Drug design

Drug design, rational drug design, rational design
An inventive process of finding new medications based on the knowledge of a biological target. The drug is most commonly an organic small molecule that activates or inhibits the function of a biomolecule such as a protein, which in turn results in a therapeutic benefit to the patient. In the most basic sense, drug design involves the design of small molecules that are complementary in shape and charge to the biomolecular target with which they interact and therefore will bind to it.

You may think of the true meaning of drug design as ligand design (i.e., design of a small molecule that will bind tightly to its target).

Typically a drug target is a key molecule involved in a particular metabolic or signaling pathway that is specific to a disease condition or pathology or to the infectivity or survival of a microbial pathogen.

Some approaches attempt to inhibit the functioning of the pathway in the diseased state by causing a key molecule to stop functioning. Drugs may be designed that bind to the active region and inhibit this key molecule.

Another approach may be to enhance the normal pathway by promoting specific molecules in the normal pathways that may have been affected in the diseased state.

These drugs were designed so as not to have serious side effects caused by interacting with other important “off-target” molecules or antitargets that may be similar in appearance to the target molecule.

Ligand-based drug design

Also called indirect drug design which relies on knowledge of other molecules that bind to the biological target of interest. These other molecules may be used to derive a pharmacophore model that defines the minimum necessary structural characteristics a molecule must possess in order to bind to the target. In other words, a model of the biological target may be built based on the knowledge of what binds to it, and this model in turn may be used to design new molecular entities that interact with the target.

An example of a pharmacophore model of the benzodiazepine binding site on the GABAA receptor. White sticks represent the carbon atoms of the benzodiazepine diazepam, while green represents carbon atoms of the nonbenzodiazepin. Red and blue sticks are oxygen and nitrogen atoms that are present in both structures. The red spheres labeled H1 and H2/A3 are, respectively, hydrogen bond donating and accepting sites in the receptor, while L1, L2, and L3 denote lipophilic binding sites.

Note: Mild inhibition of neuronal firing by drugs acting at the GABAA receptor causes a reduction of anxiety in the patient (an anxiolytic effect) while more pronounced inhibition induces general anesthesia.
Structure-based drug design

Also called direct drug design which relies on knowledge of the three dimensional structure of the biological target obtained through methods such as x-ray crystallography or NMR spectroscopy. If an experimental structure of a target is not available, it may be possible to create a homology model of the target based on the experimental structure of a related protein. Using the structure of the biological target, candidate drugs that are predicted to bind with high affinity and selectivity to the target may be designed using interactive graphics and the intuition of a medicinal chemist or various automated computational procedures to suggest new drug candidates.

The information about the structural dynamics and electronic properties about ligands increased with more information concerning 3D structures of biomolecular targets. Current methods for structure-based drug design can be divided roughly into two categories.

1) “finding” ligands for a given receptor using database search a large number of potential ligand molecules are screened to find those fitting the binding pocket of the receptor. The key advantage of database searching is that it saves synthetic effort to obtain new lead compounds.

2) “building” ligands Ligand molecules are built up within the constraints of the binding pocket by assembling small pieces in a stepwise manner. These pieces can be either individual atoms or molecular fragments. The key advantage of such a method is that novel structures, not contained in any database, can be suggested.

Requirements of rational drug discovery

The traditional methods of drug discovery rely on trial-and-error testing of chemical substances on cultured cells or animals, and matching the apparent effects to treatments. Due to the complexity of the drug design process, two terms of interest are still serendipity and bounded rationality.

Rational drug design begins with a hypothesis that modulation of a specific biological target may have therapeutic value. In order for a biomolecule to be selected as a drug target, two essential pieces of information are required.

The first is evidence that modulation of the target will have therapeutic value. This knowledge may come from, for example, disease linkage studies that show an association between mutations in the biological target and certain disease states.

The second is that the target is “drugable”. This means that it is capable of binding to a small molecule and that its activity can be modulated by the small molecule.

