

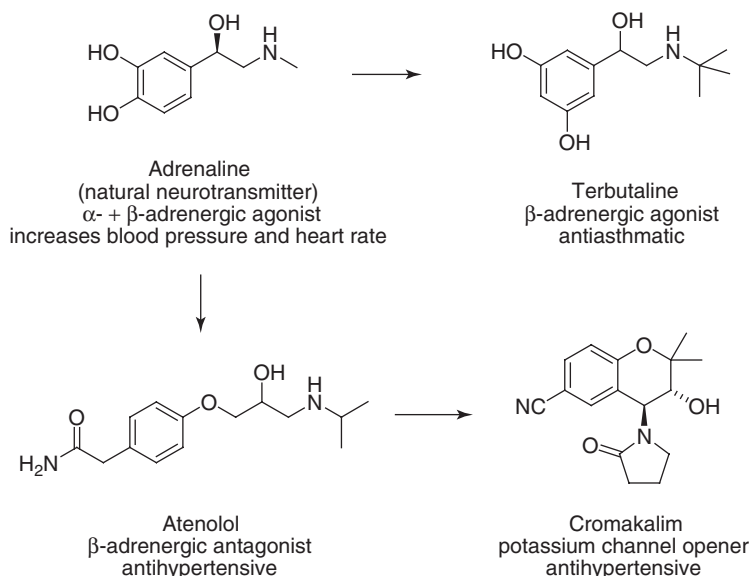
Part I

Introduction

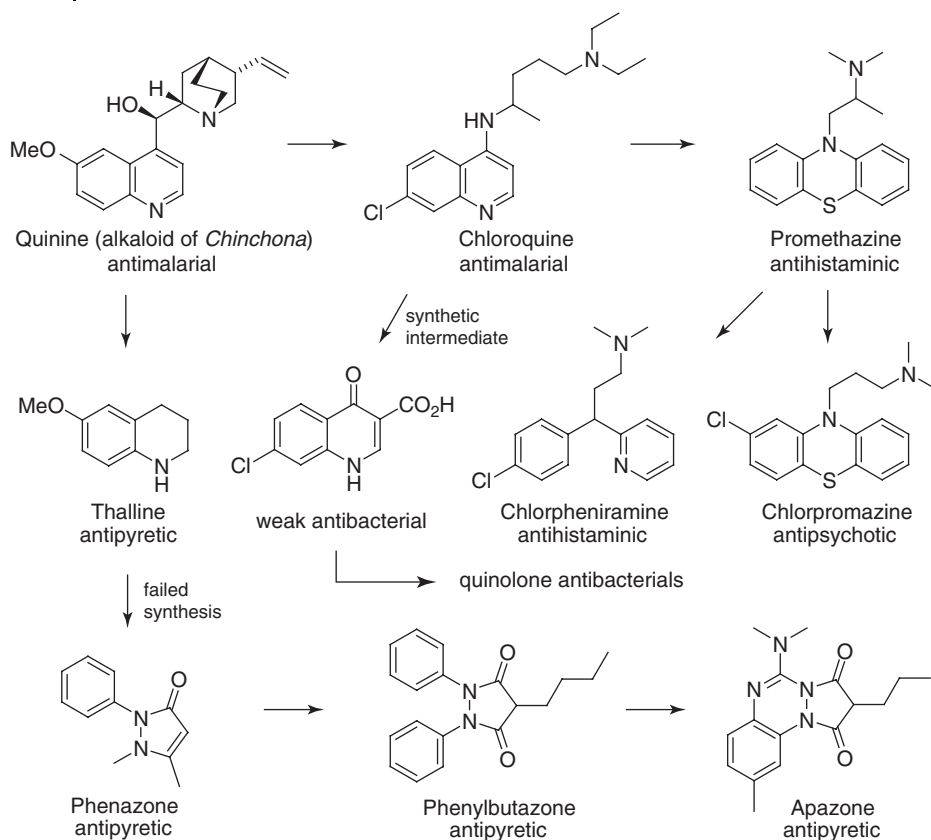
1

The Drug Discovery Process

Drugs are compounds that interact selectively with certain proteins in the human body and, thereby, suppress or activate biochemical pathways or signal transmission. Although the structures of modern drugs hardly allow to guess their origins, these were mostly natural products, discovered empirically, and used for centuries [1]. While synthesizing and evaluating new structural analogs of known hormones or natural drugs, new therapeutic applications often emerged. A fruitful starting point for the development of new drugs has in fact often been an old drug [2]. Illustrative examples of “drug evolution” are shown in Scheme 1.1.



