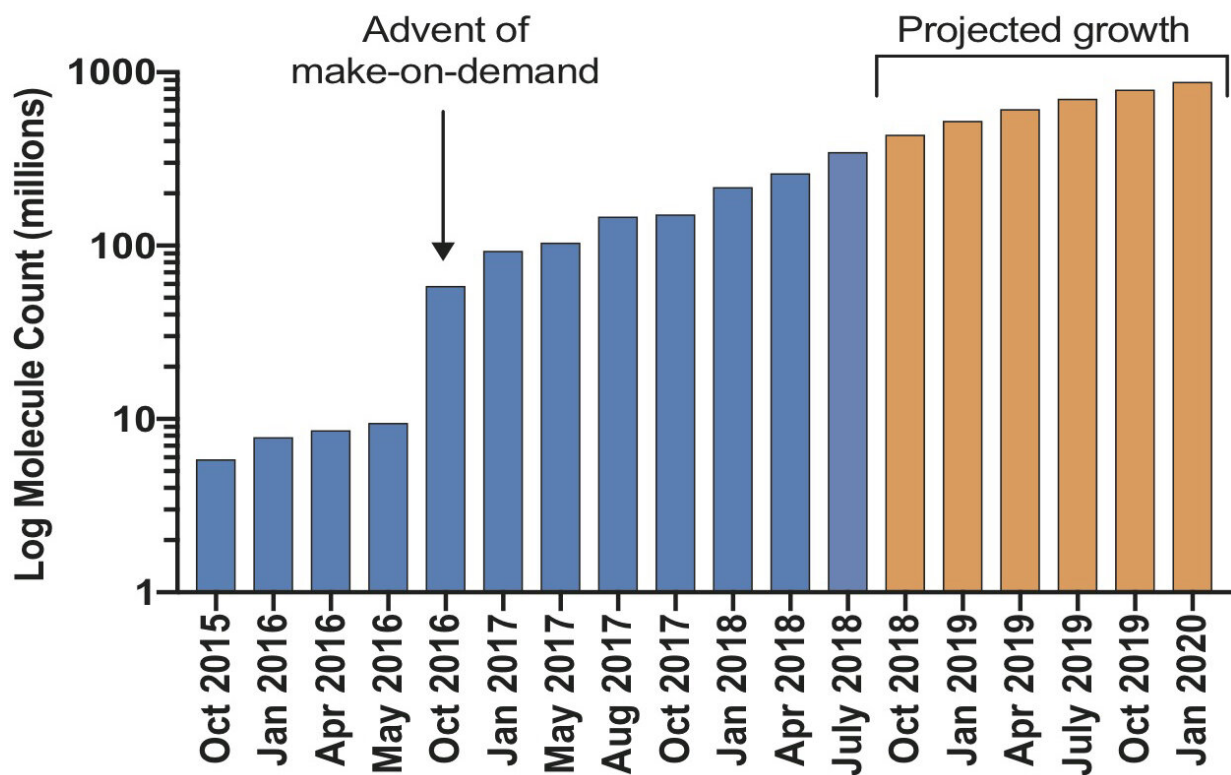# Mega docking library poised to speed drug discovery

February 6 2019

A virtual library of make-on-demand molecules available for drug discovery is expected to exceed 1 billion compounds by next year. Credit: Bryan Roth, M.D., Ph.D., of the University of North Carolina (UNC) Chapel Hill, Brian Shoichet, Ph.D., and John Irwin, Ph.D., of the University of California San Francisco, and colleagues.

Researchers have launched an ultra-large virtual docking library expected to grow to more than 1 billion molecules by next year. It will expand by 1000-fold the number of such "make-on-demand" compounds readily available to scientists for chemical biology and drug discovery. The larger the library, the better its odds of weeding out inactive "decoy" molecules that could otherwise lead researchers down blind alleys. The project is funded by the National Institutes of Health.

