
1

DRUG DISCOVERIES BY GENE FAMILY

KAREN E. LACKEY

I am fascinated by the notion that a small molecule with a fraction of the molecular weight and size of a comparatively enormous biological protein can change the activity of that protein in directed ways. For years, the scientific community dedicated to drug discovery has developed a knowledge base of overwhelming proportions to hone these interactions of small molecules to seek out the intended target and seemingly avoid all other proteins in its path. We expect these molecules to enter the body in a convenient manner, interrupt the disease pathology, and quietly exit the body upon job done. While there are certainly successes out there, some of which are described in this book, there are so many more unmet medical needs that require our sense of urgency.

When the human genome was solved, the expectations were that the targets to affect, avoid, or modulate human diseases would be tackled with drugs emerging at an unprecedented rate. Sometimes, I believe, the human genome project merely uncovered just how much we do not know and how far we need to go to understand the role of genes and protein products in normal and diseased tissues. Approximately 22,000 genes can be organized into the major protein classes they code and that are related by common structural and protein sequence features, called gene families, some of which are depicted in Fig. 1.1 (Hopkins and Groom, 2002).

1.1 GENERAL DRUG DISCOVERY COMPONENTS

Some areas of drug discovery methods transcend all gene families and will be briefly described here for general background to the subsequent chapters that are

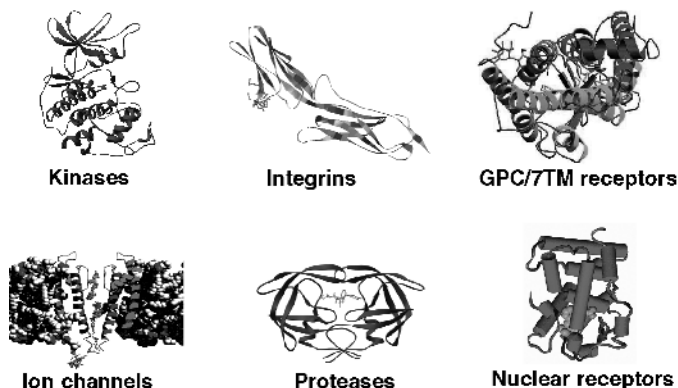


FIGURE 1.1 The gene families of proteins are classified by function and common structure motifs as can be seen by the representative structures for kinases, integrins, GPCRs, ion channels, proteases, and nuclear receptors. (See the color version of this figure in the Color Plates section.)

dedicated to the larger gene families. Broadly speaking, genomics is the study of all genes in an organism including the sequence, structure, function, and regulation. When referring to genetics, it is typically the inherited variation of genes that is being considered and is one component of genomics. Figure 1.2 summarizes the gene expression process with the general concepts of molecular biology where the rapid technology advances have made a major impact over the last few decades providing significant value to drug screening (Watson et al., 2003). In the nucleus, the DNA is transcribed into messenger RNA, which is transported to the cytoplasm. Translation occurs when transfer RNA (tRNA) reads the genetic code where three base pairs code for each amino acid of a polypeptide chain that forms the protein

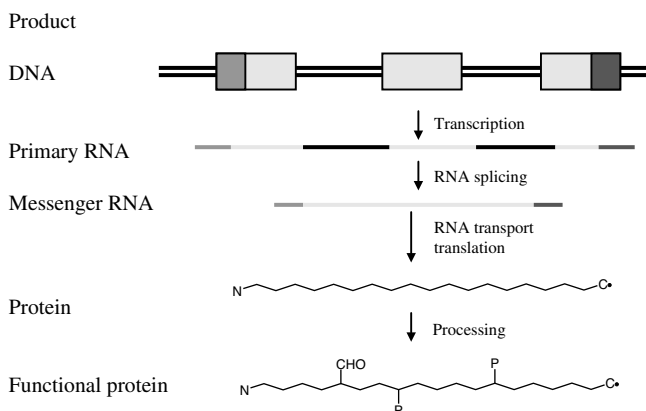


FIGURE 1.2 The normal process and products of gene expression.

product. Protein products are the biological machinery where the normal functions are carried out, such as activation, inactivation, scaffolding, and signaling in a healthy cell. Drug intervention can occur at any part of this process, as well as in the modulation of the activity of the proteins produced.