Scheme 1.1 Examples of the development of new drugs from old drugs.



Scheme 1.1 (Continued).

To initiate a drug discovery program with hormones, natural products, or old drugs as leads has a number of advantages: the biochemical concept is already proven (the compound “works”), the target is “druggable,” and, importantly, the lead structure has acceptable or at least promising PK/ADME (pharmacokinetics/absorption, distribution, metabolism, and excretion) properties; otherwise, it would not work.

High-throughput screening of compound collections only rarely provides leads for new drugs [3]. Notable exceptions include the dihydropyridine calcium channel blockers, benzodiazepines, and sulfonamide antibacterials. These important drug classes resulted from testing drug-unrelated chemicals.

New proteins are constantly being discovered, and many of them are potential targets for therapeutic intervention. It must be kept in mind, though, that only few drugs have been successfully developed from scratch, starting with a biochemical hypothesis. The odds for succeeding are higher with a lead that works *in vivo* or by optimizing or exploiting a side effect of an old drug.

Today, the development of a new drug usually comprises the following stages:

1) **Discovery**

A target protein is selected, for example a receptor or an enzyme. If no lead structure is available, high-throughput screening of suitable, leadlike compounds may yield some weak ligands (hits). Systematic structural modification of these, supported by *in vitro* assays, may provide a lead, that is, a compound with an unambiguous dose–response relationship at the target protein. Further structural modifications aim at improving potency at the target, selectivity (low affinity to other proteins), water solubility, pharmacokinetics (PK, oral bioavailability, half-life, CNS penetration, etc.), therapeutic index (LD_{50}/ED_{50}), and general ADMET properties (absorption, distribution, metabolism, excretion, and toxicity). In addition to *in vitro* assays, the medicinal chemist will need guidance by more time- and compound-consuming *in vivo* pharmacology.

2) **Preclinical development**

Once a development candidate has been chosen, the following steps will be initiated: chemical scale-up of the selected compound, formulation, stability studies, more detailed metabolic, toxicological, and PK studies (only in animals, typically rodents, dogs, pigs, or primates).

3) **Clinical development**

Phase I	Healthy volunteers; determination of the suitable dose for humans
Phase II	First studies with a small group of patients
Phase III	Extended clinical trials

Thus, the fateful selection of a development candidate, which will either fail or succeed during the expensive preclinical or clinical development, is already taken in the discovery phase. Success in drug development is, therefore, primarily dependent on the medicinal chemist, on his ability to design and prepare the compound with the desired biological properties. No matter how hardworking and talented the members of the preclinical and clinical development team are, if the medicinal chemist has not delivered the right compound, the whole project will fail. And the later the recognition of the failure, the larger the costs.

Thus, pharmaceutical companies should allocate significant resources to the hiring and training of their medicinal chemists.

1.1

Pharmacokinetics–Structure Relationship

During the discovery phase of a new drug, two different, mutually independent sets of properties of the compound must be optimized: (i) potency and selectivity at the

target protein and (ii) ADME and toxicity. The most critical ADME/PK parameters are as follows:

