DETECTION, PURIFICATION, AND ANALYSIS:

Detection, analysis, and purification of combinatorial libraries places high demands on existing analytical techniques because:
(a) the quantities to be analyzed are very small, sometimes pico moles of material.
(b) The analysis should be nondestructive, to allow recovery of the compound if possible.
(c) The methods must be suitable for rapid, parallel analysis—analysis cannot be the rate-limiting step in the procedure.

No single analytical technique can fit all the requirements, so usually some "hyphenated" analytical techniques are used, for example, high-performance liquid chromatography with a mass spectrometer detection system (HPLC-MS).

Chromatography is usually the first step in the analysis of a combinatorial mixture. If we start with solid-phase chemistry, we chemically cleave the compounds from the support and filter off the beads, giving a solution containing the compounds we synthesized.
If the solution contains just a single compound, we might use a spectrophotometer, to measure infrared (IR) and ultraviolet (UV) absorbance or fluorescence directly, or even nuclear magnetic resonance (NMR) spectroscopy, to determine the structure of the compound in solution.
If the solution contains a mixture of compounds, one must separate them before determining their structures. HPLC is a standard approach. The components in the mixture are detected by some optical method (UV absorption. fluorescence, refractive index. etc.) that gives rise to peaks on a graphical readout.
Sometimes, the output from the column is passed into a spectrophotometer or mass spectrometer to generate a spectrum for each fraction of the output. These spectra can be interpreted to determine the structure of the compound that caused a given peak.

It is also possible to use much larger chromatographic columns and run preparative HPLC to separate up to several milligrams of material for further analysis or biological assay.

Chromatographic separations and analyses can be fully automated. Thus, a chemist can place all the reaction vessels, microtiter plates, etc. from a combinatorial experiment into racks and use a robotic system to draw samples, inject them into the HPLC, and collect the data output into computer files or databases—all without further intervention from the chemist (except to wash the dishes!).
For this reason, speed and solvent handling are special concerns with combinatorial experiments. One approach that has been adopted to speed up analyses and reduce the amount of solvent that must be consumed is, supercritical fluid chromatography (SFC), their characteristic:

1. The solvent is not a common organic solvent such as acetone or ethanol. Instead, it is a pressurized gas like CO_2 that evaporates from the output, leaving pure compound behind.
2. Another advantage of SFC is speed: since the solvent molecules are small, diffusion is rapid, and separations take place in about half the time of ordinary HPLC separations or less.
3. Finally, the amount of "solvent" that is consumed is significantly lower with SFC.

A disadvantage is that certain compounds may not separate as well under SFC as under HPLC.

IR spectroscopy is often applied in combinatorial chemistry with special consideration, these are:

1. Since IR light can be reflected from materials, one can analyze resin beads directly, without cleaving the products from them.
2. Since the loading of product on any given bead is very small, usually computer-enhanced methods like Fourier transform IR (FTIR) are needed to enhance the very small spectral signal from one or a few beads.
3. Interestingly, the shape of the beads has been found to affect the IR spectra results, and flattened rather than spherical beads give stronger IR signals.

NMR spectroscopy; their characteristic are:

1. Gives more structural information than IR or UV spectroscopy, but it has traditionally not been nearly as sensitive.

2. Compounds are normally cleaved from solid support before analysis by NMR, since NMR on solid resin or on resin swollen by solvent give broadened peaks and low resolution.
3. A type of NMR called magic angle spinning NMR, in which the sample is inserted into the magnetic field at an angle of about 55^0, reduces the peak broadening and has been used to analyze swollen polymer beads directly.

NMR techniques that have been used to analyze combinatorial mixtures include:

1. Magic angle spinning NMR.
2. Various "two-dimensional" (2D) NMR techniques that use multiple magnetic fields.
3. HPLC-NMR.
4. Capillary electrophoresis coupled to NMR (CE-NMR).
5. NMR to detect the binding of drugs to receptors to identify active agents. This latter technique has been termed SAR with NMR.

Mass spectrometry (MS):

1. The technique most widely used for combinatorial library analysis.
2. The measurements can he made on resin beads directly.
3. A wide range of compounds can be analyzed.
4. MS analyses can be highly automated.
5. A very important use of MS in combinatorial chemistry is in quality control of combinatorial libraries. (Only MS provides the sensitivity and versatility to perform this checking with both solid-phase and solution-phase libraries.)