• Once a suitable biomolecule target has been identified, the target is normally cloned and expressed.
• The expressed target is then used to establish a screening assay.
• In addition, the three-dimensional structure of the target may be determined.
• The search for small molecules that bind to the target is begun by screening libraries of potential drug compounds. This may be done by using the screening assay (a “wet screen”).
• If the structure of the target is available, a virtual screen may be performed of candidate drugs.

Ideally the candidate drug compounds should be “drug-like”, that is they should possess properties that are predicted to lead to oral bioavailability, adequate chemical and metabolic stability, and minimal toxic effects. Several methods are available to estimate druglikeness such as Lipinski’s Rule of Five and a range of scoring methods such as Lipophilic efficiency.
The iterative process of structure-based drug design

The process of structure-based drug design is an iterative one (see Figure) and often proceeds through multiple cycles before an optimized lead goes into phase I clinical trials. The first cycle includes the cloning, purification and structure determination of the target protein or nucleic acid by X-ray crystallography, NMR, or homology modeling. Using computer algorithms, compounds or fragments are positioned into a selected region of the structure. These compounds are scored and ranked based on their steric and electrostatic interactions with the target site, and the best compounds are tested with biochemical assays. In the second cycle structure determination of the target in complex with a promising lead from the first cycle, one with at least micromolar inhibition in vitro, reveals sites on the compound that can be optimized to increase potency. Additional cycles include synthesis of the optimized lead, structure determination of the new target/lead complex, and further optimization of the lead compound. After several cycles of the drug design process, the optimized compounds usually show marked improvement in binding and, often, specificity for the target.

Active site identification of the target biomolecule

Active site identification is the first step
It analyzes the protein to find the binding pocket, derives key interaction sites within the binding pocket, and then prepares the necessary data for Ligand fragment link. The basic inputs for this step are the 3D structure of the protein and a pre-docked ligand in PDB format, as well as their atomic properties. Both ligand and protein atoms need to be classified and their atomic properties should be defined, basically, into four atomic types:

- Hydrophobic atom: All carbons in hydrocarbon chains or in aromatic groups.
- H-bond donor: Oxygen and nitrogen atoms bonded to hydrogen atom(s).
- H-bond acceptor: Oxygen and sp2 or sp hybridized nitrogen atoms with lone electron pair(s).
- Polar atom: Oxygen and nitrogen atoms that are neither H-bond donor nor H-bond acceptor, sulfur, phosphorus, halogen, metal, and carbon atoms bonded to hetero-atom(s).

The space inside the ligand binding region would be studied with virtual probe atoms of the four types above so the chemical environment of all spots in the ligand binding region can be known. Hence we are clear what kind of chemical fragments can be put into their corresponding spots in the ligand binding region of the receptor.
Characterization of the binding pocket

(a) Catalytic Function

(b) Regulatory mechanism

(c) Cleft Detection

(h) Interaction Energy

\[ E = \sum E_{\text{vdW}} + \sum E_{\text{el}} + \sum E_{\text{hb}} + \ldots \]

(d) Geometrical Analysis

(e) Flexibility

(g) Surface Characteristics

\[ \nabla \cdot (c \nabla \phi) - \kappa^2 \phi + 4\pi \rho = 0 \]

(f) Conservation

Five examples of AMP, ATP, and NAD to show the diversity of binding pocket shapes

Not every ligand atom contacts a protein atom and thus leaves space between parts of the ligand and the protein. The space is partially occupied by crystallographic observable water molecules. The reconstructed pocket shape shown as a black coloured mesh, the ligand shown in varicolour and the oxygen atoms of the water molecules shown as green coloured spheres.
Computer-aided drug design

Computer-aided drug design uses computational chemistry to discover, enhance, or study drugs and related biologically active molecules. The most fundamental goal is to predict whether a given molecule will bind to a target and if so how strongly.

Molecular mechanics or molecular dynamics are most often used to predict the conformation of the small molecule and to model conformational changes in the biological target that may occur when the small molecule binds to it. This provides semi-quantitative prediction of the binding affinity. Also, knowledge-based scoring function may be used to provide binding affinity estimates. These methods use linear regression, machine learning, neural nets or other statistical techniques to derive predictive binding affinity equations by fitting experimental affinities to computationally derived interaction energies between the small molecule and the target.