The discovery of drugs for diseases is typically selected based on medical need and research feasibility. A biological target is a component within a normal functioning cell or tissue where aberrant activity has given rise to a disease. For example, a genetic instability leads to mutations in cell signaling that can cause cancer. The target would become the biological component of the cascade to stop the unwanted or unchecked signal. A thorough understanding of disease pathology and/or the underlying genetics underpins the successful selection of targets. In some cases, genetic information can be used in patient selection in a field called pharmacogenomics. Whether the target is selected via genetic factors or by determining an ideal intervention point in the cellular protein activities, the ability to modulate the intended target using a small molecule is called tractability or druggability. The assessment of the tractability is shown in an oversimplified scheme in Fig. 1.3. Only a portion of the protein products of the human genome and infectious agents can be modulated by a small molecule or biological product (vaccine, antibody, etc.), and thus named the druggable genome (Imming et al., 2006). One way to increase the capability and understanding of biological target modulation involves organizing the protein products of the gene into families based on similar structure and function. This strategy allows teams of scientists to characterize the biology of the drug targets and build a knowledge base relevant to diseases that can readily be transferred from one member of the gene family to another.

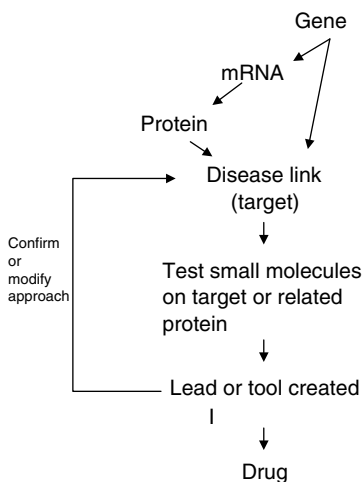


FIGURE 1.3 The process of target selection and validation is linked with assessing tractability or druggability.

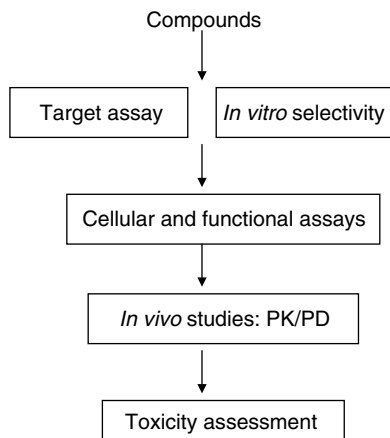


FIGURE 1.4 A generic compound evaluation pathway for a drug discovery project.

It seems that target validation (defined as the ability to effectively treat the disease via modulation of the intended biological target) is an ongoing process from the time of target selection all the way until the desired activity is observed in a patient. The drug discovery industry continuously challenges the data through a series of increasingly more complex biological evaluations. A compound evaluation pathway is typically used to assess the ability of the small molecules and to assist with decision making, as summarized in Fig. 1.4. The labels in the diagram also help with nomenclature used in the field to describe the stages of identifying small molecules that interact with the biological target of interest.

Treatment strategies from the molecular knowledge of a disease are used to prioritize targets and drug intervention points. Sometimes, it is necessary to create test compounds for multiple points in a particular disease pathology to determine the one that is the most tenable for a treatment of a patient. It could be that targeting one component of a pathway provides a more tractable target (i.e., a knowledge base exists for the gene family) or that using a test molecule in the assays that mimic the disease shows that it is insufficient to be an effective medicine. It still remains elusive to always predict which parts of a disease pathology are the most likely to be the best for producing the desired effects. Sometimes it is the wrong target, and sometimes the problem lies with compound's activity profile. In early stages of research when little is known about the disease at the molecular level, it is very difficult to discern between the two types of failures.

To begin the drug discovery process, the protein target is isolated and a screen is devised to test specific compounds, focused sets, and/or enormous sets of compounds of over 500,000 compounds often referred to as high-throughput screening or diversity screening. A compound must be potent toward the intended biological target, and the affinity is determined by an *in vitro* assay. *In vitro* assays are designed to measure antagonism (inhibition) of activity to stop an unwanted activity or agonism to provide more of an activity that has been lost or subdued by a

disease. There are many mechanisms that will be considered in the following chapters to achieve complete or partial antagonism or agonism that are specific to a gene family. Fundamental to all of them is a need to be able to assess the activity at the intended target in a way that allows a chemist to compare the inherent activity to either increase or decrease it depending on the drug intervention strategy. The screens often reflect the biological activity of the target, for example, by measuring the inhibition of an enzymatic reaction. Screens are also built to measure binding affinity—how tight the compound sticks to the target, with the assumption that by binding to certain parts of a protein, the target can be modulated. Some targets cannot be screened as an isolated assay and must be screened in a cellular setting. Throughout each of the following chapters, details regarding the type of assays that are used in each gene family of targets and reasons for the choices will be discussed.

