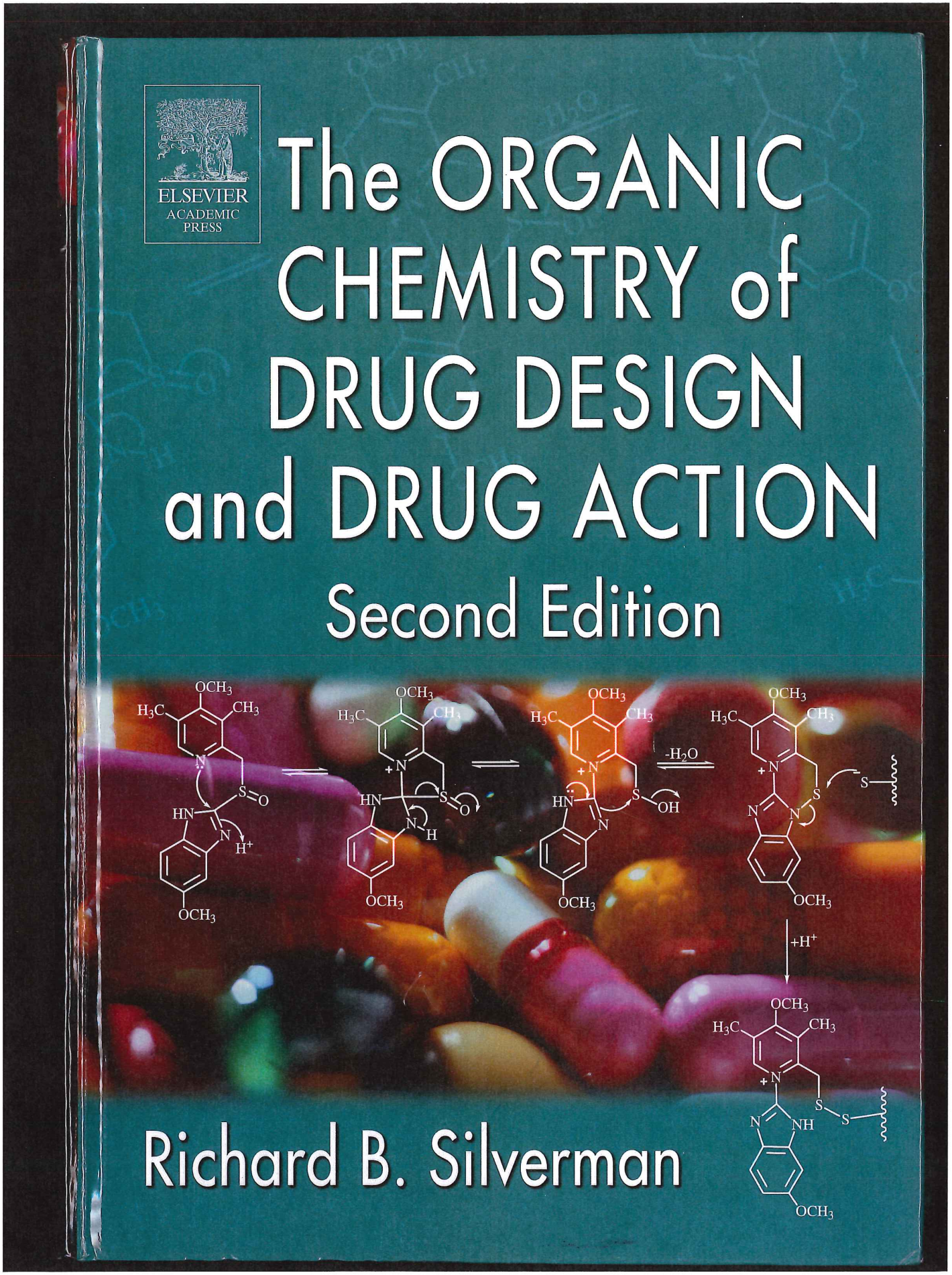


The Organic Chemistry of Drug Design and Drug Action

Second Edition

Richard B. Silverman
Northwestern University
Department of Chemistry
Evanston, Illinois

ELSEVIER
ACADEMIC
PRESS

Amsterdam • Boston • Heidelberg • London • New York • Oxford
Paris • San Diego • San Francisco • Singapore • Sydney • Tokyo

Senior Publishing Editor Jeremy Hayhurst
Project Manager Simon Crump
Editorial Assistant Desiree Marr
Marketing Manager Linda Beattie
Cover Design Eric DeCicco
Composition Newgen Imaging Systems (P) Ltd
Printer Maple-Vail Book Manufacturing Group

Elsevier Academic Press
200 Wheeler Road, Burlington, MA 01803, USA
525 B Street, Suite 1900, San Diego, California 92101-4495, USA
84 Theobald's Road, London WC1X 8RR, UK

This book is printed on acid-free paper. ♾

Library of Congress Cataloging-in-Publication Data
Application submitted

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 0-12-643732-7

For all information on all Academic Press publications
visit our Web site at www.academicpressbooks.com

Printed in the United States of America
04 05 06 07 08 9 8 7 6 5 4 3 2

Chapter 2

Drug Discovery, Design, and Development

2.1 Drug Discovery

Drug discovery is a very time-consuming and expensive process. Estimates of the average time required to bring a drug to the market range from 12–15 years at an average cost of about $800 million. For approximately every 10,000 compounds that are evaluated in animal studies, 10 will make it to human clinical trials in order to get 1 compound on the market. The clinical trials consist of three phases prior to drug approval: phase I (generally a few months to a year and a half) evaluates the safety, tolerability (dosage levels and side effects), pharmacokinetic properties, and pharmacological effects in 20–100 healthy volunteers; phase II (about 1–3 years) assesses the effectiveness of the drug, determines side effects and other safety aspects, and clarifies the dosing regimen in a few hundred diseased patients; and phase III (about 2–6 years) is a larger trial with several thousand patients in clinics and hospitals that establishes the efficacy of the drug and monitors adverse reactions from long-term use. Once the new drug application (NDA) is submitted to the Food and Drug Administration (FDA), it can be several months to several years before it is approved for commercial use. Phase IV studies are considered to be the results found with a drug that has already been allowed onto the drug market and is in general use. Drug candidates (or *new chemical entities*, NCE, as they are

often called) that fail late in this process result in huge, unrecovered financial losses for the company. This is why the cost to purchase a drug is so high. It is not that it costs that much to manufacture that one drug, but that the profits are needed to pay for all of the drugs that fail to make it to market after large sums of research funds have already been expended.

In general, drugs are not discovered. What is more likely discovered is known as a *lead* compound. The lead is a prototype compound that has a number of attractive characteristics, such as the desired biological or pharmacological activity, but may have other undesirable characteristics, for example, high toxicity, other biological activities, absorption difficulties, insolubility, or metabolism problems. The structure of the lead compound is modified by synthesis to amplify the desired activity and to minimize or eliminate the unwanted properties to a point where a *drug candidate*, a compound worthy of extensive biological, pharmacological, and animal studies, is identified; then a *clinical drug*, a compound ready for clinical trials, is developed. Prior to an elaboration of approaches to lead discovery and lead modification, two common drugs discovered without a lead are discussed.

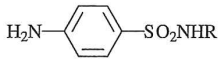
2.1.A Drug Discovery without a Lead

A.1 Penicillins

In 1928 Alexander Fleming noticed a green mold growing in a culture of *Staphylococcus aureus*, and where the two had converged, the bacteria were lysed.[1] This led to the discovery of penicillin, which was produced by the mold. Actually, Fleming was not the first to make this observation; John Burdon-Sanderson had done so in 1870, ironically also at St. Mary's Hospital in London, the same institution where Fleming made the rediscovery![2] Joseph Lister had treated a wounded patient with *Penicillium*, the organism later found to be the producer of penicillin (although the strains discovered earlier than Fleming's strain did not produce penicillin, but, rather, another antibiotic, mycophenolic acid). After Fleming observed this phenomenon, he tried many times to repeat it without success; it was his colleague, Dr. Ronald Hare,[3,4] who was able to reproduce the observation. It only occurred the first time because a combination of unlikely events all took place simultaneously. Hare found that very special conditions were required to produce the phenomenon initially observed by Fleming. The culture dish inoculated by Fleming must have become accidentally and simultaneously contaminated with the mold spore. Instead of placing the dish in the refrigerator or incubator when he went on vacation as is normally done, Fleming inadvertently left it on his lab bench. When he returned the following month, he noticed the lysed bacteria. Ordinarily, penicillin does not lyse these bacteria; it prevents them from developing, but it has no effect if added after the bacteria have developed. However, while Fleming was on vacation (July to August) the weather was unseasonably cold, and this provided the particular temperature required for the mold and the staphylococci to grow slowly and produce the lysis. Another extraordinary circumstance was that the particular strain of the mold on Fleming's culture was a relatively good penicillin producer, although most strains of that mold (*Penicillium*) produce no penicillin at all. The mold presumably came from the laboratory just below Fleming's where research on molds was going on at that time.

Although Fleming suggested that penicillin could be useful as a topical antiseptic, he was not successful in producing penicillin in a form suitable to treat infections. Nothing more was done until Sir Howard Florey at Oxford University reinvestigated the possibility of producing penicillin in a useful form. In 1940 he succeeded in producing penicillin that could be administered topically and systemically,[5] but the full extent of the value of penicillin

was not revealed until the late 1940s.[6] Two reasons for the delay in the universal utilization of penicillin were the emergence of the sulfonamide antibacterials (sulfa drugs, **2.1**; see Chapter 5, Section 5.4.B.2, p. 254) in 1935 and the outbreak of World War II.