Semi-empirical, ab initio quantum chemistry methods, or density functional theory are often used to provide optimized parameters for the molecular mechanics calculations and also provide an estimate of the electronic properties (electrostatic potential, polarizability, etc.) of the drug candidate that will influence binding affinity.

Ideally the computational method should be able to predict affinity before a compound is synthesized and hence in theory only one compound needs to be synthesized. The reality however is that present computational methods are imperfect and provide at best only qualitatively accurate estimates of affinity. Therefore in practice it still takes several iterations of design, synthesis, and testing before an optimal molecule is discovered. On the other hand, computational methods have accelerated discovery by reducing the number of iterations required and in addition have often provided more novel small molecule structures.

Drug design with the help of computers may be used at any of the following stages of drug discovery:
1. hit identification using virtual screening (structure- or ligand-based design)
2. hit-to-lead optimization of affinity and selectivity (structure-based design, QSAR, etc.)
3. lead optimization: optimization of other pharmaceutical properties while maintaining affinity

In order to overcome the insufficient prediction of binding affinity calculated by recent scoring functions, the protein-ligand interaction and compound 3D structure information are used to analysis.

Computation aide-cluster analysis

1. Using P-L interactions to cluster docked compounds
2. Selecting representative compounds for in vivo test
3D-QSAR

3D-QSAR (Quantitative structure–activity relationship models) refers to the application of force field calculations requiring three-dimensional structures, e.g., based on protein crystallography or molecule superimposition. It uses computed potentials, e.g., the Lennard-Jones potential, rather than experimental constants and is concerned with the overall molecule rather than a single substituent. It examines the steric fields (shape of the molecule), the hydrophobic regions (water-soluble surfaces), and the electrostatic fields.
The created data space is then usually reduced by a following feature. The following learning method can be any of the already mentioned machine learning methods, e.g., support vector machines. An alternative approach uses multiple-instance learning by encoding molecules as sets of data instances, each of which represents a possible molecular conformation. A label or response is assigned to each set corresponding to the activity of the molecule, which is assumed to be determined by at least one instance in the set (i.e., some conformation of the molecule).

Rule of thumb to evaluate druglikeness

Lipinski’s rule of five also known as the Pfizer’s rule of five or simply the Rule of five (RO5) is a rule of thumb to evaluate druglikeness or determine if a chemical compound with a certain pharmacological or biological activity has properties that would make it a likely orally active drug in humans. The rule was formulated by Christopher A. Lipinski in 1997, based on the observation that most medication drugs are relatively small and lipophilic molecules.
The rule describes molecular properties important for a drug’s pharmacokinetics in the human body, including their absorption, distribution, metabolism, and excretion (“ADME”). However, the rule does not predict if a compound is pharmacologically active.
The rule is important to keep in mind during drug discovery when a pharmacologically active lead structure is optimized step-wise to increase the activity and selectivity of the compound as well as to insure drug-like physicochemical properties are maintained as described by Lipinski’s rule. Candidate drugs that conform to the RO5 tend to have lower attrition rates during clinical trials and hence have an increased chance of reaching the market.

- Not more than 5 hydrogen bond donors (nitrogen or oxygen atoms with one or more hydrogen atoms)
- Not more than 10 hydrogen bond acceptors (nitrogen or oxygen atoms)
- A molecular mass less than 500 daltons (g/mole)
- An octanol-water partition coefficient log $P$ not greater than 5

The rules have spawned many extensions, for example the following:
- Partition coefficient log $P$ in −0.4 to +5.6 range
- Molar refractivity from 40 to 130
- Molecular weight from 180 to 500
- Number of atoms from 20 to 70 (includes H-bond donors [e.g., OH's and NH's] and H-bond acceptors [e.g., N's and O's])
- Polar surface area no greater than 140 Å²
Screening and design

The process of finding a new drug against a chosen target for a particular disease usually involves high-throughput screening (HTS), wherein large libraries of chemicals are tested for their ability to modify the target. For example, if the target is a novel G-protein coupled receptor, compounds will be screened for their ability to inhibit or stimulate that receptor. If the target is a protein kinase, the chemicals will be tested for their ability to inhibit that kinase.

