**Combinatorial Chemistry:** molecular diversity

*Synthesis and Applications of Small Molecule Libraries.*


pharmaceutical industry- drug discovery

\[
\text{Lead Cmpd} \xrightarrow{\text{optimization (synthesis)}} \text{Drug} \xrightarrow{10 \text{ years}} \text{$400 - 800$ million}
\]

Lead identification: literature (open & patent) 
nature (natural products)

Careful optimization of a lead structure via chemical synthesis 
“methyl-ethyl-butyl-futile game”

Number of marketable drugs per compounds that undergo preliminary biological testing: 
\[
\frac{1}{10,000} \quad \text{Rational drug design} \\
\quad \text{Combinatorial chemistry}
\]

However, they are the natural substrates for many enzymes and receptors (drug targets) and have well-defined conformation: excellent lead compounds

Enzymes: converts a substrate to a distinct product
Receptor: binds a ligand (no reaction), causing a chain of physio-chemical events leading to a pharmacological response.

Agonist: substance that interacts (binds) with a receptor and elicits an observable response
Antagonist: substances inhibits the affect of an agonist, but has no biological activity of its own
Example of molecular diversity:
tetra-peptide: \( \text{H}_2\text{N-A-B-C-D-CO}_2\text{H} \)
consider only the 20 natural amino acids (L-series)
\( 20^4 = 160,000 \) different tetra-peptides!

now include the 19 D-amino acids (20 L + 19 D = 39)
\( 39^4 = 2.3 \) million different tetra-peptides!!

now include 20 unnatural amino acids
\( 59^4 = 12 \) million different tetra-peptides!!!!

Combinatorial chemistry: method by which a family (library) of related compounds (structurally & synthetically) can be prepared and evaluated (screened)

For multi-step synthesis, one must use solid-phase synthetic approach in order to expedite purification of intermediates

\[ \text{Split synthesis (mixture libraries)} \]
Why not?

Reactivity of the coupling reaction may be different and this could bias the library

Split synthesis approach, all compounds are equally represented (each coupling is individually controlled)
Deconvolution of the library via biological activity and sub-library re-synthesis

Biological activity of the library

select library H

Lead Cmpd
Problems:
- deconvolution of the library can be labor intensive
- can be fooled by low concentrations of highly active compounds
- activity is dependent upon the compound and its concentration
- activity observed is the the combined activity of the entire library

Advantage:
- can synthesize a very large number of compounds very quickly and relatively easily depending on the chemistry.

Encoded Libraries
Encoded libraries (con’t)

CODE: 1= A 2= B 3= C 4= D 5= E 6= F 7= G 8= H 9= I

Cleave the marker and analyze the code. Each unique marker codes for a unique peptide sequence.

Synthesis of an encoded library

Code (1-10%) library

1) linker- AA1-NH-FMOC
2) linker-code1-NH-BOC

library region and coding region must be orthogonally protected
Reading the code

selectively remove the
code from the resin

hydrolyze and analyze

“On-Bead” Assay for Receptor Binding
using a Fluorescently Labeled Receptor

1) Add labeled receptor labelled with fluorescent tag
2) allow to bind to ligands on the beads
3) wash to remove unbound receptor
Spatially Addressable, Parallel Synthesis

Spatially Addressable, Parallel Synthesis (con’t)
Spatially addressable, parallel libraries

Advantage:

• synthesizing pure compounds, no need to deconvolute the library

Disadvantage

• libraries tend to be much smaller

Diversomer apparatus

ACE Inhibitors

Asp–Arg–Val–Tyr–Ile–His–Pro–Phe–His–Leu → Val–Ile–His–Asn

angiotensinogen

aspartyl protease

renin

Asp–Arg–Val–Tyr–Ile–His–Pro–Phe–His–Leu

angiotensin I

(little biological activity)

Zn²⁺ protease

angiotensin converting enzyme (ACE)

Asp–Arg–Val–Tyr–Ile–His–Pro–Phe

angiotensin II

(vasoconstriction → high blood pressure)

Proteases: catalyzes the hydrolysis of peptide bonds

1. Serine protease
2. Cysteine protease
3. Aspartyl protease
4. Zinc (metallo) protease

ACE: zinc protease (no x-ray or NMR structure), compared to carboxypeptidase or thermolysin
important catalytic groups: Glu-270, His-196, His-69 (catalytic triad)
Carboxypeptidase: General base mechanism

Nucleophilic mechanism (acyl enzyme complex)

ACE inhibitor Library

240 cmpds, each are multiple stereoisomers

$K_i \sim 160 \text{pM (3x better than Captopril}$
Parallel Synthesis of a Benzodiazepine Library

**2-aminobenzophenones**

<table>
<thead>
<tr>
<th>2-aminobenzophenones</th>
<th>α-amino acids</th>
<th>alkylating agents</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Diagram" /></td>
<td><img src="image" alt="Diagram" /></td>
<td><img src="image" alt="Diagram" /></td>
</tr>
</tbody>
</table>

2 (2-aminobenzophenone) x 12 (amino acids) x 8 (alkyl halides) = 192 compounds

**Parallel Synthesis of a Benzodiazepine Library**

Cholecystokinin (CCK)

CCK-8: Asp-Tyr(SO3H)-Met-Gly-Trp-Met-Asp-Phe-NH2

![Diagram](image)
**Peptide Mimetics or Peptidomimetics**

Desirable pharmacological properties
1. metabolic stability
2. good bioavailability
3. high affinity ($K_i \sim 10^{-9}$) and specificity for a given receptor or enzyme
4. minimal side effects

Natural ligands for receptors or enzymes are often peptides: peptides are good lead structures but poor drug candidates

---

**Peptidomimetics Strategies**
1. Amino acid modification
2. Dipeptide analogues (isosteres and TS analogues)
3. Peptide backbone modifications
4. Secondary structure mimics

**Amino Acid Modifications**: changes (restricts) conformational degrees of freedom of the peptide
Dipeptide analogues

Asp-Arg-Val-Tyr-Ile-His-Pro-Phe-His-Leu

ACE Inhibitor

Transition state analogues for a protease:

Statine

Transition State

active site water

Statine
TS analoges via combinatorial synthesis

Peptide Backbone Modifications: Increase biological half-life
Retro-Inverso Isomers: Change from L- to D-amino acids and inverse N to C sequence

Interesting idea, but the retro-inverso isomers usually showed much weaker binding. However, they were significantly more resistant to enzymatic digestion.
Amide Bond Isosteres:

\[
\begin{align*}
\text{Amide} & \quad \leftrightarrow \quad \text{Acid} \\
\text{Acid} & \quad \leftrightarrow \quad \text{Alcohol} \\
\text{Alcohol} & \quad \leftrightarrow \quad \text{Ether}
\end{align*}
\]

Secondary Structure Mimics:

main 2° structures: \(\alpha\)-helix, \(\beta\)-sheet, \(\beta\)-turns, \(\gamma\)-turns

\(\alpha\)-helix mimic


“Ala-scan” determine which residues are needed for binding

Replace with a non-peptide scaffold
Ter-phenyl as an $\alpha$-helix mimic

“stapled” $\alpha$-helices
β-amino acids

β-Sheet mimics

β-Turn Mimics:
Combinatorial synthesis of linchpin β-turn mimic

Evaluated for Somatostatin receptor binding

\[ n = 1,2 \times R_2 = 34 \text{ AA} \times R_3 = 10 = 1292 \beta\text{-turn mimics} \]