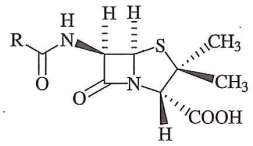


sulfa drugs
2.1

No studies related to the pharmacology, production, and clinical application of penicillin were permitted until after the war to prevent the Germans from having access to this wonder drug. Allied scientists who were interrogating German scientists involved in chemotherapeutic research were told that the Germans thought the initial report of penicillin was made just for commercial reasons to compete with the sulfa drugs. They did not take the report seriously.

The original mold was *Penicillium notatum*, a strain that gave a relatively low yield of penicillin. It was replaced by *Penicillium chysogenum*,[7] which had been cultured from a mold growing on a grapefruit in a market in Peoria, Illinois!

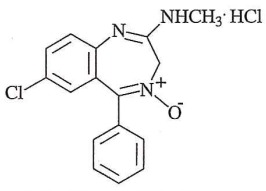
For many years debate raged regarding the actual structure of penicillin (**2.2**),[8] but the correct structure was elucidated in 1944 with an X-ray crystal structure by Dorothy Crowfoot Hodgkin (Oxford); the crystal structure was not actually published until 1949.[9] Several different penicillin analogs (R group varied) were isolated early on; only two of these early analogs (**2.2**, R = $PhOCH_2$, penicillin V; and **2.2**, R = CH_2Ph, penicillin G) are still in use today.



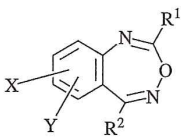
penicillin V (R = $PhOCH_2$)
penicillin G (R = CH_2Ph)
2.2

A.2 Librium

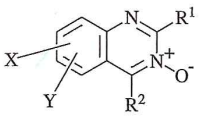
The first benzodiazepine tranquilizer drug, chlordiazepoxide HCl [7-chloro-2-(methylamino)-5-phenyl-3*H*-1,4-benzodiazepine 4-oxide; **2.3**; Librium], was discovered serendipitously.[10] Dr. Leo Sternbach at Roche was involved in a program to synthesize a new class of tranquilizer drugs. He originally set out to prepare a series of benzheptoxdiazines (**2.4**), but when R^1 was CH_2NR_2 and R^2 was C_6H_5, it was found that the actual structure was that of a quinazoline 3-oxide (**2.5**).



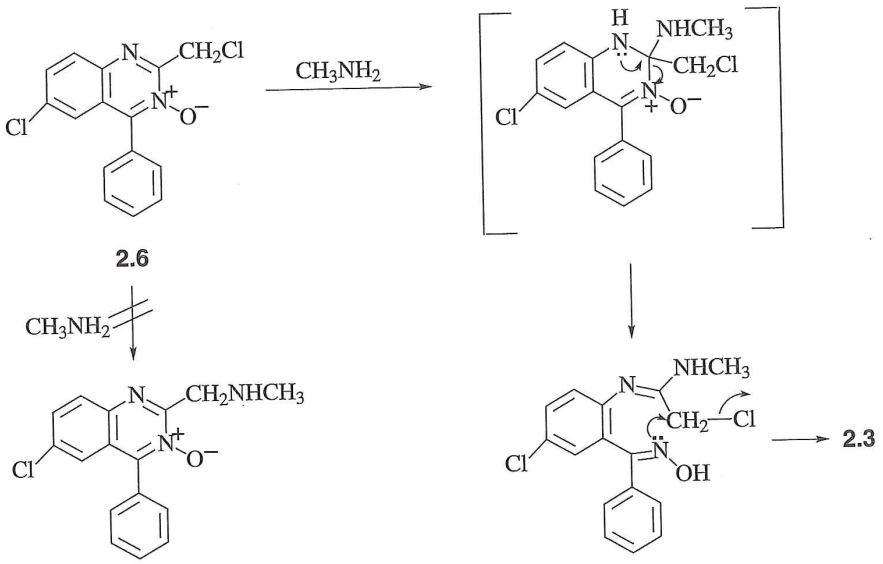
chlordiazepoxide HCl
2.3



2.4



2.5



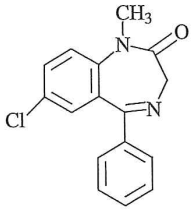
Scheme 2.1 ▶ Mechanism for formation of Librium.

However, none of these compounds gave any interesting pharmacological results. The program was abandoned in 1955 in order for Sternbach to work on a different project. In 1957 during a general laboratory cleanup a vial containing what was thought to be **2.5** (X = 7-Cl, $R^1 = CH_2NHCH_3$, $R^2 = C_6H_5$) was found and, as a last effort, was submitted for pharmacological testing. Unlike all of the other compounds submitted, this one gave very promising results in six different tests used for preliminary screening of tranquilizers. Further investigation revealed that this compound was not a quinazoline 3-oxide, but was instead the benzodiazepine 4-oxide, **2.3**, presumably produced in an unexpected reaction of the corresponding chloromethyl quinazoline 3-oxide (**2.6**) with methylamine (Scheme 2.1). If this compound had not been found in the laboratory cleanup, all of the negative pharmacological results would have been reported for the quinazoline 3-oxide class of compounds, and benzodiazepine 4-oxides may not have been discovered for many years to come.

The examples of drug discovery without a lead are relatively few in number. The typical occurrence is that a lead compound is identified, and its structure is modified to give, eventually, the compound that goes to the clinic.

2.1.B Lead Discovery

Penicillin V and Librium are, indeed, two important drugs that were discovered without a lead. However, once they were identified, they then became lead compounds for second-generation analogs. A myriad of penicillin-derived antibacterials have been synthesized as a result of the structure elucidation of the earliest penicillins. Diazepam (**2.7**, Valium) was synthesized at Roche even before Librium was introduced on to the market; this drug was derived from the lead compound Librium and is almost 10 times more potent than the lead.



diazepam
2.7

The initial difficulty arises in the discovery of the lead compound. Several approaches can be taken to identify a lead. The first requirement for all of the approaches is to have a means to assay compounds for a particular biological activity, so that researchers can tell when a compound is active. A *bioassay* (or *screen*) is a means of determining in a biological system, relative to a control compound, if a compound has the desired activity, and, if so, what the relative potency of the compound is. Note the distinction between the terms *activity* and *potency*. *Activity* is the particular biological or pharmacological effect (for example, antibacterial activity or anticonvulsant activity); *potency* is the strength of that effect.

Some screens are *in vitro* tests, for example, the inhibition of an enzyme or antagonism of a receptor; others are *in vivo* tests, for example, the ability of the compound to prevent an induced seizure in a mouse. In general, the *in vitro* tests are quicker and less expensive. Currently, *high-throughput screens* (HTS),[11] very rapid and sensitive *in vitro* screens initially developed about 1989–1991, that now can be carried out robotically in 1536- or 3456-well titer plates on small (submicrogram) amounts of compound (dissolved in submicroliter volumes) are becoming universally used. With these ultra-high-throughput screening approaches, it is possible to screen 100,000 compounds in a day! As we will see below, combinatorial chemistry (see Section 2.2.E.5, p. 34) can supply huge numbers of compounds in a short period of time, which, theoretically, should provide an increased number of *hits*, i.e., compounds that elicit a predetermined level of activity in the bioassay and, therefore, provide more leads. According to Drews,[12] the number of compounds assayed in a large pharmaceutical company in the early 1990s was about 200,000 a year; that number rose to 5–6 million during the mid-1990s, and by the end of the 1990s it was >50 million! However, the increase in the assay rate did not result in a commensurate increase in research productivity, as measured by new compounds entering the market. Of course, it can take 12–15 years for a drug to reach the market, so productivity in the early part of the 21st century should provide a more accurate ruler for success of drug discovery changes made at the end of the 20th century. Currently, HTS appears to have resulted in an increase in the number of hits, but this may be because more lipophilic compounds, which may have more drug-like properties (see Section 2.2.F.2, p. 53), can be tested by dissolving them in dimethylsulfoxide (DMSO) rather than in water. Nonetheless, it is not yet clear if this increase in hit rate is translating into a much greater number of leads and development compounds.[13]

An exciting approach for screening compounds that might interact with an enzyme in a metabolic pathway was demonstrated by Wong, Pompliano, and coworkers for the discovery of lead compounds that block bacterial cell wall biosynthesis (as potential antibacterial agents).[14] Conditions were found to reconstitute all six enzymes in the cell wall biosynthetic

pathway so that incubation with the substrate for the first enzyme leads to the formation of the product of the last enzyme in the pathway. Then by screening compounds and looking for the buildup of an intermediate, it is possible to identify not only compounds that block the pathway (and prevent the formation of the bacterial cell wall), but to determine which enzyme is blocked (the buildup of an intermediate means that the enzyme that acted on that intermediate was blocked).

Compound screening also can be carried out by electrospray ionization mass spectrometry[15] (the technique for which John Fenn received the Nobel prize in 2002) and by NMR spectrometry.[16] Tightly bound noncovalent complexes of compounds with a macromolecule (such as a receptor or enzyme) can be observed in the mass spectrum. The affinity of the *ligand* (a small molecule that binds to a receptor) can be measured by varying the collision energy and determining at what energy the complex dissociates. This method also can be used to screen mixtures (a library) of compounds, provided they have different molecular masses and/or charges, so the m/z for each complex with the biomolecule can be separated in the mass spectrometer. By varying the collision energy, it is possible to determine which test molecules bind to the biomolecule best. The ^{1}H NMR method exploits changes in either relaxation rates or diffusion rates of small molecules when they bind to a macromolecule. This method also can be used to screen mixtures of compounds to determine the ones that bind best.