Another important function of HTS is to show how selective the compounds are for the chosen target. The ideal is to find a molecule which will interfere with only the chosen target, but not other related targets. To this end, other screening runs will be made to see whether the “hits” against the chosen target will interfere with other related targets - this is the process of cross-screening. Cross-screening is important, because the more unrelated targets a compound hits, the more likely that off-target toxicity will occur with that compound once it reaches the clinic.

It is very unlikely that a perfect drug candidate will emerge from these early screening runs. It is more often observed that several compounds are found to have some degree of activity, and if these compounds share common chemical features, one or more pharmacophores can then be developed. At this point, medicinal chemists will attempt to use structure-activity relationships (SAR) to improve certain features of the lead compound:

- increase activity against the chosen target
- reduce activity against unrelated targets
- improve the druglikeness or ADME properties of the molecule.

This process will require several iterative screening runs, during which, it is hoped, the properties of the new molecular entities will improve, and allow the favoured compounds to go forward to in vitro and in vivo testing for activity in the disease model of choice.

Once a lead compound series has been established with sufficient target potency and selectivity and favourable drug-like properties, one or two compounds will then be proposed for drug development. The best of these is generally called the lead compound, while the other will be designated as the “backup”.

High throughput screening
The affinity of protein-ligand binding

The dissociation constant is commonly used to describe the affinity between a ligand \((L)\) (such as a drug) and a protein \((P)\) i.e. how tightly a ligand binds to a particular protein. Ligand-protein affinities are influenced by non-covalent intermolecular interactions between the two molecules such as hydrogen bonding, electrostatic interactions, hydrophobic and Van der Waals forces. They can also be affected by high concentrations of other macromolecules, which causes macromolecular crowding.

The formation of a ligand-protein complex \((C)\) can be described by a two-state process:

\[
C \rightleftharpoons P + L
\]

The corresponding dissociation constant is defined as:

\[
K_d = \frac{[P][L]}{[C]}
\]

The smaller the dissociation constant, the more tightly bound the ligand is, or the higher the affinity between ligand and protein. For example, a ligand with a nanomolar (nM) dissociation constant binds more tightly to a particular protein than a ligand with a micromolar (M) dissociation constant.

Drugs can produce harmful side effects through interactions with proteins for which they were not meant to or designed to interact. Therefore much pharmaceutical research is aimed at designing drugs that bind to only their target proteins (Negative Design) with high affinity (typically 0.1-10 nM) or at improving the affinity between a particular drug and its in-vivo protein target (Positive Design).

New drug application

I. preclinical studies: to determine a drug’s ultimate safety profile

Drugs may undergo pharmacodynamics (what the drug does to the body) (PD), pharmacokinetics (what the body does to the drug) (PK), ADME (absorption, distribution, metabolism, and excretion, and describes the disposition of a pharmaceutical compound within an organism), and toxicity testing through animal testing.

Bioavailability (BA) is a subcategory of absorption and is the fraction of an administered dose of unchanged drug that reaches the systemic circulation, one of the principal pharmacokinetic properties of drugs. By definition, when a medication is administered intravenously, its bioavailability is 100%.

Animal testing

The information collected from these studies is vital so that safe human testing can begin. Typically, in drug development studies animal testing involves two species. The most commonly used models are murine and canine, although primate and porcine are also used.

II. clinical trials: The trials are typically conducted in three phases:

Phase 1: The drug is tested in a few healthy volunteers to determine if it is acutely toxic.

Phase 2: Various doses of the drug are tried to determine how much to give to patients.

Phase 3: The drug is typically tested in double-blind (In a double-blind experiment, neither the participants nor the researchers know which participants belong to the control group, as opposed to the test group.), placebo controlled trials to demonstrate that it works. Sponsors typically confer with FDA prior to starting these trials to determine what data is needed, since these trials often involve hundreds of patients and are very expensive.

