TIMELINE

A brief history of novel drug discovery technologies

Leland J. Gershell and Joshua H. Atkins

Laypersons, researchers and clinicians alike speak of the biotechnology revolution with excitement. Media coverage of new breakthroughs in medicine often have the public and the investment community on the edge of their seats, eager for the next blockbuster drug to cure everything from high cholesterol levels to cancer. In this perspective, we examine some of the more popularized and influential new technologies in drug discovery and assess their relative impact on the actual attainment of new therapeutics.

The past few years have seen the advent of numerous new technologies and paradigms for drug discovery (TIMELINE). Along with the emergence of each novel method came a promise to generate better drugs in greater numbers while containing or even reducing costs. These techniques have been pursued at least to some extent at most pharmaceutical companies. During the same timeframe, industrial management has come under great pressure from shareholders and other investors to increase the productivity of drug discovery and development to meet annual growth expectations. For many diseases, the most obvious approaches to cures have been tried and have often failed. The challenge now is for scientists to attack major diseases with fresh ingenuity. The broader swathes of intellectual property coverage around lead structures, as well as revenue losses from marketed drugs coming off patent, accentuate this demand. In order to obtain a competitive advantage, companies have invested in the newer discovery platforms. A few successes have indeed been registered for each of these technologies, but unrealistic early assessments on the impact of these innovations put forth by the financial community set the stage for sector-wide declines in life science equity valuations. Senior management and scientists in the industry might have also been overly optimistic, as few would now argue that the fruit of these newer efforts have met their initial expectations.

Power in (combinatorial) numbers

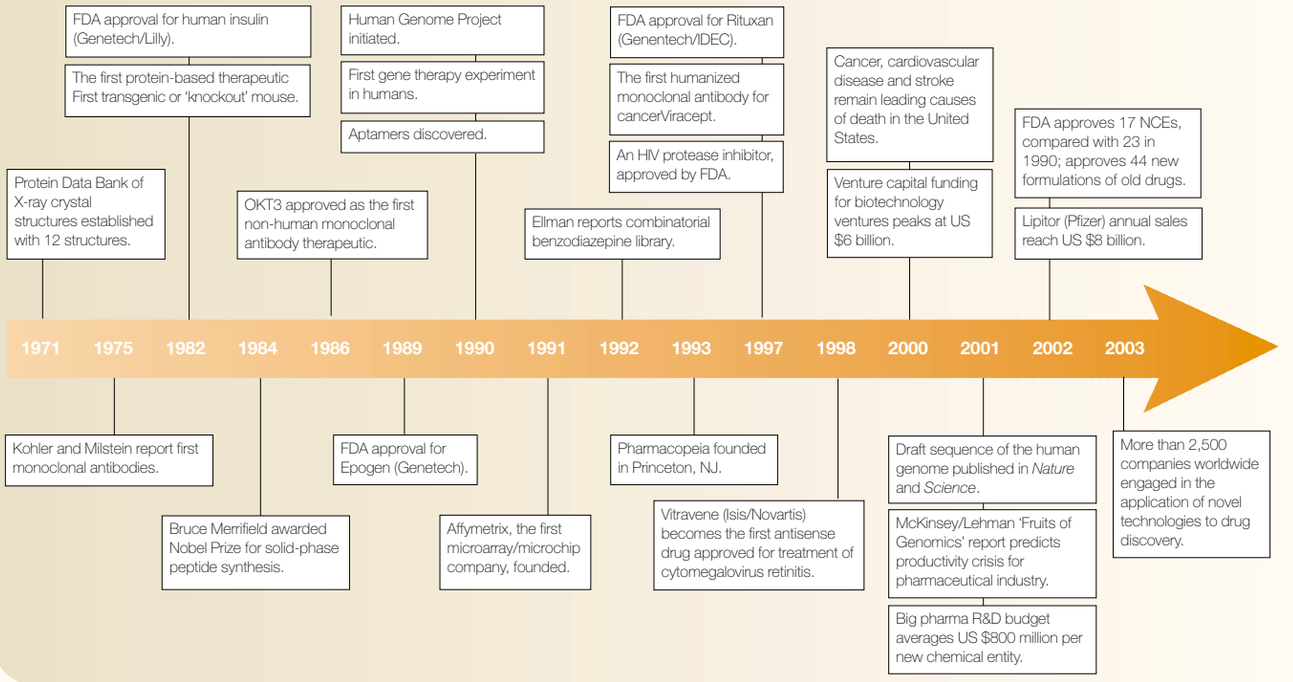
Pharmaceutical companies have traditionally relied on nature's bounty in their search for new medicinal agents. After selecting an agent from the archives of natural sources, for example the *Penicillium* mould from which penicillin was derived, scientists would embark on a time-consuming and often frustrating search for what Ehrlich called a "magic bullet". Positive activity, which could be represented by the ability to kill microbes or cancer cells for instance, would then be followed through serial purification of the crude extracts to give narrower and narrower isolates until the active principle was obtained. Once isolated, the promising leads were usually structurally complex, only available in small quantities, and expensive to purify and formulate into medicaments. These leads, usually against unknown targets and with unspecified mechanisms of action, were handed over to a team of medicinal chemists who worked to make simpler, more bioavailable compounds based on the core structure of the lead compound.

"For many diseases, the most obvious approaches to cures have been tried and have often failed. The challenge now is for scientists to attack major diseases with fresh ingenuity."