The compound evaluation pathway is designed to enable decisions on what compounds to synthesize. Designing compounds based on the data that are generated in each of the assays used in a screening scheme involves computational chemistry, structural biology, *in silico* predictions, and mathematical analyses. Different gene families rely on specialized design tools in combination with physicochemical assays. The overall process is an iterative learning cycle in which medicinal chemists generate ideas for structural changes to simultaneously optimize activities across all of the assays employed. Chemists working on a gene family system use their knowledge of the target class assisted by computational techniques to design compounds to find biological activity. Sets of compounds, often called arrays, are created to test hypotheses regarding structure–activity relationships (SAR). Where there is more protein structure information or a long history of designing active compounds, the sets of compounds are generally small as the chemist hones in on the best molecule. When less is known or available to guide the design, much larger sets of compounds are generally synthesized (>500). Taking advantage of special pockets or distinct binding interactions can help to narrow the scope of what the drug will modulate. By focusing on gene families, methods to design selectivity for one family member compared with another are discussed. It is noteworthy that building in this selectivity still requires extensive toxicity testing to verify minimal off-target effects.

After a screening event, the data are analyzed to determine if any compounds provide the readout intended. The sets may be tested in a general format to cast as wide a net to find as many active starting points as possible. Testing many compounds at a single concentration is often used. Anything that provides a hint of activity is typically called an assay “hit.” The hits are analyzed to check that they are truly active and interesting to work on. They are often narrowed to only include chemical series that will be further worked on via synthetic modifications. Depending on the gene family, the hit validation can be as simple as generating a value that is called the effective dose to give a 50% response. The reason this value is used is because it makes comparing compounds much more reliable based on using a reproducible point on the curve. It is expressed as a concentration or a pXC_{50} (the negative log of the IC_{50} for better numeric comparisons of activity).

It is also important to understand the error range of the data. Often replicates are averaged to provide values for comparison among series of chemically related compounds and that gives a fairly good representation of what activity can be used for analyses. However, to assess how significant an improvement or decrease in activity is, it is important to get an understanding of the variability of the data. Often, pED_{50} values are useful for preventing overinterpretation of the data. For example, if you compare compound A with an IC_{50} value of 5 nM with compound B that is 20 nM, you might be tempted to say the compound A is fourfold more active than B. However, if the average range of the assay data on repeat tests is ± 0.2 log units, then the compounds would be considered equipotent. The data interpretation aspects are different among the various gene families, given the state of the art of the assays systems, and some discussion of the reliability of the data is covered in each chapter. Also, the term high-content lead series is often used to describe a chemical series that meets many of the criteria for modulating the target of interest and often requires the parallel use of several hit generation and follow-up techniques (Bleicher et al., 2003).

The activity against an isolated target in an *in vitro* assay is often verified or further evaluated in a cellular environment. In cases where it is not possible to create an assay for a desired target, whole cells or more complex systems are used to find active small molecules that incorporate the desired cellular outcome, or phenotype, as the assay system. These assays are typically lower throughput (less compounds are screened because the assay is either more difficult to run or too expensive) but more information rich. Also, it might be necessary to run control assays to ensure that the activity that is observed is related to the desired biological target modulation. For many drug discovery projects, it is also important to test for selectivity against other similar targets or important conflicting biological pathways. Each of the gene families has specific selectivity challenges that will be discussed for greater understanding of how to design a compound that is effective for the disease area without unwanted or intolerable side effects. For example, often screening is done on the human protein target, but the activity must also be checked on the animal orthologue (the protein found in the animal that is analogous to the human version) to support the *in vivo* animal studies of efficacy and toxicity.