(Phase 4): These are post-approval trials that are sometimes a condition attached by the FDA to the approval.
Factors influencing bioavailability

The absolute bioavailability of a drug, when administered by an extravascular route, is usually less than one (i.e., $F < 100\%$). Various physiological factors reduce the availability of drugs prior to their entry into the systemic circulation. Whether a drug is taken with or without food will also affect absorption; other drugs taken concurrently may alter absorption and first-pass metabolism, intestinal motility alters the dissolution of the drug and may affect the degree of chemical degradation of the drug by intestinal microflora. Disease states affecting liver metabolism or gastrointestinal function will also have an effect.

Physical properties of the drug (hydrophobicity, pKa, solubility)

The drug formulation (immediate release, excipients used, manufacturing methods, modified release – delayed release, extended release, sustained release, etc.)

Whether the formulation is administered in a fed or fasted state

Gastric emptying rate

Circadian differences

Interactions with other drugs/foods:
  - Interactions with other drugs (e.g., antacids, alcohol, nicotine)
  - Interactions with other foods (e.g., grapefruit juice, pomelo, cranberry juice, brassica vegetables)

Transporters: Substrate of efflux transporters (e.g. P-glycoprotein)

Health of the GI tract

Enzyme induction/inhibition by other drugs/foods:
  - Enzyme induction (increased rate of metabolism), e.g., Phenytoin induces CYP1A2, CYP2C9, CYP2C19, and CYP3A4
  - Enzyme inhibition (decreased rate of metabolism), e.g., grapefruit juice inhibits CYP3A → higher nifedipine concentrations

Individual variation in metabolic differences
  - Age: In general, drugs are metabolized more slowly in fetal, neonatal, and geriatric populations
  - Phenotypic differences, enterohepatic circulation, diet, gender

Disease state
  - E.g., hepatic insufficiency, poor renal function

Life cycle of HIV

HIV belongs to the class of viruses called retroviruses, which carry genetic information in the form of RNA. HIV infects T cells that carry the CD4 antigen on their surface. When HIV infects its target cell it requires fusion of the viral and cellular membranes. The first step is the interaction between envelope proteins of the virus (gp120, gp41) and specific host-cell surface receptors (e.g. CD4 receptor) on the target cell. Then the virus binds to the chemokine coreceptors CXCR4 or CCR5, resulting in conformational changes in the envelope proteins. This fusion creates a pore through which the viral capsid enters the cell. Following entry into the cell the RNA of the virus is reverse-transcribed to DNA by the first virally encoded enzyme, the reverse transcriptase. The viral DNA enters the nucleus where it is integrated into the genetic material of the cell by the integrase, a second virally encoded enzyme. Activation of the host cell leads to the transcription of the viral DNA into mRNA. The mRNA is then translated into viral proteins and the third virally encoded enzyme, namely HIV protease, is required to cleave a viral polypeptide precursor into individual mature proteins. The viral RNA and viral proteins assemble at the surface of the cell into new virions. The virions bud from the cell and are released to infect other cells. All infected cells are eventually killed because of this extensive cell damage, from the destruction of the host's genetic system to the budding and release of virions.
Structure and genome of HIV

The complete nucleotide sequence of HIV-1 (6) shows a relatively simple retrovirus whose genome consists of three open reading frames (ORF), gag, pol and env. The gag ORF contains structural proteins such as capsid, nucleocapsid, and matrix, whereas regulatory proteins are encoded in the multiply spliced env ORF. The HIV-1 genome encodes only three unique enzymes, all located within the pol ORF. These enzymes, reverse transcriptase (RT), integrase, and protease (PR), have all become targets for drug discovery.

The first AIDS drugs to be identified were nucleoside inhibitors of RT, discovered and developed long before the structure of RT itself was solved. However, the development of newer RT-targeted drugs, nonnucleoside inhibitors, is closely coupled to structural investigations of enzyme complexes.