Once the screen is developed, a variety of approaches can be taken to obtain a lead. As we will see below, the typical lead compound for a receptor or enzyme is the natural ligand for the receptor or substrate for the enzyme. Another good source of lead compounds is marketed drugs.[17] In this case the target will generally be well established, and the lead structure will be known to bind well to the target and to have good absorption properties. The main stumbling block to the use of marketed drugs as leads may be patent issues for commercialization. If the target macromolecule is not known or if no new leads have come from a marketed drug, other approaches can taken.

B.1 Random Screening

In the absence of known drugs and other compounds with desired activity, a random screen is a valuable approach. *Random screening* involves no intellectualization; all compounds are tested in the bioassay without regard to their structures. Prior to 1935 (the discovery of sulfa drugs), this was essentially the only approach; today this method is still an important approach to discover drugs or leads, particularly because it is now possible to screen such huge numbers of compounds rapidly with HTSs. This is the lead discovery method of choice when nothing is known about the receptor target.

The two major classes of materials screened are synthetic chemicals and natural products (microbial, plant, and marine). An example of a random screen of synthetic and natural compounds was the "war on cancer" declared by Congress and the National Cancer Institute in the early 1970s. Any new compound submitted was screened in a mouse tumor bioassay. Few new anticancer drugs resulted from that screen, but many known anticancer drugs also did not show activity in the screen used, so a new set of screens was devised that gave more consistent results. In the 1940s and 1950s, a random screen of soil samples by various pharmaceutical companies in search of new antibiotics was undertaken. However, in this case, not only were numerous leads uncovered, but two important antibiotics, streptomycin and the

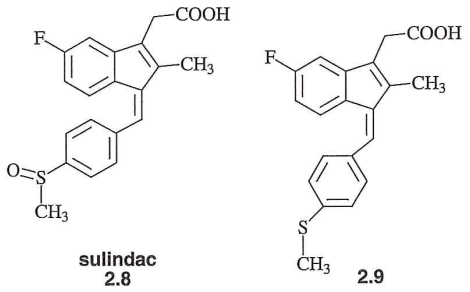
tetracyclines, were found. Screening of microbial broths, particular strains of *Streptomyces*, was a common random screen methodology prior to 1980.

B.2 Nonrandom (or Targeted or Focused) Screening

Nonrandom screening, also called *targeted* or *focused screening*, is a more narrow approach than is random screening. In this case, compounds having a vague resemblance to weakly active compounds uncovered in a random screen, or compounds containing different functional groups than leads, may be tested selectively. By the late 1970s, the National Cancer Institute's random screen was modified to a nonrandom screen because of budgetary and manpower restrictions. Also, the single tumor screen was changed to a variety of tumor screens because it was realized that cancer is not just a single disease.

B.3 Drug Metabolism Studies

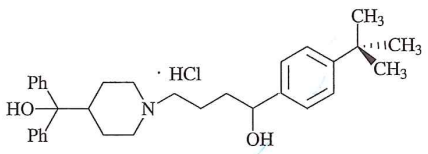
During drug metabolism studies (see Chapter 7) *metabolites* (drug degradation products generated *in vivo*) that are isolated are screened to determine if the activity observed is derived from the drug candidate or from a metabolite. For example, the anti-inflammatory drug sulindac (**2.8**, Clinoril) is not the active agent; the metabolic reduction product, **2.9**, is responsible for the activity.[18]



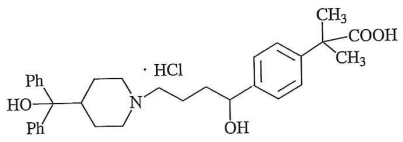
sulindac
2.8

2.9

The nonsedating antihistamine terfenadine hydrochloride (**2.10**, Seldane) was found to cause an abnormal heart rhythm in some users who also were taking certain antifungal agents, which were found to block the enzyme that metabolizes terfenadine. This caused a buildup of terfenadine, which led to the abnormal heart rhythms. However, a metabolite of terfenadine, fexofenadine hydrochloride (**2.11**, Allegra), was also found to be a nonsedating antihistamine, but it can be metabolized even in the presence of antifungal agents. This, then, is a safer drug. Metabolites can be screened for other activities as well.



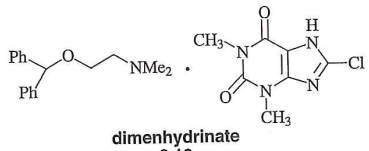
terfenadine HCl
2.10



fexofenadine HCl
2.11

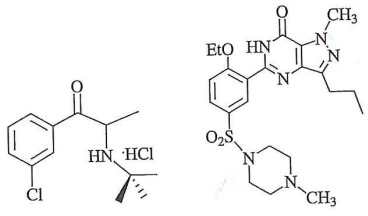
B.4 Clinical Observations

Sometimes a drug candidate during clinical trials will exhibit more than one pharmacological activity; that is, it may produce a side effect. This compound, then, can be used as a lead (or, with luck, as a drug) for the secondary activity. In 1947 an antihistamine, dimenhydrinate (**2.12**, Dramamine) was tested at the allergy clinic at Johns Hopkins University and was found also to be effective in relieving a patient who suffered from car sickness; a further study proved its effectiveness in the treatment of seasickness[19] and airsickness.[20] It then became the most widely used drug for the treatment of all forms of motion sickness.



dimenhydrinate
2.12

There are other popular examples of drugs derived from clinical observations. Bupropion hydrochloride (**2.13**), an antidepressant drug (Wellbutrin), was found to help patients stop smoking and is now the first drug marketed as a smoking cessation aid (Zyban). The impotence drug sildenafil citrate (**2.14**; Viagra) was designed for the treatment of angina and hypertension by blocking the enzyme phosphodiesterase-5, which hydrolyzes cyclic guanosine monophosphate (cGMP), a vasodilator that allows increased blood flow.[21] In 1991 sildenafil went into phase I clinical trials for angina. In phase II clinical trials, it was not as effective against angina as Pfizer had hoped, so it went back to phase I clinical trials to see how high of a dose could be tolerated. It was during that clinical trial that the volunteers reported increased erectile function. Given the weak activity against angina, it was an easy decision to try to determine its effectiveness as the first treatment for erectile dysfunction. Sildenafil works by the mechanism for which it was designed as an antianginal drug, except it inhibits the phosphodiesterase in the penis (phosphodiesterase-5) instead of the heart (Figure 2.1). Sexual stimulation causes release of nitric oxide in the penis.



bupropion HCl
2.13

sildenafil citrate
2.14

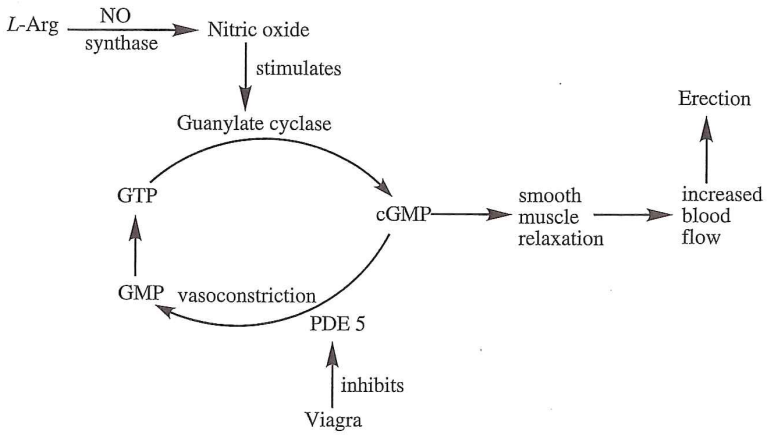


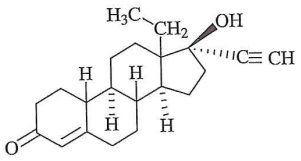
Figure 2.1 ▶ Mechanism of action of sildenafil (Viagra)

Nitric oxide is a second messenger molecule that stimulates the enzyme guanylate cyclase, which converts guanosine triphosphate to cGMP. The vasodilator cGMP relaxes the smooth muscle in the *corpus cavernosum*, allowing blood to flow into the penis, thereby producing an erection. However, phosphodiesterase-5 (PDE 5) hydrolyzes the cGMP, which causes vasoconstriction and the outflow of blood from the penis. Sildenafil inhibits this phosphodiesterase, preventing the hydrolysis of cGMP and prolonging the vasodilation effect.

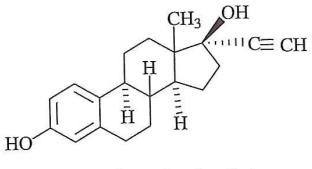
B.5 Rational Approaches to Lead Discovery

None of the above approaches to lead discovery involves a major rational component. The lead is just found by screening techniques, as a by-product of drug metabolism studies, or from clinical investigations. Is it possible to *design* a compound having a particular activity? Rational approaches to drug design now have become the major routes to lead discovery. The first step is to identify the cause for the disease state. Many diseases, or at least the symptoms of diseases, arise from an imbalance (either excess or deficiency) of particular chemicals in the body, from the invasion of a foreign organism, or from aberrant cell growth. As will be discussed in later chapters, the effects of the imbalance can be corrected by antagonism or agonism of a receptor (see Chapter 3) or by inhibition of a particular enzyme (see Chapter 5); foreign organism enzyme inhibition or interference with DNA biosynthesis or function are important approaches to treat diseases arising from microorganisms and aberrant cell growth (see Chapter 6).

