Drug Development

DOI: 10.1002/anie.201404761

Advancing the Drug Discovery and Development Process

*K. C. Nicolaou**

academic–industrial partnerships · biological assays · biological targets · drug design · organic synthesis

Dedicated to Professor Madeleine Joullié on the occasion of her 87th birthday

The current state of affairs in the drug discovery and development process is briefly summarized and then ways to take advantage of the ever-increasing fundamental knowledge and technical knowhow in chemistry and biology and related disciplines are discussed. The primary motivation of this Essay is to celebrate the great achievements of chemistry, biology, and medicine and to inform and inspire students and academics to enter the field of drug discovery and development while, at the same time, continue to advance the fundamentals of their disciplines. It is also meant to encourage and catalyze multidisciplinary partnerships between academia and industry as scientists attempt to merge their often complementary interests and expertise to achieve new improvements and breakthroughs in their respective fields, and the common goal of applying them to the discovery and invention of new and better medicines, especially in areas of unmet needs.

1. Introduction

The recent successes of the pharmaceutical enterprise are undeniable and extraordinary. Admirable as it is, the drug discovery and development process is one of the most challenging and difficult human endeavors, for it has to balance efficacy in health benefits with safety at an appropriate therapeutic index. This process is often a matter of life and death for patients; their cures are in the hands of scientists and clinicians who discover, develop, and administer medications for prevention, management, and cure of disease, injuries, and other disorders. Emerging from organic synthesis at the end of the nineteenth century,[1] as marked by the introduction of Aspirin, modern medicine has changed the world[2] and, in many ways, how we live and die. Aided by discoveries in biology and chemistry, modern pharmaceutical industry has made enormous contributions to society by continuously providing new medicines, diagnostics, and disease-preventing agents. Despite these impressive advances, however, the pharmaceutical industry still faces enormous scientific and financial challenges, with some watchers of the industry believing it is encountering an unprecedented crisis. It is apparent that continuous changes and improvements are both inevitable and needed. But how to bring about these changes? From the scientific and technical points of view, and because of its magnitude and complexity, this project should be viewed as an ongoing "Grand Challenge." Indeed, it will take a major transdisciplinary approach involving clinicians, biologists, medicinal and synthetic organic chemists, X-ray crystallographers and other structural biologists, chembioinformaticians, computational experts, and logicians, among others, working collaboratively and synergistically toward improved paradigms for drug discovery and development to bring about a substantial change. Strategic and resource aspects of the process also need to be continuously reevaluated by management, and modified accordingly for optimized productivity and cost. The prospects for success, however, are high, as the causes of the failures in the drug discovery and development process are understood, for the most part, at least by those responsible for discovering and developing drug candidates. Indeed, and much to their credit, biomedical researchers appear to be cognizant not only of the problems associated with some of the current practices, but also of possible solutions and improvements of the drug discovery and development process. This Essay is meant to inspire and motivate, especially those in academia, to think about how to use their expertise to contribute to the drug discovery and development process. Indeed, the freedom and flexibility offered by academia are conducive to risky ideas that can be pursued in collaboration with the expert drug discoverers in the pharmaceutical and biotechnology industries for optimal success. Thus, academic–industrial partnerships may provide a unique platform for advancing the art and science of drug discovery and development.

[*] Prof. K. C. Nicolaou
Department of Chemistry, Rice University
6100 Main Street, Houston, TX 77005 (USA)
E-mail: kcn@rice.edu

2. Current State of Affairs in the Drug Discovery and Development Process

Pressures from the sales of generic drugs and the high attrition of drug candidates are currently plaguing the pharmaceutical industry while its leaders are scrambling for new models and paradigms to improve the situation.[3–24] Recent analyses[11] reveal some stunning, if not disturbing statistics. The cost of developing a drug as of 2010 stood at approximately 1.8 billion US dollars (excluding target identification and validation and overhead costs; perhaps a range between 1–2 billion US dollars may be more descriptive) and rising. Clinical trials accounted for 63% of the total expenditures, while the cost of preclinical drug discovery and development was estimated to be only 32% of the total cost. The duration of the process from target validation to approval was on average 13.5 years. Success rates (probability of the success of drug candidates entering the clinical pipeline/Phase I) were estimated at 7% for small molecules and 11% for biologics (attrition rates of 93% and 89%, respectively).[11] It is clear that disproportionate resources are expended on late-stage development (i.e., clinical trials) and post-approval activities (e.g., marketing, litigation), as opposed to early-stage discovery and preclinical development.

The post-penicillin period was a golden era for the pharmaceutical industry with many drugs being approved steadily and at increasing rates until the recent notable sluggish achievement of drug approvals. Indeed, the global number of drugs approved annually during the period 1981–2013 did not increase significantly as expected (see Figure 1).

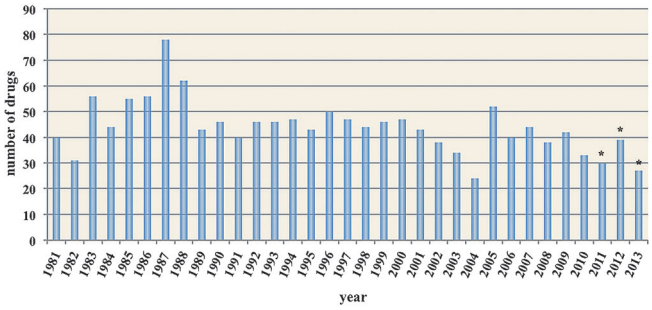
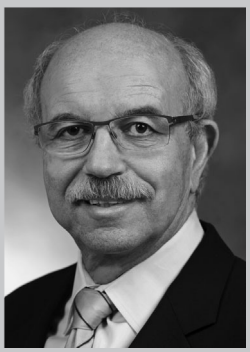


Figure 1. Number of all new approved drugs during the 1981–2013 time period (globally, modified from Ref. [49]). *Data for 2011, 2012, and 2013 are from the FDA[71] (global data not availiable as of this writing).



K. C. Nicolaou, born in Cyprus and educated in the UK and USA, holds the Harry C. and Olga K. Wiess Chair in Chemistry at Rice University. The impact of his work in chemistry, biology, and medicine flows from his contributions to chemical synthesis as described in over 750 publications. His commtiment to chemical education is reflected in his book series Classics in Total Synthesis and monograph Molecules That Changed the World, and his training of hundreds of graduate students and postdoctoral fellows.

Surprisingly to some, this phenomenon occurred despite the impressive strides made recently in chemistry and biology, the two major disciplines behind the process. It is particularly disappointing that the human genome project has not as yet had the expected impact on drug discovery, as measured by the number of drug approvals (see Figure 1), although there is no denying its beneficial impact on science and its future potential. Disappointingly, other developments that started in the 1990s, such as the combinatorial chemistry and high-throughput screening of random compound libraries, also failed to impact dramatically the drug discovery and development process despite their early promises. It is interesting to note that the advent of random combinatorial chemistry in the late 1980s coincided with the downsizing of natural products chemistry that had proved so productive in the preceding era and had been sparked by penicillin's success. To the causes of the recent slowdown in drug approvals must also be added the fact that "the low-hanging fruits" (e.g., diseases associated with known pathogenesis, druggable biological targets, predictive in vitro and in vivo assays, and reliable clinical endpoints) have already been picked, and the realization that those remaining are becoming increasingly more challenging to reach. The blame of failure to deliver better drug candidates, however, cannot entirely be placed on these developments. Rather, it appears that the actual design of the synthesized molecules during the lead discovery and optimization phase of the process in past eras was sometimes misguided, a fact recognized and pointed out by medicinal chemists and other biomedical researchers. Indeed, a series of recent reviews and commentaries convincingly argue the case for improvements and new directions in the practices of drug design of the last few decades.[3–24]

