Exercise 5.8.3

Use AlphaFold3 from this link to determine the structure of a complex between a protein and a dsDNA from just their sequence. Then compare the computed structure with a very similar X-ray structure. This example comes from a Nature paper that used AlphaFold to predict the structure of the complex. Here is the reference:

Abramson, J., Adler, J., Dunger, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024). https://doi.org/10.1038/s41586-024-07487-w

Protein: AMP-binding protein-catabolite gene activator from Rhizobium meliloti (strain 1021) (Ensifer meliloti) (Sinorhizobium meliloti), uniprot: Q92SD2. Here is the protein sequence:

MAEVIRSSAFWRSFPIFEEFDSETLCELSGIASYRKWSAGTVIFQRGDQGDYMIVVVSGRIKLSLFTPQGRELMLRQHE
AGALFGEMALLDGQPRSADATAVTAAEGYVIGKKDFLALITQRPKTAEAVIRFLCAQLRDTTDRLETIALYDLNARVAR
FFLATLRQIHGSEMPQSANLRLTLSQTDIASILGASRPKVNRAILSLEESGAIKRADGIICCNVGRLLSIADPEED

dsDNA sequence: Here are sequences of both DNA strands (5'-3'):

5'-CTAGGTAACATTACTCGCG-3' (19 mer) and 5'-GCGAGTAATGTTAC-3' (14 mer)

In AlphaFold 3, input the entry type, copies (number of molecules in the structure), and the sequences, as shown in the figure below.

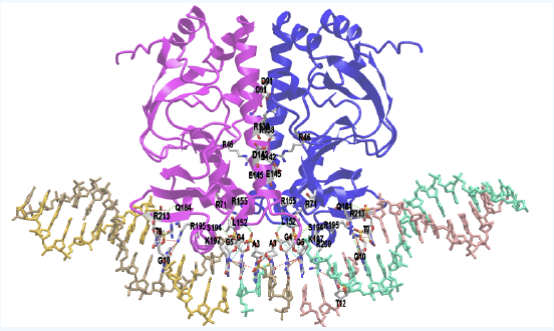
In the actual crystal structure, a protein dimer binds the dsDNA

Answer

iCn3D results: Download this PNG file for upload into iCn3D to see a rendered image of the protein. **IMPORTANT**: If the file opens as an image in a new browser window, right-click the image and save the file to download it!

Open iCn3D **File, Open, iCn3D PNG appendable** and browse for the file in your download folder.

Here is a static image of the calculated structure:



The figure below shows an interactive iCn3D model of the actual X-ray crystal structure of the capsid trimer from Clr-cAMP-DNA complex (7PZB) for comparison. Note that cAMP could not be included in the AlphaFold 3 predicted

structure of the complex using the version available.

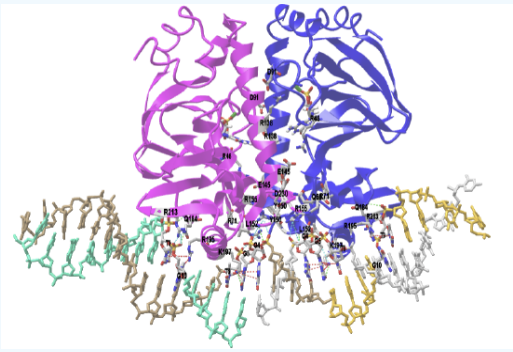


Figure: X-ray crystal structure of the capsid trimer from **Clr-cAMP-DNA complex (7PZB)**. (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...YZJitoVd7Y2LAA

.

The program can be used to determine the structure of protein complexes as well as in this example from Chapter 4.14, the human sperm proteins and egg protein complex predicted by AlphaFold.

5.8.4: Designing target proteins de novo (ex. making novel binders) with RFDiffusion

The figure below shows an interactive iCn3D model of the X-ray crystal structure of human thrombin (3U69). Thrombin is a serine protease. It cleaves other proteins using its active site serine as a nucleophile. Other examples of serine proteases are the gut proteases chymotrypsin and trypsin, which are used in digestion. Thrombin cleaves a limited repertoire of proteins, mostly involved in blood clotting and its control. A main substrate is the protein fibrinogen, which, after cleavage of small peptides, forms fibrin that associates to form a fibrin clot. In addition to binding at the active site, specific substrates like fibrinogen interact with an anionic "exosite" on thrombin near the active site. This additional binding site limits the specificity of thrombin to specific substrates.

The figure below shows an interactive iCn3D model of human thrombin (3U69). Three active site residues (catalytic triad) involved in cleaving fibrinogen and other clotting proteins are shown in CPK-colored sticks. The blue side chains and surface areas represent key positively charged residues in "Exosite 1" of thrombin, involved in binding the anionic region of fibrinogen and a modulator of thrombin activity, thrombomodulin.