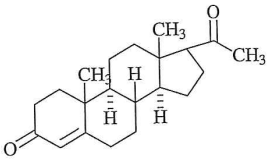
Once the relevant biochemical system is identified, initial lead compounds then become the natural receptor ligands or enzyme substrates. For example, lead compounds for the contraceptives (+)-norgestrel (**2.15**, Ovral) and 17α-ethynylestradiol (**2.16**, Activella) were the steroidal hormones progesterone (**2.17**) and 17β-estradiol (**2.18**). Whereas the steroid hormones **2.17** and **2.18** show weak and short-lasting effects, the oral contraceptives **2.15** and **2.16** exert strong progestational activity of long duration.



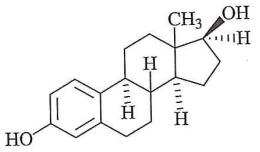
norgestrel
2.15



17α-ethynyl estradiol
2.16

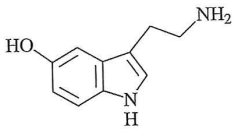


2.17

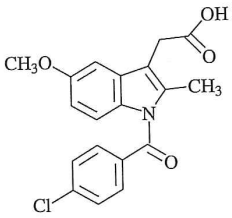


2.18

At Merck it was believed that serotonin (**2.19**) was a possible mediator of inflammation. Consequently, serotonin was used as a lead for anti-inflammatory agents, and from this lead the anti-inflammatory drug indomethacin (**2.20**, Indocin) was developed.[22]



2.19



indomethacin
2.20

The rational approaches are directed at lead discovery. It is not possible, with much accuracy, to foretell toxicity and side effects, anticipate transport characteristics, or predict the metabolic fate of a drug. Once a lead is identified, its structure can be modified until an effective drug is obtained.

2.2 Lead Modification: Drug Design and Development

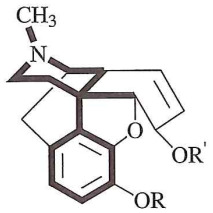
Once your lead compound is in hand, how do you know what to modify in order to improve the desired pharmacological properties?

2.2.A Identification of the Active Part: The Pharmacophore

Interactions of drugs with receptors, known as *pharmacodynamics*, are very specific (see Chapter 3). Therefore, only a small part of the lead compound may be involved in the appropriate receptor interactions. The relevant groups on a molecule that interact with a receptor and are responsible for the activity are collectively known as the *pharmacophore*. The other atoms in the lead molecule, sometimes referred to as the *auxophore*, may be extraneous.

Some of the atoms, of course, are essential to maintain the integrity of the molecule and hold the pharmacophoric groups in their appropriate positions. Some of these extraneous atoms, however, may be interfering with the binding of the pharmacophore, and those atoms need to be excised from the lead compound. Other atoms in the auxophore may be dangling in space within the receptor and are neither binding to the receptor nor preventing the pharmacophoric atoms from binding. Although these atoms appear to be innocuous, it is important to know which atoms these are, because these are the ones that can be modified without loss of potency. As we will see later, there are other aspects to lead modification that are as important as increasing binding to the target receptor, such as *pharmacokinetics* (absorption, distribution, metabolism, and excretion or ADME). Modification of the atoms that are not interfering with binding could be very important to solving pharmacokinetics problems.

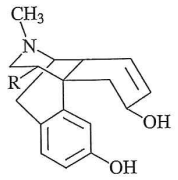
By determining which are the pharmacophoric groups and which are the auxophoric groups on your lead compound, and of the auxophoric groups, which are interfering with lead compound binding and which are not detrimental to binding, you will know which groups must be excised and which you can retain or modify as needed. One approach in lead modification to help make this determination is to cut away sections of the lead molecule and measure the effects of those modifications on potency. Consider this artificial example of how this might be done. Assume that the addictive analgesics morphine (**2.21**, R = R′ = H), codeine (**2.21**, R = CH_3, R′ = H), and heroin (**2.21**, R = R′ = $COCH_3$) are the lead compounds, and we want to know which groups are pharmacophoric and which are auxophoric.



morphine (R = R' = H)
codeine (R = CH_3, R' = H)
heroin (R = R' = $COCH_3$)
2.21

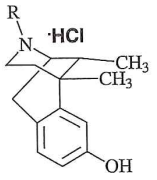
The morphine family of analgesics binds to the μ opioid receptors. The pharmacophore is known and is shown as the darkened part in **2.21**. A decrease in potency on removal of a group will suggest that it may have been pharmacophoric, an increase in potency means it was auxophoric and interfering with proper binding, and essentially no change in potency will mean that it is auxophoric but not interfering with binding.

Let's start by excising the dihydrofuran oxygen atom, which is not in the pharmacophore. This may not seem to be sensible because that atom connects the cyclohexene ring to the benzene ring; its removal will result in a change in the conformation of the cyclohexene ring and an increase in the degrees of freedom of the molecule. Excision of the dihydrofuran oxygen gives morphinan (**2.22**, R = H)[23]; the hydroxy analog, levorphanol[24] (**2.22**, R = OH, Levo-Dromoran), is three to four times *more* potent than morphine as an analgesic, but it retains the addictive properties (note that in **2.22** the cyclohexene ring conformation has not been changed for ease of comparison with **2.21**; surely, a lower energy conformer will be favored).

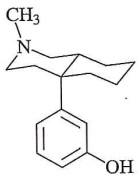


levorphanol (R = OH)
2.22

Possibly, the additional conformational mobility allowed the molecule to approximate its *bioactive conformation* (the conformation that most effectively binds to a receptor). Removal of half of the cyclohexene ring (also not in the pharmacophore), leaving only methyl substituents, gives benzomorphan (**2.23**, R = CH_3).[25] This compound shows some separation of analgesic and addictive effects; pentazocine hydrochloride (**2.23**, R = $CH_2CH{=}C(CH_3)_2$; component of Talwin) is less potent than morphine (about as potent as codeine), but has a much lower addiction liability. Remember, your goal is both to increase potency and decrease adverse effects, such as addictive properties. Although this analog is not more potent than morphine, it is less addicting. Cutting away the methylene group of the cyclohexane fused ring (**2.24**) also, surprisingly, has little effect on the analgesic activity in animal tests.

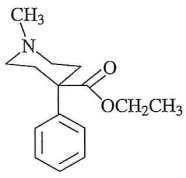


pentazocine HCl (R = $CH_2CH{=}C(CH_3)_2$)
2.23

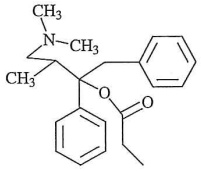


2.24

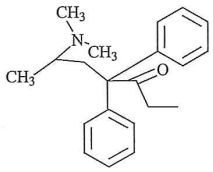
Again, this excision removes the rigidity of the parent structure. Removal of all fused rings, for example, in the case of meperidine (**2.25**, Demerol), gives an analgesic still possessing 10–12% of the overall potency of morphine.[26] Although the potency is lower, it certainly will be much easier to synthesize analogs of meperidine than of morphine. Even acyclic analogs are active. Dextropropoxyphene (**2.26**, Darvon; again note the side chain is left in a conformation to resemble the structure of morphine) is one-half to two-thirds as potent as codeine. Both morphine and dextropropoxyphene bind to the μ opioid receptor, so the activity of dextropropoxyphene can be ascribed to the fact that it can assume a conformation related to that of the morphine pharmacophore. I don't think any of us, seeing dextropropoxyphene written in a more energetically favorable conformation, would ever make the connection between this structure and that of morphine. By cutting pieces off of the lead compound, it gives you new perspectives on possible active structures, which should open up completely new scaffolds to consider in synthesis. Another acyclic analog is methadone (**2.27**, Methadose) which is as potent an analgesic as morphine; the (−)-isomer is used in the treatment of opioid abstinence syndromes in heroin abusers because it is eliminated from the body slower than morphine, allowing the body to adapt to the falling levels of drug gradually.



meperidine
2.25

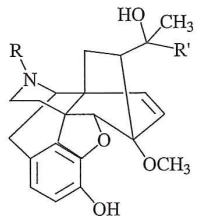


dextropropoxyphene
2.26



methadone
2.27

What if every cut in the lead produces a compound with lower potency? Then, either every excision is removing part of the pharmacophore or each cut causes a conformational change that gives a structure *less* similar to the bioactive conformation. The latter possibility is particularly relevant to such a rigid structure as in morphine. In this case, groups need to be added to the lead structure to *increase* the pharmacophore. For example, oripavine derivatives such as etorphine (**2.28**, R = CH_3, R′ = C_3H_7; Immobilon), which has a two-carbon bridge and substituent not in morphine, is 3200 times more potent than morphine[27] and is used in veterinary medicine to immobilize large animals. The related analog, buprenorphine (**2.28**, R = CH_2–cyclopropyl, R′ = *t*-Bu, double bond reduced; Buprenex) is 10–20 times more potent than morphine and has a very low level of dependence liability.



etorphine (R = CH_3; R' = C_3H_7)
2.28

Apparently, the additional rigidity of the oripavine derivatives increases the appropriate receptor interactions (see Chapter 3).