- **Plasma half-life ($t_{1/2}$):** The time required for the plasma concentration of a drug to drop by 50%. A constant half-life means that the rate of elimination of a drug is a linear function of its concentration (first order kinetics). This is never exactly the case, and $t_{1/2}$ will usually increase as the concentration of the drug declines.
- **Oral bioavailability (F):** The fraction of a drug that reaches systemic circulation after oral dosing. Oral bioavailability is determined by dividing the area under the curve (AUC) for an oral dose by the AUC of the same dose given intravenously. A low F means that either the drug is not absorbed from the gastrointestinal (GI) tract or that it undergoes extensive first-pass metabolism in the liver.
- **Plasma protein binding (pb):** Hydrophobic compounds will bind unspecifically to any hydrophobic site of a protein. For this reason, high-throughput screening often yields hydrophobic hits (which are difficult to optimize and should be abandoned). Plasma proteins, such as albumin or α -glycoproteins, may also bind to drugs and thereby reduce their free fraction in plasma, their renal excretion, their ability to cross membranes (also the blood–brain barrier (bbb)), and their interaction with other proteins (metabolizing enzymes, the target protein). Binding to plasma proteins also prevents highly insoluble compounds from precipitating upon iv dosing and helps to distribute such drugs throughout the body.

The half-life of peptides may be increased by preventing their renal excretion through enhanced binding to albumin. This can be achieved by acylating the peptide with fatty acids or other plasma protein binding compounds.

Plasma protein binding is usually determined by equilibrium dialysis or ultrafiltration. Both techniques exploit the ability of certain membranes to be permeable to small molecules but not to proteins or protein-bound small molecules. The clinical relevance of plasma protein binding has been questioned [4].

- **Volume of distribution (V):** Amount of drug in the body divided by its plasma concentration. The volume of distribution is the volume of solvent in which the dose would have to be dissolved to reach the observed plasma concentration. Compounds with small volumes of distribution (i.e., high plasma concentration) are often hydrophilic or negatively charged molecules that do not diffuse effectively into muscle and adipose tissue. Compounds strongly bound to plasma proteins will show small volumes of distribution as well. Hydrophobic and/or positively charged molecules, however, readily dissolve in fat and interact strongly with the negatively charged cell surfaces (phospholipids) and, often, have large volumes of distribution (i.e., low plasma concentrations). Experimental and computational methods have been developed to estimate the volume of distribution in humans [5].

The typical volumes of body fluids are as follows:

Plasma	0.04–0.06 l kg ⁻¹
Blood	0.07 l kg ⁻¹
Extracellular fluid	0.15 l kg ⁻¹
Total body water	0.5 l kg ⁻¹

The “ideal” volume of distribution depends on the targeted half-life and pharmacological activity. For antibiotics or antivirals directed toward intracellular pathogens, a high tissue distribution (large volume of distribution) would be desirable. For short-acting anesthetics or antiarrhythmics or for compounds with a low safety margin, smaller volumes of distribution may enable a better control of drug plasma levels.

- **Clearance (CL):** The rate at which plasma is freed of drug, the remainder of the drug diluting into the freed volume. If a constant concentration C of a drug is to be attained, the infusion rate must be $CL \times C$. CL is related to other PK parameters: $CL = \text{dose}/AUC = \ln 2 \times (V/t_{1/2})$. The liver blood flow in humans is about 25 ml min⁻¹ kg⁻¹.

With the aid of high-throughput *in vitro* assays, potency and selectivity at a given target protein can often be rapidly attained. The most promising strategy for ligand optimization is to start with a small, hydrophilic (i.e., leadlike, *not* druglike) compound identified by screening leadlike compounds at high concentrations (leads are usually smaller, more hydrophilic, and less complex than drugs and show lower affinity to proteins). Then, potency and selectivity is enhanced by systematically introducing lipophilic groups of different shapes at various positions of the hit [6]. Numerous examples of such optimizations have been reported; three examples are shown below (Scheme 1.2 [6c, 7], see also Chapter 72; discovery of losartan).