Medicinal chemistry is often viewed as a limiting factor in the creation of new drugs, as the steps involved in the preparation of just a single compound are labour-intensive. The bench chemist carrying out synthesis and purification could only be expected to produce a certain number of new compounds per year. For biochemists working with proteins, the widespread adoption of Merrifield's method for the linear solid-phase synthesis of peptides profoundly altered the nature of their work[1]. By attaching the nascent peptides to an insoluble resin, purification at each chemical step was simplified to mere rinsing, whereas the optimization of residue addition was facilitated by the use of excess reagents to increase yields. Following the development of automated peptide synthesizers, virtually any sequence could be produced in large quantities. The developments of parallel, and eventually split-and-pool, synthesis of peptides, as well as the application of this technology to nucleic-acid chemistry[2], enabled scientists to generate a much greater diversity of linear biopolymers and to keep time and labour to a minimum. Although some successes with these materials as clinical agents have been recorded, orally active, low-molecular-mass small molecules continue to be the preferred choice for drug molecules. In the interest of availing itself of greater collections of such compounds to test in assays, the pharmaceutical industry developed combinatorial synthesis approaches in which drug-like compounds with various rings and functional groups could be assembled in a rapid sequence. Since the first small-molecule combinatorial library based on a known drug scaffold was introduced in the early 1990s (REF. 3), thousands of unique libraries have been produced, mostly by companies that came to specialize in the business of producing combinatorial libraries, such as Pharmacopeia. In 1996, the CEO of Pharmacopeia, Joseph Mollica, predicted that with the new technology, chemists would boost productivity from tens of novel compounds to nearly 100,000 per year at a fraction of the original cost[4]. Since 1992, more than 1,250 combinatorial libraries have been described from both academic and industrial laboratories[5].

As the ability to generate compounds *en masse* with combinatorial methods became popular, the initial approach taken by most was to play a numbers game. Enormous investments were made by both large and small drug companies in order to purchase the specialized laboratory equipment for combinatorial synthesis. The initial, sizeable libraries of compounds were often generated as compound mixtures. These were then subjected to screens for biological activity with the hope that hits would be rapidly identified. These libraries provided researchers with millions of potential drugs, but their impact was limited by a frequent failure to identify active single entities following the de-convolution of mixtures exhibiting positive activity. A remedy was attempted by a shift towards the combinatorial synthesis of pure compounds, but this approach faced formidable challenges inherent to the chemistry involved. Natural products, such as alkaloids, macrolides, and tetracyclines,

Timeline | **Selected evolutionary landmarks in drug discovery**



which have traditionally been a fertile source of pharmaceutical leads, are characterized by complex, highly functionalized, polycyclic systems bearing multiple stereocenters. Their preparation by standard organic synthesis can require more than 20 chemical operations with almost as many purifications, which presents a daunting and inefficient task for the bench chemist. As chemists were eager to take advantage of the low-hanging fruit by making libraries guided by simpler chemistries, it is not surprising that the early libraries were composed of compounds that were not necessarily 'drug-like' in nature. Although the past two decades have seen huge advancements in methods for stereo- and enantioselective synthesis in solution, the adaptation of these methods to solid-phase chemistry has been slow, but is gradually taking hold[6]. At the same time, sophisticated techniques for stereoselective solid-phase oligosaccharide synthesis, such as those undergoing refinement at Optimer, bode well for the generation of therapeutics based on this biologically important class of molecules[7].

Chemical hurdles aside, a number of companies that tried the 'massive library' model soon abandoned it, often to re-orient their efforts toward generating smaller, focused libraries with permutations around a handful of promising scaffolds. These efforts have been aided by a better understanding of chemical versus biological structure spaces, which has generated information about biological targets that has enabled chemists to prepare libraries geared towards selected target classes[8]. Furthermore, a combinatorial approach to generating sub-libraries, in which pendant groups are selectively fixed or varied around a common core, might provide a multivariable dataset that can be evaluated during the screening process to reveal interdependent relationships of structural variations; this information can then guide subsequent hit-to-lead optimization steps[9]. This distinction is usually absent from individual experiments that examine only one parameter at a time. In this way, the mode of utility of combinatorial chemistry has shifted from 'lead discovery' to 'lead optimization'. A few companies are finding this approach productive, and now have selected compounds in clinical trials. It can be safely said that the overall quality of combi-chem libraries has improved appreciably, increasing the odds that hopes for new leads will be met with satisfaction.

The allure of the 'omics'

While the revolutions in chemistry were taking hold, advances in molecular and cell biology were also attracting attention in the pharmaceutical industry. In the late 1970s and early 1980s, recombinant DNA technology and the human genome were buzzwords, and companies such as Genentech and Amgen led the way in the nascent biotechnology market by applying the new techniques of molecular biology to basic problems in medicine. They pioneered the novel area of protein-based drugs, and developed highly successful business models on the simple scientific notion that replacing a protein deficient in a disease state would be of clinical benefit. The early successes of drugs such as Epogen and recombinant human insulin provided a tremendous stimulus to the biotech industry and lent strong support to the notion that drugs need not be small molecules, and could be made by cells in vats on a large scale without requiring tedious manipulation by chemists at the bench.

The success of the early protein-based drugs, combined with new developments in cancer biology as partially stimulated by an increased understanding of gene expression, led to the monoclonal antibody 'revolution'. The Genentech/IDEC launch of Rituxan in 1997 for non-Hodgkin's lymphoma left the biotech world and its investment community awe struck by the seemingly limitless potential of antibodies. The scientific basis of this idea was simple — proteins expressed specifically in certain diseases that were displayed on the surface of cells, or which circulated in the bloodstream, could be targeted by antibodies directed against them. These antibodies would act like a precision defence system to either neutralize the toxic molecules, or specifically deliver medications only to those tissues that bore them.

