Lectures 15-16: The Molecular Basis of Enzyme Catalysis: HIV Protease

1. The function and structure of HIV protease
   a. Introduction to proteases
   b. Discovery of HIV protease
   c. Overview of the three-dimensional structure of HIV protease

2. Chemical reactions and the energies driving them
   a. Amide bond cleavage: the reaction catalyzed by HIV protease
   b. Thermodynamics of a reaction and free energy
   c. Kinetics of a reaction and $\Delta G^\ddagger$
   d. Reaction energy diagrams
   e. Transition states, intermediates, and how to draw them

3. How enzymes accelerate chemical reactions: the case of HIV protease
   a. Catalysts alter a reaction’s kinetics, but not its thermodynamics
   b. Chemical strategies behind enzyme catalysis
      i. Proximity and orientation effects
      ii. Nucleophilicity and electrophilicity
      iii. Acid and base catalysis

4. The molecular basis of substrate specificity
   a. Trypsin substrate specificity
   b. HIV protease substrate specificity

---

Proximity Effects on Reaction Rates

\[ \text{Rate at 1 M concentrations} = 4 \times 10^{-6} \text{ M/s} \]

\[ \text{Rate} = 0.8 \text{ M/s (200,000-fold faster!)} \]

- Intramolecular reactions are much faster than intermolecular ones
Orientation Effects

(assume A and B react upon sideways collision)

- The fewer nonproductive ways two groups can be oriented, the faster they will react

Proximity and Orientation Effects in HIV Protease

- HIV protease uses hydrogen bonds to orient substrates productively
- The enzyme’s active site holds substrates (protein & water) in close proximity

HIV protease backbone (Ile 50 and Ile 50')

“attacking” water
A Network of Interactions Precisely Positions the Substrates of HIV Protease

Note: structure is of an inactive Asn 25 mutant of HIV protease complexed with a substrate.

Nucleophiles and Electrophiles

Electrophile electron-deficient, “likes electrons”

Nucleophile electron-rich, “likes nuclei”

\[
E^+ + :\text{Nu}^- \rightarrow E\text{–Nu}
\]

- Nucleophilicity and electrophilicity are kinetic parameters
- The more nucleophilic the nucleophile, or electrophilic the electrophile, the faster the reaction (by definition!)
Factors Governing Nucleophilicity

1) **More basic** molecules tend to be more nucleophilic when the nucleophilic atoms are of comparable size:

- Poor nucleophile: \( \text{pK}_a \text{ of conjugate acid} = -1.5 \)
- Good nucleophile: \( \text{pK}_a \text{ of conjugate acid} = 15.5 \)

2) **Larger atoms** (those lower on the periodic table) make better nucleophiles:

- Weaker nucleophile
- Stronger nucleophile

Electrophiles in the Molecules of Life

- Peptide hydrolysis
- Translation
- DNA polymerization
- DNA hydrolysis
- Protein phosphorylation

- The most common electrophiles in the chemistry of life are C=O and P=O
- Groups that are more electron-poor are more electrophilic and therefore react more quickly with nucleophiles
Acids and Bases Can Enhance Electrophilicity and Nucleophilicity

Better nucleophile

More electron-rich

Better electrophile

More electron-poor

Base Catalysis of Amide Hydrolysis

Weaker nucleophile: slower reaction, higher $\Delta G^\ddagger$

Better nucleophile: faster reaction, lower $\Delta G^\ddagger$

Base can accelerate amide hydrolysis by deprotonating water, increasing its nucleophilicity
Base Catalysis by HIV Protease

- HIV protease precisely positions Asp 25 to serve as a base to deprotonate water
- The enhanced nucleophilicity of deprotonated water accelerates amide hydrolysis

Acid Catalysis of Amide Hydrolysis

- Acid can accelerate amide hydrolysis by protonating the amide oxygen, increasing its electrophilicity
Acid Catalysis by HIV Protease

- HIV protease precisely positions Asp 25' to serve as an acid to protonate the substrate amide, increasing its electrophilicity and accelerating amide hydrolysis.