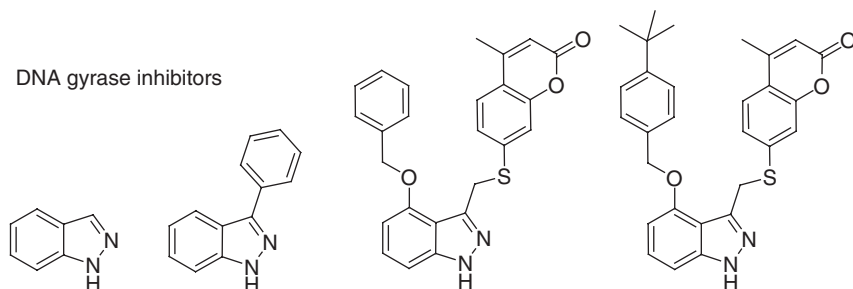
Unfortunately, the structure–activity relationship (SAR) resulting from such examples is of little use for medicinal chemists, as these only hold for specific target proteins and lack general applicability.

Far fewer examples have been reported of the optimization of PK and ADME because this requires large amounts of compound and tedious, expensive *in vivo* assays. Because few compounds reach the clinic, human data are scarce.

The ADME properties of xenobiotics are always determined by their interaction with the same set of proteins. Therefore, the study of PK–structure relationships should provide the medicinal chemist with valuable general guidelines for the optimization of leads. Structural features that strongly modulate ADME will always do so, no matter what the target protein of the drug might be.

Therefore, a thorough understanding of ADME– and PK–structure relationships should greatly facilitate the design of high-quality development candidates.

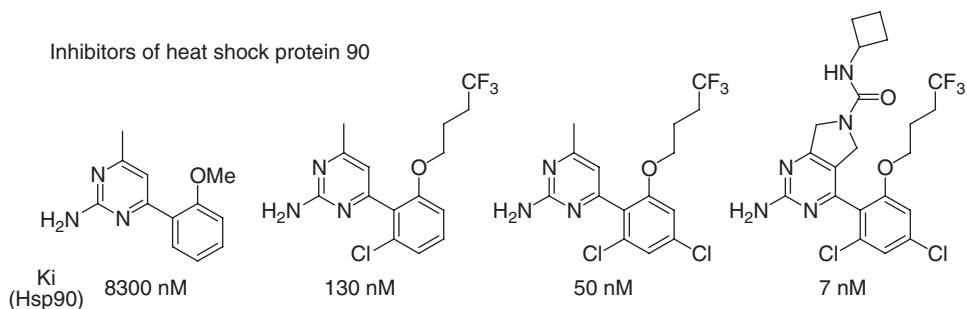
DNA gyrase inhibitors



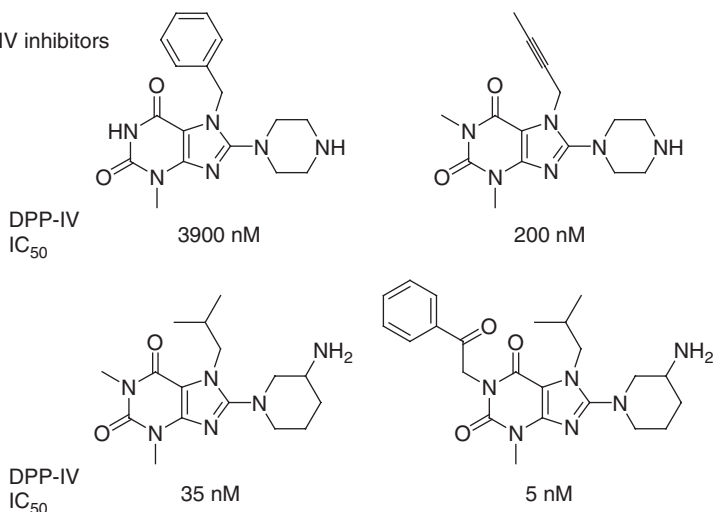
Maximal noneffective concentration:

>250 $\mu\text{g ml}^{-1}$ 8 $\mu\text{g ml}^{-1}$ 0.25 $\mu\text{g ml}^{-1}$ 0.03 $\mu\text{g ml}^{-1}$ $(K_D = 10 \text{ mM})$