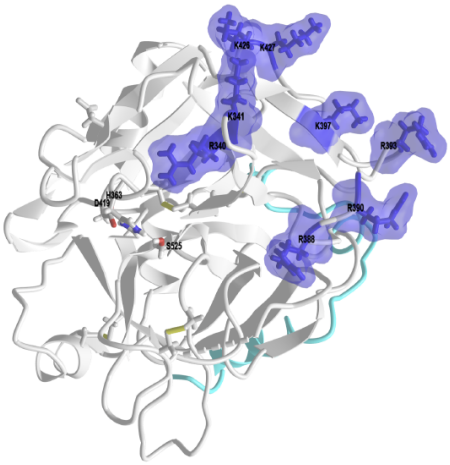
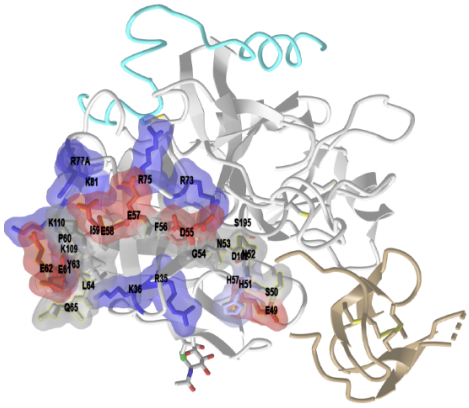


Figure: X-ray crystal structure of human thrombin (3U69). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...PkxukNXxPGuHa7

Organisms that need liquid blood for food have exploited this Exosite 1 and created peptide inhibitors that bind to it and inhibit fibrinogen binding, for example. One such organism is a leech, which secretes a protein in its saliva called **hirudin**. It has a long C-terminal anionic tail that binds in the exosite groove and inhibits the binding of fibrinogen.

The figure below shows an interactive iCn3D model of the X-ray crystal structure of the hirudin-thrombin complex (4HTC). Key anionic residues in the C-terminal tail of hirudin are shown as red sticks and red surfaces interacting with the positive residues (blue sticks and surfaces) in Exosite 1 of thrombin.



NCBI iCn3D Figure: X-ray crystal structure of the hirudin-thrombin complex (4HTC). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...EyESJzsUV85Ex5.. To change the background color on the popup choose Style, Background, Transparent

Synthetic inhibitors of thrombin that target exosite 1 of thrombin have been made. One contains a negatively charged DNA (an aptamer) covalently linked to an active site inhibitor, creating an "EXosite and ACTive site (EXACT) inhibitor".

? Exercise 5.8.4

In the next exercise, you will create a novel peptide inhibitor that targets exosite 1 of thrombin using RFDiffusion. This should be a simple example as the peptide is expected to have glutamates and aspartates, mimicking the hirudin C-terminal end.

You will run RFDiffusion from a commercial company called Neurosnap. First, set up an account (free, five models/month)

a. Go to RFDiffusion-v2
b. Separately, download this pdb file for human thrombin to your computer: 3U69
c. Open RFdiffusion-v2 and input the following data, which is shown in red. Leave the rest alone. H stands for the Heavy Chain in the PDB file.

Binder Input Chain	H
Binder Length Maximum	20
Binder Length Minimum	17
Binder ROG	false
Binding Pocket Residue End	1
Binding Pocket Residue Start	1
Fixed Residue End	20
Fixed Residue Start	10
Hotspots	H340,H341,H388,H390,H393,H397,H426,H427

Input Structure	input_structure.pdb

d. Run the program.

When completed, the output page has several tabs: Data and Visuals, Config, and Files. View the results in the Data and Visuals tab

Answer

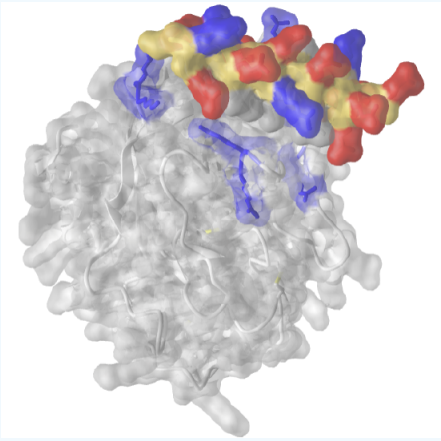
iCn3D results: Download this PNG file and upload it into iCn3D to see a rendered image of the protein. **IMPORTANT**: If the file opens as an image in a new browser window, right-click the image and save the file to download it!

Open iCn3D **File, Open, iCn3D PNG appendable,** and browse for the file in your download folder.

Note that the amino acid numbering of the hotspots changed in the output files from RFDiffusion, as shown in the table below.

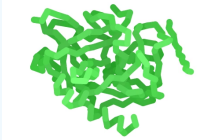
Initial PDB numbering system for Heavy Chain (RU69)	RFDiffusion numbering for Heavy Chain Hot Spots
R340	20
K341	21
R388	68
R390	70
R393	73
K397	77
K426	106
K427	107

Here is a static image:



The light chain is not shown. The heavy chain H is shown in gray. The hot spots on thrombin (H340, H341, H388, H390, H393, H397, H426, and H427) are colored by charge (blue stick and surface). The binder is shown as a gold surface with the charged side chains shown in red (negative) and blue (positive). The sequence of RFDiffusion binder is EEEEKELELLREEIEKLEKE, which has 11 glutamate (E) but also 3 lysines (K), perhaps to maintain some charge balance.

Here is an animation from the output to show how RFDiffusion produces the final peptide at the hotspot residues.



Here are the stats for the top-ranked binder

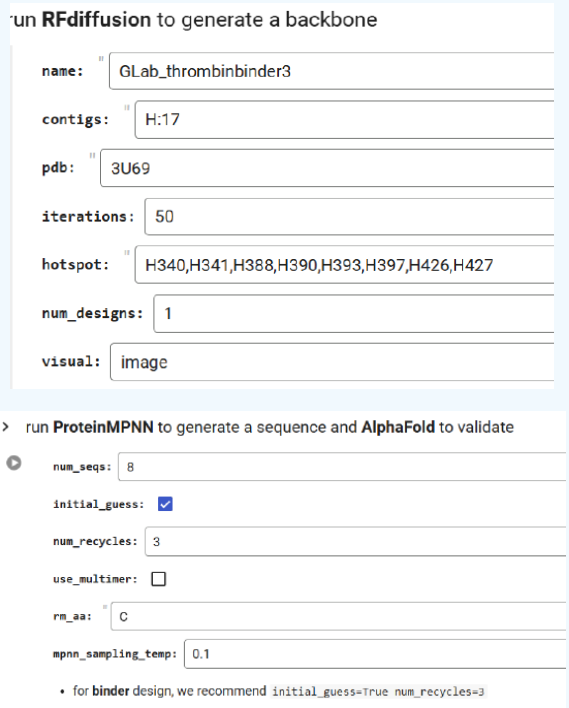
Rank	MPNN Score	RMSD	Mean pLDDT	Max PAE	pTM
1	1.14	19.69	92.98	23.48	0.91

