The drug discovery process

In the fields of medicine, biotechnology and pharmacology despite advances in technology and ever-greater understanding of biological systems, the drug discovery process can too often be a lengthy, expensive, difficult, and inefficient process. Once a compound has shown its value in testing it will begin the process of drug development prior to clinical trials. Peira reduces the research and development costs associated with the drug discovery process.

Drug discovery and design requires the identification of candidates, synthesis, characterisation, screening, and assays for therapeutic efficacy. Peira's expertise lets researchers move away from time-consuming manual procedures, towards greater automation of the research process. Our hardware and software solutions limit bias and increase the reliability of your experiments.

The five **key steps** of the drug discovery process:

1. Research & early development

This first phase of the drug development process is basic research. During this phase researchers try to understand the underlying mechanism or cause of a certain disease. Researchers look for new chemical or molecular entities that display promising activity against a particular biological target thought to be important for the disease. Other properties (including safety, toxicity, etc) and metabolic effects of the identified entities in humans are not focused on at this stage.

2. Preclinical research

Preclinical research must be completed before clinical trials (testing in humans) can start. During this stage important feasibility, iterative testing and safety data is collected. The main goal of preclinical study is to determine a product's ultimate safety profile. Products may be new or iterated medical devices, drugs and gene therapy solutions. Each class of product may undergo different types of preclinical research. For instance, drugs may undergo pharmacodynamic, pharmacokinetic, ADME, and toxicity testing through animal testing. Typically, both in vitro and in vivo tests will be performed.

3. Chemical & pharmaceutical development

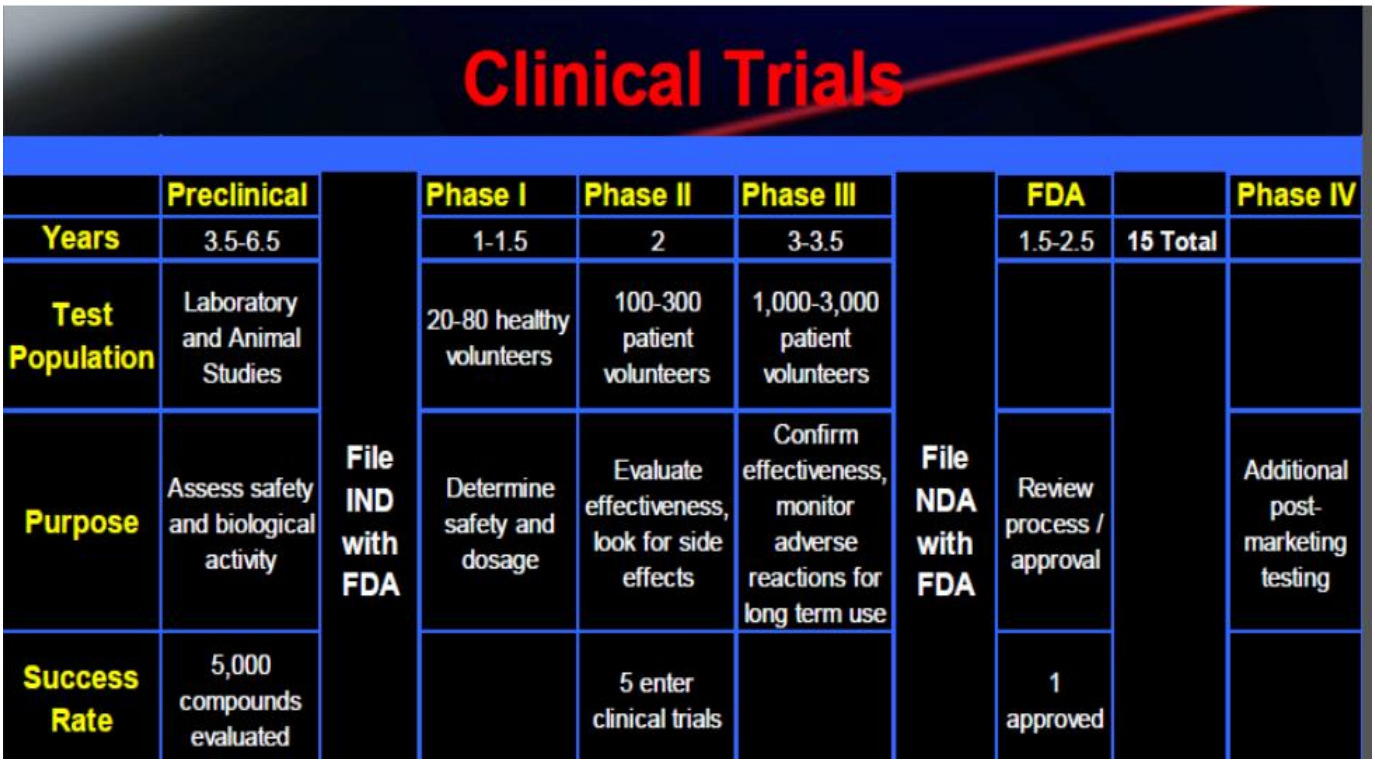
During the chemical and pharmaceutical development phase the aim is to design a quality product and its manufacturing process to consistently deliver the intended performance of the product. The information and knowledge gained from the studies and manufacturing experience provide scientific understanding to support the establishment of the design space, specifications, and manufacturing controls.