Currently, medicinal chemists have at their disposal, in addition to their experience and intuition, a number of guidelines and principles that have been developed over recent years to assist them in their endeavors as they proceed to design and optimize lead compounds to drug candidates. In most pharmaceutical companies, drug designers are also using computational models to select the best molecules for synthesis. Such computational models help them understand whether the molecules are likely to display the desired ADMET (absorption, distribution, metabolism, excretion, and toxicity) properties. In addition, they employ several structure-based drug design programs in which X-ray crystal structures help them identify the optimum small molecules to fit the targeted receptors. The optimization process involves reiterative molecular design (computer-aided or not), synthesis of the designed molecules, and biological evaluation of the synthesized compounds. Indeed, most structure–activity relationships (SARs) and other structure–property relationships are presently derived from trial and error based experimentation rather than computational chemistry or other reliable predictive methods. The first systematic guidelines to be introduced in medicinal chemistry were those delineated by Lipinski and his collaborators in their landmark papers in 2001[4] and 2004.[5] For good absorption or permeability of a compound, the so-called Lipinski rule of five (RO5) stipulates limits for certain parameters [i.e., molecular weight less than 500 Dalton (MW < 500); calculated partition

coefficient less than five ($c\log P < 5$); less than five hydrogen bond donors (HBDs < 5); and less than ten hydrogen bond acceptors(HBAs $< 10 = 2 \times 5$)]. To these parameters were later added the number of rotatable bonds (NRot, averaging around six in recent years); topological polar surface area (TPSA > 75, "3–75 rule"); and flatness as measured by the fraction of sp^3 carbons (Fsp^3, ratio of sp^3 carbons to the total number of carbons within a molecule) ($Fsp^3 < 0.47$).[13] These parameter limits were expected to impart on the compound favorable properties such as suitable lipophilicity for desired levels of absorption, solubility, cell permeability, and brain barrier penetration. These properties are important to and are usually correlated with formulation, delivery, and off-target selectivities linked to toxicity. Together with metabolism (e.g., CYP oxidation), the ADMET characteristics of a drug or its pharmacokinetic behavior need to be optimally and appropriately balanced in order for a compound to become a viable drug candidate for clinical development. For the most part, these "rules" have served medicinal chemists well in the past few years, although notable exceptions are evident. Most importantly, medicinal chemists have introduced further refinements for their drug design efforts such as "ligand efficiency" [defined as $LE = -1.4 \times \log K_i$/number of heavy atoms (atoms other than H);[25] where K_i = dissociation constant; relates binding energy per heavy atom to in vitro potency], "ligand-lipophilicity efficiency" [LLE,[26,27] also known as "lipophilic efficiency" (LipE),[28] defined as $LLE = LipE = -\log_{10}(K_i \text{ or } IC_{50}) - \log D$; relates lipophilicity to in vitro potency], "ligand-efficiency-dependent lipophilicity" (defined as $LELP = \log P/LE$),[29] the "central nervous system multiparameter optimization" algorithm [referred to as CNS MPO],[30] and "lipophilic metabolism efficiency" [defined as $LipMetE = \log D_{7.4} - \log_{10}(CL_{int,u})$, where $\log D_{7.4}$ is the log D value at pH 7.4 and $CL_{int,u}$ is the unbound intrinsic clearance in human liver microsomes; relates lipophilicity to metabolic stability].[31] These and other medicinal chemistry design parameters promise to provide additional tools for rational drug design as more data sets emerge and are exploited appropriately.

The properties of small organic molecules are, for the most part, translations of their molecular structures, the assemblies of the various structural motifs that make up their architectures. It is, therefore, not surprising that correlations of properties with certain structural motifs have been made by analysis of available data of known drugs, compounds that failed clinical trials, preclinical drug candidates, and other ligands. Matching molecular pair (MMP) analyses are becoming increasingly powerful tools for lead identification and optimization purposes as they can point to significant property adjustments by small structural changes.[32–36] MMP refers to compounds differing only in relatively small features in molecular structure (e.g., halogen vs. H, ester vs. OH, Me vs. *i*Pr). The systematic build-up of such structure–activity relationships could lead to a powerful toolbox providing correlations of structural motifs with estimates of in vitro potencies and other properties, including ADMET. Several recent reports[33–36] demonstrate the usefulness of this approach in drug discovery programs while its adoption is spreading as a consequence of its early successes. Among the most valuable general conclusions are those pertaining to lipophilicity, potency, promiscuity, and solubility. Higher lipophilicity usually leads to higher potency but also results in higher aqueous insolubility and promiscuity, both of which are liabilities for the compound. It is important to note here that lipophilic efficiency (LLE and LipE) considerations may help to understand whether potency increases are due to nonspecific lipophilic factors alone or whether specific interactions are involved. Higher numbers of aromatic, especially benzenoid, rings within the structure of a molecule increase lipophilicity, and thus potency, while at the same time lower solubility. Three aromatic rings have been suggested as the maximum number tolerable for a drug candidate, although notable exceptions exist. A better measure for this structural requirement is perhaps the Fsp^3 parameter, which takes into account the entire molecular assembly of the structure. Replacement of aromatic rings with sp^3 structural motifs is currently considered as a favorable feature for improving the properties of a compound as a drug candidate. Increasing numbers of chiral centers has also been recognized as a desirable feature within the structures of potential drug candidates.

The rules and metrics aiming to quantitatively provide guidance for drug design are not without issues, as evidenced by recent reports questioning their absolute predictivity and validity. For example, a new measure for the "drug-likeliness" of molecules has been proposed based on desirability (desirable properties). Called the quantitative estimate of "drug-likeliness" (QED), this intuitive metric reflects the distribution of molecular properties and can be used to rank candidate molecules.[37] In another more recent report, further doubts are cast on the validity of several of the so-called efficiency indices and metric rules for drug design.[38] Combining theoretical and experimental data, this study provides convincing analysis of a number of examples and concludes that, at the least, the majority of the proposed rules and metrics have to be viewed with skepticism, leaving LipE and the originally proposed Lipinski rules as the only guidelines warranting further scrutiny and use. The recent proliferation of such criteria and rules are indeed in want of critical evaluation and ranking themselves, pointing to the need for further improvement of the drug discovery and development process with regards to predictivity of properties based on molecular structure.

Intelligence gathering on known and emerging biological targets and their ligands, whether known drugs or otherwise, small or large molecules, endogenous or exogenous, is extremely important for drug discovery. Such knowledge helps scientists to understand and deconvolute the mechanism of action of both the biological targets and their ligands and provides essential information to chemists and biologists as they embark on drug discovery programs, and later during the optimization phase. A number of databases containing useful informatics on biological targets and their binding ligands already exist, and include the DrugBank database, the Therapeutic Targets Database, the U.S. FDA Orange Book (for small-molecule drugs), the Center for Biologics Evaluation and Research (CBER) website (for biological drugs), the Protein Data Bank (PDB), and the Online Mendelian

Inheritance of Man (OMIM) (for genetic diseases). Analyses of data from these databanks reveal interesting facts and trends. Among them are a) the findings that only a few hundred targets, and even fewer privileged druggable domains, account for all the approved therapeutic drugs, b) the emergence of target families [gene families, e.g., protein kinases, G protein coupled receptors (GPCRs)], and c) the recognition of the importance of drug polypharmacology (binding to and modulation of several targets).[6] In this respect, it is interesting to note that current knowledge places the number of human genes to 25000, human proteins to 200000, and human cells to 12 trillion.[39] The task of the medicinal chemist to design and synthesize a molecule that would navigate selectively to its target is enormous. The fact that it has been done so many times is a tribute to medicinal chemists and those other scientists that contributed so brilliantly to bring the state of affairs in drug discovery to its present admirable condition.

Recent analyses of drug targets and their ligands revealed further useful intelligence and insights.[16] Thus, up to 2010, 435 effect-mediating biological targets in the human genome were modulated by 989 drugs through 2242 binding interactions.[16] Classified in several groups, these targets include the families of receptors (193), enzymes (97), transporter proteins (67), and others (51). Among the latter group are enzyme-interacting proteins, structural and adhesion proteins, and ligands. The receptor group includes G protein coupled receptors (GPCRs, 82), ligand-gated ion channels (39), tyrosine kinases (22), immunoglobulin-like receptors (21), nuclear receptors (17), and other receptors (12). The enzyme category includes the families of oxidoreductases (22), transferases (21), hydrolases (43), lyases (3), isomerases (5), ligases (1), and other groups (2). Within the transporter protein class are the voltage-gated ion channels (29), other ion channels (6), solute carriers (12), active transporters (7), other transporters (3), and auxiliary transport units (10). Database analyses also revealed that from 1982 to 2010, a total of 520 drugs were approved by the FDA.[16] Derived from these studies were also the conclusions that most of these drugs operate on previously targeted human proteins, and that the rate of successful modulation of targets over the last 30 years has been stable. In the last few decades only a few new "druggable" biological targets have emerged each year. This is in contrast to the rather dramatic increase in investment and despite the impressive advances made in biology and chemistry over this period. This dissymmetry may be traced to a number of reasons, including the aforementioned "low-hanging fruits" explanation, pressures to deliver drug candidates prematurely, and temptations for temporary gains at the drug candidate optimization phase vs. long-term benefits. Most of the drugs discovered in the 1982–2010 period were small molecules and include those targeting novel biological targets.