Watson, Joseph et al. “Broadly applicable and accurate protein design by integrating structure prediction networks and diffusion generative models” bioRxiv.Org. doi: https://doi.org/10.1101/2022.12.09.519842. Neurosnap Inc. - Computational Biology Platform for Research. Wilmington, DE, 2022. https://neurosnap.ai/.

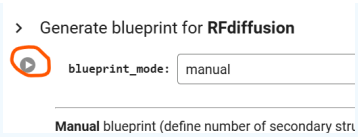
Exercise 5.8.5

Create a binder for human thrombin using Google Colab at this site (note: this link might be replaced in the future)

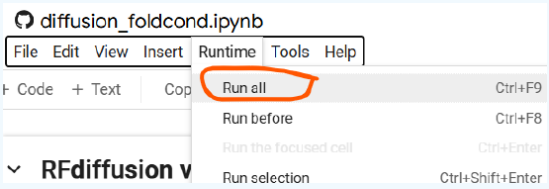
a. **Inputs**: Accept all defaults except for the sections below. For those, use the inputs shown. The goal is not to optimize the run but to show how easy it is to create a binder. The inputs allow you to create a 17 amino acid binder to the H (heavy) chain of human thrombin (3U69) at the positively charged Exosite 1 hot spots residues in the H chain, as described above.



b. **Run the program**. You can run sections individually or all at once. To run one section after the other, click the arrow for the first section as shown below, wait for it to complete, and then run the rest



The first section sets up the program and takes a few minutes to run. Alternatively, run all at once as shown in the image below.



The results should be downloaded to a .zip file in your computer's download folder.

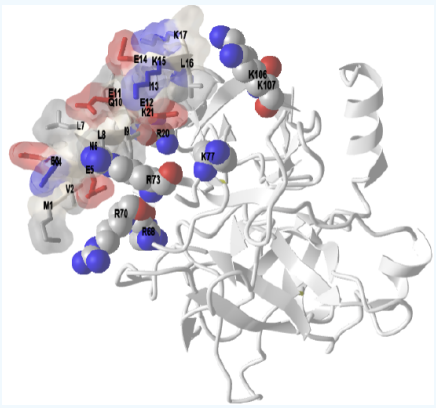
c. **Visualizing the results:** Unzip the folder and open the "best_design0.pdb file in iCn3D. Display and render your file in iCn3D.

Answer

iCn3D results: Download this PNG file and upload it into iCn3D to see a rendered image of the protein-binder complex. **IMPORTANT**: If the file opens as an image in a new browser window, right-click the image and save the file to download it!

Open iCn3D **File, Open, iCn3D PNG appendable,** and browse for the file in your download folder.

Note that the number system for the H chain changed as the 1st run. Here is a static image:



The gray chain is the heavy chain of thrombin. The hot spots on this chain are shown in spacefill, CPK colors, and labeled. The binder is shown as a transparent surface with side chains colored by charge (red = positive, blue = negative).

5.8.5: Predicting Protein Function through Structure with FoldSeek

FoldSeek can be run directly on a local PDB file or one retrieved from the PDB through this link. However, it can also be run through iCn3D, adding many more possibilities for rendering and analyses.

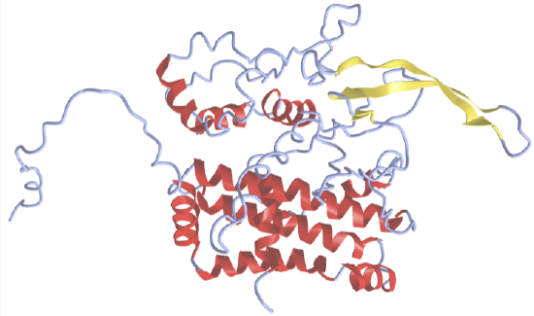
Use iCn3D to run FoldSeek

Exercise 5.8.6

We will explore the structural similarities of a newly predicted viral protein from the Wuhan insect virus from an uncertain family to the structure of all proteins in the databases using FoldSeek using iCn3D.

First, download this file: hypothetical_protein__YP_009329883__Wuhan_insect_virus_23__1923727.pdb

a. Open iCn3D
b. **File, Open File, PDB appendable**, and load the downloaded file
c. **Color, Secondary, Sheets in Yellow**. Note that the protein is predominantly alpha-helical with some beta-sheets. Here is a static view other viral protein.



Now let's use FoldSeek within iCn3D to find similar structures (not sequences) with the databases (PDB and AlphaFold).