The principal technological advancement that permitted the use of antibodies as therapeutics related not to genomics, proteomics or combinatorial chemistry, but to monoclonal technology that enabled scientists to 'humanize' mouse or other animal antibody proteins so that they would not be recognized as foreign by the immune system[10]. Monoclonal antibodies had been produced and characterized since 1975 (REF. 11), and shortly after were in development as therapies to prevent transplant rejection. However, with the advent of the new humanization technology and its subtle variations, companies such as Abgenix, Medarex and Protein Design Labs took the biotech community by storm, with pharmaceutical partnerships and new prospects being announced on a frequent basis. In fact, in 2000 nearly 25% of all new drugs in development were believed to be monoclonal antibodies, with 90 monoclonal antibodies in clinical trials[12].

In 2003, monoclonals continue to have a high profile, but with the tempered optimism that usually settles on the biomedical community as the challenges of implementation become clearer. In late 2000, analysts could not stop talking about the upcoming billion-dollar drug Xolair — a monoclonal antibody targeted at IgE for asthma, in development by Genetech, Novartis and Tanox. By the summer of 2001, the FDA was requesting additional data, indicating that the potential for patient treatment, at least initially, would be less than that originally projected. Xolair remains in development, which began more than a decade ago, and it is unlikely to see the market before 2004.

The Human Genome Project, which was launched in the late 1980s, heralded a revolution in medicine, and provided further promises for new therapeutics. It was the start of a media frenzy, with a *New York Times* piece nearly 15 years ago suggesting that "in the not-so-distant future, we can expect to walk into a physician's office for an annual physical and walk out with a blueprint of our genetic inheritance — and with the knowledge of the most likely cause of our own death"[13]. As scientists, we have learned to temper the enthusiasm of the media with our own understanding of nature's devilish complexity and redundancy. At the same time, the surge of gene and protein data in the late 1990s led to the irresistible idea that once all of the disease targets were characterized, drugs for each would eventually follow suit. However, as target validation has lagged behind the greater access to potential targets, the chances for failure have actually increased. So, the importance of sufficiently validating of a target before embarking on a drug-discovery project must be underscored, and indeed the requirement for validation has been a driving force for the development of many new techniques in drug discovery. Targeted gene disruption in whole animals, such as knockouts or transgenic models, has been increasingly used over the past decade to determine the relevance of a gene implicated in a particular pathway or phenotype[14]. However, this method has been hindered in the past by a number of factors, including the time it takes to produce and study knockout or transgenic animals, as well as the possibility that compensatory mechanisms might alter the phenotype. Companies specializing in the transgene field have begun to approach the technology more aggressively. For example, scientists at Lexicon Genetics have identified 5,000 targets for which they plan to generate knockout

Box 1 | Human immunodeficiency virus — new science and new targets

As a group, the anti-human immunodeficiency virus (HIV) therapeutics offer a shining example of how the newer drug discovery techniques have borne great commercial success. These drugs have enabled the HIV-infected patient to transform what was essentially a death sentence into a manageable, long-term disease. Although drug resistance by the virus remains an issue, options for combination therapy against multiple viral proteins has further opened the window for intervention.

HIV protease was identified in the late 1980s as a potential therapeutic target. It functions in the assembly and maturation of virus particles, and its inactivation was found to lead to noninfectious virions. The crystal structures of the HIV proteases became available shortly after, both in isolation and in complexes with numerous inhibitors. Computational studies of these inhibitor complexes were reported in the early 1990s, and their modes of binding could be investigated with the aid of molecular mechanics.

At this time, the small companies Agouron and Vertex had become dedicated to developing drugs against HIV. Both of these companies had been established in the 1980s for the express purpose of creating small-molecule drugs through rational design. Techniques such as protein crystallography, nuclear magnetic resonance and computational biochemistry would be brought to bear on building compounds 'atom-by-atom' against protein targets in disease. Nelfinavir (Viracept; Agouron) gained FDA approval in 1997 for the treatment of HIV infection, and generated US $43.6 million in sales in its first quarter on the market. Two years later, amprenavir (Agenerase; Vertex/GlaxoSmithKline) received FDA approval. There are now six protease inhibitors on the market.

In 2003, the development of new drugs for HIV infection continues, and efforts are now directed towards additional targets, including HIV-1 integrase[33]. This enzyme allows HIV DNA to integrate into human cellular DNA. Obtaining a complete X-ray crystal structure of this protein has proved more difficult than for HIV protease, but with partial structures and computer models in hand, the industry is working to apply the principles of rational and structure-based drug design to this new challenge, with novel agents lurking on the horizon.