Inhibitors of heat shock protein 90



DPP-IV inhibitors

**Scheme 1.2** Examples of the enhancement of binding affinity of small molecules to specific proteins.

1.2

The Future of Small-Molecule Drugs

The number of new drugs reaching the market has steadily declined in recent decades, despite steeply increasing R&D expenditures. Thus, while 53 new molecular entities were approved by the US Food and Drug Administration (FDA) in 1996, only 17 attained approval in 2007. Moreover, many “new drugs” are not really new but just resolved racemates, metabolites, prodrugs, or new formulations of older drugs. Thus, the proton pump inhibitor nexium, which ranged among the top 10 best-selling drugs worldwide in 2006, is just resolved omeprazole, launched 1988, and thus no true innovation.

The reasons for this decline in productivity are manifold:

- 1) Because we have gained a deep understanding of biochemistry and pharmacology, many drug discovery programs start with a target protein and a hypothetical biochemical mechanism. This requires the development of a potent, selective ligand before any proof-of-principle is possible. Because of the high complexity and redundancy of biochemical pathways, however, most hypothetical therapeutic principles will not work in animals.
- 2) Instead of testing compounds directly in animals, we rely too much on *in vitro* assays. Earlier, the discovery of new drugs was mainly based on *in vivo* assays, as only few *in vitro* assays were available. An *in vivo* assay will only yield a positive result if compounds are sufficiently soluble and lipophilic to reach the target protein and are neither metabolized too rapidly nor removed by active transport mechanisms and if the whole therapeutic concept works. Toxicity and unwanted side effects will also be rapidly recognized in an *in vivo* assay. Thus, by testing compounds directly in animals, efficacy data and a large amount of additional information are obtained. In fact, many older drugs were successfully launched before their biochemical mode of action was understood, and many successful drugs, for example, neuroleptics such as top-selling olanzapine or aripiprazole are highly unselective and would probably have been discarded by modern screening plans.

To limit oneself to *in vitro* assays or to neglect disappointing *in vivo* results (“a more potent compound should do it”) can quickly cost a lot of time, which may have been spent on a more promising project. If a compound cannot reach its target or if the underlying biochemical hypothesis is flawed, an increase in potency or selectivity will not do the job.

- 3) Advances in molecular modeling and the availability of X-ray structural analyses of proteins cocrystallized with small ligands have ignited further hopes for “*in silico*” drug discovery. Results up to now have mostly been disappointing. Molecular modeling may be a useful source of inspiration for medicinal chemists and has on occasions succeeded to improve potency and selectivity of a ligand, but the large number of failures remains unreported [8]. Although proteins are flexible and readily change their conformation and adapt to new ligands, most current molecular modeling packages keep the protein

structure fixed while docking new ligands [9] and do not automatically include water or ions in the docking process [10]. Even if faster computers and more sophisticated software would overcome these limitations, that would not be sufficient. The prediction of solubility, ADME properties, and chemical stability would also be required [11]. Progress in this field is rapid, but there is still a long way to reliable *in silico* drug development.

- 4) Development time and costs have increased significantly, because more potential side effects of drugs are continuously being discovered, requiring the identification of much more selective compounds than earlier. While the discovery of paroxetine (1992, Ferrosan) only required the synthesis and testing of ~130 compounds to optimize few parameters, today much larger numbers of compounds are prepared in most discovery projects because a larger number of parameters must be optimized. Between identification of a drug and FDA approval, long development times are often required: taxol, 1971–1992; omeprazole, 1979–1988; fluoxetine, 1972–1985 [12]. The normal time of patent protection (20 years) is, therefore, inadequate for drugs. Unless the time of patent protection for drugs is extended, prices of new drugs will continue skyrocketing.
- 5) The advent of parallel synthesis in the mid-1990s has enabled medicinal chemists to prepare many compounds quickly, using unsophisticated chemistry. Little time is left, however, to consider test results carefully and to design and prepare the right compound, no matter how difficult or automatable its synthesis. Despite the large numbers of test compounds prepared, the output of new drugs keeps falling.
- 6) Computers have invaded our offices and laboratories. Instead of training their key skills, chemists and technicians are wasting their time learning how to use databases, virtual screening software, electronic laboratory journals, and trying to keep up with the ever-growing flood of irrelevant emails. In some companies, chemists and technicians are even burdened with accounting tasks and Six Sigma drills. With all these distractions, no sustained focus is possible any longer, and we are losing our expertise in classical medicinal chemistry and organic synthesis. Today, too few hours are spent in front of the hood, and too few experiments are performed [13].