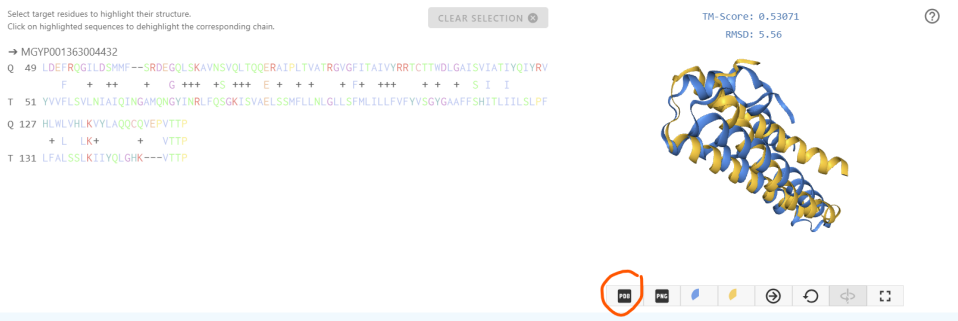
d. **File, Search Similar, FoldSeek (PDB and AlphaFold)**
e. Click Submit

A new tab with FoldSeek appears with the best statistical structural fits. The statistics shown in this table include the probability (0-1) that the structural match of the query (the protein studied) to the target (what it might resemble) is not due to chance. The other statistic is the expected or E-value, which gives the probability (0-1) that the structures (or sequences) match just by chance. The closer the E-value is to zero, the more significant the match. (See this PDB page describing E-values for sequence alignments.)

f. Which hit is best? How good is the quality of the fit?
g. If you click the blue hyperlink for the best hit, you will get the structure from the ESM database, and you can download it. The database protein structure is shown below.

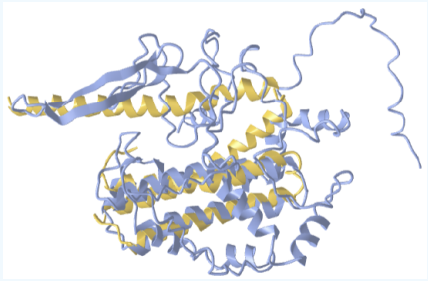
Instead, click the staggered three lines (≡) to the right to get the alignment with the hypothetical protein.

This pops up:



The query (blue, the protein studied) and the target (orange, what it might resemble)

h. Click the PDB link to download a PDB file of the two aligned protein segments. Model them in iCn3D to replicate the structure above.



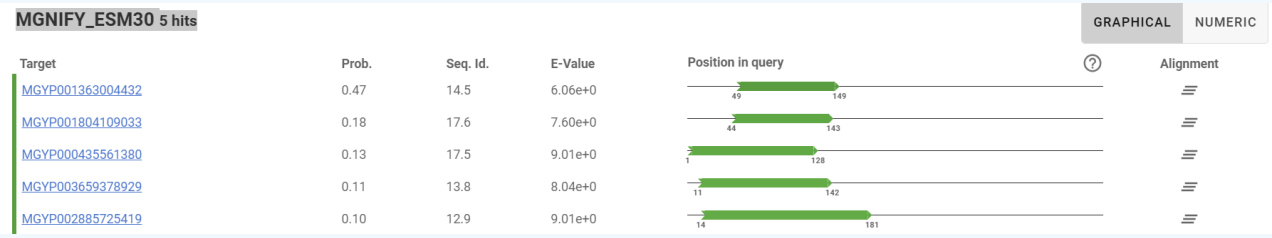
Query (hypothetical_protein__YP_009329883__Wuhan_insect_virus_23__1923727) is blue, Target (MGYP001363004432) is orange

This hypothetical viral protein structure is very dissimilar to any structure in the databases except for the helical bundles similar to protein MGYP001363004432 in the **ESM Metagenomic Atlas database.**

Answer

f. You will see that for this particular structure, there are just five "hits" in only one database, mgnify_esm30, from the ESM Metagenomic Atlas Database described above.

Here are the five hits.



Target	Prob.	Seq. Id.	E-Value	Position in query	Alignment
MGYP001363004432	0.47	14.5	6.06e+0	49–149	
MGYP001804109033	0.18	17.6	7.60e+0	44–143	
MGYP000435561380	0.13	17.5	9.01e+0	1–128	
MGYP003659378929	0.11	13.8	8.04e+0	11–142	
MGYP002885725419	0.10	12.9	9.01e+0	14–181	

None of the structural alignments are good statistically.

iCn3D results: Download this PNG file for upload into iCn3D to see a rendered image of the protein-binder complex. **IMPORTANT**: If the file opens as an image in a new browser window, right-click the image and save the file to download it!

Open iCn3D **File, Open, iCn3D PNG appendable,** and browse for the file in your download folder.

This hypothetical viral protein structure is very dissimilar to any structure in the databases except for the helical bundles similar to protein MGYP001363004432 in the **ESM Metagenomic Atlas database.**

Query (hypothetical_protein__YP_009329883__Wuhan_insect_virus_23__1923727) is blue, Target (MGYP001363004432) is orange

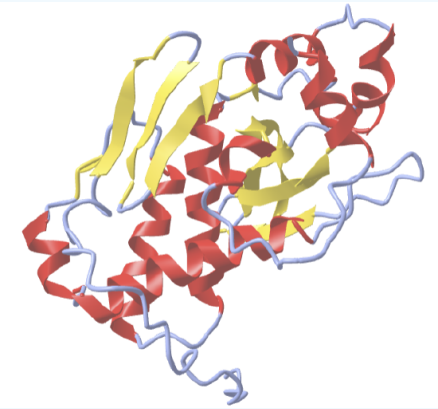
Exercise 5.8.7

Now, let's try a new viral protein structure whose function we can likely infer from the FoldSeek structural overlay.

Download this protein structure (whose function has been surmised through programs like FoldSeek: matrix_protein__YP_001531158__Marburg_marburgvirus__11269.pdb.

a. Open iCn3D
b. **File, Open File, PDB appendable**, and load the downloaded file
c. **Color, Secondary, Sheets in Yellow**. Note that the protein is predominantly alpha-helical with some beta-sheets. Here is a static view other viral protein.