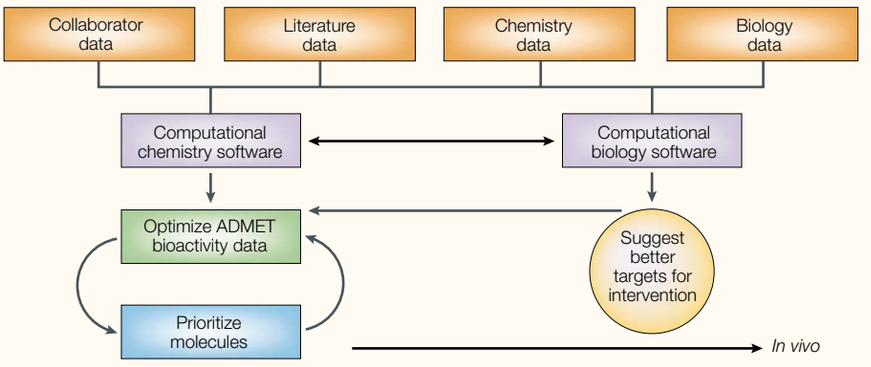


Figure 1 | **Virtual drug discovery.** Computational approaches represent a hopeful means to leverage biological and physico-chemical data to drive the structural optimization of compounds for their interaction with selected targets. Similarly, absorption, distribution, metabolism, excretion and toxicology (ADMET) algorithms could serve to weed out problematic candidates early on. Such screening of 'virtual compounds' is likely to be faster and cheaper than wet methods, a boon to an industry suffering rising costs of development. Adapted from Ekins, S., *et al.J. Comput. Aid. Mol. Des*. **16**, 381-401 (2002) © Kluwer Academic Publishers.

mice in which they can study mechanisms of disease and potential therapeutic candidates[15]. Moreover, the recent completion of the sequencing of the mouse genome might provide further insights into the functional differences between human and mouse systems, and could eventually enhance the transgenic approach to target validation and lead development.

Antisense technology, which offers a means to validate targets by preventing protein expression of a particular mRNA species, has been used with some success[16]. Newer antisense approaches, including the application of chimeric small nuclear RNA (snRNA) to control gene expression, are particularly promising[17]. Multigene disruption in single animals is an emerging technology that could augment the data gained from knockout models. In the ever-more-automated state of drug discovery technology, it is suspected that target validation will continue to remain a challenge.

One of the most realistic promises of the Human Genome Project has in fact been realized — a large increase in the number of potential drug targets. Indeed, one of the broadest changes in drug discovery efforts has been a shift towards an inductive model, in which the molecular pathways and entities believed responsible for a disease — or which at least have a causal role in the disease — form the starting point in the search for new drugs. We can expect that these heretofore unknown targets, following appropriate validation, will generate new perspectives on the treatment of diseases that have been resistant to therapies, such as neurological disorders and obesity.

Box 2 | **On the cutting edge — pharmacogenomics**

Nearly half all of drugs fail due to unforeseen toxicity or metabolism issues, and liver toxicity is the most common reason for FDA drug withdrawal or restriction (see table). Improved success in the late stages of drug development might lie in the ability to anticipate side-effects, interactions with other medications, and the structure and bioactivity of drug metabolites. As genetic variability is an important factor in determining the differences in drug responses among people, 'fingerprints' of individual human genomes might one day be used to tailor the selection of patients for clinical trials, or to select specific medications for an individual. The technology for assessing individual genetic variations in every aspect of disease, including metabolic processes, might also help drug companies to further study toxic side effects seen in a small minority of patients, thereby allowing the companies to refine dosing guidelines rather than withdrawing drugs from the market. Recently, UK-based Astex Technology reported the X-ray structure of the liver-based cytochrome p450 enzyme CYP3A4, which is believed to be responsible for the metabolism of at least 50% of therapeutics presently used. Structural knowledge of the enzymes involved in metabolism, and how they vary within populations, could inform drug discovery efforts so as to reduce the likelihood of untoward toxicity.

Drug	Company	Disease/indication	Toxicity	Action
Baycol	Bayer	High cholesterol levels	Rhabdomyolysis	Withdrawn 2001
Duract	Wyeth-Ayerst	Pain	Liver toxicity	Withdrawn 1998
Fen-phen	Wyeth-Ayerst	Obesity	Cardiac arrhythimia	Withdrawn 1998
Lotronex	GlaxoSmithKline	Irritable bowel syndrome	Ischemic colitis	Withdrawn 2000
Propulsid	Janssen	Abnormal gastrointestinal motility	Cardiac arrhythmia	Withdrawn 2000
Raplon	Organon	Anesthesia	Bronchospasm	Withdrawn 2001
Rezulin	Parke-Davis/ Warner-Lambert	Type II diabetes	Liver toxicity	Withdrawn 2000
Seldane	Hoechst	Allergy	Cardiac arrhythmia	Withdrawn 1998
Serzone	Bristol-Myers Squibb	Depression	Liver toxicity	Selective withdrawal 2003
Trovan	Pfizer	Anti-microbial	Liver toxicity	New restrictions
Zyflo	Abbott	Asthma	Liver toxicity	New restrictions

Today, however, the large majority of new drugs are still directed at a very small subset of targets — mostly G-protein-coupled receptors and other enzymes that have been well-characterized for years.

Can disease be rationally eradicated?

Although companies have been discovering drugs for over a century, it has only been in the past few decades that depictions of protein targets, sometimes with their cognate ligands bound, have become available. As the resolution of this three-dimensional information has improved greatly, companies have sought to 'rationally' design drugs from scratch that would selectively interfere with cellular function. A *Washington Post* article in 1988 carried the headline 'Computer is Drug Design's New Mortar and Pestle'[18] and highlighted new efforts in rational drug design, stimulated by research on drugs to treat infection with the human immunodeficiency virus (HIV). There was much talk of integrating chemists, X-ray crystallographers, cell biologists and others into teams that might expeditiously assimilate new data for the design of precisely targeted drugs for any disease in which a causative protein was implicated. Agouron, a small pharmaceutical company based in San Diego, demonstrated the elegance and utility of this approach with the approval of the HIV protease inhibitor nelfinavir (Viracept) in 1997. The drug was designed specifically to bind to a portion of the HIV protease enzyme, which was first characterized in 1987, using computer modelling against the X-ray crystal structure of the enzyme (BOX 1). It was a remarkable success and eventually led to the acquisition of Agouron by Pfizer. Unfortunately, outside of HIV protease inhibitors only a few other rationally designed small-molecule drugs have followed — most notably Relenza for the prevention of influenza infection. However, the simple notion that with the right technology and an integrated discovery team one could proceed *de novo* from gene, to protein and its three-dimensional structure, and finally to a drug, created a whirlwind in the pharmaceutical community. Almost all of the major pharmaceutical companies adopted this strategy to varied extents internally or in partnership.

