]Soka University, Hachioji, Japan. [257]Institute of Clinical Physiology, National Research Council, Pisa, Italy. [258]Charles Sturt University, Bathurst,

Australia. [259]Washington University, St Louis, USA. [260]University of Dundee, Dundee, UK. [261]Cincinnati Children's Hospital Medical Center, University of Cincinnati College of Medicine, Cincinnati, USA. [262]Montreal Heart Institute, Montreal, Canada. [263]Columbia University Vagelos College of Physicians and Surgeons, Columbia, USA. [264]Georgetown University, Washington, USA. [265]Duke University, Durham, USA. [266]Center for Applied Medical Research, University of Navarra, Pamplona, Spain. [267]King's College London, London, UK. [268]Precision Vaccines Program, Division of Infectious Diseases, Boston Children's Hospital, Boston, USA. [269]Fred Hutchinson Cancer Research Center, Seattle, USA. [270]Louisiana State University, Baton Rouge, USA. [271]Massey University, Palmerston North, New Zealand. [272]Wake Forest University School of Medicine, Winston-Salem, USA. [273]Central South University, Changsha, China. [274]SUNY Upstate Medical University, Syracuse, USA. [275]University of Oxford, Oxford, UK. [276]Goethe-University, Frankfurt, Frankfurt, Germany. [277]Institute for Cardiovascular Prevention (IPEK), Ludwig-Maximilians-Universität München, Munich, Germany. [278]Harvard T.H. Chan School of Public Health, Boston, USA. [279]School of Medicine, Boston University, Boston, USA. [280]University of Texas Health Science Center at San Antonio, San Antonio, USA. [281]University of Helsinki, Helsinki, Finland. [282]Wadsworth Center, NYSDOH, Albany, USA. [283]The University of Sydney, Sydney, Australia.