ENCODING COMBINATORIAL LIBRARIES:

Once we have found a mixture or sub library that shows biological activity, how do we determine exactly which structure or structures are responsible for the activity?
We can purify and analyze as described in the previous section. But if no direct analysis is available, we need to encode(شفرة) or tag(رقعة) the support or the molecules themselves, using physical or molecular "barcodes."(شفرة كومبيوترية توضع على السلع) An obvious approach that can be used only with small libraries is to physically label each vial of one-bead one-compound resin. This may be practical for a few tens of compounds, but what if we have a library of 32000 compounds, or even I000000 compounds, in a mixture?

Clearly, there is a need for a more automated means of identifying the structures that are in the library.
The most common approach to encoding solid-phase libraries is to attach a chemical tag to the resin beads as the target molecule gets synthesized. Typically, at each step in the reaction, a tag is attached that is unique for the given step.

For example, if we are creating a tripeptide and we have 10 possible amino acids at each position, we need to attach either a single tag that says the tripeptide on this bead has amino acid Ala at position 1, Phe at position 2, and Gly at position 3." or we need to attach three different tags, one for each position.

Types of chemical encoding:

1. The earliest types of chemical encoding were the attachment of oligonucleotides (usually single-strand DNA) to beads on which peptides or peptoids were being built.
Since there are 20 possible amino acids and only four nucleotide bases, enough bases must be attached at each amino acid addition step to identify properly the amino acid being attached. Although three bases are used in the DNA genetic code, it is customary(مالوف) to use up to six bases for library tagging.
For decoding, the DNA tag is amplified by use of the polymerase chain reaction (PCR). The same reaction that is used in forensic (شرعي)DNA analysis. For this reason, the chemical tag must also bear PCR primer sequences.

Two types of anchors have been used to connect the DNA tags to the solid support as shown in this figure:

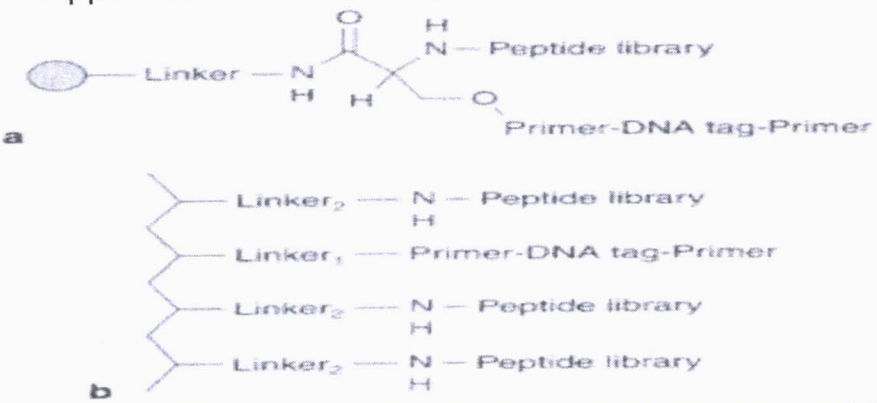


Figure 3–12 ▪ Two ways of attaching DNA tags to solid supports. **a.** A DNA tag is attached with each peptide molecule via a bifunctional serine residue. **b.** The DNA and peptide groups have separate linkers.

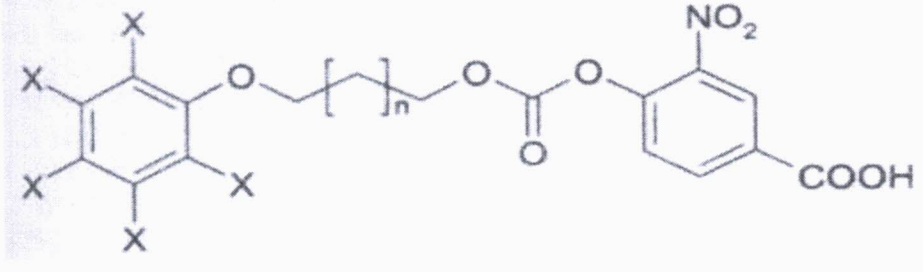
1. The growing DNA chain is attached to the α carbon of a serine group that is anchored to the solid support by a linker molecule. The growing peptide chain is attached to the serine amino group, possibly through a spacer.
2. The DNA chain has its own anchor to the solid support. In this case, fewer DNA tags are attached than the number of polypeptide molecules.