Mechanism of action

Human immunodeficiency virus (HIV) is a lentivirus that has two major species, HIV-1 which causes the majority of the epidemic, and HIV-2, a close relative whose distribution is concentrated in western Africa. HIV was identified as the causative agent of acquired immune deficiency syndrome (AIDS) and its complete genome was immediately available. HIV-2 carries a slightly lower risk of transmission than HIV-1 and infection tends to progress more slowly to AIDS. In common usage HIV usually implies HIV-1. HIV-1 protease is one of the best known aspartic proteases, and an attractive target for the treatment of AIDS. In 2009, ten protease inhibitors have reached the market for treatment against HIV but one protease inhibitor, amprenavir, was withdrawn from the market in 2004.

There are several steps in the HIV life cycle that may be interfered with, thus stopping the replication of the virus. A very critical step is the proteolytic cleavage of the polypeptide precursors into mature enzymes and structural proteins catalyzed by HIV protease. HIV protease inhibitors are peptide-like chemicals that competitively inhibit the action of the virus aspartyl protease. These drugs prevent proteolytic cleavage of HIV Gag and Pol polyproteins that include essential structural and enzymatic components of the virus. This prevents the conversion of HIV particles into their mature infectious form. Protease inhibitors can alter adipocyte metabolism causing lipodystrophy, a common side effect associated with the use of most HIV protease inhibitors.

When HIV infects a cell, reverse transcriptase copies the viral single stranded RNA genome into a double-stranded viral DNA. The viral DNA is then integrated into the host chromosomal DNA, which then allows host cellular processes, such as transcription and translation to reproduce the virus. RTIs block reverse transcriptase's enzymatic function and prevent completion of synthesis of the double-stranded viral DNA, thus preventing HIV from multiplying.

Integrase inhibitors are a class of antiretroviral drug designed to block the action of integrase, a viral enzyme that inserts the viral genome into the DNA of the host cell. Since integration is a vital step in retroviral replication, blocking it can halt further spread of the virus. Integrase inhibitors were initially developed for the treatment of HIV infection, but they could be applied to other retroviruses.
The HIV protease is a C2-symmetric homodimeric enzyme consisting of two 99 amino acid monomers. Each monomer contributes an aspartic acid residue that is essential for catalysis, Asp-25 and Asp-25'. The HIV protease has the sequence Asp-Thr-Gly, which is conserved among other mammalian aspartic protease enzymes. An extended beta-sheet region on the monomers, known as the flap, constitutes in part the substrate binding site with the two aspartyl residues lying on the bottom of a hydrophobic cavity. Each flexible flap contains three characteristic regions: side chains that extend outward (Met46, Phe53), hydrophobic chains extending inward (Ile47, Ile54), and a glycine rich region (Gly48, 49, 51, 52). Ile50 remains at the tip of the turn and when the enzyme is unliganded a water molecule makes hydrogen bonds to the backbone of Ile50 on each monomer.

HIV proteases catalyze the hydrolysis of peptide bonds with high sequence selectivity and catalytic proficiency. The mechanism of the HIV protease shares many features with the rest of the aspartic protease family although the full detailed mechanism of this enzyme is not fully understood. The water molecule seems to play a role in the opening and closing of the flaps as well as increasing the affinity between enzyme and substrate. The aspartyl residues are involved in the hydrolysis of the peptide bonds. The preferred cleavage site for this enzyme is the N-terminal side of proline residues, especially between phenylalanine and proline or tyrosine and proline.

Conformational changes of HIV protease upon inhibitor binding

The active site is covered by two symmetric flaps that change their conformation between the free and inhibited enzymes. The area adjacent to the active site is the most rigid and most highly conserved in the whole molecule (16, 35), whereas the flaps are the most dynamic. The areas leading to the flaps have been implicated in facilitating motions necessary to allow substrate entry and release.
Computational studies of the inhibitor complexes of HIV-1 protease

To understand the mode of binding and to optimize inhibitor design.

1. Energy minimization using molecular mechanics confirmed that the contribution of the main-chain atoms to the total interaction energy ranged from 56% to 68%. This has high correlation between the interaction energy and the experimentally determined IC50 constants for almost 50 inhibitor-enzyme complexes.