In the meantime, however, the numbers of biologic drugs such as monoclonal antibodies, fusion proteins, and enzymes have been steadily increasing in the last two decades, demonstrating a more than fashionable trend. These include antibodies [for example, rituximab (Rituxan, Roche; binds to CD20 B-lymphocyte antigen, used to treat rheumatoid arthritis, multiple sclerosis, and other autoimmune diseases), adalimumab (Humira, Abbott; binds to TNFα, used to treat rheumatoid arthritis and other autoimmune diseases) and trastuzumab (Herceptin, Genentech; used against HER2 positive breast cancer)] and antibody drug conjugates (ADCs) with cytotoxic drugs as payloads for targeted cancer chemotherapy [e.g., brentuximab vedotin (Adcetris, Seattle Genetics and Millenium/Takeda; used against advanced Hodgkin's lymphoma) and trastuzumab emtansine (Kadcyla, Genentech/Roche; used against late-stage HER2 positive breast cancer)]. Biologics will continue to be on the rise as drugs and drug candidates. Indeed, a recent report from America's Biopharmaceutical Research Companies[39] lists 907 biologic drug candidates (antisense, cell therapy, gene therapy, monoclonal antibodies, recombinant proteins, vaccines and others) in clinical development (Phases I–III and pending application for approval, see Figure 2). Targeting

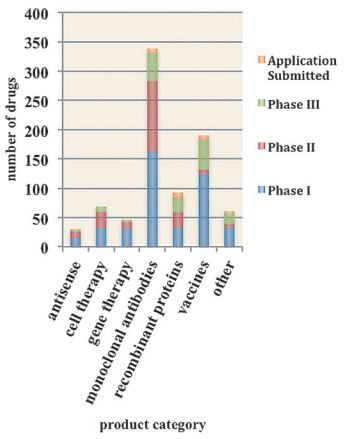


Figure 2. Number of biologics (total 907) in clinical development (phases I–III and approval process) by product category (2013 report from America's Biopharmaceutical Research Companies).[39]

more than 100 diseases (cancer, 38; infectious diseases, 176; autoimmune diseases, 71; cardiovascular diseases, 58; and others, including diabetes and digestive, genetic, neurologic, and respiratory disorders), these drugs promise to push the frontiers of science and medicine to new domains and advance healthcare to new heights. Identifying which drugs will help which patients and following up with personalized medicines is clearly the new paradigm in medicine and will certainly contribute to the improvement of the drug discovery and development process and better healthcare for the patients. While biologics are currently being hotly pursued, we should not allow the success of any given modality to swing the pendulum too far in one direction and certainly not away from small molecules and natural products. Indeed the complementarity of each approach should be exploited and viewed as a strength, for each approach has its own advantages and disadvantages.

Another powerful trend, that of multitargeting drugs,[40,41] has emerged over the last few years as marked by the introduction of imatinib (Gleevec, Novartis). Initially targeted against the mutant BRC-ABL kinase and used for the

treatment of chronic myelogenous leukemia (CML), imatinib was later found to bind to several kinases. Another impressive example of a multikinase inhibitor drug is sorafenib (Nexavar, Bayer). Approved by the FDA in 2005 for the treatment of advanced renal cell carcinoma, this drug is now employed for a number of other cancer types. This approach of multitarget drug mechanism of action is becoming a new paradigm in drug discovery. Network pharmacology and polypharmacology,[40,41] two relatively new directions in biomedical research, aim at understanding and exploiting this new strategy of molecular targeting for the treatment of disease.[6] The overarching new paradigm of both small-molecule drugs and biologics aims at more targeted strategies to cure disease without collateral damage that often leads to undesired side effects due to off-target promiscuity.

Despite all of the new information on structural motifs and the properties they impart, and technological advantages in organic synthesis, the drug discovery and development process still fails to realize gains in the number of drug candidates successfully crossing the finish line of clinical trials and approval. To be fair to the medicinal chemists, we should note that drug candidates fail not only due to deficiencies in their molecular structures but also, and most importantly, because of lack of full understanding of the pathogenesis and biology of the disease. It has been estimated that, on average, 10000 compounds are synthesized and tested before one of them makes it to the clinic as an approved drug. It is also interesting to note that analyses of several databases suggest that a "typical" medicinal chemistry molecule (not necessarily an actual molecule)[13] has a molecular weight in the range of 350–550, log P between 3.5–5.5, and TPSA of 60–90 $Å^2$; it possesses 0–2 chiral centers, 30–50 % of its carbon atoms are in the sp^3 configuration, and it contains a biaryl bond linking a fused aromatic system and another ring (with one of the rings being a benzenoid). The typical molecule is also likely to contain a "solubilizing" group such as morpholine or piperazine bridged through a linker to an aryl ring, an amide, and an aromatic ring carrying a fluoride or chloride residue. This "typical" molecule would have most likely been synthesized in four to six steps that included an amide bond formation, a deprotection step (most probably a Boc removal from a structural motif introduced from a commercially available building block), and a palladium-catalyzed cross-coupling reaction (most likely a Suzuki reaction).[13] It is also of interest that the average potency of approved drugs is around 20 nM.

Synthetic organic chemists have made impressive strides in their science and medicinal chemists have performed admirably in applying some of the emerging technologies in organic synthesis to their drug discovery efforts, matching the enormous advances made by biologists and clinicians in their domains. And yet a number of menacing diseases such as Alzheimer's and certain types of cancer remain untreatable. To be sure, scientists and clinicians are capable and poised to undertake the challenge of improving even further the drug discovery and development process by systematic diagnostic and corrective actions through collaborative efforts and new strategic initiatives that could bring within reach cures for some of the remaining intransigent diseases.

3. Advancing the Drug Design and Development Process

As convincingly argued by medicinal chemists and other pharmaceutical experts, the art and science of the drug discovery and development process needs changes and new paradigms.[4–24] However, due to the immense complexity of the drug discovery process, the response to this challenge can only be slow, under current conditions, despite the issues and uncertainties associated with the pharmaceutical industry. This somewhat paradoxical state of affairs becomes even more puzzling if we consider the modern instrumentation and technologies that could be deployed to address the remaining challenges. These sharp tools and powerful technologies include computer power and computational methods, chembioinformatics, organic synthesis, genomics, biological assays, animal models (when appropriately predictive), and cognitive science. Among the possible explanations for this slow, rather than decisive move toward new paradigms of drug discovery and development, the more likely reasons are perhaps the current pressures to deliver drug candidates in shorter and shorter times, considerations of cost in manufacturing the drug if approved (which provides resistance to employ costly materials and modern synthetic technologies), and lack of appreciation of the enormous long-term medical and economic benefits to be derived from such improvements (a phenomenon that leads, in turn, to favoring instead short-term and temporary gains).

Improvements in the classical drug discovery and development process (see Figure 3, center; main pipeline indicated by red arrow) may come from recent and pending advances in chembioinformatics, computational methods and computer modeling (Figure 3, top), and chemistry and biology (Figure 3, bottom). Thus, strengthening and encouraging integration of intelligence gathering and processing using modern computer power, computational methods, cognitive science, and continuously updated databanks could provide a major boost to the theoretical and chembioinformatics components of the drug discovery and development process, while major innovations may be derived from modern chemical, biological, and pharmacological developments. The latter should include a better understanding and validation of biological targets,[42] epigenetics,[43] diagnostic biomarkers, and clinical endpoints, new and improved biological and pharmacological in vitro and in vivo assays, wider applications of modern organic synthesis strategies and methods, novel structural motifs[44–46] and compound libraries, and more predictive early-phase clinical trials.

Facilitated by the ever-increasing power of computers, computational methods and cognitive science, continuously updated databanks, and useful programs for mining them rapidly (i.e., "google-like" efficiency and speed) should become routine and accessible to biologists and chemists alike as they attempt to carry out their respective tasks. Databanks of biological targets and biological target–ligand matched pairs (TLMPs), matching molecular pairs (MMPs),[32–36] biomarkers, and biological assays could provide crucial intelligence and assistance for the target identification and validation and lead identification and optimization