Enzymes Can Catalyze Reactions in Ways That Simple Acids and Bases Cannot

- Enzymes can simultaneously use acidic and basic groups that, in a flask, would wander and neutralize each other.
HIV Protease Active Site in Action

Note: locations of hydrogen atoms are usually inferred (and often not shown)

HIV Protease Catalysis Summary

Proximity and orientation effects

Acid catalysis

Base catalysis

Tetrahedral intermediate

Acid catalysis

Base catalysis
Lectures 15-16: The Molecular Basis of Enzyme Catalysis: HIV Protease

1. The function and structure of HIV protease
   a. Introduction to proteases
   b. Discovery of HIV protease
   c. Overview of the three-dimensional structure of HIV protease

2. Chemical reactions and the energies driving them
   a. Amide bond cleavage: the reaction catalyzed by HIV protease
   b. Thermodynamics of a reaction and free energy
   c. Kinetics of a reaction and $\Delta G^\ddagger$
   d. Reaction energy diagrams
   e. Transition states, intermediates, and how to draw them

3. How enzymes accelerate chemical reactions: the case of HIV protease
   a. Catalysts alter a reaction’s kinetics, but not its thermodynamics
   b. Chemical strategies behind enzyme catalysis
      i. Proximity and orientation effects
      ii. Nucleophilicity and electrophilicity
      iii. Acid and base catalysis

4. The molecular basis of substrate specificity
   a. Trypsin substrate specificity
   b. HIV protease substrate specificity

Protease Substrate Specificity

To N-terminal end of substrate

Enzyme specificity pockets recognize the specific amino acid residues surrounding the bond to be hydrolyzed

To C-terminal end of substrate
Trypsin: A Digestive Protease that Cleaves Substrates Containing Lys or Arg

Basis of Trypsin Substrate Specificity

- The presence of anionic Asp 189 in the S1 site causes a strong preference for P1 to be a cationic Lys or Arg
HIV Protease-Substrate Interactions

The HIV Protease Active Site: A Closer Look

Note: structure is of an inactive Asn 25 mutant of HIV protease complexed with a substrate.
HIV Protease Substrate Selectivity

The 10 sites cleaved by HIV protease:

<table>
<thead>
<tr>
<th>Substrate Cleavage Site</th>
<th>P5</th>
<th>P4</th>
<th>P3</th>
<th>P2</th>
<th>P1</th>
<th>P1'</th>
<th>P2'</th>
<th>P3</th>
<th>P4</th>
<th>P5'</th>
</tr>
</thead>
<tbody>
<tr>
<td>MA-CA</td>
<td>Val</td>
<td>Ser</td>
<td>Gin</td>
<td>Asn</td>
<td>Tyr</td>
<td>Pro</td>
<td>Ile</td>
<td>Val</td>
<td>Gin</td>
<td>Asn</td>
</tr>
<tr>
<td>CA-p2</td>
<td>Lys</td>
<td>Ala</td>
<td>Arg</td>
<td>Val</td>
<td>Leu</td>
<td>Ala</td>
<td>Glu</td>
<td>Ala</td>
<td>Met</td>
<td>Ser</td>
</tr>
<tr>
<td>p2-NC</td>
<td>Pro</td>
<td>Ala</td>
<td>Thr</td>
<td>Ile</td>
<td>Met</td>
<td>Met</td>
<td>Gin</td>
<td>Arg</td>
<td>Gly</td>
<td>Asn</td>
</tr>
<tr>
<td>NC-nt1</td>
<td>Gly</td>
<td>Arg</td>
<td>Gin</td>
<td>Ala</td>
<td>Asn</td>
<td>Phe</td>
<td>Leu</td>
<td>Gly</td>
<td>Lys</td>
<td>Ile</td>
</tr>
<tr>
<td>p1-p6</td>
<td>Arg</td>
<td>Pro</td>
<td>Gly</td>
<td>Asn</td>
<td>Phe</td>
<td>Leu</td>
<td>Gin</td>
<td>Ser</td>
<td>Arg</td>
<td>Pro</td>
</tr>
<tr>
<td>TF-PR</td>
<td>Val</td>
<td>Ser</td>
<td>Phe</td>
<td>Asn</td>
<td>Phe</td>
<td>Pro</td>
<td>Gin</td>
<td>Ile</td>
<td>Thr</td>
<td>Leu</td>
</tr>
<tr>
<td>AutoP</td>
<td>Pro</td>
<td>Gin</td>
<td>Ile</td>
<td>Thr</td>
<td>Leu</td>
<td>Trp</td>
<td>Lys</td>
<td>Arg</td>
<td>Pro</td>
<td>Leu</td>
</tr>
<tr>
<td>PR-RT</td>
<td>Cys</td>
<td>Thr</td>
<td>Leu</td>
<td>Asn</td>
<td>Phe</td>
<td>Pro</td>
<td>Ile</td>
<td>Ser</td>
<td>Pro</td>
<td>Ile</td>
</tr>
<tr>
<td>RT-RH</td>
<td>Gly</td>
<td>Ala</td>
<td>Glu</td>
<td>Thr</td>
<td>Phe</td>
<td>Tyr</td>
<td>Val</td>
<td>Asp</td>
<td>Gly</td>
<td>Ala</td>
</tr>
<tr>
<td>RH-IN</td>
<td>Ile</td>
<td>Arg</td>
<td>Lys</td>
<td>Ile</td>
<td>Leu</td>
<td>Phe</td>
<td>Leu</td>
<td>Asp</td>
<td>Gly</td>
<td>Ile</td>
</tr>
</tbody>
</table>