2. Free energy perturbation calculations coupled with molecular dynamics help to explain the differences between the binding constants of similar inhibitors, to analyze HIV-1 protease mutants and their affinity to different inhibitors, or to aid molecular modeling and drug design.

3. Docking based on the crystal structures of the inhibitor complexes of HIV-1 protease:
   (i) Docking methods and algorithms were tested using the structural data and experimental characteristics and by comparison with de novo constructed inhibitors
   (ii) Examined empirical free energy as a target function in docking and design (in contrast to the calculated interaction energy)
   (iii) Design of a variety of new inhibitors, some of which became clinical candidates.

FDA approved inhibitors of HIV-1 protease:

- Saquinavir (Ki: 0.12 nM)
- Indinavir (Ki: 0.6 nM)
- Kivexinavir (Ki: ~2.0 nM)
- Ritonavir (Ki: 0.12 nM)
- Lopinavir (Ki: 1.3 pM)
- Nelfinavir (Ki: 1.3 pM)
- Tipranavir (Ki: 5.0 pM)
- Amprenavir (Ki: 0.6 nM)
- BMS-232632 (Ki: 0.3 nM)
**HIV reverse transcriptase**

The HIV reverse transcriptase (RT) enzyme is responsible for RNA-dependent DNA polymerization and DNA-dependent DNA polymerization. The p66 subunit contains 560 amino acids, whereas the p51 subunit contains only the first 440 residues. Although the amino acid sequence of p51 is identical to the first 440 residues of the p66 subunit, it adopts a markedly different structural conformation. The p66 subunit contains the DNA-binding groove and the active site; the p51 subunit displays no enzymatic activity and functions as a scaffold for the enzymatically active p66 subunit. The p66 subunit has subdomains including the fingers, palm, and thumb subdomains that participate in polymerization, and the connection and RNaseH subdomains.

The enzyme is captured in register with the nucleic acid template (yellow) / primer (orange) and incoming dNTP shown in yellow spacefill mode. The enzyme active site which consists of three catalytic aspartates, D110, D185, and D186 is shown in white spacefill mode. The p51 monomer is shown in grey. The p66 monomer is colored as follows: fingers (cyan), palm (green), thumb (red), connection (blue), RNaseH (purple). RNaseH has an active site which is responsible for degrading the RNA template from the RNA-DNA hybrid created during reverse transcription. The binding cleft is configured so that the nucleic acid contacts both the polymerase and the RNaseH active sites; these are located about 17 or 18 bp apart.

**Drugs targeting reverse transcriptase (RT)**

RT inhibitors can be divided into two general classes. The first to be discovered were compounds that act as terminators of chain elongation. These analogs of the nucleoside substrates bind in the substrate-binding site and can inhibit both HIV-1 and HIV-2 RT. Another class of RT inhibitors, nonnucleoside inhibitors (NNIs), are specific to a pocket that is found in the vicinity of the active site in HIV-1 RT but does not exist in HIV-2 RT.
Development of drug resistance

Because retroviral RT has no editing function, transcription errors during nucleic acid replication are very common, and the viral pool contains species with all conceivable mutations. The presence of drugs provides a powerful selection pressure for virus modifications that produce lower susceptibility to such compounds. Rapid appearance of drug-resistant HIV species was considered a major obstacle in the development of newer therapies, such as PR inhibitors or NNIs. It is now clear that resistance can be minimized, both by combining NNIs with other inhibitors and by starting therapy with high concentrations of the drugs.

The use of sufficiently high doses and combination therapies have been quite successful in delaying or overcoming resistance. Drug-resistant mutations of HIV PR are considered so important that resistance studies now precede any attempts to introduce such compounds into clinical practice.