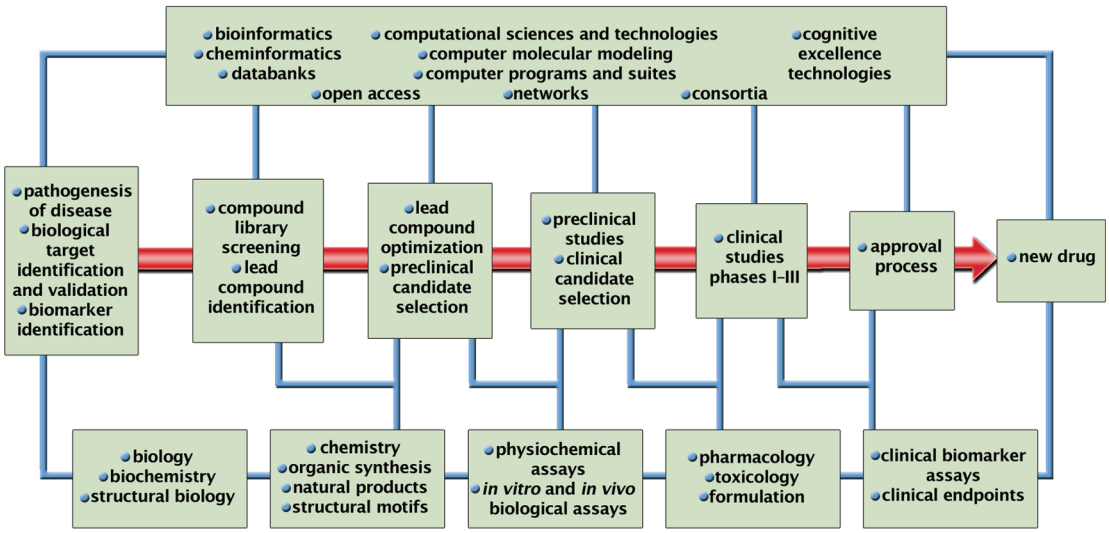


Figure 3. The drug discovery and development process (middle) with its auxiliary arms (top and bottom).

processes (Figure 3). Similarly, continuously updated databanks of selected drugs, drug candidates (failed or not), natural products, compound libraries, and structural motifs should be developed and used in the lead identification phase through in silico and in vitro screening against the biological target. Molecular modeling and docking programs should also be strengthened and employed to assist in the lead identification and optimization processes. The use of computational suites and upgraded programs for calculating ADMET properties should be intensified and spread in academic groups embarking on drug discovery programs as part of the recent surge in such activities as a result of initiates from funding agencies and industrial institutions. Such computational tools should become routine components of the drug design process as complementary inputs to creativity, imagination, and intuition. In several pharmaceutical companies this practice is already routinely employed with considerable success by their drug designers. Exemplary among the various computational and modeling packages are those included in the Schrödinger Software Suite Quickprop (for predicting ADMET properties) and Glide (for molecular docking).[47] More precise and expanded rules and guidelines are needed for predicting compound properties, including those of molecules larger than those traditionally considered suitable drug candidates. With regards to the molecular weight criteria, several factors are of importance, including the medical indication, the method of administration, and the structural motifs that make up the molecule. Several important and long-lasting drugs of natural origin such as Taxol, amphotericin B, and vancomycin, for example, lie outside the rule of five and, in that respect, they are inspiring.[48–50] Networks for collecting and sharing knowledge, cognitive science methods, and tools for drug discovery and development should be established and supported. Not withstanding the challenges involved regarding secrecy and intellectual property, such sharing of knowledge and resources will certainly result in a more productive, expedient, and cost effective drug discovery and development process.

While medicinal chemists can be empowered by theoretical and computational methods, their endeavors can also benefit from the changes and improvements in their own past practices, particularly at the lead identification and optimization stages. Starting with superior lead compounds will certainly accelerate and improve the optimization process to yield better drug candidates. To ensure this, the quality of compound libraries for biological screening must be continuously improved and so should their accessibility, especially to academic investigators. Since an excess of aromatic moieties has been identified as a source of undesirable properties,[9,10] its presence in the library compounds should be avoided or limited. Selected ligands of proven value, however, may constitute acceptable lead compounds as a start, provided the optimization process includes early modifications to achieve a better balance of structural motifs, and as a consequence lead to more favorable properties. Natural products should be integral parts of such compound libraries due to their past success as drugs, lead compounds, and biological tools.[2,48–50] While the numbers vary from year to year, the percentage of the natural product (NP), natural botanical (NB), and natural derived (ND) of total global drug approvals average between 30–50% (see Figure 4) (not to mention drugs inspired by natural products).[49] The three-dimensional structures of natural products and their wealth of chiral centers should serve as an inspiration and motivation for drug designers. The dimensionality and chirality of biological receptors and the fact that natural products have evolved along and against such biomolecules explains their diverse, potent, and often selective biological properties.[31,32] Employing them and molecules like them[48–52] as leads and introducing some of their structural features in drug designs makes good sense and should be a complementary approach to the currently employed drug design practices. The complexity of the molecule should not be a deterrent since most natural products can be obtained from natural sources, biotechnology processes, or organic synthesis. Given its recently acquired power, the latter approach is particularly

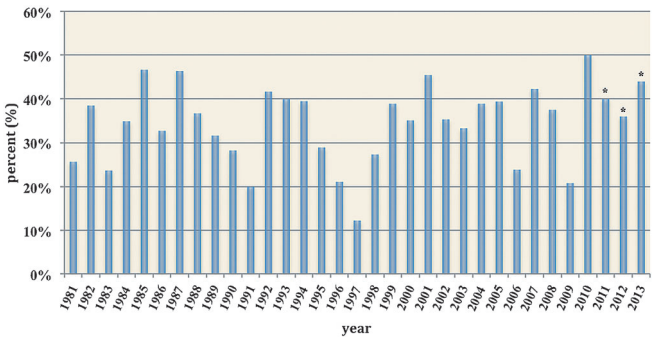


Figure 4. Percentage of natural (N), natural botanical (NB) and natural-derived (ND) drugs approved in the 1981–2013 time period (globally, modified from Ref. [49]). *Data for 2011, 2012, and 2013 are from the FDA[72–74] (global data not available as of this writing).

suited for scarce compounds that are difficult to obtain from laboratory cultures, especially those endowed with extremely high potencies. Such molecules would be needed in smaller quantities than traditional drugs and could be deployed, for example, for rare diseases and as payloads for antibody drug conjugates (ADCs).[53,54] It might be true that natural products chemistry requires longer-term plans and higher initial investments, but in the long run, the endeavor pays off as demonstrated by its rich and glorious history. The advent of the ADCs has signaled a renaissance in natural products chemistry, providing strong justification for its reinvigoration through new directions and funding initiatives. Investigating the vast unexplored kingdoms of living creatures on land and sea using high-tech instruments and modern biotechnologies, and the broad biological screening of newly discovered natural products are some of the initiatives that could pay off handsomely on a steady basis once set in motion. Academic investigators are ideal for undertaking such efforts. They should be enabled to pursue them through academic–industrial partnerships, government funding, and other resources.

A number of structural motifs have been recognized as imparting beneficial drug properties to the molecule and should be adopted for incorporation into optimized compounds more often.[44–46] In addition to saturated and chiral moieties, halogens and halogen-containing residues are known to improve pharmacological properties.[55–57] Strategically placed within a molecule, fluorine[55,57] and chlorine,[56,57] in particular, may impart beneficial properties, such as higher target affinity (through halogen bonding) and metabolic stability, to the molecule. Bromine and iodine[56,57] are less frequently employed, although they are endowed with higher potential for halogen-bonding interactions that may lead to improved selectivity and other desirable features. The adoption of iodine residues (especially on aromatic scaffolds), however, should be made with caution because of potential toxic effects, depending on the dosage (i.e. interference with thyroid hormone receptors). New and improved methods for incorporating such halogen residues are desirable. Moreover, the entire range of modern organic synthesis methods should be employed in the search for the best possible drug candidates, even if this choice may mean higher initial investment of time and effort, for in the end improved drug candidates and better and more drugs are likely to emerge. Here it is important to note that many published synthetic methods from academic laboratories are often not medicinal-chemistry friendly, or at least have not been tested with druglike molecules in mind. In this regard, academic–industrial partnerships[58] are important and necessary in order to refine the synthetic methods to include applications relevant to drug discovery. Parenthetically, it is ironic that a number of powerful synthetic reactions have led medicinal chemistry astray in some instances and have not augmented its record of success as measured by the number of successful drug candidates. Certain palladium-catalyzed cross-coupling reactions are a striking example of this phenomenon which is clearly due to the practical and expedient manner by which these admirable processes lead to "optimized" drug candidates. As amply documented, the majority of these so-called "druglike" molecules end up failing in the clinic, often for reasons of low solubility and high promiscuity in their binding preferences resulting from the excess of aromatic rings.[9,10] This is not to detract from the awesome power and usefulness of these reactions whose employment has recently revolutionized organic synthesis and its applications. The data, however, suggest more caution and wisdom for their use in medicinal chemistry.