2. "Binary" approach (if DNA tags cannot be used, one can label beads by using a "binary" approach(مزدوج رقم ثنائي).
Suppose we are building a tripeptide with four possible amino acids at each position. We can use binary digits to encode which amino acid is at a given position us follows (each binary "number" is read from the right) (Table -I):

TABLE 3–1 Binary Encoding of a Tripeptide, Using 18 Possible Tags

Amino Acid	Position 1	Position 2	Position 3
Ala	00 00 00	00 00 00	00 00 00
Phe	00 00 01	00 01 00	01 00 00
Gly	00 00 10	00 10 00	10 00 00
Lys	00 00 11	00 11 00	11 00 00

Thus, using 18 different tags (3 x 6) we can encode for any of the (4^3 = 64) members of the library.
For example, if the product is Ala-Gly-Lys, the encoding would be 00000000100011 0000.
3. It is common to use poly halogenated aromatic compounds as tags, such as



Where X represents some combination of halogen atoms. The halogens make the tags show up clearly in MS analysis of the mixture, and by varying the chain length a, the tag can be made flexible enough not to interfere with attachment of the product.

4. Isotopically labeled peptides and dyes.

When it is not possible to use chemical tags, one must physically label the solid or liquid support itself.
Types of physically label:
1. Radiofrequency encoding, in which tiny microchips are added to the resin or to the solution phase. As various reactions are conducted to generate the products, at each step a radiofrequency signal is stored in the microchip.
This signal can be recalled to identify the sequence of reactions that generated the product.
2. Laser optical encoding is yet another approach, in which the solid support consists of a ceramic chip covered with a polypropylene—polystyrene polymer solid phase. The barcode pattern is actually burned into the ceramic at each step in the reaction and is decoded visually with use of a microscope.
3. Finally, one can embed semiconductor particles into the solid phase that fluoresce at different wavelengths. These are called "quantum dots" by their manufacturer.

HIGH-THROUGHPUT SCREENING (HTS):(حاجز،غربال عالي)

Without the ability to screen libraries rapidly for activity, there would be no combinatorial chemistry. Fortunately, the biologists are just as adept at developing rapid high-throughput assays as the chemists are at generating structures.
(HTS): The process for rapidly assessing the activity of samples from a combinatorial library or other compound collection, usually done by running parallel assays in plates of 96 or more wells. A screening rate of 100.000 assays per day is termed ultrahigh-throughput screening.
HTS is an extremely broad topic, encompassing enzymes, organelles, cells, various tissues, whole organs, and even whole animal testing, via cassette dosing. This section briefly discusses only a part of the role of HTS in drug discovery, with emphasis on a few recent developments.

Successful HTS programs integrate(يدمج، يوحد) several activities, including:

1. Target identification (genomics and molecular boilogy groups).
2. Reagent preparation (protein expression and Purification groups).
3. Compound management (تعامل) (information management group).
4. Assay development (biologist and pharmacologist).
5. High-throughput library screening (biologists and chemists). Formerly, these activities were handled separately, and multiple handoffs of samples were involved.

It is becoming more common to integrate the activities and share expertise. This increases efficiency of the screening process.

Another route to increasing efficiency is a move to higher-density screening platforms (خطط ، برامج).
The standard layout for HTS has been a 96-well micro titer plate (12 x 8). Denser formats, up to 1,536 wells per plate, are increasingly being used. This requires advances in liquid handling, precision of detection, and laboratory automation.
One of the first activities in developing a HTS assay is selecting the target. About 500 targets are currently being used by drug companies. Of these:

1. Cell membrane receptors, mostly G-protein—coupled receptors make up the largest group (about 45% of the total).
2. Enzymes make up the next largest group (28%).
3. Hormones (11%).
4. Unknowns (7%).
5. Ion channels (5%).
6. Nuclear receptors (2%).
7. DNA (2%).

It is expected that the annotation (حاشية) of the human genome will add additional targets. Although the rate of this addition is not known.

In fact, much HTS involves screening compounds that are part of the corporate storehouse of compounds synthesized in the past, or they may be a library purchased from a vendor(بائع). Such libraries usually consist of microtiter plates containing frozen or dried samples of compound—perhaps only micrograms per well.

The size of such libraries may range from a few thousand compounds to nearly a million.
The cost of completely screening such a library against just a single assay may amount to over $300,000, so such large scale screens are conducted rather infrequently, compared with routine day-to-day screens. It has been estimated that one must screen at least 120,000 "quality" compounds (i.e.. diverse drug-like structures) to discover a single-lead series for a therapeutically sound target.

As discussed above in the section on pooling strategies, one can reduce the screening effort by pooling groups of structures and running assays on mixtures of compounds.
This also led to:

1. Conserves reagents and biological material.
2. Have smaller storage requirements.
3. Requires fewer personnel.

There are potential problems with pooling:

1. A number of factors limit the number of different compounds we can test in a given well, including ionization, reactivity, and solubility.
2. Compounds can enter a screening program in a nonrandom order, such that a given assay plate may have compounds that are highly similar structurally. This may give rise to false-positive hits. False-negative hits are less likely to arise from pooling.
3. The use of replicates— compounds from the same series— in a given assay.