The evolution of faster methods for generating protein structures has paved the way for companies to specialize in the high-throughput structural determination of target proteins with arrays of putative inhibitors. Two biotech ventures with high-profile management and ample capital have come to the fore — Structural Genomix and Syrrx. These

companies are leveraging new technology in rapid, high-throughput protein crystallization with advances in computer modelling programs, similar to those used at Agouron, to better understand binding modes and then to use this information to design superior ligands. It remains to be seen whether advances in biophysical chemistry are yet robust enough for these predictive technologies to translate generally into successful drug candidates. There already exist substantial databases of both public and proprietary crystal structures of validated clinical targets, and yet few drugs seem to be reaching the market as a result. One case in point is British Biotech's experience with marimastat, a potent small-molecule inhibitor of matrix metalloproteinases (MMPs), a class of enzymes known to be involved in a range of diseases from arthritis to cancer[19]. The drug failed numerous clinical trials despite the strong rational basis of its design. Moreover, other companies have continued to explore the fertile ground of MMP inhibitors, but to date, despite more than 15 available crystal structures, large numbers of preclinical and clinical candidates and tremendous resources, no drugs aimed toward these enzymes have successfully reached the market. This experience serves as a lesson that although structure-based design has evolved into a viable method for the discovery of inhibitors, there is no escaping the manifold physiological requirements put on a compound to qualify as a real drug. There is little doubt that rational drug design will remain an important force in drug development, but success in this effort still requires extensive empirical experimentation and remains subject to the surprises and challenges inherent in all approaches to drug development.

Virtual screening

The ability to model and predict the binding sites of proteins on the one hand, coupled with large databases of small-molecule compounds on the other, has encouraged the development of software that iteratively screens compounds *in silico* (FIG. 1). These tools have become more integral to the search for new drugs as computing power has increased and become cheaper. The potential of this approach has already been demonstrated by the identification of several inhibitors and antagonists. Although these efforts might serve to accelerate the preclinical drug discovery process, their usefulness remains limited by the predictive value of the compound 'descriptors', as well as by the differences between molecular structures that can be conceived versus those that can be readily synthesized. Although diversity is crucial, it is generally more desirable to design an 'informed' virtual library that contains synthesizable and drug-like compounds than it is to generate a library that maximally samples diversity space with billions of random compounds. This is in agreement with recent studies showing that within the current pharmacopeia, there are only about 10,000 drug-like compounds[20], and that the diversity of shapes in the set of known drugs is quite low[21]. Computational teams in pharmaceutical companies often embark on the search for drug candidates for molecular targets that already have established lead compounds; in this case, they are attempting to find structures that are sufficiently different so as to circumvent outside patents. Although appealing, virtual screens might be less productive than anticipated in this regard because the chances of finding a true lead — that is, a high-potency inhibitor — are far lower than those of simply finding a hit that represents a much earlier point in the optimization process. Moreover, efforts by some to mine compound databases for molecules that show similarity to the fingerprint of a given lead (but for which the two-dimensional representation is quite different, in terms of what is relevant to the nature of composition patents), are often met with disappointment in activity assays because too small a sample set of resulting compounds is tested. In any case, the great number of algorithms and software products available for virtual screening both attest to the hopes for this technology as well as to the fact that the pharmaceutical community has yet to settle on defined standards.

Siliconization and automation

In parallel with the development of rational drug design, combinatorial chemistry, and a variety of genome and proteome projects, there has been a growing fascination with silicon chips and robotics. This was stimulated, in part, by the increasing reliance of drug discovery on computer technology and simulation, but even more so by the need to screen the large number of drugs and targets created by other breakthrough developments. Silicon chips have provided some degree of automation and high-throughput and have also reduced the amounts of test reagents needed, saving both laboratory time and money. Investigations once considered to be unrealistic because of the sheer number of individual experiments that