4. Clinical research

In this phase clinical trials are conducted to collect safety and efficacy data for new drugs. These trials can only take place once there is adequate information about the quality of the product, its non-clinical safety and once health authority approval has been granted. As positive safety and efficacy data are gathered, the number of patients is typically increased. Clinical trials can vary in size from a single centre in one country to multi-centre trials in multiple countries. After conclusion of the clinical trials the drug will be submitted for regulatory approval, for example with the Food and Drug Administration in the US or with the European Medicines Agency.

5. Chemical & pharmaceutical production

Once a new drug has been approved by the regulatory agencies such as the Food and Drug Administration in the US a full scale manufacturing plant will be built based on the scientific understanding gathered during the chemical and pharmaceutical development phase.



LEAD DISCOVERY

10.1 Approaches to Searching for Hits
- Traditional Library Screening
- Fragment-Based Screening
- Virtual Screening

10.2 Filtering Hits to Leads
- Pharmacodynamics and Pharmacokinetics
- Biological Assays
- Lipinski's Rules and Related Indices
- Final Concerns for Promotion of a Hit to a Lead

10.3 Special Cases
- Serendipity
- Clinical Observations
- Natural Products

Once a target, normally an enzyme or receptor, has been established and an assay for activity has been developed, the medicinal chemistry team must discover, find, and make compounds that interact with the target. Through the screening process, some compounds emerge with sufficient activity to warrant further investigation. The active compounds are then examined against a number of criteria, including complexity and anticipated pharmacokinetic behavior. Compounds that satisfy the selection criteria are called leads and advanced for further optimization of activity, selectivity, and biological behavior. Occasionally, leads are found through other methods, such as serendipity or clinical observations. This chapter describes techniques of discovering active compounds through screening and selecting the most promising compounds as leads. The overall process is collectively known as lead discovery.

10.1 Approaches to Searching for Hits

The most common tool for discovering hits is library screening. The library may consist of traditional compounds with potentially high activity molecules, smaller fragments of less activity, or even virtual molecules tested through molecular modeling simulations.

Traditional Library Screening

The goal of screening of a library, in whole or in part, is to discover compounds with modest activity against a target. The active compounds discovered through a screen are called *hits*. The threshold for activity varies based on the target, but hit-level activity is typically 1 mM or lower. Targets are normally enzymes or receptors, so the term *activity* refers to an *IC*50 or *EC*50 value.