Pressures to deliver drug candidates on increasingly shorter timelines are responsible to a large measure for the selection of reliable and versatile reactions to rapidly assemble families of compounds for intelligence gathering (SARs) and optimization purposes. Despite these constraints, however, practitioners of the art of drug discovery have shown courage and begun to adopt a wider range of newer and more exotic reactions in their efforts to move beyond the traditional "flatland" molecules[9,10] by incorporating more saturated structural motifs and chiral centers in their designed and synthesized structures. They ought to venture even further into the awesome panoply of modern organic synthesis and employ its tools to construct more diverse and sophisticated molecules that would include more than the rather few chiral centers (usually one or two) of past drug candidates. It is indeed fortunate that today the state of the art of organic synthesis is capable of taking us much beyond this "typical" and restrictive model. Drug designers and medicinal chemists should be given the opportunity to explore the full range of strategies and technologies currently available in their quest for high-quality drug candidates and better drugs. At the same time, new and improved synthetic methods should be sought in order to enable expeditions into new molecular diversity through their practicality and efficiency.[51,52,59] New and improved reactions are also needed to access novel structural motifs[46] to replace some of the traditional aromatic systems, especially benzenoid rings, in order to achieve overall dimensionality and improved pharmacological properties. Adoption of new techniques and instrumentation should be encouraged by demonstrating their power and convenient use. Included among them are microwave and flow techniques and automation systems such as those so effectively employed to synthesize peptides and oligonucleotides. The efficiency of peptide and phosphate

bond-forming reactions and the iterative nature of their applications to construct oligomeric structures from similar structural units are responsible for the success of machines and automation in accessing such systems. Intensifying collaborative efforts between synthetic organic chemists and engineers should lead to innovations beyond the amide and phosphate bonds and would allow the introduction of instrumentation and automation into the laboratory for accommodating other reactions. Indeed, it is surprising that so many of the essential experimental techniques employed in laboratories of organic synthesis today are very similar, if not the same, as those employed one or even two generations ago, in contrast to the rather dramatic increase of useful reactions for molecular construction and improvements in analytical instrumentation in the same period.

Improvements in the biology and pharmacology inputs of the drug discovery and development are also possible and imminent, and should be incorporated. Continued advances in sequencing and understanding normal and disease-associated human genomes are guiding current drug discovery and development programs directed toward novel drugs (drugs acting through new mechanisms). This is particularly evident in the cancer area[20] where certain gene mutations were identified as "drivers" for disease. Among the several challenges in applying such knowledge is the inability to find drugs against these driver mutations ,which have often been proven "undruggable," and the lack of predictive models for robust predictions regarding clinical efficacy and safety. With regard to the first challenge, "druggability" should be considered a relative and not an absolute term. Protein kinases were initially viewed as "undruggable" targets until persistence and innovation proved them otherwise. Further innovations in organic synthesis and the willingness to invest time and effort should turn so-called "undruggable" into druggable targets. Such difficult and somewhat risky targets can make ideal projects for academic work or academic–industrial partnerships, which are evidently on the rise.[58] The other important issue, that of predicting clinical efficacy and safety at the preclinical phase of discovery through more robust pharmacological models, is also showing strong indications of gathering momentum as pharmacologists and clinicians intensify their efforts to develop new and improved biological assays and animal models.[20] Incidentally but most importantly, Sir James Black, for example, favored and championed such powerful assays using animals and animal-derived organs. Improved and new in vitro biological assays to predict the in vivo behavior of compounds could provide crucial intelligence for the optimization process. A number of such assays are currently employed, including metabolism (CYP) and heterologously expressed human voltage-gated potassium channel subfamily H member 2 (KCNH2, commonly known as hERG) assays. Additional parallel in vitro assays to test compounds against numerous relevant molecular biological targets are now standard in most pharmaceutical companies and should become routine in academia as well through appropriate collaborations and open-access precompetitive consortia. Phenotypic screening in cells should be promoted and adopted more routinely than in the past, due to its recognized importance and value. This paradigm has already been implemented in several pharmaceutical companies employing multiwell formats, acoustic-wave-based distribution, and other methods. This practice would avoid pitfalls and save time and effort as well as cost, as attempts to optimize one property often result in loss in gains already made in other properties in the sequential process.

The emergence of cancer stem cells (CSCs) as primary drivers of tumor growth, perpetuation, and drug resistance provides new opportunities for progress in cancer chemotherapy.[60] Obliterating CSCs has the exciting possibility to eradicate cancer from patients, provided personalized medicines can be discovered and developed. To succeed, the cancer stem cell paradigm requires the identification and characterization of subtype cancer stem cells, through biomarkers or other unique biological targets, and development of drugs that can specifically seek them out and destroy them. Coupled with the patient-derived xenografts (PDXs) model, this paradigm stands at the cutting edge of cancer biology, drug discovery, and personalized medicine. Its complexity challenges pathologists, surgeons, biologists, and chemists and their institutions to navigate through the multi- and trans-disciplinary collaborations needed for success.[60]

Reducing the safety-related attrition of drug candidates through in vitro pharmacological profiling is becoming a must, especially in big pharma companies where experience and case studies proved invaluable and pathpointing. The recent collaborative intelligence gathering and sharing of information by AstraZeneca, GlaxoSmithKline, Novartis, and Pfizer on this topic is admirable and exemplary.[21] According to their recommendations, molecular optimization of preclinical drug candidates should not only be guided by potency (primary effect) considerations but also by intelligence gathering from parallel pharmacological assays to identify off-target effects (secondary effects) and safety hazards.[21] The road to preclinical candidacy should also include hazard elimination and mechanistic understanding of both the primary and secondary effects of the molecule being groomed before it is declared a clinical candidate. A core panel of biological targets (i.e., receptors, ion channels, enzymes, and transporters) selected for their potential safety liability are recommended for embedding into the screening process along with the primary target. The additional cost for such assays could be, however, an issue, and cost–benefit judgments have to be made as to the extent of their employment. Data from those assays can then be used to guide the drug design efforts of the medicinal chemists so as to avoid costly surprises at a later stage. These in vitro studies may also impact other downstream activities such as further screening against a broader panel of targets and in vivo studies in order to gain further mechanistic and safety insights before final decisions on drug candidacy are made. A panel of 44 targets has been identified by this consortium in their initial recommendation, with the expectation of further expansion of the panel later as knowledge of new targets with safety liabilities emerges.[21] In addition to detecting off-target binding of drug candidates, such assay panels can provide useful information about newly discovered bioactive molecules having an unknown mechanism of action. This proposition is of particular relevance to academic investigators who

frequently discover molecules active at the cellular level (e.g., cancer cells, bacteria, etc.) but do not know the mechanism of action at the molecular level [i.e., biological target(s)]. Open access precompetitive consortia with as many biological target assays (such as the NIH "High-Throughput Screening" initiative) would be of great value to research programs in chemical biology and drug discovery at the precompetitive stage, especially in academic institutions.

Interestingly, the first drugs to be approved for use against new targets in recent times are those emerging from research directed toward orphan or rare diseases.[61,62] This phenomenon, which could be explained by the faster track status usually granted to drug candidates for these disorders, and by the fact that rare diseases are often associated with distinct biomarkers or single point mutations, is significant in that it offers unique research opportunities for discovery and innovation leading to new paradigms that could be later applied to more common diseases. Continuous mining of the drug discovery and development literature and systematically storing data on biological targets, and drugs and other ligands associated with them, in open access databases is of great importance to researchers in the field, especially in academia and small biotechnology companies. Computational[63] and cognitive technologies[64] should be at the forefront of these chembioinformatics initiatives as we strive to turn the art of drug design to a more precise and predictive science. The latter technologies are based on relevant data banks and constellatory thinking and are being deployed as computational tools to assist decision making on a variety of issues; in principle, they could be applied to drug design. Indeed, our collective knowledge on drugs and impressive advances in the biological, chemical, and clinical sciences should be more synergistically and wisely exploited and, together with computational methods, brought to bear on the grand challenge of improving the drug discovery and development process, preferably on a commonly shared basis (Figure 3, top).