Box 3 | Congestive heart failure — still looking for better drugs

Mortality rates for congestive heart failure (CHF) remain high, with most estimates reporting a mortality of 50% within five years of diagnosis. Early models of CHF suggested that a reduction in blood flow to the kidneys as a result of poor cardiac output induces compensatory mechanisms that retain salt and water. This yields a volume-overloaded state, with the result that peripheral vasoconstriction to maintain blood pressure becomes excessive. Two of the more commonly prescribed drugs approved to treat CHF — the β-adrenoceptor antagonist carvedilol (Coreg; GlaxoSmithKline) and the angiotensin-converting enzyme inhibitor enalapril (Vasotec; Merck) — were originally approved as antihypertensives. They do happen to antagonize the sympathetic and neurohormonal elements of CHF, however, and have proven beneficial for CHF patients. At the same time, our current understanding of CHF has matured into a complex, multidimensional model in which numerous targets have been identified for therapeutic intervention. Potential new therapeutics include endothelin-receptor antagonists, neutral-endopeptidase inhibitors (NEPIs), angiotensin-receptor antagonists and aldosterone antagonists[34]. Drugs in some of these categories are already on the market as antihypertensives, including valsartan (Diovan; Novartis), spironolactone (Aldactone; Searle) and bosentan (Tracleer; Actelion/Genentech). Bristol-Myers Squibb recently had the novel drug omapatrilat (Vanlev) in the NEPI class, but it was withdrawn from clinical trials due to unforeseen side effects. However, even if Vanlev had been approved, it had previously been shown that Vanlev was no more effective for CHF than established therapies. Drug manufacturers are now aggressively pursuing label expansions for existing antihypertensives to treat CHF as the potential market opportunities are enormous. Regardless, the fact remains that none of the above medications has been shown to substantially reduce the overall course of CHF in most patients. Unfortunately, heart transplantation remains the only definitive intervention.

Adenosine receptors are now CHF targets of current interest at companies such as Aderis Pharmaceuticals and CV Therapeutics, but it remains to be seen whether drugs directed towards the adenosine pathway will have a superior clinical effect. As the link between basic science and clinical reality is often tenuous, many hypotheses to explain CHF have ultimately failed when applied to the quest for better drugs.

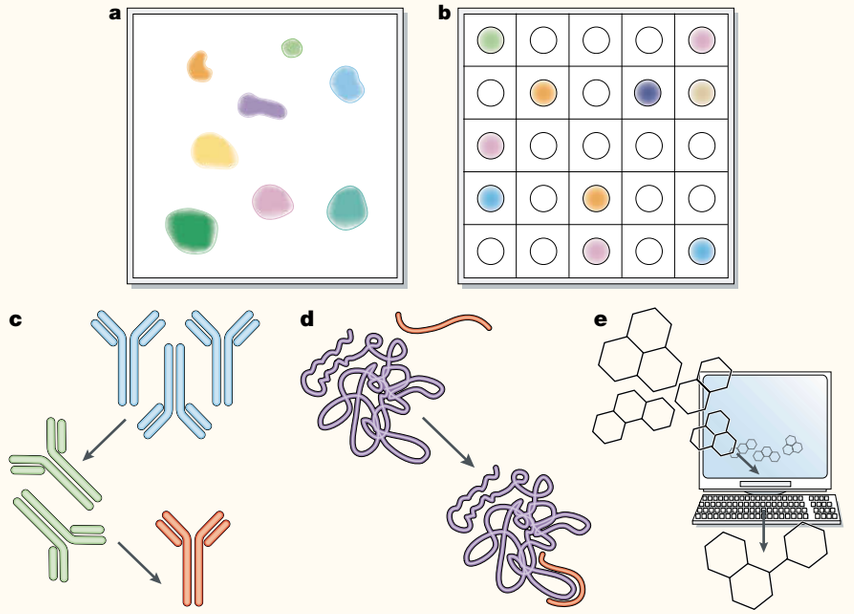


Figure 2 | **Emerging technologies and their developers. a** | Automated two-dimensional gel electrophoresis for the high-throughput purification of complex mixtures of proteins at very low concentrations[35] (Amersham Biosciences, Applied Biosystems and Millipore). **b** | Tissue microarrays[36], which are microchips coated with diverse types of normal and disease human tissues for the high-throughput validation of targets and the evaluation of efficacy, distribution and toxicity of drugs in the context of native complex biochemical milieu (Beecher Instruments, Biolog, Diomeda Life Sciences and Invitrogen). **c** | Directed evolution, which involves the engineering of protein/antibody therapeutics and microorganisms by controlled manipulation of cellular machinery through strategic, directed mutagenesis and chemical gene synthesis[37] (Applied Molecular Evolution, Maxygen and Morphotek). **d** | Aptamer technology. Aptamers are functional nucleic acids that bind specific regions of proteins to modulate their function *in vivo*, and have applications in target validation, therapeutics and diagnostics[38] (Aptanomics, Archemix, Gilead Sciences and Somalogic). **e** | Predictive and *in silico* absorption, distribution, metabolism, excretion and toxicity platforms. Simulation software, microarrays, and bioinformatics mining and processing technologies are used to optimize the performance of clinical drug candidates in late-stage clinical trials and beyond[27] (Amedis, Arqule, Cyprotex, Schrödinger and Tripos).

would be involved are now carried out in days, or even hours. The original Affymetrix concept of diagnostic chips for every pathology has yet to be realized, but technologies such as GeneChip arrays have been generously applied, particularly in the area of differential nucleic-acid expression in health and disease[22,23]. The recent shift of focus to proteins, with the technological challenges inherent in the production of protein-, cell- and tissue-coated silicon chips, as well as in the optimization of sensitive and specific detection systems, has slowed the progress of this technology[24]. There also remain challenges in the area of quality control and correlation to *in vivo* data. At this stage, silicon technologies are primarily tools for generatin interesting data for early stage screening and mechanism-based studies, but as the technology comes of age there is tremendous potential for applications to every aspect of the development process[25]. To date, there are no drugs yet approved that have come directly out of merged efforts between combinatorial chemistry and high-throughput screening[9], although inhibitors to thrombin, the cathepsins, and p38 MAP kinase are presently in clinical trials.