Data generated from evaluating compounds in cell-based assays are important to integrate with any isolated enzymatic or binding assay. A compound's activity is affected by cellular context that includes multiprotein complexes, cell-type specific pathways, branching pathways, intracellular localization, and multiple target isoforms or states (phosphorylated, open, closed, etc.). Of course, compensatory bypass or feedback loop mechanisms cannot be detected in an isolated target assay and thus allow the cell assay to be used as a way to judge the tractability of the drug intervention point. In this way, cellular assays can better measure efficacy, not just potency. As a generalized screening tool, cell-based assays can also be set up to select for only permeable, nonmetabolized, and nontoxic compounds. This method of identifying hits and optimizing compounds can be useful in limiting the number of chemical series. On the contrary, it automatically eliminates potentially potent compound interactions that may need properties such as solubility or permeability

built in. Each target class strategy of molecular design balances these aspects of cellular screens based on the typical properties or needs of the drug discovery research efforts.

The SAR is the study of the correlation between the biological activity and the chemical structure. There can be SAR at every level of the screening cascade, and it guides the entire process of designing an effective medicine. The knowledge derived from SAR forms an understanding of interaction features such that predictions can be made on what kinds of molecules will retain activity and which ones will not within a chemical series. There are nearly always compromises that can be made in potency to add solubility and to create an improved therapeutic index (ratio of the dose required for biological efficacy to the dose at which unacceptable toxicity occurs). All along the project path, it is important to be aware of the SAR to know where on the molecule changes can be made to adjust for newly uncovered risks or to be aware of what cannot be changed without losing all target affinity.

This book should help the reader know what types of compounds to synthesize in each gene family, and examples are provided to understand the scope of the generic properties of the small molecules. Each family of targets appears to depend on a few major design techniques based on the properties of the biological targets. However, paramount to the success of creating a drug candidate with optimized gene family protein interactions is the ability to incorporate drug developability characteristics. Requirements of developability properties are different for the variety of intended routes of administration (e.g., oral, inhaled, and intravenous) and typically cover absorption, distribution, metabolism, and excretion, which are abbreviated as ADME properties. There is a whole branch of science supporting the necessary ADME features and components that affect them related to disease pathology and normal body functions. Although the details are beyond the scope of the target class design focus, these developability properties must be optimized while generating potent interactions with the intended biological targets.

Developability, metabolism, and pharmacokinetics (DMPK) encompass many types of assays that are used to determine the *in vivo* properties of potential drug molecule. While the emphasis in this book focuses on designing molecules for specific biological targets within gene families, the ultimate design of the compounds must take into consideration the properties that allow the compound to get to the tissue or site of action. *In vitro* tests to measure solubility at various pH levels can often be done to predict absorption levels or the handling properties of the drug product. The activity of compounds is measured in a variety of *in vitro* assays with CYP isoforms (e.g., CYP3A4 or CYP2D6) to predict which compounds might have better or worse *in vivo* properties based on predictions on metabolism of the parent molecule.

The principal evaluation of *in vivo* properties comes by testing representative compounds in rodent models looking for data to determine half-life, bioavailability, volume of distribution, and clearance of the compound. A thorough discussion of this topic is beyond the scope of the book, but it is important to note that each target class discussed may require certain types of compound classes for proper interactions to modulate the target. These features may offer particular challenges in

designing an effective medicine and will be pointed out in the chapters that follow. A high-level example, though, is to look at the properties of many drugs that modulate nuclear receptors. They tend to be lipophilic in nature and, in contrast, the kinase compounds that interact at the ATP binding domain are frequently “flat” heterocycles with inherently poor solubility. These features are addressed in drug discovery research projects and it helps to have a knowledge base in the area to facilitate the appropriate design focus.

More recently, *in silico* techniques have made huge advances in the area of predictive properties. Many software packages are available on the computer desktop for either straightforward analyses or more complex analyses for a trained computational chemist. They are used to identify potential issues with active series identified from a screening event to help prioritize the hits with the highest likelihood of generating high-quality compounds. The *in silico* methods are also used to prioritize a synthetic plan by taking chemistry that is designed to functionalize one or more sites on a central scaffold and create the structures by building the final products as a virtual set (existing only as an electronic collection of structures). The virtual compounds are then put through predictive models in order to select the subset of compounds that would be predicted to retain the desired activity. For example, if the intended target is in the central nervous system (CNS), then an *in silico* method can be used to virtually screen a large set of possible compounds for CNS penetration property predictions. The purpose of such an exercise is to reduce the number of compounds required for synthesis, thus reducing time, chemical reagent costs and waste, and biological screening costs. To predict oral absorption using *in silico* predictions, the focus is on oral bioavailability (how much drug reaches the blood compared to the original dose) and predicting interactions with metabolizing enzymes (e.g., enzymes in the liver designed to rid the body of toxins might degrade a compound rapidly or too slowly).