Here is a static image:



e. Now repeat the procedure above and use FoldSeek within iCn3D (as described in Exercise 6 above) to find proteins of similar structure and presumably function.
f. Model one of the top hits in iCn3D

Answer

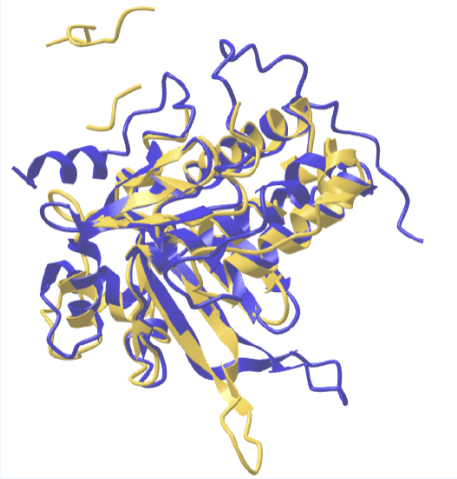
The optimal target is VP24 of the Marburg virus (Uniprot ID: **Membrane-associated protein VP24 - P35256**). Both query and target proteins are from the Marburg virus, in the Filioviridae family, including the Ebola virus. These are deadly viruses that cause hemorrhagic fevers and have high death rates. Here is one example of an almost-perfect hit.

Target	Description	Scientific Name	Prob.	Seq. Id.	E-Value	Position in query	Alignment
4or8-assembly1_B	Crystal structure of Marburg virus VP24	Marburg virus - Musoke, Kenya, 1980	1.00	94.1	9.90e-32		

The query (blue, the protein studied) and the target (orange, what it might resemble)

iCn3D results: Download this PNG file for upload into iCn3D to see a rendered image of the protein-binder complex. **IMPORTANT**: If the file opens as an image in a new browser window, right-click the image and save the file to download it!

Open iCn3D **File, Open, iCn3D PNG appendable,** and browse for the file in your download folder. A static image is shown below.



Query (matrix_protein__YP_001531158__Marburg_marburgvirus__11269) is blue, Target (VP24 of the Marburg virus) is orange

From the Uniprot page for the VP24 protein, FoldSeek can be run to determine which proteins in the database are similar. In this case, VP24 is the query and FoldSeek finds targets of similar structure.

The best hits are in the PDB database:

Target	Description	Scientific Name	Prob.	Seq. Id.	E-Value

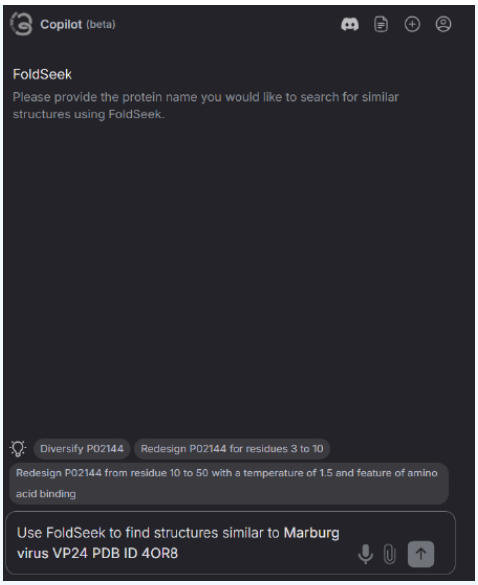
Target	Description	Scientific Name	Prob.	Seq. Id.	E-Value
4or8-assembly1_A	Crystal structure of Marburg virus VP24	Marburg virus - Musoke, Kenya, 1980	1.00	100	1.03e-46
4or8-assembly1_B	Crystal structure of Marburg virus VP24	Marburg virus - Musoke, Kenya, 1980	1.00	94.7	4.12e-37
6ehm-assembly1_C	Model of the Ebola virus nucleocapsid subunit from recombinant virus-like particles	Ebola virus - Mayinga, Zaire, 1976	1.00	37	7.36e-19
4u2x-assembly3_C	Ebola virus VP24 in complex with Karyopherin alpha 5 C-terminus	Ebola virus - Mayinga, Zaire, 1976	1.00	36.5	6.94e-19
3vne-assembly1_A	Structure of the ebolavirus protein VP24 from Sudan	Sudan ebolavirus	1.00	35.5	3.25e-18
4d9o-assembly1_A	Structure of ebolavirus protein VP24 from Reston	Reston ebolavirus - Reston	1.00	36	1.55e-16
3vnf-assembly1_A	Structure of the ebolavirus protein VP24 from Sudan	Sudan ebolavirus	1.00	34.8	1.75e-16
4d9o-assembly1_B	Structure of ebolavirus protein VP24 from Reston	Reston ebolavirus - Reston	1.00	35	3.84e-15

Running FoldSeek using 310 Copilot

? Exercise 5.8.8

Now run FoldSeek using 310 Copilot. This program is part of a suite of commercial programs from Open AI. These programs allow users to input questions and data as sentences (much like Chatbots like ChatGPT, Claude, Gemini, etc.) to address complex questions in biology and biochemistry.

a. Click try it now. Input the query shown in the bottom of the figure below. IN this case you are asking to find structures similar to the VP24 of the Marburg using the PDB ID = 4or8



Answer

Here are the results:

name	subject id	evalue	tmscore
4or8_1	4or8 B	1.0	1.0
4or8_2	4or8 A	0.922	0.9277
4or8_3	6ehm C	0.821	0.8276

As of 11/21/24, you can't upload a local pdb file (as we did with iCn3D above) and run FoldSeek within Copilot.