The probability of technical success manifests itself most decisively during clinical trials which are revealing in terms of safety (phases I/II) and efficacy (phases II/III) in human patients. Provided the biological target has been validated beyond doubt, drug candidates fail for reasons of poor oral bioavailability and pharmacokinetics or toxicity that apparently were not predicted by ADMET studies and animal models. With regards to toxicity, monoclonal antibodies have an advantage over small molecules in that they are less likely to exhibit off-target effects due to their high specificity, although their potential long-term effects have yet to be determined. Overall, predictive ADMET methods are improving and the clinical failures traced back to these properties are decreasing as a result. On the other hand, the targeting of poorly understood and insufficiently validated targets, lack of predictive animal models, higher efficacy/safety margins, and stricter guidelines imposed by regulatory authorities appear to be the main reasons for late-stage failures in the clinic. To lower the clinical attrition of drug candidates, better understanding and improved validation of targets before the start of the drug discovery process, increased knowledge on biomarkers, and proof of principle studies at phase I clinical trials are needed (see Figure 3, bottom). Biomarkers are of great importance for personalized medicine. Identification of such biological entities is, therefore, of extreme urgency in developing effective and safe drugs for patients whose disease is associated with such diagnostic markers. Diagnostic tools for such markers should be developed and be ready for phase II/III clinical trials and immediately after launch so that they can be used to identify and optimally benefit patients. Imatinib (Gleevec) and trastuzumab (Herceptin), for example, are associated with biomarkers that guide their use for the best-suited patients. Gaining understanding of the role of the biological target, its mechanism of action, and clinical consequences is a prerequisite to its validation (see Figure 3). The target validation process is improving as a result of progress in genomics, chemical validation tools (e.g. potent inhibitors/binders, irreversible inhibitors), gene silencing technologies (shRNA, CRISPR), new biological assays, and animal models, including patient-derived xenografts (PDXs). Further research and development in the latter area in particular could help establish a wide range of PDX animal models to be used in preclinical evaluations, particularly of cancer drug candidates, as powerful clinical predictors. A number of vendors are already offering collections of genetically annotated tumors for PDX model testing. Such closer to the clinic evaluations should result in lower attrition rates of drug candidates, and lead to better drugs, while at the same time save money at the more costly clinical phase of the drug discovery and development process, especially as we move into personalized medicine.[20,22] The establishment of biomarkers and surrogate end points (e.g., measurement of glucose or cholesterol in diabetes and high cholesterol disease, respectively), if not clinical end points early on in clinical trials (phase I), are also becoming highly desirable practices as they could provide red flags. Shifting attrition from phase III to phase II or, even more importantly, phase I could result in significantly lower costs. These savings could be funneled back to fuel the efforts for target validation and compound lead discovery and optimization. This strategic shift will lead to more and better drug candidates, resulting in overall improvement of the drug discovery and development process. Hasty advancement of candidates into phase III clinical trials should be avoided despite temptations of perceived temporary gains. Phase II clinical trials have had the highest rate of failure in the clinical drug development phase. While drug candidates are more likely to succeed in phase I clinical trials than not (> 50 % success rate), data for the 2008–2010 period showed a success rate of around 18 % in phase II. Of these remaining drug candidates, about 50 % fail due to lack of efficacy, 19 % due to safety issues, and 29 % because of strategic reasons. The latter category often refers to failures to differentiate the candidates from approved drugs for the same indication, whether they act through the same or different mechanisms.[18] These statistics suggest that the drug discovery and development process needs to improve its predictivity models for clinical efficacy and place more emphasis on clinical biomarkers.

Collaborative knowledge creation and sharing could be another important paradigm shift in attempts to overcome the current status as academia[65] and industry are joining forces

to improve the drug discovery and development process (Figure 3, top). Indeed, a number of initiatives have recently emerged in which information and resources are shared among the participants, accelerating the pace of research. This precompetitive sharing of information and resources, including chembioinformatics, compound libraries, and biological assays, could be of enormous importance, especially to academic groups and small biotechnology companies that often do not have sufficient resources and capabilities to establish and exploit such tools.[11,12] The newly established Academic Drug Discovery Consortium (ADDC), for example, aims to bring together and facilitate the work carried out in the increasing number of academic centers in the U.S. and other countries.[65] Sharing special websites, compound libraries, and biological assays through such organizations, especially when they include big pharma and biotechnology companies as well as government institutions, such as the NIH (USA), could prove crucial to the success of the pharmaceutical enterprise as it strives to find new paradigms and models to sustain itself as a productive and profitable industry. Some experts even suggest that the traditional "fully integrated pharmaceutical company" should be replaced with a "fully integrated pharmaceutical network" as the new model in order to save cost and improve productivity.[11]

A light of hope for increased productivity in the drug discovery and development process appears to be on the horizon as signaled by the upward trend of first-action drug approval rates by the FDA in the last few years (Figure 5).[66]

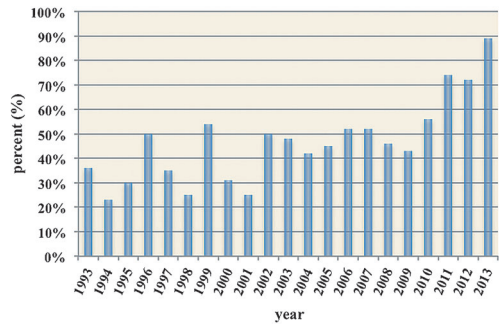


Figure 5. Percentage New Molecular Entity (NME) and New Biologic Entity (NBE) first-action approval rates per year during the 1993–2013 time period.[66,74]

Starting in 2009, this increase in first action approvals may mean (for the optimists) a true signal that improvements in the drug discovery and development process are beginning to pay off, or (for the pessimists) it may simply be a reflection of FDA relaxing its approval criteria as a consequence of politics associated with the Affordable Care Act which, coincidentally, was passed by the U.S. Congress in the same year as the first action approval rates began their ascent (temporary or not?). Be that as it may, the long vision surge for advances and improvements must continue at an accelerated pace and with renewed vigor and determination.

4. Conclusion

It is clear to those practicing and watching the art and science of drug discovery that the pharmaceutical enterprise is facing complex and challenging issues.[67–70] Indeed, it is both timely and prudent to mobilize the relevant scientific community toward integrating and advancing the process through collaborative and transdisciplinary efforts that would include both academic and industrial groups. As academic–industrial partnerships are being formed and intensified, more friendly and open-access information and tools, especially to academic investigators, are needed in order to facilitate their drug discovery and development research activities. Empowered by such intelligence and resources, and through strategic alliances and consortia, biomedical investigators are poised to develop new paradigms that could enable and improve the drug discovery and development process for the benefit of all stakeholders of the healthcare system and its foundations. Furthermore and most importantly, enhanced academic–industrial partnerships[58] could lead, if structured properly to maintain academic freedom and protect intellectual property, to improved education and training of those students who are destined for the pharmaceutical and biotechnology sectors.

Ideally, the biological target should be fully validated. Ideally, computer-aided drug design should be of sufficient predictive power such that a molecule with appropriate pharmacological properties can be designed to bind selectively to a specific receptor with low nanomolar or lower potencies. Ideally, organic synthesis should be able to efficiently deliver a more or less complex compound selectively and efficiently. Ideally, preclinical models should be able to predict the clinical efficacy and safety of a drug candidate. These ideals are certainly not currently within sight. What is certain, however, is that the foundations for progress toward these idealistic goals are here today. With sufficient will and investment, robust methods and algorithms could be devised that would set in motion the necessary multi-, inter-, and transdisciplinary partnerships between academia and industry to ensure a focused and successful response to this grand challenge. With the currently available and emerging knowledge and tools, such a concentrated effort should lead, in the not too distant future, to a success rate for drug candidates of higher than 50% and beyond. For substantial improvements in the drug discovery and development enterprise to occur, a number of paradigm shifts and resource reallocations are needed.[20] Among the tasks, tools, and initiatives to be undertaken en route toward that goal may be the following (some of these are already in place in big pharma and related consortia but are not fully open access; their listing below should be informative to students and newcomers to the field, especially those from academia and new start-up companies):

1. Continue to develop better methods for understanding of the pathogenesis and biology of disease at the molecular level.
2. Continue to develop and make available (through consortia and other organizations) biological and pharmacological assays for all important and new biological