The perfect drug that fails in the clinic

All too frequently, a compound that works well in the laboratory and in preclinical trials, and even in small patient samples, meets with disaster when applied to larger study groups (BOX 2). In fact, it has been estimated that at least 50% of development candidates fail because of problems with absorption, distribution, metabolism, elimination and toxicity (ADMET), and about half of drugs on the market suffer some kind of ADMET problem[26]. Even after many collective years of knowledge accumulated by scientists in the industry, it remains a real challenge to predict these pharmacokinetic properties from just the structure of a compound[27]. Given the disproportionate failure of drugs during clinical trials, a great deal of effort has focused on developing better predictive models of how compounds will act in the body. Ensuring that a compound meets ADMET standards can involve hundreds of tests in cell-based assays and animal models, some of which require large quantities of compound. The process represents a significant bottleneck, as combinatorial chemistry and high-throughput screening make available only small quantities of compounds for screening. One approach that has met with some success has been taken by Deltagen, and involves the use of transgenic mice bred to express specific genes involved in drug toxicity or, in the case of cancer, specific drug-resistance genes[28,29]. A number of tissue-based systems that use cell lines have been developed for the determination of absorption and distribution properties established (for example, Caco-2 for intestinal absorption), but there is still a dearth of assays for predictive toxicity. Also, data generation tends be relatively slow in comparison with the rate of compounds processed by high-throughput activity screens. This might be the main reason why ADMET profiling is not applied earlier in the drug discovery process, that is, before many compounds have been excluded based on activity alone. A more progressive approach has been the adoption of parallel processing, in which potential drugs are put through ADMET screens at every step of the development process to reduce surprises in the 'end-game'[30]. This is particularly applicable to the design of combinatorial libraries[15]. However, this approach demands that the basic parameters of drug metabolism be well-defined, which, generally, they are not, despite Lipinski's 'rule-of-five'[31,32].

The cutting-edge solution to this problem is the emergence of ADMET prediction software as developed by companies such as ArQule and Schrödinger. Their products are based on predictive models that allow the virtual screening of compounds according not only to parameters such as membrane permeability and lipid solubility, but to the actual chemistry of metabolism of the cytochrome p450 system. Although highlighted by the press as a new drug development technology, these programs are essentially an outgrowth of quantitative structure–activity relationship (QSAR) approaches that have been around for some 30 years and to which the latest computational methods, as well as accumulated basic and clinical knowledge, are now being applied. In any case, these promising methods often suffer from poor validation and inaccuracies, although access to larger databases of known compounds and their

properties has allowed the software developers to offer greater reliability in their latest releases[32]. Overall, the most successful technologies will provide *in silico* predictive information that will identify the most effective balance among therapeutic efficacy, potency, bioavailability and toxicity.

Conclusions

The past two decades have witnessed remarkable progress in our understanding of the genetic and mechanistic basis of disease, the central importance of intercellular communication in the disease process and the potential success of highly targeted therapeutics. There has been both a rapid explosion in the amount of data available to researchers trying to create new drugs, and a parallel rise of new technologies in screening and automation to allow scientists to manage and sort the incoming data. In the course of achieving these goals, new challenges have been identified. In particular, the frequent and highly unpredictable failure of drugs in human trials is increasingly recognized as a roadblock to greater productivity in drug development. The full impact of the numerous genome and proteome projects has yet to be felt with regard to the long-promised development of genotype- or phenotype-specific therapeutics, and predictive technologies for drug disposition in the body remain at a primitive level. Overall, there are indications that productivity in the pharmaceutical industry is decreasing, that recent technologies have not reduced the cost of drug discovery and that the means for creating data have outpaced the ability to interpret and apply the data. Indeed, to date, the biotech revolution has made only a tangential contribution to the management and treatment of the major diseases facing society, such as congestive heart failure (BOX 3), cancer, asthma and sepsis.

As with many developments in medicine, optimistic predictions and hyperbole contribute to the development of unreasonable expectations for new therapeutics. Despite substantial technological advancement, however, it remains difficult to imagine a day without the spectre of diseases such as cancer, congestive heart failure and Alzheimer's disease. A number of newer technological avenues, having already shown some promise, are now being travelled with hopes for success (FIG. 2). Although we might look forward to groundbreaking therapies in the future, as a general rule, experience wins over expectation.

AGW BioStrategy, 55 West 26 Street, Suite 10D, New York, New York 10010, USA. Correspondence to L.J.G. e-mail: consultant@agwgroup.com