Virtual screening techniques are often used for the gene family target interactions. Chemical descriptors form the basis of the structure input and often account for the nature of the compounds, such as molecular size, polarity, hydrogen bonding capacity, lipophilicity, and acidity/basicity. As a medicinal chemist builds knowledge in each of the target classes, the data from virtual and measured screening are used together to form the project strategy in creating a drug.

Lipophilicity is measured by $\log P$, which is a partition coefficient that describes the ratio of the concentrations of a compound between two phases: octanol and water. Why this matters in drug design is due to the fact that the cell membrane permeability (ability of compounds to pass through the cell membrane to reach intracellular targets) requires more lipophilic nature, which has a tendency to be less soluble in blood/aqueous media. Conversely, hydrophilic compounds tend to be water soluble and amenable to blood transport, but are poorly cell permeable. This measure of lipophilicity requires a balance of design between the various effects to create a drug that gets to the site of action and retains the key functional groups necessary for target interaction for the desired biological effect. A calculated $\log P$ value, designated $\text{clog } P$, has become fairly easy to calculate and is often used to prioritize which of the compounds should be synthesized given the ideal range

of 2–4 (for oral drugs), where the larger the number the greater the lipophilicity (Wenlock et al., 2003).

Solubility is an important physical property for a drug for both biological effectiveness and drug product formulation and processing. Solubility testing can be done in a variety of conditions to mimic neutral aqueous, gastric fluid simulation or intestinal fluid simulation, or any pH necessary for drug delivery. For oral drugs need to be well absorbed by the body, unless they are designed to specifically combat a GI tract disease for which systemic exposure is unnecessary (e.g., certain bacteria infections of the gut). Understanding the solubility characteristics of series can also assist in interpreting the biological data from screening. Care should be taken to ensure that the compounds are fully solubilized in the test solution to study the SAR trends.

In silico methods such as structure-based design and pharmacophore modeling are best covered under the specific target class chapter. When the three-dimensional structure of an enzyme or receptor is known, the information can be used to determine the interactions between it and a bound compound. Computer models can be generated to mimic other similar molecules bound to the active site so that designing can be based on maximizing interactions (while taking into account all of the other properties). It is often referred to as finding where the molecule can tolerate a substitution that would allow the synthetic chemist to attach functional groups to improve the properties without destroying the positive interactions with the ligand/protein. Crystallography can be used to design ideal drug molecules and will be highlighted in the gene families where it currently has the greatest impact.

Often, the three-dimensional structure of a protein target is not known, but using binding or catalytic activity of compounds from screening events, a pharmacophore model can be built. A high-level explanation follows: A set of compounds that are known to have activity for the intended target can be used to gain insight into the steric and electronic structural features necessary for a compound to bind. Even if they are poorly active, a model can be built to identify the specific features that garner any activity. Iterative searching and testing in the *in vitro* assay allows the scientists to refine the pharmacophore model and hone in on the key features for better and better affinity. Also, models can be built based on the similarity of proteins or protein folds from one member of a protein family to another. Often, the models help scientists understand how changes in the small molecule affect the activity observed in the assay and designs incorporate better interactions through appending functional groups.

In typical drug discovery programs, assays and *in silico* predictions such as solubility, *in vitro* metabolic stability, and cell permeability are used as filters for determining ideal pharmacodynamic and pharmacokinetic properties (Leeson et al., 2004). There seems to be an ongoing discussion in the medicinal chemistry field over the value of rules in drug design, but most scientists agree that trends can be derived for general predictions (Leeson and Springthorpe, 2007). However, within a given chemical series, there can be anomalies and those should be considered when using assays for guidance in progressing compounds through more complex and resource-intensive biological evaluation systems. Specific examples of screens

for *in vitro* developability assays include solubility in multiple solvents (neutral aqueous, simulated gastric fluid, and simulated intestinal fluid), Caco-2 permeability, p450 enzyme assays, and microsomal metabolism studies using methods similar to published reports (Guo and Shen, 2004; Li 2004). It is invaluable to have training sets of compounds to investigate correlations between *in vitro* assays designed to predict drug-like characteristics and actual *in vivo* data within chemical series.