5.8.6: Exploring protein complexes with FoldSeek Multimer

In Exercise 3 above, you used AlphaFold to predict the structure of protein complexes (in that particular example, a protein:DNA complex) from sequences. By analogy, FoldSeek Multimer can find the 3D structures of target complexes from the 3D structure of a known (query) complex. In short, AlphaFold can do large-scale sequence-to-structure comparisons, while FoldSeek Multimer can do large-scale structure-to-structure comparisons. Let's try an example with a simple complex, the hepatitis A virus C3 proteinase (1HAV), a homodimer of an A and B chain.

? Exercise 5.8.9

Go to the FoldSeek Multimer server

a. Load Accession 1hav and run the program.
b. Copy/Snip the top 4 results. Compare the 1st and 4th results
c. Click the alignment icons = for each and take a screen snips. =

Answer

b. Here is a snip of the top results.

Complex		Chain						
qTM	tTM	Chain paring	Scientific Name	Prob.	Seq. Id.	E-Value	Position in query	Alignment
0.34	0.13	B ➔ ProtVar_P83110_Q96RQ3_A	Homo sapiens	1.00	10.2	2.13e-7		
0.34	0.19	B ➔ ProtVar_O43464_Q99988_A	Homo sapiens	1.00	11.4	1.31e-5		
0.34	0.23	A ➔ ProtVar_P61009_P83110_B	Homo sapiens	1.00	9	5.00e-6		
0.34	0.16	A ➔ LevyLab_P83110_V1_2_relaxed... B ➔ LevyLab_P83110_V1_2_relaxed...	Homo sapiens Homo sapiens	1.00 1.00	8.2 9.5	9.99e-6 6.91e-7		

The top results show high-quality alignments of one chain with a human protein, but that human protein is not part of a similar dimer. Result 4 shows good alignment of both chains of the query with a human dimer.

Note the label for the top hit: B ➔ ProtVar_P83110_Q96RQ3_A. The P83110 is the Uniprot number for human serine protease HTRA3.

c. Hit 1 - alignment for just 1 chain of the query and target

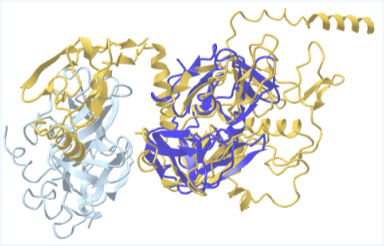
Query (blue, the protein studied) and the target (orange, what it might resemble)

The icons underneath the blue/orange models are explained below.

- PDB: select to download the combined file and model in iCn3D
- blue: toggle between the entire query and the aligned structure
- orange: toggle between the entire target and the aligned structure. Note that in this case, the target (human protein) is much bigger than the query (viral) protein so the alignment is only between part of the human protein.

iCn3D results: Download this PNG file and upload it into iCn3D to see a rendered image of the protein-binder complex. **IMPORTANT**: If the file opens as an image in a new browser window, right-click the image and save the file to download it!

Open iCn3D **File, Open, iCn3D PNG appendable,** and browse for the file in your download folder. A static image is shown below.

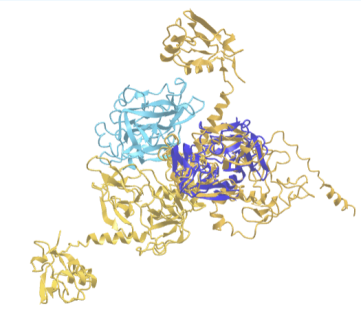


Query (blue, the protein studied) and the target (orange, what it might resemble)

Hit 4: Alignment of two chains (A and B) for the query and target

iCn3D results: Download this PNG file for upload into iCn3D to see a rendered image of the protein-binder complex. **IMPORTANT**: If the file opens as an image in a new browser window, right-click the image and save the file to download it!

Open iCn3D **File, Open, iCn3D PNG appendable,** and browse for the file in your download folder. A static image is shown below.

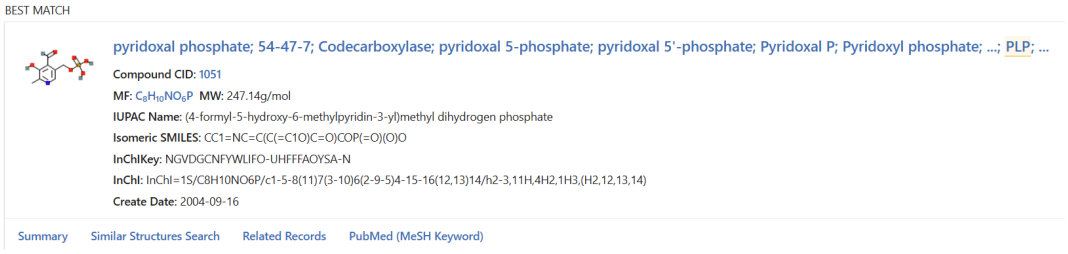


5.8.7: Docking of Small Molecules (Ligands)

Predicting binding interactions between small molecules and proteins is at the heart of the pharmaceutical industry. Previously, docking software for small molecule ligands and proteins was proprietary, costly, and difficult to run. Now, docking can be done online. In the next exercises, you will use three programs: 310 Copilot DiffDock (from a commercial company), DiffDock through Neuroapp, and SwissDock (Molecular Modeling Group, University of Lausanne, and the SIB Swiss Institute of Bioinformatics).