If only one representative of a given series is present. The chance of missing that series as a possible lead series is greater than if multiple members are present. Therefore, it is common to include several members of each series in a given assay when possible.

To be effective, a given compound must dissolve completely in the assay medium. It is common to add a small amount (1%) of dimethyl sulfoxide (DMSO) to the assay to assist solvation. The best concentration of compound to use is somewhat debatable.

High concentrations (10 μM and above) often lead to more false positives than screening at a low concentration (3 μM). The reason for this may be nonspecific binding at the higher concentration.

Just as there are several ways to detect and identify members of a combinatorial library, there are many ways so measure activity in HTS assays. Any such method must be:

1. Accurate.
2. Reproducible.
3. Have a high signal, to-noise ratio (S/N).

Typically, the result of HTS is a qualitative (yes/no) or semi quantitative one (high-medium-low), rather than a precise value (e.g.. KI_{50} or LD_{50}).
The methods for detection in HTS fall into the categories of nonradiometric and radiometric.

Nonradiometric methods include absorbance, fluorescence, and luminescence(صفة ذو علاقة بالتالق) spectroscopy. Enzyme assays are a common example.
Radiometric methods include filtration and scintillation (يومض، يطلق شررا) proximity assay (SPA). These assays use radioisotopes. So safe storage and handling are of concern.

In filtration assay, a radioactive substrate bound to a capture group is cleaved by its enzyme, removing the radioactivity from the capture group. The mixture is filtered through special filter paper that the capture group sticks to, but everything else passes through. A scintillation fluid is added, and the radioactivity of the filter is measured. The degree to which the radioactivity is retained measures the strength of the inhibition (Fig. 3-I3a).

SPA is a newer, simpler method (Fig. 3-I3b). We start with the same radioactive substrate, which may not necessarily need a capture group. The enzyme and potential drug are added, causing the cleavage of the substrate to some degree.
Now, instead of filtering, a special resin bead coated with a scintillant—a compound that fluoresces in the near presence of the radioactive substrate (near being about 20 μM)—is added to the mixture. The lysed and unlysed substrate binds to the beads, and if the radioactive part of the substrate is still attached, the bead will fluoresce. If not, the radioactive parts of the substrate floating in the solution will be too far from the beads to cause any fluorescence.
The presence of fluorescence implies that the test compound inhibited the enzyme.

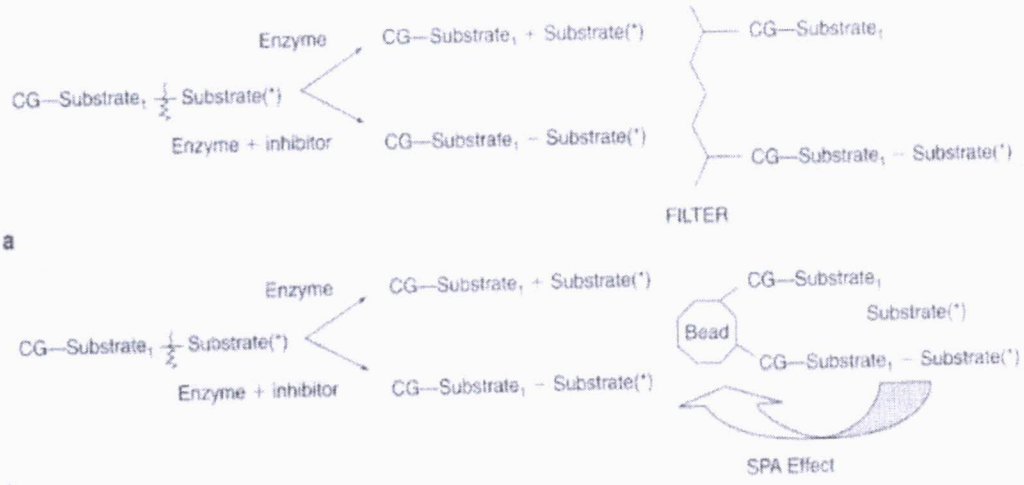


Figure 3–13 ■ Comparison of filtration and scintillation proximity assays. **a.** In filtration assay, the enzyme, substrate, and inhibitor are mixed; the uninhibited enzyme splits the radioactive portion (*) off the substrate, and filtering the mixture, followed by measuring the radioactivity of the filter, tells how much inhibition has occurred. **b.** In SPA, the same mixture is treated with resin beads containing a scintillant that fluoresces only in close proximity to the radioactive source. Any radioactivity that was split off by the enzyme does not need to be filtered in SPA.