The entire discussion on protein family affinity and modulation coupled with incorporating the desired parameters for delivering the small molecule to the site of action is only a fraction of the drug discovery process. Most researchers will agree that the more the quality is put into the compounds in the early stages of identifying a potential drug candidate, the more likely it will be successful in reaching a patient in need. However, as can be seen by the remarkable financial investments (Goozner, 2004) made in the field, I still believe that the combination of knowing which intervention point to modulate (target selection and validation), achieving selective target interactions with no off-target activities, meeting all of the drug-like criteria for the intended route of administration, and being able to synthesize sufficient, pure quantities of material in the correct form is a formidable task to expect of one small molecule. However, there are success stories out there. We owe it to ourselves to understand how to modulate every potential biological target to enable successful drug discovery for the patients still in need.

1.2 FURTHER READING FOR EXPERT KNOWLEDGE

1. *Textbook of Drug Design and Discovery*, 3rd ed., edited by Povl Krosggaard-Larsen, Tommy Liljefors, and Ulf Madsen, Taylor & Francis, Inc., New York, 2002.

This book is a thorough guide to drug discovery with a compilation of chapters going into details on mechanics of small molecule–protein interactions, kinetics, and development of pharmacophore models and descriptors. A terrific explanation of quantitative SAR with physicochemical parameters and how to interpret the data is also included. A general explanation of receptor-based targets in multiple gene families is provided, with subsequent chapters discussing in more detail on ion channels and therapeutic outcomes from modulation of targets. Creating and interpreting data from imaging studies with labeled compounds provides insights into understanding drug distribution and why so much effort goes into understanding physicochemical properties of molecules. The chapter on enzymes with mechanism and kinetics of inhibition offers many drug discovery approaches to these important biological reactions. More advanced topics in drug design include prodrugs (a precursor to the active parent drug typically done to mask an unwanted property of the compound and dependent upon the human metabolism), peptides and their mimics, and the use of metals in modulating biological targets that are dependent on them for their normal processes. The textbook ends with specific examples of drug discovery in the research area of antiviral and anticancer agents where many of the concepts discussed throughout the

book are shown in practice of selecting targets in the disease pathology and determining how to modulate them for a beneficial effect.

2. *Biopharmaceutical Drug Design and Development*, edited by Susanna Wu-Pong and Yongyut Rojanasakul, Humana Press, Inc., Totowa, NJ, 1999.

It provides excellent descriptions of using biotechnology in generating very specific target modulation and the unique properties of these biologicals in creating useful drugs to combat diseases. It is an easy to read scientific book that offers a marked contrast to small molecule design in the gene family focused molecular design. Vaccines and recombinant human products, for example, have excellent specificity but offer difficult challenges to manufacture, deliver to the site of action in the body, and management of stability issues throughout the entire process of “bench to bedside.” Many technology advances have been made and continue to be made to provide effective medicines in this class of drugs, and this book provides a foundation of knowledge to appreciate the more expert literature.

3. *Guidebook on Molecular Modeling in Drug Design*, edited by N. Claude Cohen, Academic Press, San Diego, CA, 1996.

There are many computational approaches needed to operate in the potential chemical space that can be reached with synthetic chemistry to maximize interactions with the intended biological target(s). This guidebook delves into the details of the tools used with a variety of molecular capabilities and applies them to examples in ligand generation. The complexity of drug–biological target interactions that take into account pharmacophores, desolvation, and binding properties is discussed from the perspective of computer-assisted drug discovery and modeling. While the book is over 10 years old, it has a good foundation for anyone interested in understanding the breadth of contributions from computational chemistry to the design of drugs.

4. *Protein–Ligand Interactions. From Molecular Recognition to Drug Design*, edited by H.-J. Böhm and G. Schneider, in: *Methods and Principles in Medicinal Chemistry*, Vol. 19, edited by R. Mannhold, H. Kubinyi, and G. Folkers, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2003, ISBN 3-527-30521-1.

This book is a must-have expert reference for any serious medicinal chemist who wants to understand the intimate interactions of a small molecule and a large biological protein. From the list of abbreviations provided in the front of the book to the well thought-out order of chapters describing details of small molecule modulation properties, the readers will, at a minimum, understand the nomenclature of the business partners they work with and, at best, will utilize the appropriate principles that affect the SAR of the drug they are hoping to discover. Most of the formulas are broken down into the important features of the design tools, so even if you were not going to use the calculated approaches described, the lessons that are included add insights for the medicinal chemist. The combination of computational chemistry and compound screening explanations is useful for understanding all of the factors that go into compound design at the molecular recognition detailed level.