The inputs for a docking protein include the protein (typically the PDB ID) and the ligand, which can be represented as the actual structure but more often a code in the SMILES or InChi formats. For these exercises, you will dock the small molecule pyridoxal phosphate to a low molecular weight protein tyrosine phosphatase, a protein that cleaves phosphorylated tyrosines in proteins.

The input representation for PLP can be obtained through PubChem shown below.



Docking using 310 Copilot

? Exercise 5.8.10

Docking using Copilot - DiffDock

a. Go to PubChem and get the SMILES representation for PLP.
b. Open 310 Copilot
c. Input this text: Run a docking experiment with the protein PDB ID = 5JNR and the small molecule SMILES=CC1=NC=C(C(=C1O)C=O)COP(=O)(O)O
d. When the docked structure is shown, select Download PDB
e. Save as 310CoPilotDock5JNR_PLP.pdb

Now model the results in iCn3D as follows:

f. Open iCn3D
g. **File, Open File, PDB appendable**, and load 310CoPilotDock5JNR_PLP.pdb

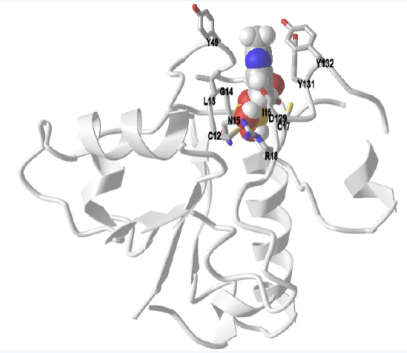
h. **Color, Secondary, Sheets in Yellow**. Note that the protein is predominantly alpha-helical with some beta-sheets. Here is a static view of the viral protein.

Answer

iCn3D results: Download this PNG file and upload it into iCn3D to see a rendered image of the protein-binder complex. **IMPORTANT**: If the file opens as an image in a new browser window, right-click the image and save the file to download it!

Open iCn3D **File, Open, iCn3D PNG appendable,** and browse for the file in your download folder.

Here is a static view.

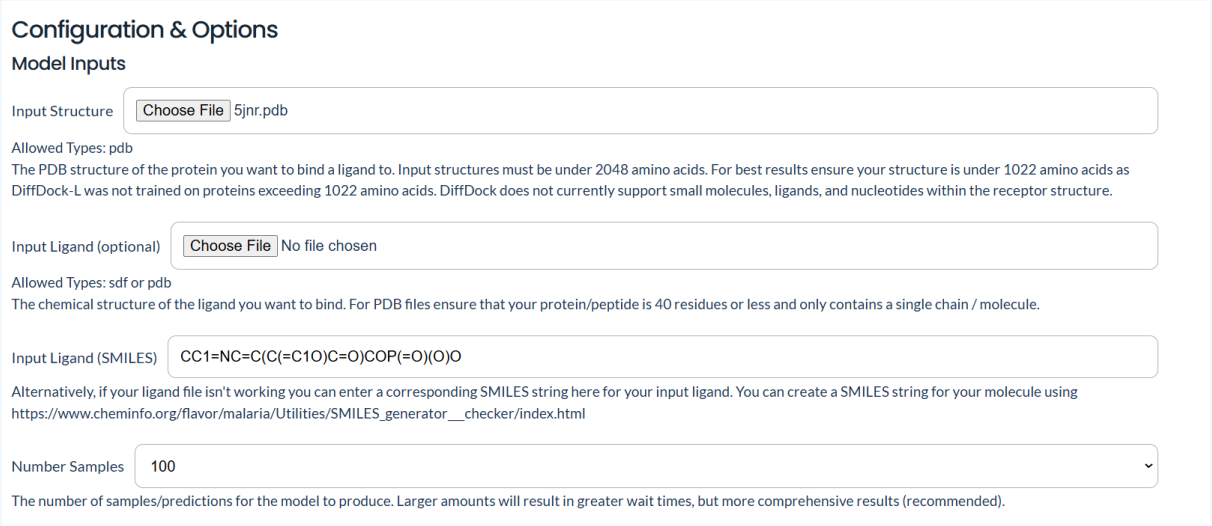


Docking using DiffDock in Neurosnap

? Exercise 5.8.11

Now, use DiffDock through Neurosnap to dock the ligand and protein.

a. Complete the menu as shown below.



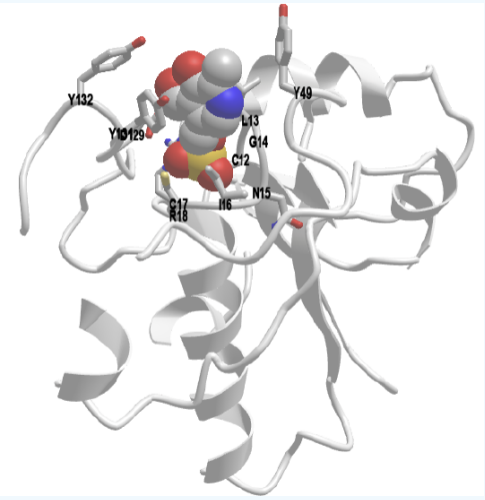
The output folder contains a separate PDB file and 100 SDF files for the various ligand "poses." Display them in iCn3D.

a. Open iCn3D
b. **File, Open File, SDF**, and choose the top-ranked SDF file (rank1_confidence-0.09.sdf)
c. File, Open File, PDB Appendable, and choose the proteins_no_ligands pdb file.

d. Render it as you see fit

Answer

Here is a screen snap of the docked structure.



iCn3D results: Download this PNG file for upload into iCn3D to see a rendered image of the protein-binder complex. **IMPORTANT**: If the file opens as an image in a new browser window, right-click the image and save the file to download it!

