Drug Discovery

DOI: 10.1002/anie.201204625

Challenges for Academic Drug Discovery

*William L. Jorgensen**

drug discovery · high-throughput screening · lead optimization · pharmacology · virtual screening

Background

There is much discussion of opportunities and need for drug discovery in research universities to help fill the gap created by the reduction of medicinal chemistry activities at pharmaceutical and biotechnology companies. Such translational science is also being promoted by funding agencies, such as the US National Institutes of Health (NIH), to the point where translational plans are common and often necessary in grant applications. In this atmosphere, it is difficult for universities to not become more involved in drug discovery. However, the melding of the traditional focus on basic research with more-applied activities presents numerous challenges.

The path to a drug begins with identification of a molecular target, whose regulation is expected to have useful therapeutic impact (Scheme 1). Relevant biological assays, such as for enzyme inhibition or modulation of an intracellular process, are then needed to identify regulators. The focus here is on small organic molecules, as they represent the largest class of marketed drugs.[1] The next step is to find active compounds, "hits", by executing the assays on chosen compounds. The test compounds may arise in many ways. They may be found in large, general compound libraries or in more focused libraries, which have demonstrated prior success for the target class. If the target has known inhibitors or substrates, test compounds may be selected that bear similarities in structure and/or properties to these leads; this is referred to as "ligand-based design". Pharmacophore models may be generated from such structure/activity data and used to screen compound libraries. Alternatively, if a crystal structure of the biomolecular target is available, more focused "structure-based design" can be pursued to seek compounds complimentary to the binding site.[2] In the latter case, it is also possible to perform virtual screening by computationally docking libraries of commercially available compounds into the structure;[3] promising compounds are then purchased and assayed.

target → assay → hit → lead → preclinical → GLP toxicity → IND → clinical → NDA/drug

Scheme 1. Steps in drug discovery. NDA: new drug application.

The resultant hits almost invariably have modest activity, for example, at micromolar concentrations, while the average activity of marketed drugs is in the low nanomolar range.[4] Greater activity can improve target selectivity and reduce drug dosages. Evolution of the hit to the potent agent is the process of lead optimization, which requires much synthetic effort in preparation of numerous analogues of the hit. Attention must also be paid to the chemical nature of the hits and analogues to avoid reactive functionalities and potential problems with pharmacologically important properties, including solubility in water and cell permeability.[5] Suitable leads can advance to preclinical development, which includes identification of undesirable off-target activities (secondary pharmacology), such as inhibition of metabolic enzymes, neurological, adrenergic, and hormone receptors, or ion channels, evaluation of toxicity and bioavailability in limited rodent studies, and, if possible, confirmation of efficacy in an animal model of the target disease.

Success with these hurdles would permit proceeding to the much larger, fully regulated toxicity studies in two animal species, often rats and dogs. These studies are required for approval of an investigational new drug (IND) application by national regulatory agencies, including the US FDA (Food and Drug Administration). The studies often require a few kilograms of the active pharmaceutical ingredient (API), which can be prepared by contract research organizations (CROs). API synthesis and two-species testing must be performed using good laboratory practice (GLP) procedures, if the results are intended to support an application to the FDA. Award of the IND permits initiation of human clinical trials, which proceed through three phases testing safety in healthy volunteers, and safety and efficacy in small and then large groups of patients with the target disease.

Successes

Successful drug discovery is becoming a more common event in academia. A recent survey found that during 1970–2009 there were approvals of 93 new small-molecule-drug applications from research at US public-sector research institutions (PSRIs), which include universities, research hospitals, non-profit institutes, and US federal laboratories, especially the NIH.[6] Among the 93 approvals, 40 were for