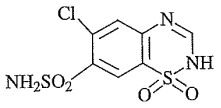
The activity and potency of a molecule are related to the interactions of the pharmacophoric groups with groups on the receptor (see Chapter 3, Section 3.2.B). The binding constants of 200 drugs and potent enzyme inhibitors were used by Andrews and coworkers[28] to calculate the average binding energies of common functional groups; these energies can be used to determine how well a new molecule binds to its receptor. If the test molecule has a measured binding energy that is lower than the calculated average value, it suggests that the molecule contains groups that do not interact with the receptor (are not in the pharmacophore). These groups, then, could be excised without loss of potency, giving a simplified lead for further structural modification. This *Andrews analysis* was carried out on a highly substituted lead compound, leading to a more simple analog structure that was modified to give molecules with enhanced potency.[29] If the test compound has a binding energy greater than the calculated average value, then the molecule may bind differently than suspected, leading to enhanced binding interactions. This indicates that manipulation of functional groups is an important lead modification approach.

2.2.B Functional Group Modification

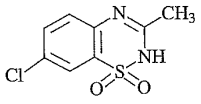
The importance of functional group modification is demonstrated by **2.29**. The antibacterial agent, carbutamide (**2.29**, R = NH_2), was found to have an antidiabetic side effect; however, it could not be used as an antidiabetic drug because of its antibacterial activity, which could lead to bacterial resistance (see Chapter 5, Section 5.2, p. 231). The amino group of carbutamide was replaced by a methyl group to give tolbutamide (**2.29**, R = CH_3; Orinase) and in so doing the antibacterial activity was eliminated from the antidiabetic activity.

tolbutamide (R = CH_3)
2.29

In some cases, an experienced medicinal chemist knows what functional group will elicit a particular effect. Chlorothiazide (**2.30**, Aldocor) is an antihypertensive agent that has a strong diuretic effect as well. It was known from sulfanilamide work that the sulfonamide side chain can give diuretic (increased urine excretion) activity (see Section 2.2.C below). Consequently, diazoxide (**2.31**, Hyperstat) was prepared as an antihypertensive drug without diuretic activity.



chlorothiazide
2.30



diazoxide
2.31

Obviously, a relationship exists between the molecular structure of a compound and its activity. This phenomenon was first realized about 135 years ago.

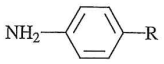
2.2.C Structure–Activity Relationships

In 1868 Crum-Brown and Fraser,[30] suspecting that the quaternary ammonium character of curare, a potent poison known since the 16th century that was used on arrowheads, may be responsible for its muscular paralytic properties (it blocks the action of the excitatory neurotransmitter acetylcholine on muscle receptors), examined the neuromuscular blocking effects of a variety of simple quaternary ammonium salts and quaternized alkaloids in animals. From these studies they concluded that the physiological action of a molecule was a function of its chemical constitution. Shortly thereafter, Richardson[31] noted that the hypnotic activity of aliphatic alcohols was a function of their molecular weight. These observations are the basis for future *structure–activity relationships* (SARs).

Drugs can be classified as being structurally specific or structurally nonspecific. *Structurally specific drugs*, which most drugs are, act at specific sites, such as a receptor or an enzyme. Their activity and potency are very susceptible to small changes in chemical structure; molecules with similar biological activities tend to have common structural features. *Structurally nonspecific drugs* have no specific site of action and usually have lower potency. Similar biological activities may occur with a variety of structures. Examples of these drugs are gaseous anesthetics, sedatives and hypnotics, and many antiseptics and disinfectants.

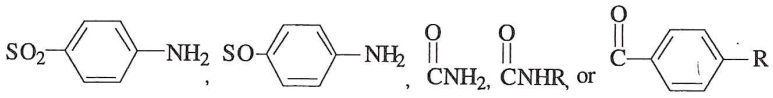
Even though only a part of the molecule may be associated with its activity, a multitude of molecular modifications could be made. The hallmark of SAR studies is the synthesis of as many analogs as possible of the lead and their testing to determine the effect of structure on activity (or potency). Once enough analogs are prepared and sufficient data accumulated, conclusions can be made regarding structure–activity relationships. Unfortunately, ease of synthesis, rather than cogent rationales, is often the guiding force behind the choice of analogs made.

An excellent example of this approach came from the development of the sulfonamide antibacterial agents (sulfa drugs). After a number of analogs of the lead compound sulfanilamide (**2.1**, R = H; AVC) were prepared, clinical trials determined that compounds of this general structure exhibited diuretic and antidiabetic activities as well as antimicrobial activity. Compounds with each type of activity eventually were shown to possess certain structural features in common. On the basis of the biological results of greater than 10,000 compounds, several SAR generalizations were made.[32] Antimicrobial agents have structure **2.32** (R = SO_2NHR' or SO_3H).



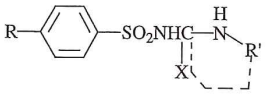
2.32

In **2.32**, (1) the amino and sulfonyl groups on the benzene ring should be para; (2) the anilino amino group may be unsubstituted (as shown) or may have a substituent that is removed *in vivo*; (3) replacement of the benzene ring by other ring systems, or the introduction of additional substituents on it, decreases the potency or abolishes the activity; (4) R may be any of the alternatives shown below, but the potency is reduced in most cases; (5) N′-monosubstitution



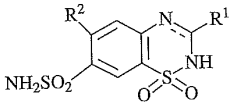
($R = SO_2NHR'$) results in more potent compounds, and the potency increases with heteroaromatic substitution; and (6) N′-disubstitution ($R = SO_2NR'_2$), in general, leads to inactive compounds.

Antidiabetic agents are compounds with structure **2.33**, where X may be O, S, or N incorporated into a heteroaromatic structure such as a thiadiazole or a pyrimidine or in an acyclic structure such as a urea or thiourea. In the case of ureas, the N-2 should carry as a substituent a chain of at least two carbon atoms.[33]

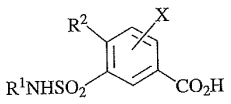


2.33

Sulfonamide diuretics are of two general structural types: hydrochlorthiazides (**2.34**) and the high ceiling type[34] (**2.35**). The former compounds have 1,3-disulfamyl groups on the benzene ring and R^2 is an electronegative group such as Cl, CF_3, or NHR.



2.34



2.35

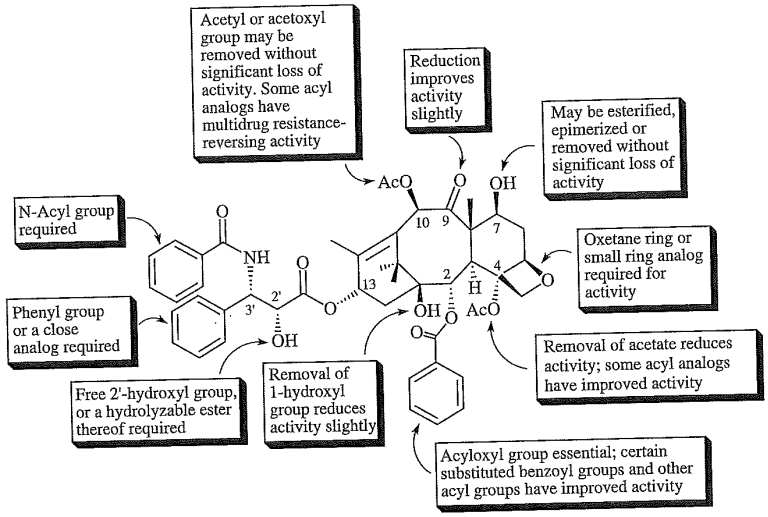
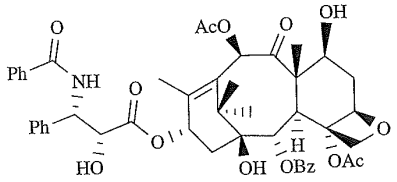


Figure 2.2 ► SAR for paclitaxel (Taxol)

The high ceiling compounds contain 1-sulfamyl-3-carboxy groups. Substituent R^2 is Cl, Ph, or PhZ, where Z may be O, S, CO, or NH and X can be at position 2 or 3 and is normally NHR, OR, or SR.[35]

A more recent example of a SAR is that of the natural product anticancer drug paclitaxel (**2.36**, Taxol), which was the first anticancer compound found to act by promoting the assembly of tubulin into microtubules, thereby blocking mitosis.[36] After a large number of modifications were introduced,[37] many SAR (actually, structure–*potency* relationship) conclusions could be made (Figure 2.2). A common way to track the structural changes is with the use of *molecular activity maps*, structural drawings of a lead compound annotated to show where in the molecule specific structural changes affect activity or potency measured in a single bioassay.



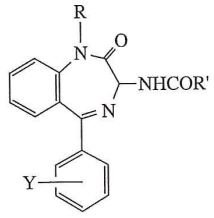
paclitaxel
2.36

For large lead molecules that have been extensively modified, these maps concisely summarize a huge number of facts relating structures with their activities and potencies. Effects could include the abolishment of an activity, unexpected toxicity, or large changes in potency. These maps may depict the results of a long-lasting drug discovery effort involving numerous chemists (and the biologists who do the screens). Their main virtues are that they can prevent your coworkers from synthesizing analogs that have already been made and tested, and they may direct the chemists' creativity to unexplored regions of the lead compound, yielding novel structural changes.