Combination therapies (highly active antiretroviral therapy, HAART) using different inhibitors promise the best clinical outcome. However, it is not clear whether it is better to use combinations of different drugs from the same family or drugs belonging to different classes. On one hand, combinations such as ritonavir-saquinavir, nelfinavir-saquinavir, or ritonavir-indinavir combine two similar drugs with distinct resistance patterns and, especially in the case of ritonavir, with different metabolisms. On the other hand, combinations of indinavir or nelfinavir with nevirapine, or indinavir plus efavirenz, assure that the development of resistance will require mutations in two different enzymes, making resistance less likely.

In Western countries, drug treatment is reducing AIDS to a manageable and treatable long-term disease. However, even with all the drugs already on the market, it is clear that the serious nature of the AIDS pandemic and the limitations of the therapies will make it necessary to continue drug development. Until a safe, effective vaccine against HIV has been found, it will be necessary to introduce new therapies and combinations of drugs to counteract the development of resistant variants. The understanding of drug-target interactions on the molecular level, coupled with extensive studies using the techniques of molecular biology, are of great help in achieving rapid success.

Other pharmaceutical interest targets

<table>
<thead>
<tr>
<th>Target</th>
<th>Involved in</th>
<th>Role</th>
</tr>
</thead>
<tbody>
<tr>
<td>Epidermal growth factor receptor (EGF receptor)</td>
<td>19 pathways</td>
<td>Non-small cell lung cancer</td>
</tr>
<tr>
<td>Protein kinase C beta II (PRKC βII)</td>
<td>19 pathways</td>
<td>Non-small cell lung cancer</td>
</tr>
<tr>
<td>Vascular endothelial growth factor receptor 2 (VEGFR2)</td>
<td>Three pathways</td>
<td>VEGF signaling pathway</td>
</tr>
<tr>
<td>BCL2 B-cell CLL/lymphoma 2 (BOI2)</td>
<td>Eight pathways</td>
<td>Small cell lung cancer</td>
</tr>
<tr>
<td>P-glycoprotein (P-gp)</td>
<td>Transporters</td>
<td>ABC transporters</td>
</tr>
<tr>
<td>Estrogen receptor alpha (ERα)</td>
<td>Five pathways</td>
<td>Role of ERBB2 in signal transduction and oncology</td>
</tr>
<tr>
<td>BCR-ABL kinase domain</td>
<td></td>
<td>CML patients, who are BCR-ABL positive</td>
</tr>
</tbody>
</table>
Examples of rational drug design

The first unequivocal example of the application of structure-based drug design leading to an approved drug is the carbonic anhydrase inhibitor dorzolamide, which was approved in 1995. Dorzolamide (trade name Trusopt) is a carbonic anhydrase inhibitor. It is an anti-glaucoma agent by decreasing the production of aqueous humour. It is optically applied in the form of eye drops which is used to lower increased intraocular pressure in open-angle glaucoma and ocular hypertension.

| 5-HT3 antagonists                      | effective in controlling the nausea and vomiting produced by cancer chemotherapy |
| Acetylcholine receptor agonists        | in treatment of Alzheimer's disease and schizophrenia |
| Angiotensin receptor blockers          | lowering blood pressure, treatment of obesity |
| Cannabinoid receptor antagonists       | potential therapeutic applications in the treatment of HIV infections for chemotherapy-induced vomiting |
| CCR5 receptor antagonists              | a class of anti-migraine drugs |
| NK1 receptor antagonists               |                                                                 |
| Triptans                               |                                                                 |
| Bcr-Abl tyrosine kinase inhibitors     | the first-line therapy for most patients with chronic myelogenous leukemia (CML). |
| Cyclooxygenase 2 inhibitors            | anti-inflammation |
| Dipeptidyl peptidase-4 inhibitors      | a potent treatment for type 2 diabetes |
| HIV protease inhibitors:               | used in the treatment of human immunodeficiency virus (HIV) |
| Non-nucleoside reverse transcriptase   |                                                                 |
| inhibitors                             |                                                                 |
| Proton pump inhibitors                 | treatment of choice for acid-related diseases to relief chronic pain |
| TRPV1 antagonists                      | in the treatment of small intestine ulcers |
| Renin inhibitors                       | in the treatment of various type of cancers |
| c-Met inhibitors                       |                                                                 |