molecular targets (e.g., receptors, enzymes, ion channels, etc.).
3. Develop on a continuous basis and make available cellular and related assays (e.g. cancer cells, stem cells, bacteria, fungi, viruses, etc.).
4. Develop on a continuous basis and make available biological and pharmacological assays for animal tissues and organs and animal models as well as patient-derived xenografts (PDXs).
5. Develop on a continuous basis and make available systematic parallel rather than sequential biological and pharmacological evaluation of potential drug candidates during the optimization phase.
6. Develop on a continuous basis and make available a database of biological target-selected drugs/ligands as an intelligence tool to identify lead compounds for homologous newly emerging biological targets.
7. Develop and update on a continuous basis and make available universal rules and guidelines for drug design in order to improve predictivity of potency and ADMET properties.
8. Develop on a continuous basis and make available molecular matching pairs (MMPs) data for drug design.
9. Develop on a continuous basis and make available cognitive excellence technologies as a tool for decision making in biological target–ligand matching and drug design practices.
10. Develop systematically novel structural motifs to replace excess aromatic moieties within drug candidates as a means to achieve improved pharmacological properties.
11. Expand molecular space for chemical biology and drug discovery studies through discovery and development of new synthetic methods and strategies.
12. Expand applications of organic synthesis into the domain of biologics including conjugation, linkers, large organic molecules, polypeptides, and oligonucleotides.
13. Expand applications of chemical synthesis into the domain of nanomedicine including nanoparticles and other special polymeric materials.
14. Promote (e.g. through organized government and private funding initiatives) the isolation of novel natural products and subject them to broad biological screening against biological targets and cellular assays.
15. Employ modern organic synthesis technologies and genetic engineering techniques to produce rare but potent natural products as lead compounds and drug candidates for conjugation to antibodies and other selective delivery systems.
16. Place more emphasis on creativity and imagination at the discovery and optimization phase rather than shortcuts, unrealistic deadlines, and pressures that inevitably lead to most-likely-to-fail, rather than -succeed clinical candidates.
17. Place more emphasis and resources on the early phases of the drug discovery and development process, namely target identification and validation, lead identification and optimization, and preclinical development in order to shift attrition earlier in the process.
18. Promote and establish inter- and transdisciplinary partnerships between academic and industrial groups for competitive (proprietary) and precompetitive (open access) research, especially in challenging areas such as "undruggable" targets, complex natural and designed molecules, novel methods and structural motifs, and new methods for biological target identification.
19. Adopt and promote phenotypic screening in cells (i.e. high content screening) and chemogenomics/proteomics to define novel biological targets and map out biological networks.
20. Maintain a balance among modalities in drug discovery (especially biologics vs. small molecules and natural products), for each has its advantages and disadvantages.

Based on these forward-looking guidelines, perhaps the most crucial for the advancement of the art and science of drug discovery are a) the efforts to better understand the pathogenesis and biology of disease so as to identify and validate the right biological target(s) against which to develop the drug candidates; b) identification and exploitation of the molecules of living nature as biological tools, lead compounds, and drug candidates; c) improved drug design employing modern organic synthesis and computational techniques; d) higher predictivity biological assays and animal models; and e) strategic decisions to address the above priorities, including the shifting of more resources to the early stages of the drug discovery and development process and adoption of long-term vision and practices based on science and technology rather than short-term financial gains.

This list of practices can be extended when opportunities arise as new paradigms in drug design and discovery emerge and evolve. Hopefully one day we will be able to press a button on a computer and a structure will appear on the screen of a potential drug candidate that we can feed into another machine, press another button, and obtain the pure drug candidate with its full analytical data the same way we press a switch on a remote control today and a television monitor comes on with a crystal-clear picture. Should we be able to gather the courage and resources to start moving toward these "Utopia" scenarios, we would be able to enjoy the wisdom and benefits, certain to be derived during the voyage even if we do not reach our "Ithaca" in our lifetime. The drug discovery and development process should be viewed as undergoing continuous emergence and evolution within itself, from itself, and into itself, meaning that any advances and improvements will have to come from the lessons of the process, from its practitioners, and for its benefit. The tremendous advances made in molecular, structural and cell biology, organic synthesis, chembioinformatics, computational science, and cognition should serve as the locomotive and the inspiration to charge forward toward advancing the drug discovery and development process in a meaningful way, for the benefits to science and society are clearly worth the investment of talent and resources.

I wish to express my gratitude and many thanks to Rice University and the Cancer Prevention and Research Institute of Texas for financial support of our educational and research

programs. I am also indebted to Dr. Antonia F. Stepan (Pfizer Worldwide Research and Development) and Dr. Gregory D. Vite (Bristol-Myers Squibb) for their encouragement and insightful and valuable comments and suggestions on the content of this essay.

Received: April 28, 2014
Published online: July 13, 2014

[1] K. C. Nicolaou, *Angew. Chem.* **2013**, *125*, 141 – 157; *Angew. Chem. Int. Ed.* **2013**, *52*, 131 – 146.
[2] K. C. Nicolaou, T. Montagnon, *Molecules That Changed the World*, Wiley-VCH, Weinheim, **2008**, p. 366.
[3] G. Wess, M. Urmann, B. Sickenberger, *Angew. Chem.* **2001**, *113*, 3443 – 3453; *Angew. Chem. Int. Ed.* **2001**, *40*, 3341 – 3350.
[4] C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney, *Adv. Drug Delivery Rev.* **2001**, *46*, 3 – 26.
[5] C. A. Lipinski, *Drug Discovery Today Technol.* **2004**, *1*, 337 – 341.
[6] J. P. Overington, B. Al-Lazikani, A. L. Hopkins, *Nature* **2006**, *5*, 993 – 996.
[7] P. D. Leeson, B. Springthorpe, *Nature* **2007**, *6*, 881 – 890.
[8] J. D. Hughes, J. Blagg, D. A. Price, S. Bailey, G. A. DeCrescenzo, R. V. Devraj, E. Ellsworth, Y. M. Fobian, M. E. Gibbs, R. W. Gilles, N. Greene, E. Huang, T. Krieger-Burke, J. Loesel, T. Wager, L. Whiteley, Y. Zhang, *Bioorg. Med. Chem. Lett.* **2008**, *18*, 4872 – 4875.
[9] T. J. Ritchie, S. J. F. Macdonald, *Drug Discovery Today* **2009**, *14*, 1011 – 1020.
[10] F. Lovering, J. Bikker, C. Humblet, *J. Med. Chem.* **2009**, *52*, 6752 – 6756.
[11] S. M. Paul, D. S. Mytelka, C. T. Dunwiddie, C. C. Persinger, B. H. Munos, S. R. Lindborg, A. L. Schacht, *Nat. Rev. Drug Discovery* **2010**, *9*, 203 – 214.
[12] J. S. Altshuler, E. Balogh, A. D. Barker, S. L. Eck, S. H. Friend, G. S. Ginsburg, R. S. Herbst, S. J. Nass, C. M. Streeter, J. A. Wagner, *Sci. Transl. Med.* **2010**, *2*, 52cm26.
[13] S. D. Roughley, A. M. Jordan, *J. Med. Chem.* **2011**, *54*, 3451 – 3479.
[14] A. F. Stepan, 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] W. P. Walters, J. Green, J. R. Weiss, M. A. Murcko, *J. Med. Chem.* **2011**, *54*, 6405 – 6416.
[16] M. Rask-Andersen, M. S. Almén, H. B. Schiöth, *Nature* **2011**, *10*, 579 – 590.
[17] P. D. Leeson, S. A. St-Gallay, *Nature* **2011**, *10*, 749 – 765.
[18] a) J. Arrowsmith, *Nat. Rev. Drug Discovery* **2011**, *10*, 328 – 329; b) J. Arrowsmith, *Nat. Rev. Drug Discovery* **2011**, *10*, 87.
[19] S. Orchard, B. Al-Lazikani, S. Bryant, D. Clark, E. Calder, I. Dix, O. Engkvist, M. Forster, A. Gaulton, M. Gilson, R. Glen, M. Grigorov, K. Hammond-Kosack, L. Harland, A. Hopkins, C. Larminie, N. Lynch, R. K. Mann, P. Murray-Rust, E. Lo Piparo, C. Southan, C. Steinbeck, D. Wishart, H. Hermjakob, J. Overington, J. Thornton, *Nat. Rev. Drug Discovery* **2011**, *10*, 661 – 669.
[20] W. R. Sellers, *Cell* **2011**, *147*, 26 – 31.
[21] J. Bowes, A. J. Brown, J. Hamon, W. Jarolimek, A. Sridhar, G. Waldron, S. Whitebread, *Nat. Rev. Drug Discovery* **2012**, *11*, 909 – 922.
[22] M. De Palma, D. Hanahan, *Mol. Oncol.* **2012**, *6*, 111 – 127.
[23] T. Tsukamoto, *ACS Med. Chem. Lett.* **2013**, *4*, 369 – 370.
[24] K. Mullane, R. J. Winquist, M. Williams, *Biochem. Pharmacol.* **2014**, *87*, 189 – 210.
[25] A. L. Hopkins, C. R. Groom, A. Alex, *Drug Discovery Today* **2004**, *9*, 430 – 431.
[26] A. R. Leach, M. M. Hann, J. N. Burrows, E. J. Griffen, *Mol. BioSyst.* **2006**, *2*, 429 – 446.
[27] T. Ryckmans, M. P. Edwards, V. A. Horne, A. M. Correia, D. R. Owen, L. R. Thompson, I. Tran, M. F. Tutt, T. Young, *Bioorg. Med. Chem. Lett.* **2009**, *19*, 4406 – 4409.
[28] P. D. Leeson, B. Springthorpe, *Nat. Rev. Drug Discovery* **2007**, *6*, 881 – 890.
[29] G. M. Keserü, G. M. Makara, *Nat. Rev. Drug Discovery* **2009**, *8*, 203 – 212.
[30] T. T. Wager, X. Hou, P. R. Verhoest, A. Villalobos, *ACS Chem. Neurosci.* **2010**, *1*, 435 – 449.
[31] A. F. Stepan, G. W. Kauffman, C. E. Keefer, P. R. Verhoest, M. Edwards, *J. Med. Chem.* **2013**, *56*, 6985 – 6990.
[32] A. G. Leach, H. D. Jones, D. A. Cosgrove, P. W. Kenny, L. Ruston, P. MacFaul, J. M. Wood, N. Colclough, B. Law, *J. Med. Chem.* **2006**, *49*, 6672 – 6682.
[33] C. E. Keefer, G. Chang, G. W. Kauffman, *Bioorg. Med. Chem. Lett.* **2011**, *19*, 3739 – 3749.
[34] E. Griffen, A. G. Leach, G. R. Robb, D. J. Warner, *J. Med. Chem.* **2011**, *54*, 7739 – 7750.
[35] S. Schultes, C. de Graaf, H. Berger, M. Mayer, A. Steffen, E. E. J. Haaksma, I. J. P. de Esch, R. Leurs, O. Kramer, *MedChemComm* **2012**, *3*, 584 – 591.
[36] J. Boström, A. Hogner, A. Llinas, E. Wellner, A. T. Plowright, *J. Med. Chem.* **2012**, *55*, 1817 – 1830.
[37] G. R. Bickerton, G. V. Paolini, J. Besnard, S. Muresan, A. L. Hopkins, *Nat. Chem.* **2012**, *4*, 90 – 98.
[38] M. D. Shultz, *Bioorg. Med. Chem. Lett.* **2013**, *23*, 5980 – 5991.
[39] America's Biopharmaceutical Research Companies. *Biologics: Medicines in Development.* Pharmaceutical Research and Manufacturers of America, **2013**, http://www.phrma.org.
[40] J. L. Medina-Franco, M. A. Giulianotti, G. S. Welmaker, R. A. Houghten, *Drug Discovery Today* **2013**, *18*, 495 – 501.
[41] a) P. Ciceri, S. Müller, A. O'Mahony, O. Fedorov, P. Filippakopoulos, J. P. Hunt, E. A. Lasater, G. Pallares, S. Picaud, C. Wells, S. Martin, L. M. Wodicka, N. P. Shah, D. K. Treiber, S. Knapp, *Nat. Chem. Biol.* **2014**, *10*, 305 – 312; b) B. K. Shoichet, *Nat. Chem. Biol.* **2014**, *10*, 244 – 245.
[42] B. J. Schmidt, J. A. Papin, C. J. Musante, *Drug Discovery Today* **2013**, *18*, 116 – 127.
[43] A. Mai, *ChemMedChem* **2014**, *9*, 415 – 417.
[44] J. A. Burkhard, G. Wuitschik, M. Rogers-Evans, K. Müller, E. M. Carreira, *Angew. Chem.* **2010**, *122*, 9236 – 9251; *Angew. Chem. Int. Ed.* **2010**, *49*, 9052 – 9067.
[45] A. F. Stepan, K. Karki, W. S. McDonald, P. H. Dorff, J. K. Dutra, K. J. DiRico, A. Won, C. Subramanyam, I. V. Efremov, C. J. O'Donnell, C. E. Nolan, S. L. Becker, L. R. Pustilnik, B. Sneed, H. Sun, Y. Lu, A. E. Robshaw, D. Riddell, T. J. O'Sullivan, E. Sibley, S. Capetta, K. Atchison, A. J. Hallgren, E. Miller, A. Wood, R. S. Obach, *J. Med. Chem.* **2011**, *54*, 7772 – 7783.
[46] H.-X. Dai, A. F. Stepan, M. S. Plummer, Y.-H. Zhang, J.-Q. Yu, *J. Am. Chem. Soc.* **2011**, *133*, 7222 – 7228.
[47] Glide version 6.1, Schrödinger, LLC, New York, NY, 2013. See also R. A. Friesner, R. B. Murphy, M. P. Repasky, L. L. Frye, J. R. Greenwood, T. A. Halgren, P. C. Sanschagrin, D. T. Mainz, *J. Med. Chem.* **2006**, *49*, 6177 – 6196.
[48] K. C. Nicolaou, F. Roschangar, D. Vourloumis, *Angew. Chem.* **1998**, *110*, 2120 – 2153; *Angew. Chem. Int. Ed.* **1998**, *37*, 2014 – 2045.
[49] D. J. Newman, G. M. Cragg, *J. Nat. Prod.* **2012**, *75*, 311 – 335.
[50] A. Bauer, M. Brönstrup, *Nat. Prod. Rep.* **2014**, *31*, 35 – 60.
[51] S. Wetzel, R. S. Bon, K. Kumar, H. Waldmann, *Angew. Chem.* **2011**, *123*, 10990 – 11018; *Angew. Chem. Int. Ed.* **2011**, *50*, 10800 – 10826.
[52] S. L. Schreiber, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 6699 – 6702.