In-House Libraries

As medicinal chemists synthesize molecules during their day-to-day research, small samples of new compounds are submitted for inclusion in the company's compound library. Over the course of years and decades of research, a compound library steadily grows. A library will reflect the areas of research that have contributed to the collection. A company that has historically been strong in researching b-lactam antibiotics (**10.5** and **10.6**) would have a very different library from a company with strength in estrogen receptor binding compounds (**10.7** and **10.8**) (**Figure 10.1**). Once combinatorial chemistry became recognized in the 1990s as a method for making large numbers of

molecules, most pharmaceutical companies hired teams of chemists to create collections of molecules to augment the company's existing library. Samples from natural sources may also be included in a library.
When pharmaceutical companies merge, their libraries merge as well. Through a library, the purchasing company gains a tangible chemical record of the research of the acquired firm. In 1995, Glaxo acquired Wellcome. An area of strength for Wellcome was antiviral research. Wellcome's products at the time included acyclovir (**10.9**) and zidovudine (AZT, **10.10**), nucleoside analogues with activity against the herpes simplex and human immunodeficiency viruses, respectively (**Figure 10.2**). Today, Glaxo, now operating under the name GlaxoSmithKline (GSK), still maintains a strong presence in treatments for viral infections. One recently developed drug is lamivudine (**10.11**), an anti–human immunodeficiency virus nucleoside analogue.

Compound libraries may be bought and sold individually. After the dissolution of the Soviet Union in 1991, laboratories that had been formerly well funded by the Soviet government suddenly became essentially broke. As a means of generating funds, some research groups began to sell portions of their in-stock compounds. The samples were readily purchased by Western companies, including the pharmaceutical industry. The value of the compounds depends on the novelty of the structures and their purity. Specs, founded in 1987 and based in Delft, The Netherlands, purchases compounds from all over the world, mostly from academic laboratories. These compounds are added to Specs' existing library and in turn sold to interested companies. A company can search the Specs library and purchase promising compounds or the entire collection. Those compounds become the outright property of the purchasing company. If Specs has a sufficient
amount of a given compound, the company will sell samples of each compound many times over. The amount sold for each compound may be only 0.5 to 1.0 mg.

Out-Sourced Libraries
Just as a library can be purchased, a library can also be rented. The owner of the library typically enters into an agreement with a drug company. The drug company pays to access and test the compound in a screen. If the compound eventually results in the discovery of a new drug, the owner of the library may receive a bonus. One company built on this type of business model was Pharmacopeia, Inc., of Cranbury, New Jersey.
Pharmacopeia was founded in the early 1990s by W. Clark Still of Columbia University and Michael H. Wigler of Cold Spring Harbor Laboratory. Still and Wigler were early pioneers in the development of combinatorial chemistry for pharmaceutical development purposes.3 As of 2007, Pharmacopeia claimed to have used the Still and Wigler techniques to prepare a library of over 7.5 million compounds—a massive number that is far greater than the library of a typical major drug company. With a library of this size, Pharmacopeia entered licensing agreements with drug companies. In each partnership, Pharmacopeia brought a large library for discovering hits and an ability to make additional compounds as needed for lead optimization. In turn, the pharmaceutical company provided expertise in screening compounds, performing clinical trials, and marketing a drug. If a compound provided by Pharmacopeia became a marketed drug, Pharmacopeia shared in the revenues from the drug's sales. Over the years, as its own resources grew, Pharmacopeia shifted its business plan more toward developing drugs independently, without the involvement of an outside pharmaceutical company. In 2008, Pharmacopeia was purchased

by Ligand Pharmaceuticals of La Jolla, California. Ligand now owns the full chemical library of Pharmacopeia.

Mycosynthetix of Hillsborough, North Carolina, out-sources a large library of fungal broths. The screening of products from fungi is attractive because the number of different species of fungi is astonishingly large and essentially unexplored. Any biological activity discovered through screening compounds from fungi is almost certainly previously unknown and therefore more easily protected through patents.