Key: MA = matrix; CA = capsid; NC = nucleocapsid; TF = trans frame peptide; PR = protease; AutoP = autoproteolysis (self-cleaving) site; RT = reverse transcriptase; RH = RNase H; integrase = IN

- HIV protease recognizes more substrate amino acids than trypsin, but without a strong preference at any one position

HIV Protease Specificity: P1 and P3

- HIV protease’s structure can explain some aspects of its substrate specificity
- P1 = large and hydrophobic; complements S1 = Gly 49
- P3 = polar or charged; S3 contains a bound water
Key Points: HIV Protease & Enzyme Catalysis

- HIV protease catalyzes polyprotein amide bond hydrolysis
- Thermodynamics reflect the difference in energy between reactants and products, as measured by $\Delta G^\circ_{\text{rxn}}$
- Kinetics reflect reaction rates, determined by $\Delta G^\ddagger$
- Enzymes lower $\Delta G^\ddagger$ by using a variety of chemical strategies to create a precise transition state-stabilizing active site environment
- Enzymes use proximity and orientation effects to increase the concentration of substrates, increasing rates of reactions
- Enzymes use acid and base catalysis to enhance the nucleophilicity or electrophilicity of reactants
- Specificity arises from protein-substrate interactions
Lectures 17-18: The molecular basis of drug-protein binding:

HIV protease inhibitors

1. Drug development and its impact on HIV-infected patients
2. Energetic dissection of a small molecule binding to a protein
   a. Enthalpy changes upon binding
   b. Entropy changes upon binding
3. Case studies of saquinavir and ritonavir, two small-molecule HIV protease inhibitors
   a. Fill hydrophobic pockets with hydrophobic groups
   b. Provide complementary hydrogen bond donors and acceptors
   c. Mimic the transition state of a reaction
   d. Maximize the rigidity of the drug
   e. Displace bound water molecules

Impact of Anti-HIV Drugs

- 1990s: anti-HIV drugs transform HIV infection from a short death sentence to a chronic (but very serious) illness
- 13 FDA-approved drugs inhibit HIV reverse transcriptase; 9 drugs inhibit HIV protease (first approved December, 1995)
- Mortality rate of U.S. patients with advanced AIDS:
  - 29% per year in 1995
  - 9% per year in mid-1997
- 1997-2003: Death rate from AIDS in Europe falls 80%
- Gains primarily attributed to combination therapy involving HIV protease inhibitors + other antiretroviral agents
Drug Development is Very Difficult

- Total cost to develop a drug = ~$1 billion + ~10-15 years

Successful Drugs Must Satisfy Many Chemical and Biological Requirements

1) Potency (affinity) + $K_{eq} = K_a = 1 + K_d$

2) Specificity (toxicity, immunogenicity)

3) Bioavailability

4) Biostability

5) Economics