The above examples provide strong evidence to support the notion that a correlation does exist between structure and activity, but does each structure interact with only one receptor and lead to only one activity?

2.2.D Privileged Structures and Drug-Like Molecules

Evans and coworkers first introduced the term *privileged structures* for certain molecular scaffolds that appear to be capable of binding to multiple receptor targets, and, consequently, with appropriate structure modifications, could exhibit multiple activities.[38] The Merck group used benzodiazepines (**2.37**) as the example of this phenomenon, which was earlier mentioned by Ariëns and coworkers without referring to them as privileged structures.[39] A number of other privileged structures are known.[40]



2.37

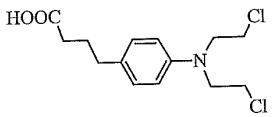
The commonality of molecular features in a variety of drugs was apparent by the revelation that only 32 scaffolds describe half of all known drugs.[41] Likewise, a small number of moieties account for a large majority of the side chains found in drugs.[42] The average number of side chains per molecule is four. If the carbonyl side chain is ignored, then 73% of the side chains in drugs are from the top 20 most common side chains. On the basis of what is known about privileged structures and common scaffold structure, Ajay and coworkers proposed that *drug-likeness* is a possible inherent property of some molecules,[43] and this property could determine which molecules should be selected for screening. They used a set of one- and two-dimensional parameters in their computation and were able to predict correctly more than 90% of the compounds in the Comprehensive Medicinal Chemistry (CMC) database.[44] Another computational approach to differentiate drug-like and nondrug-like molecules using a scoring scheme was developed[45] that was able to classify correctly 83% of the compounds in the Available Chemicals Directory (ACD)[46] and 77% of the compounds in the World Drug Index (WDI).[47] A variety of other approaches have been taken to identify drug-like molecules.[48]

There also are many nondrug-like molecules that show up as active compounds in screens, but later are demonstrated to be *false positives* (inactive compounds that appear to be active

in the screen). These compounds initially appear to interact with a variety of receptors and are known as *promiscuous binders* (*promiscuous antagonists* if they antagonize many receptors or *promiscuous inhibitors* if they inhibit many enzymes). However, they are nondrug-like because they act at a site different from that of the natural ligand or substrate, show little relationship between structure and activity, and have poor selectivity for a specific receptor or enzyme. As a result, much time is wasted following up on the activity of these compounds, later to find out that they are inactive. Shoichet and coworkers showed that many of these compounds are actually aggregates of molecules, and it is the aggregate that produces the false-positive activity.[49] If the conditions of the screen are changed, these aggregates dissociate, and the individual molecules can be shown to be inactive.

2.2.E Structure Modifications to Increase Potency and the Therapeutic Index

How do you know what molecular modifications to make in order to fine tune the lead compound? The preceding section makes clear that structure modifications are the keys to activity and potency manipulations. After years of structure–activity relationship studies, various standard molecular modification approaches have been developed for the systematic improvement of the *therapeutic index* (also called the *therapeutic ratio*), which is a measure of the safety of a drug as determined from the ratio of the concentration of a drug that gives undesirable effects to that which gives desirable effects. The therapeutic index can be determined by any method that measures undesirable and desirable drug effects, but often it is taken as the dose-limiting toxicity versus the desirable pharmacological effect in an animal model (preferably humans). For example, the therapeutic index could be the ratio of the LD_{50} (the lethal dose for 50% of the test animals) to the therapeutic ED_{50} (the effective dose that produces the maximum therapeutic effect in 50% of the test animals); a toxic ED_{50} (the dose that produces toxicity in 50% of the test animals) may substitute for the LD_{50}. The larger the therapeutic index, the greater the margin of safety of the compound. In other words, you would like to have to administer gram quantities of the drug before any undesirable effects are observed, but administer only milligrams of the drug to attain the desirable effects. There is no specific minimum value for a therapeutic index that must be attained before a drug can be approved; it depends on the disease that is being treated and whether other therapies are already available. A low therapeutic index is tolerable for lethal diseases, such as cancer or AIDS (maybe even as low as 1–5), especially if no other treatment is available or if the side effect is minor compared with the treatment benefit. For less threatening diseases, therapeutic indices on the order of 10–100 may be reasonable. As an example, the therapeutic index for the antitumor agent chlorambucil (**2.38**, Leukeran) is 23.[50] The Merck Index is a good source for obtaining LD_{50} data for drugs in animals.



chlorambucil
2.38

A number of structural modification methodologies follow.

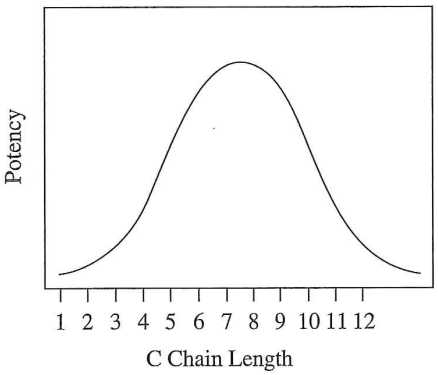


Figure 2.3 ► General effect of carbon chain length on drug potency

E.1 Homologation

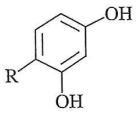
A *homologous series* is a group of compounds that differ by a constant unit, generally a CH_2 group. As will become more apparent in Section 2.2.F.2, p. 53, biological properties of homologous compounds show regularities of increase and decrease. For many series of compounds, lengthening of a saturated carbon side chain from one (methyl) to five to nine atoms (pentyl to nonyl) produces an increase in pharmacological effects; further lengthening results in a sudden decrease in potency (Figure 2.3). In Section 2.2.F.2.b, p. 55, I show that this phenomenon corresponds to increased lipophilicity of the molecule to permit penetration into cell membranes until its lowered water solubility becomes problematic in its transport through aqueous media. In the case of aliphatic amines, another problem is micelle formation, which begins at about C_{12}. This effectively removes the compound from potential interaction with the appropriate receptors. One of, if not the, earliest examples of this potency versus chain length phenomenon was reported by Richardson,[51] who was investigating the hypnotic activity of alcohols. The maximum effect occurred for 1-hexanol to 1-octanol; then the potency declined on chain lengthening until no activity was observed for hexadecanol.

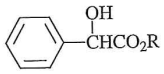
A study by Dohme *et al.*[52] on 4-alkyl-substituted resorcinol derivatives showed that the peak antibacterial activity occurred with 4-*n*-hexylresorcinol (see Table 2.1), a compound now used as a topical anesthetic in a variety of throat lozenges. Funcke *et al.*[53] found that the peak spasmolytic activity of a series of mandelate esters occurred with the *n*-nonyl ester (see Table 2.1).

E.2 Chain Branching

When a simple lipophilic relationship is important, as described above, then chain branching lowers the potency of a compound because a branched alkyl chain is less lipophilic than the corresponding straight alkyl chain as a result of larger molar volumes and shapes of branched compounds. This phenomenon is exemplified by the lower potency of the compounds having isoalkyl chains in Table 2.1. In this case, pharmacokinetics would be the overriding factor for potency. However, another explanation for lower potency with branching could be pharmacodynamics; chain branching may interfere with receptor binding. For example, phenethylamine ($PhCH_2CH_2NH_2$) is an excellent substrate for monoamine oxidase, but

TABLE 2.1 ▶ Effect of Chain Length on Potency. Antibacterial activity of 4-*n*-alkylresorcinols and spasmolytic activity of mandelate esters

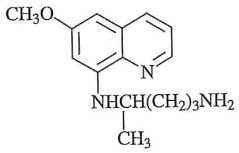




R	Phenol coefficient	% Spasmolytic activity[a]
methyl	—	0.3
ehtyl	—	0.7
n-propyl	5	2.4
n-butyl	22	9.8
n-pentyl	33	28
n-hexyl	51	35
n-heptyl	30	51
n-octyl	0	130
n-nonyl	0	190
n-decyl	0	37
n-undecyl	0	22
i-propyl	—	0.9
i-butyl	15.2	8.3
i-amyl	23.8	28
i-hexyl	27	—

[a] Relative to 3,3,5-trimethylcyclohexanol, set at 100%.

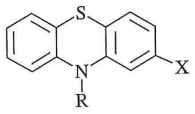
α-methylphenethylamine (amphetamine) is a poor substrate. Primary amines often are more potent than secondary amines, which are more potent than tertiary amines. For example, the antimalarial drug primaquine phosphate (**2.39**, Primaquine) is much more potent than its secondary or tertiary amine homologs.



primaquine
2.39

Major pharmacological changes can occur with chain branching or homologation. Consider the 10-aminoalkylphenothiazines (**2.40**, X = H). When R is $CH_2CH(CH_3)N(CH_3)_2$ (promethazine HCl; Phenergan), antispasmodic and antihistaminic activities predominate. However, the straight-chain analog **2.40** with R being $CH_2CH_2CH_2N(CH_3)_2$ (promazine) has greatly reduced antispasmodic and antihistaminic activities, but sedative and tranquilizing activities are greatly enhanced. In the case of the branched chain analog **2.40** with R equal

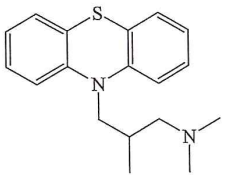
to $CH_2CH(CH_3)CH_2N(CH_3)_2$ (trimeprazine) (next larger branched-chain homolog), the tranquilizing activity is reduced and antipruritic (anti-itch) activity increases. This indicates that multiple receptors are involved, and branching or homologation can cause the molecule to bind more or less well to the receptors responsible for antispasmodic activity, antihistamine activity, tranquilizing activity, or antipruritic activity.