Fragment-Based Screening

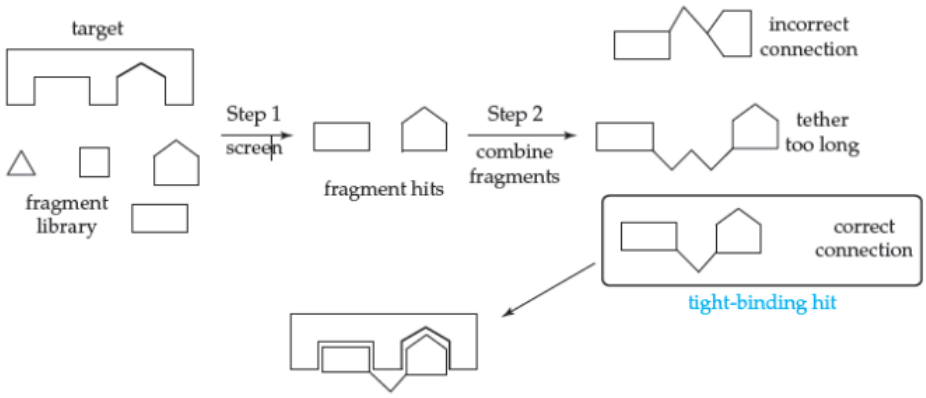
Fragment libraries are no different than traditional compound libraries except molecules in a fragment library are smaller. Fragments have a molecular weight of only 120 to 250 g/mol. Limiting molecular weight dramatically decreases diversity in the library. Far fewer molecules are required to sample the molecular space of a 250 MW library than a 400 or 500 MW one. Smaller molecules have fewer potential sites for intermolecular binding than larger molecules. Therefore, small molecules rarely bind as strongly as larger compounds. For example, a hit from a typical combinatorial library may show activity (*K*D, *IC*50, *K*i, *EC*50) at concentrations of 1 mM or lower. In contrast, a hit in a fragment library may be selected with activities of around 1 mM, which is a 1,000-fold difference in activity. Remember that a larger *IC*50 value implies weaker enzymeinhibitor binding.4,5

By itself, a single fragment with 1 mM binding is not very interesting. However, if multiple fragments are known to bind near the same site on a target, then the fragments can sometimes be connected to form a single strongly binding hit (**Scheme 10.2**). The key to discovering hits through fragment-based screening requires two steps. First binding fragments must be discovered. Second, the fragments must be properly connected and rescreened to discover a hit. Proper connection of the fragments can be a challenge. The tether between the fragments must be the correct length and placed appropriately. Successful examples of fragment-based hit discovery involve targets with two or more Active site pockets, each of which can accommodate a fragment-sized group. Because fragment connection requires a spatial understanding of how a fragment and target interact, fragment-based screening methods need to have a three-dimensional model of the target. Visual models arise from NMR, x-ray crystallography, and molecular modeling data which were discussed in the opening section of Chapter 9. In x-ray crystallography, x-ray structures of fragments bound to a target provide both the site and position of binding. With quality structural information to guide the drug discovery group, determining the ideal linker length and position is a much easier task.

Recent research performed in the laboratory of George Whitesides at Harvard University suggests that the only poor choice for a linker is one that is too short to span the distance between fragment binding sites. Linkers that are longer than necessary simply fold upon themselves to bring the fragments closer for binding a target. This research implies that initial linkers to tether fragments should be longer rather than shorter. The optimal length can be determined in a subsequent study.

The following three Case Studies demonstrate the use of fragment-based screening to discover leads.

SCHEME 10.2 Hit development through fragment-based screening



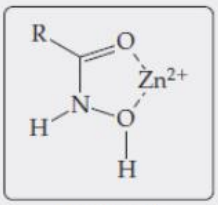
CASE STUDY

Inhibition of Stromelysin[7,8]

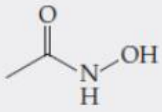
Stromelysin is a zinc-dependent protease that is responsible for breaking down and re-forming connective tissues, including collagen. Dysfunctional activity of stromelysin and related enzymes is associated with arthritis and tumor activity. An effective inhibitor of stromelysin may serve as a treatment for these conditions. With these facts in mind, a discovery team at Abbott used NMR to screen a fragment library for stromelysin binding. The technique of using NMR to discover hits from fragment libraries is often called *SAR by NMR*. SAR, or *structure-activity relationship*, is a term normally associated with the lead optimization process, not lead discovery. However, as has already been mentioned in this chapter, lead discovery and optimization are not completely independent processes.