doi:10.1038/nrd1064

1. Merrifield, R. B. Solid-phase syntheses (Nobel lecture). *Angew. Chem*. **97**, 801–812 (1985).
2. Elmblad, A. *et al.* Synthesis of mixed oligodeoxyribonucleotides following the solid phase method. *Nucleic. Acids Res*. **10**, 3291–3301 (1982).
3. Bunin, B. A. & Ellman, J. A. A general and expedient method for the solid-phase synthesis of 1,4-benzodiazepine derivatives. *J. Am. Chem. Soc.* **114**, 10997–10998 (1992).
4. Day, K. Testing a combination punch against disease. *The Washington Post (Washington)* (April 14, 1996).
5. Dolle, R. E. Comprehensive survey of combinatorial library synthesis: 2000. *J. Comb. Chem*. **3**, 477–517 (2001).
6. Arya, P. *et al.* Toward high-throughput synthesis of complex natural product-like compounds in the genomics and proteomics age. *Chem. Biol*. **9**, 145–156 (2002).
7. Sears, P. & Wong, C.-H. Toward automated synthesis of oligosaccharides and glycoproteins. *Science* **291**, 2344–2350 (2001).
8. Wess, G. How to escape the bottleneck of medicinal chemistry. *Drug Discov. Today* **7**, 533–535 (2002).
9. Auld, D. S. *et al.* Targeting signal transduction with large scale combinatorial collections. *Drug Discov. Today* **7**, 1206–1213 (2002).
10. Hughes, D. Therapeutic antibodies make a comeback. *Drug Discov. Today* **3**, 439–442 (1998).
11. Kohler, G. & Milstein, C. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature* **256**, 495–497 (1975).
12. To, A. *et al.* Monoclonal antibodies: present and future. *Credit Suisse First Boston Corporation* (2000).
13. Schmeck, Jr, H. M. Gene studies emerging as key engine of science. *The New York Times* (September 6, 1988).
14. Debouck, C. & Metcalf, C. The impact of genomics on drug discovery. *Annu. Rev. Pharmacol. Toxicol.* **40**, 193–208 (2000).
15. Zambrowicz, B. P. & Sands, A. T. Knockouts model the 100 best-selling drugs — will they model the next 100? *Nature Rev. Drug. Discov*. **2**, 38–51 (2003).
16. Schmitt, C. A. & Lowe, S. W. Apoptosis and chemoresistance in transgenic animal models. *J. Mol. Med.* **80**, 137–146 (2002).
17. De Angelis, F. G. *et al.* Chimeric snRNA molecules carrying antisense sequences against the splice junctions of exon 51 of the dystrophin pre-mRNA induce exon skipping and restoration of a dystrophin synthesis in Δ 48–50 DMD cells. *Proc. Natl Acad. Sci. USA* **99**, 9456–9461 (2002).
18. Specter, M. Computer is drug design's new mortar and pestle. *Washington Post* (December 17, 1988).
19. Whittaker, M. *et al.* Design and therapeutic application of matrix metalloproteinase inhibitors. *Chem. Rev*. **99**, 2735–2776 (1999).
20. Lipinski, C. A. Drug-like properties and the causes of poor solubility and poor permeability. *J. Pharmacol. Toxicol. Methods* **44**, 235–249 (2000).
21. Bemis, G. W. & Murcko, M. A. Properties of known drugs. 1. Molecular frameworks. *J. Med. Chem*. **39**, 2887–2893, (1996).
22. Golub, T. R. *et al.* Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **286**, 531–537 (1999).
23. Scherf, U. *et al.* A gene expression database for the molecular pharmacology of cancer. *Nature Genet*. **24**, 236–244 (2000).
24. Kodadek, T. Protein microarrays: prospects and problems. *Chem. Biol*. **8**, 105–115 (2001).
25. Gerhold, D. L. *et al.* Better therapeutics through microarrays. *Nature Genet*. **32**, 547–551 (2002).
26. Hodgson, J. ADMET — turning chemicals into drugs. *Nature Biotechnol*. **19**, 722–726 (2001).
27. Ekins, S. & Rose, J. *In silico* ADME/Tox: the state of the art. *J. Mol. Graph. Model* **20**, 305–309 (2002).
28. Schmitt, C. A. & Lowe, S. W. Apoptosis and chemoresistance in transgenic animal models. *J. Mol. Med.* **80**, 137–146 (2002).
29. Zhang, J. *et al.* Modulation of acetaminophen-induced hepatotoxicity by the xenobiotic receptor CAR. *Science* **298**, 422–424 (2002).
30. Tarbit, M. H. & Berman, J. High-throughput approaches for evaluating absorption, distribution, metabolism and excretion properties of lead compounds. *Curr. Opin. Chem. Biol.* **2**, 411–416 (1998).
31. Lipinski, C. A. *et al.* Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **23,** 3–25 (1997).
32. Beresford, A. P. *et al.* The emerging importance of predictive ADME simulation in drug discovery. *Drug Discov. Today* **7**, 109–116 (2002).
33. Craigie, R. HIV integrase, a brief overview from chemistry to therapeutics. *J. Biol. Chem*. **276**, 23213–23216 (2001).
34. Klapholz, M. Emerging therapies in heart failure. *Cardiol. Spec. Ed*. **7**, 71–74 (2001).
35. Michels, D. A. *et al.* Fully automated two-dimensional capillary electrophoresis for high sensitivity protein analysis. *Mol. Cell. Proteomics* **1**, 69–74 (2002).
36. Perkel, J. M. Tissue microarrays: advancing clinical genomics. *The Scientist* **21**, 39 (2002).
37. Cornish, V. & Tao, H. Y. Milestones in directed enzyme evolution. *Curr. Opin. Chem. Biol.* **6**, 858–864 (2002).
38. James, W. Nucleic acid and polypeptide aptamers: a powerful approach to ligand discovery. *Curr. Opin. Pharmacol*. **1**, 540–546 (2001).

Acknowledgements

The authors wish to acknowledge helpful discussions with the associate consultants and advisors of AGW BioStrategy.

Online links

DATABASES

The following terms in this article are linked online to:

Cancer.gov: http://www.cancer.gov/cancer_information/
Non-Hodgkin's lymphoma

LocusLink: http://www.ncbi.nlm.nih.gov/LocusLink/
P38 MAP kinase

Online Mendelian Inheritance in Man:
http://www.ncbi.nlm.nih.gov/Omim/
Asthma | Congestive heart failure

FURTHER INFORMATION

5Z.com: http://www.5z.com

Bio.com: http://www.bio.com

The National health Museum:
http://www.accessexcellence.org

Access to this interactive links box is free online.