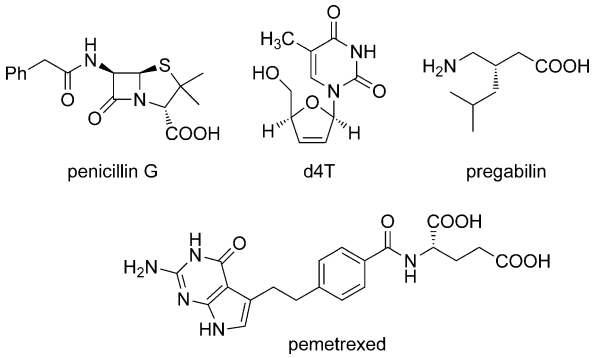
[*] Prof. W. L. Jorgensen
Department of Chemistry, Yale University
New Haven, CT 06520-8107 (USA)
E-mail: william.jorgensen@yale.edu

new chemical entities (NCEs) as opposed to new formulations, combinations, methods of treatment, or synthesis. 28 of the 40 originated at universities. For the subperiod 1990–2007, 64 of the 483 NCEs approved by the FDA originated from PSRIs. The overall rate of NCE approvals from PSRIs is accelerating and is now about three to four per year.[6,7] During the 40-year period, there were also approvals for 36 biologics and 15 vaccines from PSRIs. Today, about half of drug discovery activities is occurring at PSRIs and biotechnology companies.[7] It is also the case that the drugs arising from academia have been emphasizing treatment of cancer and infectious diseases, which is consistent with the funding patterns of the NIH.[6]

When there has been success in academia, it has often stemmed from pursuit of basic science coupled with alert observations or clever insights rather than large-scale preparation of analogues. A classic example is the isolation of penicillin (Scheme 2) by Alexander Fleming at St. Mary's Hospital Medical School in 1928. At Yale, William Prusoff began research on nucleoside analogues as antiviral drugs in the 1950s. In view of the emergence of the HIV crisis in the 1980s and the success of AZT, Prusoff and Tai-shun Lin were led to consider additional thymidine analogues as potential anti-HIV agents. This included compounds, which, like AZT, had been evaluated for possible anti-cancer activity in the 1960s, but were not found to be potent. In their studies, they identified d4T as another thymidine analogue that was effective against HIV. Yale University licensed the method-of-treatment patent to Bristol–Myers Squibb, and d4T (Zerit) became the fourth FDA-approved anti-HIV drug.

Antonín Holý from the Academy of Sciences of the Czech Republic has also been very successful in the development of nucleoside analogues as antiretroviral drugs, including the anti-HIV agent tenofovir. Tenofovir is used along with the cytidine analogue emtricitabine, which was discovered by Woo-Baeg Choi, Dennis Liotta, and Raymond Schinazi at Emory University, in the important combination therapies Truvada and Atripla.

The anti-convulsant drug pregabalin (Lyrica; Scheme 2) provides a fascinating case. Its discovery by Richard Silverman at Northwestern University involved significant serendipity and perseverance.[8] Pursuit of the original biomolecular target led to compounds that were effective in mice; however, it was several years before the true mechanism of action was elucidated. There are also examples that involved a more systematic medicinal chemistry approach, for example, the discovery of the anti-cancer agent pemetrexed (Alimta; Scheme 2) by E. C. Taylor at Princeton in collaboration with Eli Lilly.

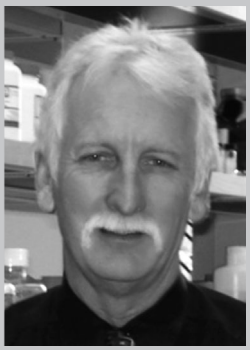


Scheme 2. Some drugs stemming from academic research.

The potential rewards are, of course, great through licensing fees that can amount to hundreds of millions of dollars. The big payoffs are normally for licensing composition-of-matter patents, which cover the specific compounds that are to become drugs. Some obstacles can be addressed at each stage of the academic drug discovery process. They are presented herein for awareness to new adventurers and to consider solutions.

Target Selection

The investigations of physiological processes, diseases, their molecular and cellular mechanisms, identification of biomolecular targets, and the development of appropriate assays are in the realm of biologists and biochemists. Academia excels here, owing to the hundreds of faculty members engaged in such studies at most large US universities, especially those with medical schools. A critical issue is selection of an appropriate target such that its modulation should provide an unambiguous, therapeutically significant response that is not prone to rapid circumvention, that is, induction of resistance. Novelty and druggability of the target are important. Selection of targets that could lead to first-in-class drugs is highly desirable. For example, academic work on kinase inhibitors that bind to the ATP site may now be risky in view of the large volume of prior work, competition from pharmaceutical companies, and the extensive patent coverage that needs to be navigated. Work on classes of targets that have proved challenging, owing to off-target activities, toxicities, or difficulties in discovering potent, selective inhibitors should also be approached with caution. Targets with relatively polar and/or flat binding sites can also be problematic. Just because it is one's favorite biology and potentially disease-related does not mean that it is viable for productive modulation with drugs.