Screening by NMR requires a ^{15}N-labeled target, stromelysin in this case, that has a well-characterized two-dimensional ^{15}N-^{1}H spectrum. In a labeled protein, each amino acid has a ^{15}N-^{1}H bond on the backbone of the protein. Each ^{15}N-^{1}H bond gives a separate signal in the NMR spectrum. If the spectrum is characterized, then the signals corresponding to amino acids in the active site are located. Addition of a fragment, if it binds the target, will cause changes in the spectrum. If the spectral changes occur with signals known to be associated with the active site, then the fragment likely binds at the active site. Once two active site binding fragments are identified, linkers can be designed to connect them.

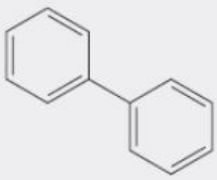
Before the Abbott study, two important facts about stromelysin were already known. First, the Zn^{2+} ion in the active site binds weakly to hydroxamic acids (Figure 10.3). So, acetohydroxamic acid (**10.12**) was chosen as the first fragment with a modest K_D of 17 mM. Second, the active site of stromelysin was known to



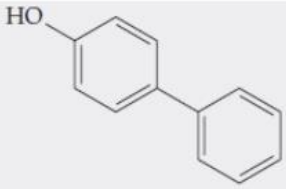
binding of hydroxamic acids to Zn^{2+}



acetohydroxamic acid
MW 75
10.12
K_D = 17 mM



biphenyl
10.13



4-hydroxybiphenyl
MW 170
10.14
K_D = 0.28 mM

FIGURE 10.3 Fragments with modest binding to stromelysin

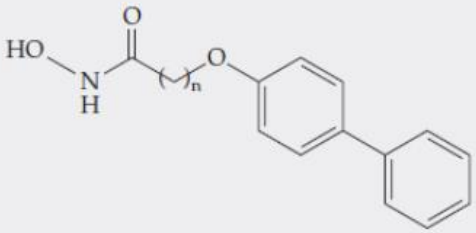
contain a hydrophobic binding site. A fragment library of biphenyl (**10.13**) and related compounds was selected to test binding in the hydrophobic pocket. 4-Hydroxybiphenyl (**10.14**) emerged as a promising fragment with a K_D of 0.28 mM.

Based on additional NMR information, the active site of stromelysin was modeled, and the relative orientations of fragments **10.12** and **10.14** were approximated in the binding pocket. Synthesis of a series of compounds (**10.15–10.18**), a combination of both fragments with different linker lengths, generated a set of micromolar inhibitors (Figure 10.4). The most potent was **10.16**. Note that both **10.12** and **10.14** are fragment hits with low molecular weights of 75 and 170. Compound **10.16**, as a combination of two fragments, shows an IC_{50} activity that is typical for a traditional hit. Structure **10.19** was subsequently found to be a strongly binding compound with an IC_{50} of 15 nM (0.015 μM).

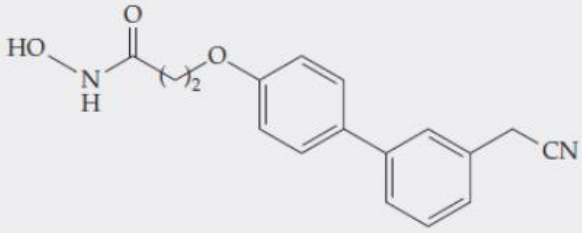
The effect of combining fragments can be examined energetically. K_D values are equilibrium constants and related to the standard free energy of association/binding (ΔG°_{bind}) by Equation 10.1.