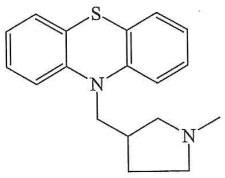
promethazine (R = $CH_2CH(CH_3)N(CH_3)_2$)
promazine (R = $CH_2CH_2CH_2N(CH_3)_2$)
trimeprazine (R = $CH_2CH(CH_3)CH_2N(CH_3)_2$)
2.40

E.3 Ring-Chain Transformations

Another modification that can be made is the transformation of alkyl substituents into cyclic analogs, which often does not affect potency greatly. Consider the promazines again (**2.40**). Chlorpromazine (X = Cl, R = $CH_2CH_2CH_2N(CH_3)_2$; Thorazine) and **2.40** (X = Cl, R = $CH_2CH_2CH_2N$⬠) are equivalent as tranquilizers in animal tests. The branched methyl group and one of the dimethylamino methyl groups of trimeprazine (**2.41**, Vallergan or Temaril) could be connected to give methdilazine (**2.42**, Dilosyn), which have similar antipruritic activity in man.



trimeprazine
2.41



methdilazine
2.42

However, a ring-chain transformation could have an important pharmacokinetic effect, such as to increase lipophilicity or decrease metabolism, which could make the drug more effective *in vivo*. Also, by connecting substituents into a ring, pharmacodynamic properties could be enhanced by constraining the groups into a particularly favorable conformation. Of course, it also could constrain the molecule into an unfavorable conformation, and potency could drop!

Different activities can result from a ring-chain transformation as well. For example, if the dimethylamino group of the tranquilizer chlorpromazine is substituted by a methylpiperazine ring (**2.40**, X = Cl, R = $CH_2CH_2CH_2N$⬡NCH_3; prochlorperazine), antiemetic (prevents nausea

TABLE 2.2 ▶ Classical Isosteres

1. Univalent atoms and groups
   a. CH_3 NH_2 OH F Cl
   b. Cl PH_2 SH
   c. Br *i*-Pr
   d. I *t*-Bu

2. Bivalent atoms and groups
   a. $-CH_2-$ $-NH-$ $-O-$ $-S-$ $-Se-$
   b. $-COCH_2R$ $-CONHR$ $-CO_2R$ $-COSR$

3. Trivalent atoms and groups
   a. $-CH=$ $-N=$
   b. $-P=$ $-As=$

4. Tetravalent atoms
   a. $-\overset{|}{\underset{|}{C}}-$ $-\overset{|}{\underset{|}{Si}}-$
   b. $=C=$ $=\overset{+}{N}=$ $=\overset{+}{P}=$

5. Ring equivalent
   a. $-CH=CH-$ $-S-$ (e.g., benzene, thiophene)
   b. $-CH=$ $-N=$ (e.g., benzene, pyridine)
   c. $-O-$ $-S-$ $-CH_2-$ $-NH-$ (e.g., tetrahydrofuran, tetrahydrothiophene, cyclopentane, pyrrolidine)

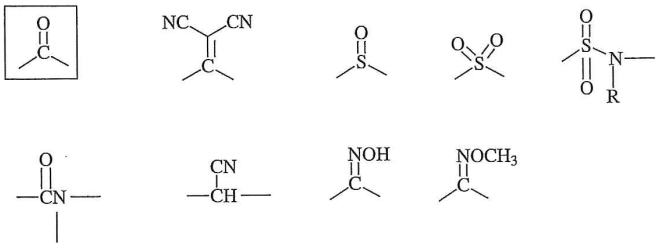
and vomiting) activity is greatly enhanced. In this case, however, an additional methylamino group is also added, which may have an effect in changing the activity.

E.4 Bioisosterism

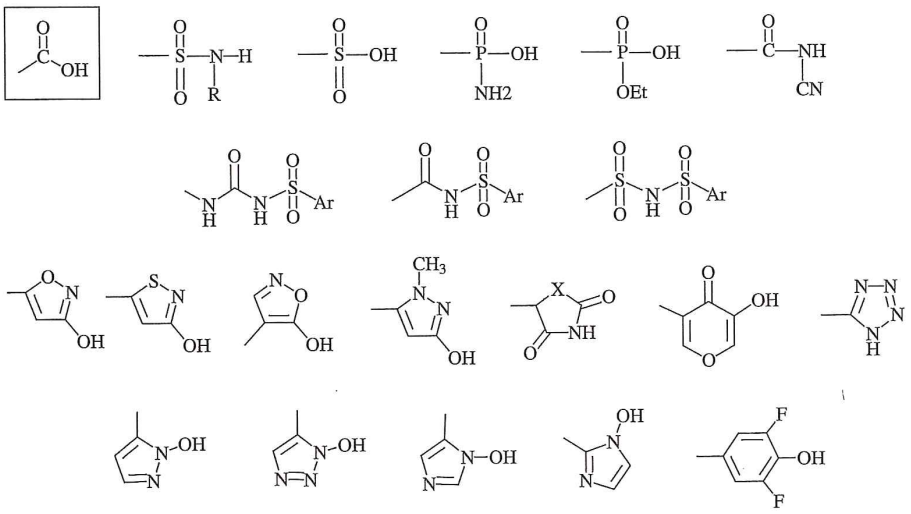
Bioisosteres are substituents or groups that have chemical or physical similarities, and which produce broadly similar biological properties.[54] Bioisosterism is an important lead modification approach that has been shown to be useful to attenuate toxicity or to modify the activity of a lead, and may have a significant role in the alteration of pharmakinetics of a lead. There are classical isosteres[55,56] and nonclassical isosteres.[57] In 1925 Grimm[58] formulated the *hydride displacement law* to describe similarities between groups that have the same number of valence electrons, but may have a different number of atoms. Erlenmeyer[59] later redefined isosteres as atoms, ions, or molecules in which the peripheral layers of electrons can be considered to be identical. These two definitions describe *classical isosteres*; examples are shown in Table 2.2. *Nonclassical bioisosteres* do not have the same number of atoms and do not fit the steric and electronic rules of the classical isosteres, but do produce similar biological activity. Examples of these are shown in Table 2.3. Ring-chain transformations also can be considered to be isosteric interchanges. There are hundreds of examples of compounds that differ by a bioisosteric interchange;[60] some examples are shown in Table 2.4.

TABLE 2.3 ▶ Nonclassical Isosteres

Carbonyl group



Carboxylic acid group



Amide group

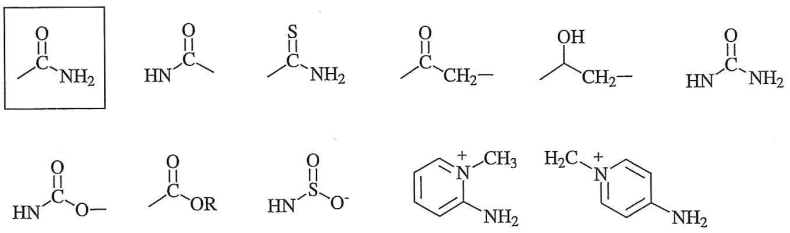
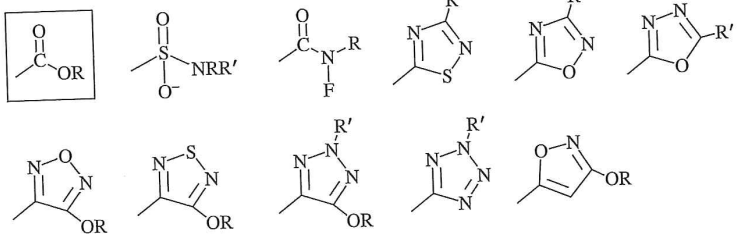
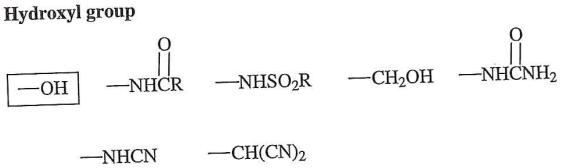