[53] A. M. Thayer, *Chem. Eng. News* **2014**, *92*, 13 – 20, January 20 issue.
[54] R. V. J. Chari, M. L. Miller, W. C. Widdison, *Angew. Chem.* **2014**, *126*, 3872 – 3904; *Angew. Chem. Int. Ed.* **2014**, *53*, 3796 – 3827.
[55] K. Müller, C. Faeh, F. Diederich, *Science* **2007**, *317*, 1881 – 1886.
[56] a) P. Metrangolo, H. Neukirch, T. Pilati, G. Resnati, *Acc. Chem. Res.* **2005**, *38*, 386 – 395; b) *Halogen Bonding*: *Fundamentals and Applications* (Eds.: P. Metrangolo, G. Resnati), Springer, Berlin, **2008**.
[57] M. Z. Hernandes, S. M. T. Cavalcanti, D. R. M. Moreira, W. F. de Azevedo, Jr., A. C. L. Leite, *Curr. Drug Targets* **2010**, *11*, 303 – 314.
[58] K. C. Nicolaou, *Angew. Chem.* **2014**, *126*, 4822 – 4823; *Angew. Chem. Int. Ed.* **2014**, *53*, 4730 – 4731.
[59] K. C. Nicolaou, C. R. Hale, C. Nilewski, H. A. Ioannidou, *Chem. Soc. Rev.* **2012**, *41*, 5185 – 5238.
[60] S. A. Williams, W. C. Anderson, M. T. Santaguida, S. J. Dylla, *Lab. Invest.* **2013**, *93*, 970 – 982.
[61] L. M. Jarvis, *Chem. Eng. News* **2013**, *91*, 10 – 23.
[62] C. Klein, *Hematology* **2009**, 344 – 350.
[63] W. L. Jorgensen, *Angew. Chem.* **2012**, *124*, 11848 – 11853; *Angew. Chem. Int. Ed.* **2012**, *51*, 11680 – 11684.
[64] See: The Parmenides Eidos Visual Reasoning Language and the Cognitive Excellence Program at http://www.Parmenides-Foundation.org.
[65] B. S. Slusher, P. J. Conn, S. Frye, M. Glicksman, M. Arkin, *Nat. Rev. Drug Discovery* **2013**, *12*, 811 – 812.
[66] J. Jenkins, CDER New Drug Review: 2012 Update. *U. S. Food and Drug Administration*, http://www.fda.gov.
[67] H. Wild, D. Heimbach, C. Huwe, *Angew. Chem.* **2011**, *123*, 7590 – 7592; *Angew. Chem. Int. Ed.* **2011**, *50*, 7452 – 7453.
[68] H. Waldmann, *Angew. Chem.* **2012**, *124*, 6388 – 6389; *Angew. Chem. Int. Ed.* **2012**, *51*, 6284 – 6285.
[69] T. Hoffmann, R. Metternich, *Angew. Chem.* **2012**, *124*, 8800 – 8801; *Angew. Chem. Int. Ed.* **2012**, *51*, 8670 – 8671.
[70] S. Laufer, U. Holzgrabe, D. Steinhilber, *Angew. Chem.* **2013**, *125*, 4164 – 4168; *Angew. Chem. Int. Ed.* **2013**, *52*, 4072 – 4076.
[71] A. Mullard, *Nat. Rev. Drug Discovery* **2014**, *13*, 85 – 89.
[72] Center for Drug Evaluation and Research, 2011: Novel New Drugs. *U. S. Food and Drug Administration*, http://www.fda.gov.
[73] Center for Drug Evaluation and Research, 2012: Novel New Drugs. *U. S. Food and Drug Administration*, http://www.fda.gov.
[74] Center for Drug Evaluation and Research, Novel New Drugs: 2013 Summary. *U. S. Food and Drug Administration*, http://www.fda.gov.