$$\Delta G^\circ_{bind} = -2.3RT \log K_A = -2.3RT \log \frac{1}{K_D} = +2.3RT \log K_D \quad (10.1)$$

The standard free energy of binding for **10.12** with a K_D of 17 mM is −2.4 kcal/mol. Compound **10.14**, with a K_D of 0.28 mM, has a binding energy of −4.9 kcal/mol. Bringing both fragments together with an appropriate tether should provide a binding energy of −7.3 kcal/mol (−2.4 + −4.9). Instead, the energy of binding for **10.16** is −8.9 kcal/mol, 1.6 kcal/mol lower than expected. The difference is likely attributable to a combination of two factors: weak binding by the carbons in the tether and/or changes to the fragments by adding the tether. For example, capping the OH of **10.14** forms an ether. The nonpolar ether of **10.16** may complement the binding site better than the polar phenol of **10.14**.



	n	MW	IC_{50} (μM)
10.15	1	243	3.9
10.16	2	257	0.31
10.17	3	271	110
10.18	4	285	100



10.19
IC_{50} = 0.015 μM

FIGURE 10.4 Inhibitors of stromelysin

Sample Calculation Determining a Binding Energy from an IC_{50} Value

PROBLEM Determine the standard free energy of binding of compound **10.19** based on its IC_{50} of 0.015 μM.

SOLUTION Binding energies are simple to determine with Equation 10.1. The hardest part may be determining the proper value of R, the gas constant, to use in the equation. The value of R depends on the desired units for ΔG°. We have been using kcal/mol. For these units, we need to use $R = 0.00199$ kcal/mol · K. We will use $T = 298$ K. As always, be careful with the units on K_D and IC_{50} values. Always convert them to molarity. Therefore, in place of 0.015 μM, we need 1.5×10^{-8} M. With all the details handled, the calculation is fairly straightforward. Keep in mind that logarithmic operations are unitless, so the molarity units on IC_{50} disappear in the calculation. The ΔG°_{bind} calculates as −10.7 kcal/mol.

$$\Delta G^{\circ}_{\text{bind}} = +2.3RT \log K_D \qquad (10.1)$$

$$\Delta G^{\circ}_{\text{bind}} = 2.3 \times 0.00199 \frac{\text{kcal}}{\text{mol} \cdot \text{K}} \times 298\ \text{K} \times \log\left(1.5 \times 10^{-8} \frac{\text{mol}}{\text{L}}\right)$$

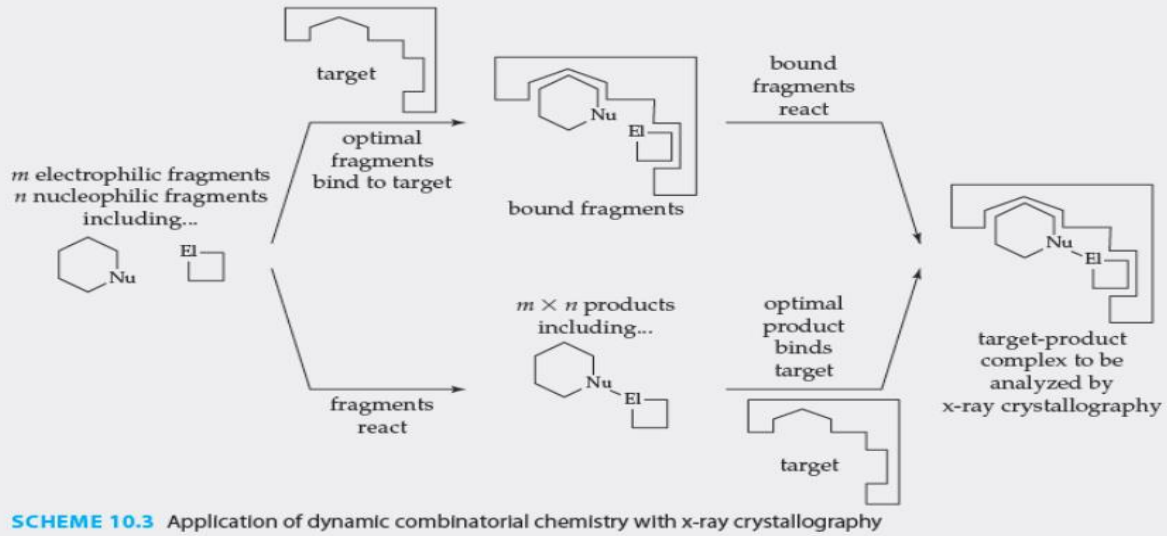
$$\Delta G^{\circ}_{\text{bind}} = -10.7 \frac{\text{kcal}}{\text{mol}}$$