William L. Jorgensen was born in New York City; his father was Danish and his mother was American. After graduating from Princeton and Harvard, he spent 15 years on the faculty at Purdue University. In 1990, he moved to Yale University where he is Sterling Professor of Chemistry. The research interests in his group span from quantum and statistical mechanics to organic synthesis and computer-aided drug discovery. The group's advances in understanding molecular recognition evolved into modeling protein–ligand binding and inhibitor design. Current activities emphasize de novo drug design and synthesis for anti-HIV, anti-cancer, and anti-inflammatory agents. Jorgensen is a member of the International Academy of Quantum Molecular Science, American Academy of Arts and Sciences, and US National Academy of Sciences.

High-Throughput and Virtual Screening

A central problem for high-throughput screening (HTS) in universities is that it may not exist or, if it does, the collections of screening compounds by industrial standards are very small. A university screening center may be lucky to have one or two hundred thousand compounds in its library versus several million at pharmaceutical companies. The university's compounds have also likely been obtained from commercial vendors, such as Maybridge and ChemBridge, so the diversity and proprietary nature of the compounds are limited. Drugs can be discovered without HTS. However, if HTS is to be used, then substantial investment in the screening libraries, associated automation equipment, and their management is unavoidable. Universities need to evolve their own proprietary libraries from their drug-discovery endeavors. This requires adequate incentives consistent with the added costs to encourage research groups to deposit their compounds. Collaboration between universities on HTS could be considered to avoid redundancy for both compound collections and equipment; consortia might be possible as in the case of high-field NMR spectroscopy and electron microscopy. The NIH Molecular Libraries Program has been an attempt in this direction; its strength seems to have been the implemention of assays in an HTS format. In any event, intellectual property (IP) and confidentiality issues become more complex in the consortia mode. In any mode, IP and credit issues need to be addressed in advance for cases where a library compound from Professor X emerges as a key hit in a screen for faculty member Y.

A related issue is data management. Universities need to utilize compound registration systems as in all pharmaceutical companies to record the activity results and any property data for all compounds that are purchased or synthesized. This task has largely remained at the whim of the individual research groups with the structure and activity data often loosely organized in spreadsheets. However, there is much value added for proper organization of the structure/activity data. It can provide a basis for data mining and similarity searching, a proprietary library for virtual screening, support for IP claims, and a means to avoid redundant syntheses. Owing to confidentiality issues, which are harder to control in universities, research groups may not be able to register compounds before the IP has been protected.

Assuming hits are obtained, validation by synthesis and repetition of the assays is needed. It is also important to avoid assay "frequent hitters"; these are compounds that commonly have reactive functional groups, which can covalently modify assay components, or that form aggregates owing to poor solubility.[9] Given multiple, diverse, attractive hits, decisions are needed on the chemical series that will be subjected to lead optimization. Novelty of the core structure is desirable and requires knowledge of the journal and patent literature. The issue is less significant for a first-in-class target, since the literature should be less extensive. Molecular complexity and associated synthetic accessibility are considerations. Structural simplicity, small size, and high ligand efficiency (potency per non-hydrogen atom) are attractive features.[10,11] These notions are reflected in the popularity of fragment screening.[12] Large size and complexity are undesirable, because synthesis of analogues is more challenging, it is difficult to have high ligand efficiency, optimization for increased potency may add further to the molecular weight, and there are more opportunities for metabolic liabilities.[4] If possible, it is desirable to pursue two or three lead series in parallel. Invariably, some will progress much better than others.