TABLE 2.3 ▶ *Continued*

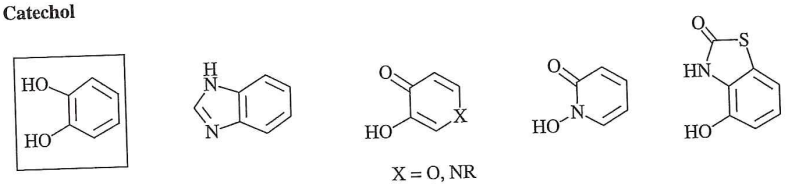
Ester group



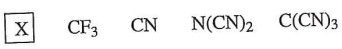
Hydroxyl group



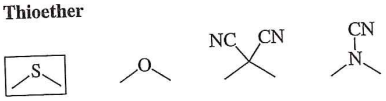
Catechol



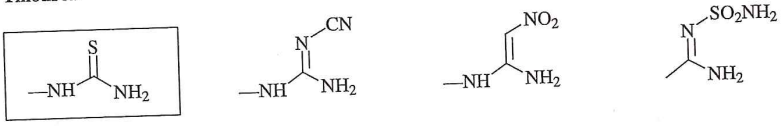
Halogen



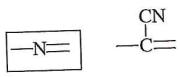
Thioether



Thiourea



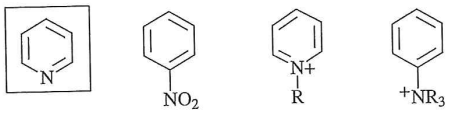
Azomethine



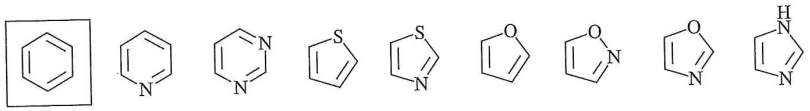
Continued

TABLE 2.3 ▶ *Continued*

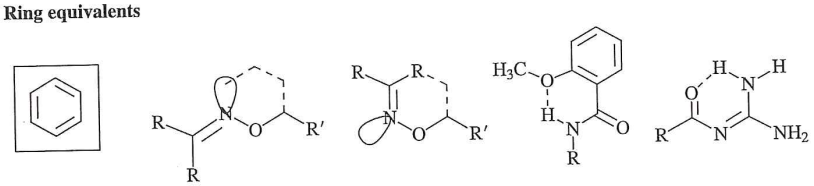
Pyridine



Benzene



Ring equivalents

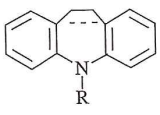


Spacer group

Hydrogen

H F

Bioisosterism, however, also can lead to changes in activity or potency. For example, if the sulfur atom of the phenothiazine neuroleptic drugs (**2.40**) is replaced by $-CH{=}CH-$ or $-CH_2CH_2-$ bioisosteres, then dibenzazepine antidepressant drugs (**2.43**) result. A change in enzyme affinity also can be observed. For example, when the thiazolone ring in a series of anti-inflammatory compounds selective for the cyclooxygenase-2 isozyme over the cyclooxygenase-1 isozyme (see Chapter 5, Section 5.5.B.2.b, p. 280) was substituted by an oxazolone ring (i.e., the S was replaced by O), the selectivity for the two isozymes reversed.[61]

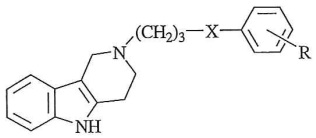


2.43

It is actually quite surprising that bioisosterism should be such a successful approach to lead modification. Perusal of Table 2.2, and especially of Table 2.3, makes it clear that in making a bioisosteric replacement, one or more of the following parameters will

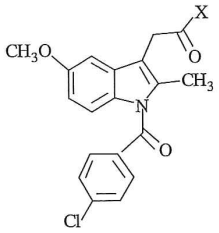
TABLE 2.4 ▶ Examples of Bioisosteric Analogs

1. **Neuroleptics (antipsychotics)**

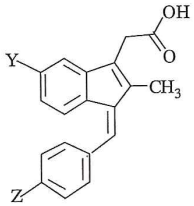


X =  or CHCN

2. **Anti-inflammatory agents**



X = OH (indomethacin)
= NHOH
= 



Y = CH_3O Z = Cl
Y = F Z = SCH_3 (sulindac)

3. **Antihistamines**

R—X—$(CH_2)_n$—Y

X = NH, O, CH_2
Y = N $(CH_3)_2$ (n = 2)

 (n = 1)

 (n = 1, 2)

change: size, shape, electronic distribution, lipid solubility, water solubility, pK_a, chemical reactivity, and hydrogen bonding. Because a drug must get to the site of action, then interact with it (see Chapter 3), bioisosteric modifications made to a molecule may have one or more of the following effects:

1. *Structural.* If the moiety that is replaced by a bioisostere has a structural role in holding other functionalities in a particular geometry, then size, shape, and hydrogen bonding will be important.

2. *Receptor interactions*. If the moiety replaced is involved in a specific interaction with a receptor or enzyme, then all of the parameters except lipid and water solubility will be important.
3. *Pharmacokinetics*. If the moiety replaced is necessary for absorption, transport, and excretion of the compound, then lipophilicity, hydrophilicity, pK_a, and hydrogen bonding will be important.
4. *Metabolism*. If the moiety replaced is involved in blocking or aiding metabolism, then the chemical reactivity will be important.

It is because of these subtle changes that bioisosterism is effective. This approach allows the medicinal chemist to tinker with only some of the parameters in order to augment the potency, selectivity, and duration of action and to reduce toxicity. Multiple alterations may be necessary to counterbalance effects. For example, if modification of a functionality involved in binding also decreases the lipophilicity of the molecule, thereby reducing its ability to penetrate cell walls and cross other membranes, the molecule can be modified at a different site with a lipophilic group to increase absorption. But where can these bioisosteric replacements be made? A pharmacophore study (Section 2.2.A, p. 17) could have identified those groups on the lead that could be modified without an effect on receptor binding (the scissions that led to little change in potency). Those are the positions that can be safely modified. Modifications of this sort, however, may change the overall molecular shape and result in another activity.

E.5 Combinatorial Chemistry

a. General Aspects

Combinatorial chemistry involves the synthesis or biosynthesis of *chemical libraries* (a family of compounds having a certain base chemical structure) of molecules for the purpose of biological screening, particularly for lead discovery or lead modification.[62] Typically, these chemical libraries are prepared in a systematic and repetitive way by covalent assembly of *building blocks* (various reactant molecules that build up parts of the overall structure) to give a diverse array of molecules with a common *scaffold* (the parent structure in the family of compounds). The advantage of this methodology is that it is carried out on a solid (polymeric) support, so that isolation and purification of the product of each reaction can be performed by simple filtration and washing (with a variety of solvents) of the polymeric support to which the building blocks have been attached. Because of the insolubility of the polymer, everything not attached to the polymer is removed, which allows the use of excess reagents to drive the synthetic reactions. The disadvantages of this methodology are the difficulty in scaling up the reactions and the sluggishness of reactions. An alternative strategy (*covalent scavenger technology*) is to carry out the reactions in solution with excess reagent, which is then scavenged with a polymeric-supported scavenger after the reaction is completed (Figure 2.4). In this approach, filtration removes the excess reagent attached to the scavenger polymer, leaving the product in solution.[63] Another approach is to use polymer-supported reagents with solution reactions. To avoid problems of heterogeneous polymer reactions, soluble polyethylene glycol polymers can be used.[64]

The main differences among the various combinatorial approaches are the solid support used, the methods for assembling the building blocks, the state (immobilized or in solution) and numbers (a fraction of the total library or individual entities) in which the

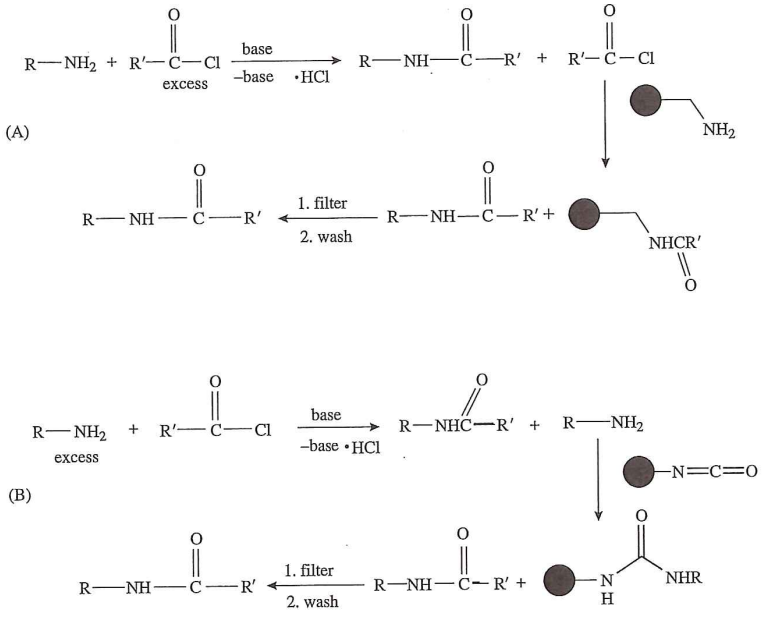


Figure 2.4 ▶ Use of polymer-bound reagents to scavenge excess reactants in a reaction

libraries are screened, and the manner in which the structures of active compounds are determined.

The number of possible different compounds in a library (N) is determined by the number of building blocks used in each step (b) and the number of synthetic steps (x). If an equal number of building blocks is used in each synthetic step, then Equation 2.1 holds. If the number of building blocks in each step is varied (e.g., b, c, and d for a three-step synthesis), then Equation 2.2 is relevant.

$$N = b^x \tag{2.1}$$

$$N = bcd \tag{2.2}$$

A combinatorial library of all of the pentapeptides that comprise the 20 commonly encoded amino acids would be $N = 20^5$ or 3.2 million different peptides. Combinational chemistry originally was used to make peptide libraries, but now is most commonly employed for the synthesis of large arrays of diverse nonpeptidic small molecules. It is only because of the discovery and development of HTS techniques that combinatorial chemistry was able to thrive. Unless there were a method to test 3 million new compounds generated in a library in a short period of time, there would be no advantage to being able to make that many compounds. Theoretical estimates conclude that there are up to 10^{180} possible drug-like molecules with molecular weights below 800, but this amount exceeds the mass of the universe. The much lower estimate of 10^{60} molecules[65] would still take over 10^{51} years to synthesize, even if a million compounds a day were prepared![66] Therefore, even combinatorial methods are inconsequential relative to the total number of compounds possible, but the belief is that it can approach the theoretical value of compounds quicker than by conventional synthetic methods. However, the diversity of molecules that can be attained by combinatorial chemistry