"To improve medications for mental illnesses, we need to screen huge numbers of potentially therapeutic molecules," explained Joshua A. Gordon, M.D., Ph.D., director of NIH's National Institute of Mental Health (NIMH), which co-funded the research. "Unbiased computational modeling allows us to do this in a computer, vastly expediting the process of discovering new treatments. It enables researchers to virtually "see" a molecule docking with its receptor protein—like a ship in its harbor berth or a key in its lock—and predict its pharmacological properties, based on how the molecular structures are predicted to interact. Only those relatively few candidate molecules that best match the target profile on the computer need to be physically made and tested in a wet lab."

Bryan Roth, M.D., Ph.D., of the University of North Carolina (UNC) Chapel Hill, Brian Shoichet, Ph.D., and John Irwin, Ph.D., of the University of California San Francisco, and colleagues, report on their findings Feb. 6, 2019 in the journal *Nature*. The study was supported, in part, by grants from NIMH, National Institute of General Medical Sciences (NIGMS), the NIH Common Fund, and National Institute of Neurological Disorders and Stroke (NINDS).

The NIH Common Fund's Illuminating the Druggable Genome (IDG) Program—launched in 2014 to catalyze research on proteins that are currently understudied and potential targets of therapeutic intervention—funded the docking library expansion.

Over the past few years, Roth, Shoichet, and colleagues have employed their virtual structure-based docking approach to uncover molecular secrets of an antipsychotic drug and LSD docked in their respective target receptors—and to create a designer painkiller that selectively targets brain analgesic circuitry without morphine's side effects.

A staggering number of potential drug-like molecules are known to exist. Yet, hundreds-of-millions to billions of diverse molecules have remained inaccessible due to limitations of existing methods used to compile molecular libraries, say the researchers. For example, their virtual structure-based docking technique, while promising, risks finding many false positives or "decoys—flaws in the model allow for molecules that appear plausible but turn out to be biologically inactive.
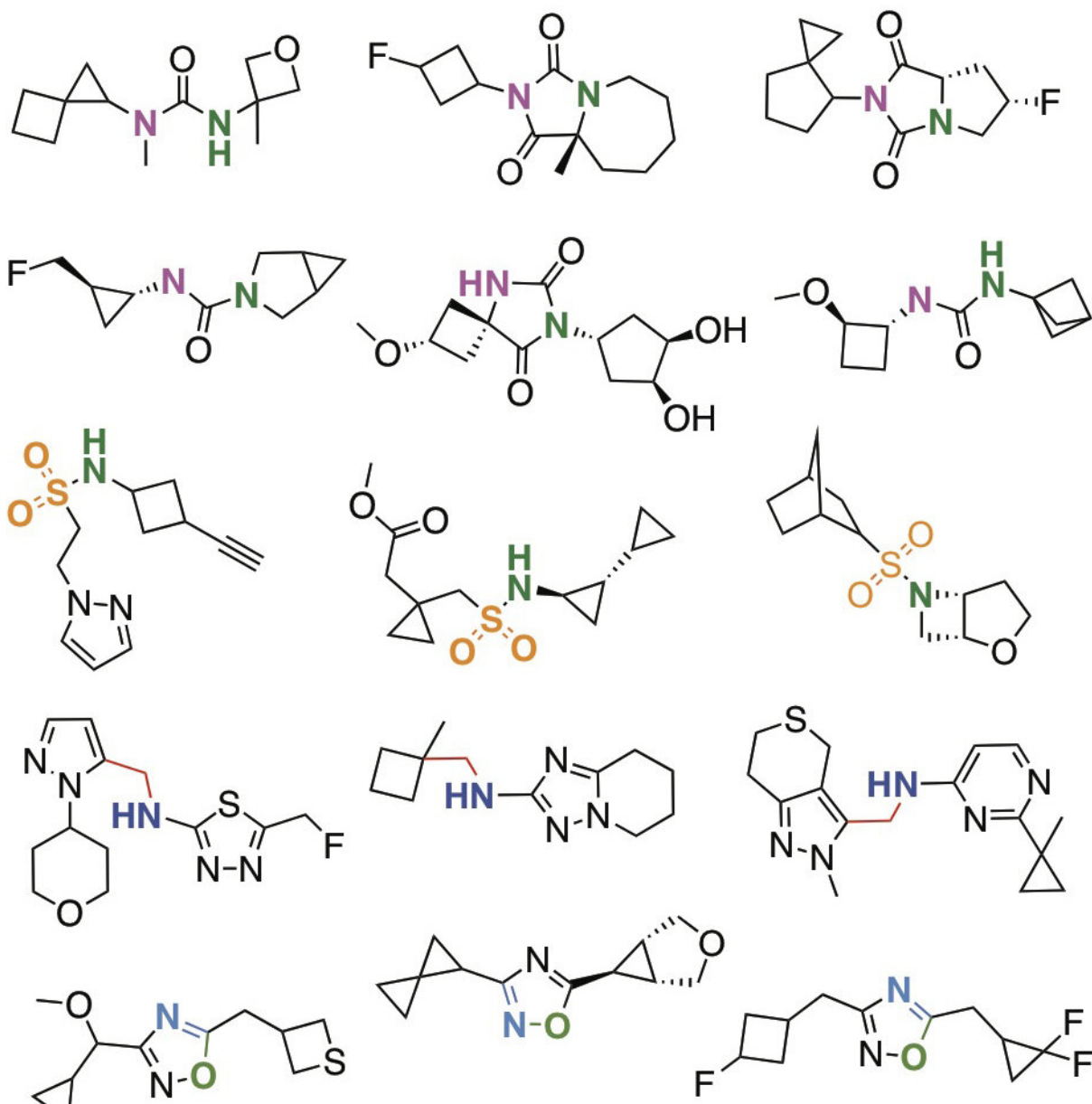
Selection of molecules discovered using the mega docking library. Credit: Bryan Roth, M.D., Ph.D., of the University of North Carolina (UNC) Chapel Hill, Brian Shoichet, Ph.D., and John Irwin, Ph.D., of the University of California San Francisco, and colleagues.