Lead Optimization

A major problem at this point is the effort and cost of lead optimization. In the pharmaceutical industry, lead optimization accounts for the dominant preclinical expense.[13] This arises from the need to synthesize numerous analogues of the hit to improve its potency, while staying mindful of its properties. Organic synthesis is labor intensive. Success is also not guaranteed. That is, a synthetic plan can be devised on paper, but when it is executed, some of the reactions do not occur as expected, the yields are poor, or separation of the product mixture is difficult. Patient reworking of some of the steps or the entire synthesis is often needed. Thus, one cannot be certain of the time it will take to synthesize any particular organic molecule. If an organization is serious about drug discovery, it is essential to engage many experienced, full-time synthetic organic chemists. At the bench level in academia, this usually means postdoctoral associates. Graduate students may participate, but it is worth remembering that the standard and potential competition are highly experienced medicinal chemists and their assistants in pharmaceutical companies, who are not distracted by courses, fellowship applications, and thesis requirements.

For compounds of average complexity in the molecular-weight range of 300–600, a diligent postdoc can be expected to generate about 50 purified compounds, ready for assay in a year. If the reaction sequences and purification are particularly straightforward, perhaps the annual yield is 100 compounds. Simple math considering salaries, materials, and overhead implies an average cost of $1000–$2000 per compound. Successful optimization to give agents with activities in the low-nanomolar range and good properties from the synthesis of only circa 200 analogues would be exceptional. Figuring in the costs for the pursuit of one unproductive series gives an estimated best-case cost for the project of $500 000.

Normally, an external grant would be needed to sponsor such an endeavor. Issues with that include the highly competitive current atmosphere, delay on the start of projects until funding arrives, and potential criticism from reviewers on the more applied nature of the work. For example, NIH study sections are dominated by biologists and biochemists, who are not easily inspired by preparative chemistry. There may also be some members from pharmaceutical and biotechnology companies, who are not always welcoming of intrusion by academicians into their domain. Their attitude can be that it is fine to help with new target identification, but leave the drug discovery (and composition-of-matter patents) to us. Thus, careful crafting of proposals is needed to combine interesting biology with interesting compound design and

synthesis. In the end, universities have cost advantages over for-profit companies, but they are not poised for rapid response to new opportunities. Pursuit of means to reduce the number of compounds that have to be synthesized and assayed is certainly desirable, and it may take advantage of the strengths in biomolecular structure determination and computational science at universities.[2]

Another, critical issue with lead optimization at universities is that there are few research groups able and willing to carry out the work. The only candidate faculty members are subsets found in departments of chemistry, medicinal chemistry, pharmacology, or pharmaceutical sciences. In comparison to the hundreds of biologists, there are on the order of ten research groups at a typical US university capable of the necessary small-molecule synthesis. The situation strikes me as similar in Europe. Furthermore, half of the synthetic chemistry groups may emphasize reaction development or synthesis and study of complex natural products; they would typically exude little enthusiasm for preparation of 100–200 analogues in a lead series with routine synthetic challenge. An additional problem is that synthetic chemists in chemistry departments are often not very knowledgeable about the pharmacology and characteristics of drugs; they are not used to participating in drug-discovery and lead-optimization projects. To begin, they need to know about frequent hitters and toxicophores[9,14] as the presence of a functional group in a natural product does not imply that it is acceptable in a drug for humans. By contrast, in most major pharmaceutical companies there are many more medicinal (synthetic) chemists than biologists. Thus, to be serious about drug discovery, universities will have to increase the number of research groups pursuing synthesis for lead optimization and/or expand medicinal chemistry resources in drug-discovery centers with experienced practitioners as directors. Otherwise, an inadequate number of compounds will be synthesized to make success in drug discovery likely.

Preclinical Studies and Technology Transfer

If all goes well, novel, potent compounds will have been discovered for a compelling biomolecular target. Patent applications and initial publications would be appropriate. The university's technology-transfer office would be engaged, and it is conceivable that the intellectual property could be licensed to a biotechnology or pharmaceutical company, which would then take over the preclinical and clinical development. The royalties would begin to flow and before long new laboratory buildings and science centers might be constructed with some of the profits. Indeed, there will be some winners, but there will be much invested in projects that do not progress, as in the pharmaceutical industry.