The advantage of SPA over filtration is that:

1. No filtering of the solution is needed, so beads can be added directly to the assay mixture in wells or test tubes.
2. Special scintillation fluid is not needed.

The beads for SPA can be engineered to attach a variety of substrate types.

Other HTS assay advances include:

1. The use of microorganisms such as bacteria and yeast.
2. The cloning and expression of mammalian receptors in microorganisms.
3. Probing protein— protein interactions.
4. Very importantly, DNA and protein arrays.

VIRTUAL (IN SILICO) SCREENING: (افتراضي،عملي)

Virtual, or in silico, screening refers to the use of computers to predict whether a compound will show desired properties or activity on the basis of its two-dimensional (2D) or three dimensional (3D) chemical structure or its physicochemical properties.

The motivation (دافع، حافز) for using virtual screening arises from:

1. The flood (فيضان، طوفان) of new structures coming from combinatorial chemistry.
2. The expense and time required to run conventional HTS.

3. The ethical concerns about using animal tissue instead of predictive models.
4. An increasing failure rate for structures coming out of combinatorial programs.

In general, a virtual screening program attempts to answer one or both of these questions:

I. Will a particular compound show sufficient binding to a known receptor?
2. Will a particular compound possess any undesirable ADMET properties?

To answer these questions, we must build

- computer models of the interaction of drugs with receptors (docking, molecular modeling, quantum mechanics) and
- models for predicting ADMET properties on the basis of chemical structure (QSAR models).

Like much of drug discovery, the virtual screening process is a cyclic one composed from:

- A collection of structures and run predictive models on them.
- Generating some subset of "best" structures.
- Test these structures in real assays or screens to see if the predictions were accurate.
- Incorporate information learned from the real assays back into our predictive models to improve them for future use.

Models for the prediction of binding arise from the field of molecular modeling. This includes molecular mechanics (predicting the 3D structure of molecules from the standpoint of the atomic nuclei) and quantum mechanics (predicting the 3D structure of molecules from the standpoint of their electrons).
We can generate fairly accurate 3D structures of stand-alone molecules with either of these approaches.

If we know the 3D structure of a receptor, we can predict whether a given compound will "fit" into the receptor with sufficiently tight binding to present normal substrates from binding. i.e... an enzyme inhibitor (or drug). Typically, this does not consider any effects of transport, metabolism, interaction with solvent. etc. Consequently, the existing models for the binding of drugs to receptors are rather crude (فض، خشن), with errors of 50% or more.
Nevertheless, for the purpose of screening, this is often adequate.

If a structure is predicted to show tight binding, it will probably be synthesized and tested in a real HTS assay, to obtain a more accurate estimate of its activity.

We can combine the prediction of binding to a receptor with the design of molecules. If we start with a set of building blocks and a known receptor, we can align (يرصف) complementary building blocks with pockets in the receptor, in conformations that maximize hydrogen bonding and other interactions. When the fragments are aligned, we can then connect the fragments with appropriately sited spacer fragments to "build" a drug molecule within the confines (حدود، حاجز) of the receptor. This approach is sometimes called *de novo* or structure based drug design.

Pfizer researchers found that since the introduction of combinatorial chemistry, chemists had tended to design larger, more complex, and more lipophilic structures than in the past. Since the structures were supposed to be designed as lead structures rather than optimized drug molecules, it was becoming more difficult to optimize them into drugs. This led Lipinski to enumerate some rules for the rejection of structures. He proposed rejecting any structures that fail two or more of the following criteria:

- Molecular weight should be <500.
- Number of hydrogen bond donors (NH. OH should be fewer than 5.
- Number of hydrogen bond acceptors t—N—. —0—. —S—) should be fewer than 10.
- Calculated log P value should be less than 5.

Since the number 5 shows up in many of these criteria, they have become known as the "rule of five." The "rule of five" is designed to be a yes/no filter for the rejection of structures.

More quantitative models designed to predict some value or levels of property have also been developed. Examples include

1. Predicting Caco-2 cell permeability (Caco-2 cells are human intestinal epithelial cells that can be layered and studied in vitro)
2. Predicting binding to cytochrome P-450.

LIpInski "rule of five": A set of criteria for predicting the oral bioavailability of a compound on the basis of simple molecular properties (molecular weight. <500: log P. <5; number of

hydrogen-bond donors. <5; and number of hydrogen-bond acceptors. <10). Typically, the criteria are applied to a library to filter structures from the library before any synthesis takes place. Any structure exceeding two or more criteria is rejected.