CASE STUDY

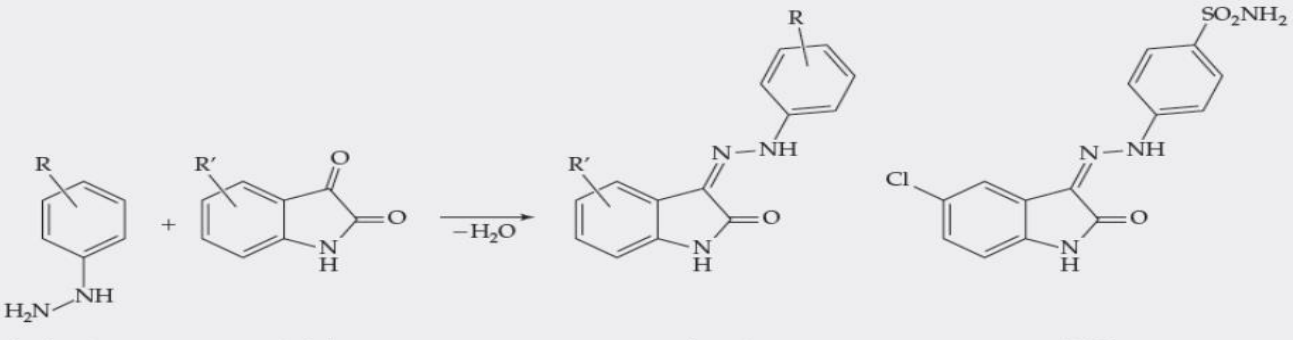
Inhibition of Cyclin-Dependent Kinase 2[9]

Researchers have been able to allow the target itself to select an ideal hit from a library of fragments and fragment products. This technique is called *dynamic combinatorial chemistry*. Fragments with complementary reactivity react in the presence of a target (Scheme 10.3). Either the target binds optimal fragments and causes them to link in situ, or the target selects the best product of a fragment reaction. In either case, the tightest-binding product occupies the active site of the target. Careful analysis of the electron-density map of the target cocrystallized with the product reveals the x-ray structure of the product. Naturally, a limitation of this method is that the target must be able to be crystallized.

An inhibitor of cyclin-dependent kinase 2, an enzyme implicated in a number of human cancers, has been developed through a simple application of dynamic



SCHEME 10.3 Application of dynamic combinatorial chemistry with x-ray crystallography



SCHEME 10.4 DCC-based discovery of a CDK2 inhibitor

combinatorial chemistry with x-ray crystallographic analysis. Kinase crystals were added to a solution of a fragment library of six arylhydrazines (**10.20**) and five oxindoles (**10.21**) (Scheme 10.4). A total of 30 condensation products (**10.22**) were possible, and all 30 had been shown to form under the conditions of the study. Furthermore, compounds of type **10.22** were previously found to bind the kinase, so the x-ray study was guaranteed to afford potent hits. The kinase crystals were isolated from the mixture and analyzed to reveal binding of one compound (**10.23**) in the active site.

Although the study of cyclin-dependent kinase 2 was a proof-of-principle example with compounds already known to show activity, the technique of dynamic combinatorial chemistry is a viable method for screening for hits from a fragment library.