Academic drug discovery has typically stopped at the lead-optimization stage or earlier, owing to the inabilities and expense of progressing to lead-optimization and preclinical studies. In the university setting it has not been clear how to pay for these costs. The fundamental modus operandi has been to collect fees (overhead) to cover the expenses associated with research, especially facilities and personnel. Universities have avoided at-risk investments, such as those characteristic of preclinical and clinical drug development. If through sponsored research a faculty member manages to discover a compound series that is licensed, this is fine as discretionary funds of the university have not been expended. There will be patent expenses for the technology-transfer office (ca. $50000 or more for patent preparation and submission), but these are the first to be repaid by any royalty income, and typically a slice of all royalty income is taken to cover expenses for unlicensed patents. However, the value added for even early preclinical studies (secondary pharmacology, metabolic stability, and initial mouse toxicity and bioavailability) is huge. Compounds without these data are of little value and interest to pharmaceutical companies. With a novel target and potent compounds that have passed these tests, the ability of the technology-transfer office to initiate productive conversations with potential licensees is dramatically enhanced.

In this regard, many technology-transfer offices need to improve their knowledge and expertise surrounding license agreements for small molecules rather than for biological IP, such as targets, assays, biomolecular structures, and diagnostics. Similarly, more experience with the preparation and background evaluation for composition-of-matter patents is needed. Significant damage can be done by filing patent applications that are not up to the standards of pharmaceutical companies. Critical damage can also be done by faculty members and their co-workers making unintended disclosures in publications, outside lectures, posters, and associated abstracts. If drug discovery is a goal, it is advisable for technology-transfer offices to include a trained medicinal chemist, with prior experience in a pharmaceutical company, on their staff. Such a person can interface well with the faculty members pursuing small-molecule drug discovery, including help to educate them on IP and pharmacological issues; he or she can also participate more effectively in the processing of patents and negotiations with pharmaceutical companies.

Few universities have the capabilities to perform the routine preclinical studies in a service mode, though individual research groups might engage in some rodent toxicity and efficacy studies, for example, with xenographs. Thus, universities need to find means to select their most valuable compounds and obtain the early preclinical information. A viable option is to out-source the work to the many CROs that now provide these services. The technology-transfer office could coordinate these activities and identify preferred providers. The costs for the secondary pharmacology, CYP450 inhibition, and microsomal and plasma stability are manageable, circa $10000, so these tests should be performed first. The rodent toxicity and bioavailability studies would follow; minimal ones would raise the total costs to about $30000, more extensive ones to $50000. The likelihood that a university would venture into the next phase, preparation of GLP material and full animal-toxicity studies, is low. A corporate partner is needed with experience in these areas, careful IP protection, IND filing, and human clinical studies. Alternatively, creation of a spinout company for these purposes might be possible, if private-investor interest can be mustered.

Conclusions

In summary, several bottlenecks for academic drug discovery can be identified, and are summarized here. For those wishing to perform HTS, weak compound collections need to be addressed by providing realistic incentives to medicinal chemistry groups to deposit their compounds or as part of the efforts of drug discovery centers. Research on and application of lower-cost alternatives, such as virtual screening, fragment screening, and screening of proprietary-focused libraries, should be pursued.[2,3] The shortage of truly medicinally oriented synthetic chemists in universities can be addressed by targeted faculty hiring, additions of research scientists in the drug discovery centers, and support of postdoctoral fellows in selected labs. Outsourcing of synthetic chemistry by universities is not compelling in view of the comparatively low cost for postdocs versus external full-time equivalents. To reduce the synthesis load, improved methods for computationally guided lead optimization should be sought.[2] It is also advisable for chemistry departments to provide more exposure to concepts in medicinal chemistry and pharmacology through appropriate course offerings; excellent textbooks are available.[15–17]

Greater interaction between research groups interested in drug discovery needs to be encouraged. Faculty members often know remarkably little about the research activities in other departments of their university. A seminar invitation should be made every semester to at least one person from another department. I have received one such invitation during my career, while giving circa 700 external, invited lectures. There should also be internal symposia and/or workshops on drug discovery. These can include discussion on how to utilize the academic setting most effectively and not just reproduce the traditional, industrial approaches. As indicated above, advantage should be taken of the state-of-the-art activities in biology, molecular structure determination, synthetic chemistry, computational science, and informatics at universities.