In their essay “The role of the medicinal chemist in drug discovery – then and now” [14], Lombardino and Lowe suggest that the efficiency of drug discovery may be improved by (i) performing more *in vivo* assays earlier in the projects; (ii) by including in each discovery team a “drug champion,” that is, an older, experienced scientist to provide historical perspective (“institutional memory”) and background information (this book can help); and (iii) by continuous training of young medicinal chemists. I would also propose a return to less molecular modeling, less software but more thinking instead, and less parallel synthesis, in particular during the lead optimization phase, where not chemistry but only structure–property relationships should be the driving force of the project.

For a number of devastating diseases (cancer, multiple sclerosis, bacterial infections with antibiotic-resistant strains, Alzheimer’s disease, etc.), no cure is

available today. Governmental “regulation” of drug prices is forcing pharmaceutical companies to focus only on high-margin areas but will ultimately preclude any commercial development of new drugs. If politicians believe that medical treatment must be free of charge then the government will also have to provide the new drugs. Only the taxpayer can be forced to finance loss-making ventures forever. However, do politicians really have incentives to improve and extend the life of pensioners even further?

The development of new drugs requires a huge and sustained effort and will only occur in a free society with strong economic incentives and a strong protection of individual property rights. As our democracies continue to degenerate toward socialistic kleptocracies, and governmental size, harassment, and plunder expand, the number of new innovative drugs will keep shrinking [15].

References

1. Newman, D.J. (2008) Natural products as leads to potential drugs: an old process or the new hope for drug discovery? *J. Med. Chem.*, **51**, 2589–2599.
2. Wermuth, C.G. (2004) Selective optimization of side activities: another way for drug discovery. *J. Med. Chem.*, **47**, 1303–1314.
3. Proudfoot, J.R. (2002) Drugs, leads, and drug-likeness: an analysis of some recently launched drugs. *Bioorg. Med. Chem. Lett.*, **12**, 1647–1650.
4. Rolan, P.E. (1994) Plasma protein binding displacement interactions – why are they still regarded as clinically important? *Br. J. Clin. Pharmacol.*, **37**, 125–128.
5. (a) Hollósy, F., Valkó, K., Hersey, A., Nunhuck, S., Kéri, G., and Bevan, C. (2006) Estimation of volume of distribution in humans from high throughput HPLC-based measurements of human serum albumin binding and immobilized artificial membrane partitioning. *J. Med. Chem.*, **49**, 6958–6971; (b) Lombardo, F., Obach, R.S., DiCapua, F.M., Bakken, G.A., Lu, J., Potter, D.M., Gao, F., Miller, M.D., and Zhang, Y. (2006) A hybrid mixture discriminant analysis–random forest computational model for the prediction of volume of distribution of drugs in human. *J. Med. Chem.*, **49**, 2262–2267; (c) Gleeson, M.P., Water, N.J., Paine, S.W., and Davis, A.M. (2006) In silico human and rat V_{ss} quantitative structure – activity relationship models. *J. Med. Chem.*, **49**, 1953–1963; (d) Lombardo, F., Obach, R.S., Shalaeva, M.Y., and Gao, F. (2002) Prediction of volume of distribution values in humans for neutral and basic drugs using physicochemical measurements and plasma protein binding data. *J. Med. Chem.*, **45**, 2867–2876; (2004) *J. Med. Chem.*, **47**, 1242–1250.
6. (a) Congreve, M., Chessari, G., Tisi, D., and Woodhead, A.J. (2008) Recent developments in fragment-based drug discovery. *J. Med. Chem.*, **51**, 3661–3680; (b) Hajduk, P.J. (2006) Fragment-based drug design: how big is too big? *J. Med. Chem.*, **49**, 6972–6976; (c) Boehm, H. *et al.* (2000) Novel inhibitors of DNA gyrase: 3D structure based biased needle screening, hit validation by biophysical methods, and 3D guided optimization. A promising alternative to random screening. *J. Med. Chem.*, **43**, 2664–2674; (d) Maly, D.J., Choong, I.C., and Ellman, J.A. (2000) Combinatorial target-guided ligand assembly: identification of potent subtype-selective c-Src inhibitors. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 2419–2424; (e) Hajduk, P.J., Meadows, R.P., and Fesik, S.W. (1997) Discovering high-affinity ligands for proteins. *Science*, **278**, 497–499; (f) Hann, M.M., Leach, A.R., and Harper, G. (2001) Molecular complexity and its impact on the probability of finding leads for drug discovery. *J. Chem. Inf. Comput. Sci.*, **41**, 856–864; (g) Oprea, T.I., Davis, A.M., Teague, S.J., and Leeson, P.D. (2001) Is there a difference between leads and drugs? A historical perspective. *J. Chem. Inf. Comput. Sci.*, **41**, 1308–1315.