To overcome this challenge, the researchers focused on molecules that

result from 130 well-characterized chemical reactions using 70,000 different chemical building blocks. Computer simulations with these molecules showed that as the size of a library grew, the ratio of "true actives" to decoys increased—just as a study's statistical power increases with a larger sample.

In the new study, the researchers examined the structure-based docking of 138 million molecules with either the D4 receptor, a key protein that mediates the actions of the brain chemical messenger dopamine, or the enzyme AmpC, which confers resistance to certain antibiotics and has proven difficult to block.

"The D4 receptor is of particular interest to NIMH because of its role in cognition and other executive functions of the brain's prefrontal cortex that are often disturbed in mental illnesses," said Laurie Nadler, Ph.D., of the NIMH Division of Neuroscience and Basic Behavioral Science, program officer for the grant supporting the D4 receptor study.

The researchers then synthesized and tested, in a lab, the top 549 molecules that virtually docked best with the D4 receptor and 44 molecules that docked best with the enzyme. These studies revealed several novel drug-like molecules that bind only to the D4 receptor (and not the closely related D2 or D3 dopamine receptors) and turned the receptor on or off. Additionally, one molecule (4163) emerged as the most potent binder of AmpC ever. A virtual molecule's docking rank predicted its actual likelihood of binding to the D4 dopamine receptor in a lab assay.

The discovery of new and potent molecules for both targets also confirmed that ultra-large libraries contain molecules better suited to a given receptor structure than smaller libraries and that virtual docking can recognize the molecules and predict the total number of expected active compounds within a library.

"This new study illustrates the potential of unbiased computational screening and molecular docking to discover new tool molecules and potential therapeutic agents, providing a rapid and robust pathway that will lead directly to novel drug treatments for mental illnesses," added Gordon.

More information: Ultra-large library docking for discovering new chemotypes. Lyu j, Wang S, Balius TE, Singh I, Levit A, Moroz YS, O'Meara MJ, Che T, Algaa E, Tolmachova K, Tolmachev AA, Shoichet BK, Roth BL, Irwin JJ. *Nature*, Feb. 6, 2019. DOI: 10.1038/s41586-019-0917-9

Provided by National Institutes of Health