Furthermore, the need for investment in early preclinical studies is pronounced. The costs can be borne by some initial investment and creation of a pool from a fraction of licensing revenues. It is also wise for the university to put a percentage of any income back into the research groups that did the work to help reimburse them for their expenditures and to encourage further efforts. An attitude that the researchers can just go get another grant for the next project is not farsighted. Rapid response to a compelling new project is possible with some resources in hand. With attention to these issues and reasonable levels of government support, academic drug discovery should become an increasingly viable activity with wide-ranging benefits.

Received: June 13, 2012
Published online: October 24, 2012

[1] "Natural Products as Sources of New Drugs over the Last 25 Years": D. J. Newman, G. M. Cragg, *J. Nat. Prod.* **2007**, *70*, 461–477.
[2] "Efficient Drug Lead Discovery and Optimization": W. L. Jorgensen, *Acc. Chem. Res.* **2009**, *42*, 724–733.
[3] "Virtual ligand screening: strategies, perspectives and limitations": G. Klebe, *Drug Discovery Today* **2006**, *11*, 580–594.
[4] "Molecular obesity, potency and other addictions in drug discovery": M. M. Hann, *Med. Chem. Commun.* **2011**, *2*, 349–355.
[5] "Generation of a Set of Simple, Interpretable ADMET Rules of Thumb": M. P. Gleeson, *J. Med. Chem.* **2008**, *51*, 817–834.
[6] "The Role of Public-Sector Research in the Discovery of Drugs and Vaccines": A. J. Stevens, J. J. Jensen, K. Wyller, P. C. Kilgore, S. Chatterjee, M. L. Rohrbaugh, *N. Engl. J. Med.* **2011**, *364*, 535–541.
[7] "The importance of new companies for drug discovery: origins of a decade of new drugs": R. Kneller, *Nat. Rev. Drug Discovery* **2010**, *9*, 867–882.
[8] "From basic science to blockbuster drug: the discovery of Lyrica": R. B. Silverman, *Angew. Chem.* **2008**, *120*, 3552–3556; *Angew. Chem. Int. Ed.* **2008**, *47*, 3500–3504.
[9] "New Substructure Filters for Removal of Pan Assay Interference Compounds (PAINS) from Screening Libraries and for Their Exclusion in Bioassays": J. B. Baell, G. A. Holloway, *J. Med. Chem.* **2010**, *53*, 2719–2740.
[10] "Molecular complexity and its impact on finding leads for drug discovery": M. M. Hann, A. R. Leach, G. Harper, *J. Chem. Inf. Comput. Sci.* **2001**, *41*, 856–864.
[11] "Thermodynamics of Ligand Binding and Efficiency": C. H. Reynolds, M. K. Holloway, *ACS Med. Chem. Lett.* **2011**, *2*, 433–437.
[12] "Recent progress in fragment-based lead discovery": M. N. Schulz, R. E. Hubbard, *Curr. Opin. Pharmacol.* **2009**, *9*, 615–621.
[13] "Pharmaceutical productivity—the imperative for new paradigms": G. W. Milne, Jr., *Annu. Rep. Med. Chem.* **2003**, *38*, 383–396.
[14] "Structural Alert/Reactive Metabolite Concept as Applied to Medicinal Chemistry to Mitigate the Risk of Idiosyncratic Drug Toxicity": A. F. Stephan, D. P. Walker, J. Bauman, D. A. Price, T. A. Baillie, A. S. Kalgutkar, M. D. Aleo, *Chem. Res. Toxicol.* **2011**, *24*, 1345–1410.
[15] R. B. Silverman, *The Organic Chemistry of Drug Design and Drug Action*, 2nd ed., Elsevier, Oxford, UK, **2004**.
[16] *Foye's Principles of Medicinal Chemistry*, 6th ed. (Eds.: T. L. Lempke, D. A. Williams), Wolters Kluwer, Baltimore, MD, **2008**.
[17] G. L. Patrick, *An Introduction to Medicinal Chemistry*, 4th ed., Oxford University Press, Oxford, UK, **2009**.