7. (a) Zehnder, L. *et al.* (2011) Optimization of potent, selective, and orally bioavailable pyrrolidinopyrimidine-containing inhibitors of Heat Shock Protein 90. Identification of development candidate 2-amino-4-(4-chloro-2-[2-(4-fluoro-1*H*-pyrazol-1-yl)ethoxy]-6-methylphenyl)-*N*-(2,2-difluoropropyl)-5,7-dihydro-6*H*-pyrrolo[3,4-*d*]pyrimidine-6-carboxamide. *J. Med. Chem.*, **54**(9), 3368–3385; (b) Eckhardt, M. *et al.* (2007) 8-(3-(*R*)-Aminopiperidin-1-yl)-7-but-3-ynyl-3-methyl-1-(4-methylquinazolin-2-ylmethyl)-3,7-dihydropurine-2,6-dione (BI 1356), a highly potent, selective, long-acting, and orally bioavailable DPP-4 inhibitor for the treatment of type 2 diabetes. *J. Med. Chem.*, **50**, 6450–6453.
8. Warren, G.L. *et al.* (2006) A critical assessment of docking programs and scoring functions. *J. Med. Chem.*, **49**, 5912–5931.
9. Cozzini, P. *et al.* (2008) Target flexibility: an emerging consideration in drug discovery and design. *J. Med. Chem.*, **51**, 6237–6255.
10. Michel, J., Verdonk, M.L., and Essex, J.W. (2006) Protein–ligand binding affinity predictions by implicit solvent simulations: a tool for lead optimization? *J. Med. Chem.*, **49**, 7427–7439.
11. Selick, H.E., Beresford, A.P., and Tarbit, M.H. (2002) The emerging importance of predictive ADME simulation in drug discovery. *Drug Discov. Today*, **7**, 109–116.
12. Wong, D.T., Perry, K.W., and Bymaster, F.P. (2005) The discovery of fluoxetine hydrochloride (prozac). *Nat. Rev. Drug Discov.*, **4**, 764–774.
13. Kubinyi, H. (2003) Drug research: myths, hype and reality. *Nat. Rev. Drug Discov.*, **2**, 665–668.
14. Lombardino, J.G. and Lowe, J.A. III (2004) The role of the medicinal chemist in drug discovery – then and now. *Nat. Rev.*, **3**, 853–862.
15. (a) Ridley, M. (2010) *The Rational Optimist. How Prosperity Evolves*, Fourth Estate, London; (b) Bastiat, F. (1996, first publishing date: 1845) The physiology of plunder, *Economic Sophisms*, Irvington-on-Hudson, The Foundation for Economic Education, Inc., New York <http://www.econlib.org/library/Bastiat/basSoph5.html#S.2>, Ch.1.