5. *Receptor-Based Drug Design*, edited by Paul Leff, Marcel Dekker, New York, 1998, ISBN 0-8247-0162-3.

This book is part of a long series of textbooks dedicated to drugs and pharmaceutical sciences. It is a useful reference book for a more in-depth look at a subset of the gene families where the biological target is a receptor. Details and insights are provided on the functions of a receptor such that thorough interpretation of assay systems can be achieved. Mechanisms of interactions range from determining binding affinity to understanding many forms of modulation. Understanding the outcome of the effect of small molecule on the receptor in an information-rich assay system should help guide the reader to more expert advice on judging the tractability of a receptor for drug discovery. While advances in assay systems for receptors have advanced over recent years along with a large increase in the number of ligand-bound protein crystal structures, this book will still provide expert details for more in-depth understanding of the reader's research project and current scientific literature.

6. *Analogue-Based Drug Design*, edited by Janos Fischer and C. Robin Ganellin, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, 2006, ISBN 3-527-31257-9.

This book is one in a series of six titles dealing with different aspects of drug discovery. It covers the general approach that a synthetic chemist can use to explore chemical space around an established active compound with the potential for differentiating the final product. Nineteen specific case studies are provided in detail to demonstrate how the approach works, the parameters used to differentiate analogues, and what the properties were of the final drug candidate.

7. *Structure-Based Drug Discovery*, edited by Harren Jhoti and Andrew Leach, Springer, Dordrecht, The Netherlands, 2007, ISBN 1-4020-4406-2.

Understanding the cellular proteins at the molecular level from three-dimensional protein structures allows researchers to design specific interactions to create potent, small molecule drug candidates. This book details the process improvements in protein crystallization and structure determination, and provides an advanced view of scaling the technologies for parallel production for higher throughput. A general guide to using structure in discovering small molecule fragments and strategies for converting the fragments into lead series is also included. Multiple methods of screening for fragments of drug-like molecules are covered including NMR and X-ray crystallography. An alternative to linking fragments together to increase affinity is discussed, whereby the screening event begins with finding the small core scaffold followed by functionalization to improve affinity. All these methods of finding ligand-efficient hits depend upon understanding the mechanisms of the interactions between the small molecules and the protein, so the chapter on biophysical methods is useful. The final chapters cover computational ways to extend the value of the structure-based design methods.

REFERENCES

- Bleicher KH, Boehm H-J, Mueller K, Alanine AI, 2003. A guide to drug discovery: hit and lead generation: beyond high-throughput screening. *Nat. Rev. Drug Discov.* 2(5):369–378.
- Goozner M, 2004. *The \$800 Million Pill*. University of California Press, Berkeley, CA.
- Guo Y, Shen H, 2004. pK_a , solubility, and lipophilicity: assessing physicochemical properties of lead compounds. *Optimization in Drug Discovery*. Humana Press, Totowa, NJ, pp. 1–17.
- Hopkins AL, Groom CR, 2002. The druggable genome. *Nat. Rev. Drug Discov.* 1:727.
- Imming P, Sinning C, Meyer A, 2006. Drugs, their targets and the nature and number of drug targets. *Nat. Rev. Drug Discov.* 5(10):821–834.
- Leeson PD, Springthorpe B, 2007. The influence of drug-like concepts on decision-making in medicinal chemistry. *Nat. Rev. Drug Discov.* 6(11):881–890.
- Leeson PD, Davis AM, Steele J, 2004. Drug-like properties: guiding principles for design—or chemical prejudice? *Drug Discov. Today: Technol.* 1(3):189–195.
- Li AP 2004. *In vitro* approaches to evaluate ADMET drug properties. *Curr. Top. Med. Chem.* 4(7):701–706.
- Watson JD, Baker TA, Bell SP, Gann A, Levine M, Losick R, 2003. *Molecular Biology of the Gene*, 5th ed. Addison Wesley, Boston, MA.
- Wenlock MC, Austin RP, Barton P, Davis AM, Leeson PD, 2003. A comparison of physicochemical property profiles of development and marketed oral drugs. *J. Med. Chem.* 46(7):1250–1256.