Open iCn3D **File, Open, iCn3D PNG appendable,** and browse for the file in your download folder.

Docking using SwissDock

? Exercise 5.8.12

Now use SwissDock to dock PLP to a different low molecular weight protein tyrosine phosphate, PDB ID = 1xww. This has a SO_4^{2-} in the active site, which must be removed before the docking. The program uses two different methods, AutoDock Vina and Cavity Prioritization. Use the second one, which seems to give better results for this protein.

a. Input the values shown in the box below and run the docking experiment. When you input the target PDB file (1xww), choose the prompts shown below.

Query

Ligand	CC1=NC=C(C(=C1O)C=O)COP(=O)(O)O
Target	1xww_modified.pdb
Method	Attracting Cavities 2.0
Date	November 23, 2024, 1:57 pm UTC

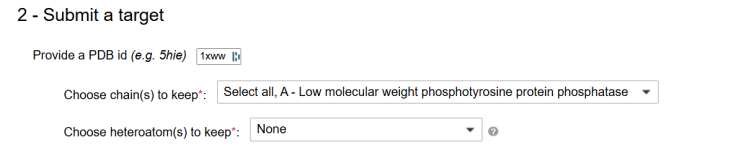
Parameters:

Box center:	13 - 0 - 18	Sampling exhaustivity:	medium	Number of RIC:	4
Box size:	20 - 20 - 20	Cavity prioritization:	buried		

If you publish these results, please, cite the following papers:

Bugnon M, Röhrig UF, Goullieux M, Perez MAS, Daina A, Michielin O, Zoete V. SwissDock 2024: major enhancements for small-molecule docking with Attracting Cavities and AutoDock Vina. *Nucleic Acids Res.* **2024**

Röhrig UF, Goullieux M, Bugnon M, Zoete V. Attracting Cavities 2.0: improving the flexibility and robustness for small-molecule docking. *J. Chem. Inf. Model.*, **2023**



b. Download the result as a zip file.

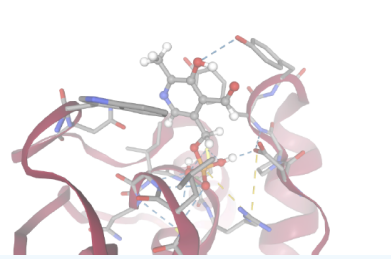
View and render the docked ligand in iCn3D. First, you must modify one of the files to view the docking results in iCn3D.

c. Choose Export Results as Zip file
d. Open the Zip folder and extract all.
e. Open the result.dock4 file in a simple text editor. Save just the first PLP coordinates by finding the 1st line with TER (terminate). Delete everything after that. Save the file as result0dock4.pdb file into the File subfolder you just extract
f. Open iCn3D
g. **File, Open File, PDB appendable**, and choose both the file you just made (results.dock4.pdb) and the receptor.pdb load it.
h. Render the image as you wish.

Answer

Here is the link to the actual SwissDock docking results page.

Here is a snip of the best docked structure.



Top Pose:

iCn3D results: Download this PNG file for upload into iCn3D to see a rendered image of the protein. **IMPORTANT**: If the file opens as an image in a new browser window, right-click the image and save the file to download it!

Open iCn3D **File, Open, iCn3D PNG appendable** and browse for the file in your download folder.

CHAPTER OVERVIEW

6: Enzyme Activity

Return to Fundamentals of Biochemistry Search Fundamentals of Biochemistry

6.1: How Enzymes Work

Learning Goals (ChatGPT, 1/30/25)

- **Understanding Reaction Energetics and Transition States**
  - Describe why uncatalyzed reactions (such as ester hydrolysis) are inherently slow due to high activation energies and the formation of high-energy, charge-separated transition states.
  - Analyze energy diagrams to identify how the transition state's energy and structure (often resembling an unstable intermediate) influence reaction rates.
- **Transition State Stabilization and Its Role in Catalysis**
  - Explain how any factor that stabilizes the charges developed in the transition state (e.g., protonation, deprotonation, or metal ion coordination) lowers the activation energy and catalyzes the reaction.
  - Apply transition state theory and Pauling's hypothesis to quantitatively compare the binding affinity of an enzyme for the substrate versus its transition state.
- **Acid-Base Catalysis Strategies**
  - Differentiate between specific acid (or base) catalysis and general acid (or base) catalysis in terms of reaction mechanism and pH dependency.
  - Illustrate, using mechanistic diagrams, how general acid catalysis (via proton donation) and general base catalysis (via proton abstraction) reduce the energy barrier in ester hydrolysis.
- **Metal Ion (Electrostatic) Catalysis**
  - Discuss how metal ions (such as Cu^{2+}, Zn^{2+}, or Fe^{3+}) stabilize negative charges in transition states through coordination, thereby enhancing reaction rates.
  - Evaluate examples (e.g., decarboxylation of β-keto acids and carbonic anhydrase) to understand the role of metal ions in enzyme catalysis.
- **Covalent (Nucleophilic) Catalysis**
  - Describe how a nucleophilic catalyst (e.g., pyridine) can form a covalent intermediate with a substrate, effectively creating a lower-energy pathway for the reaction.
  - Use energy diagrams to compare uncatalyzed versus covalently catalyzed reactions, noting the formation of additional intermediates and transition states.
- **Intramolecular Catalysis and Effective Concentration**
  - Explain how intramolecular reactions (e.g., in aspirin derivatives or phenylsuccinate) benefit from increased effective concentration, leading to dramatic rate enhancements compared to analogous intermolecular reactions.
  - Quantitatively compare intramolecular versus intermolecular rate constants to illustrate the concept of effective molarity.
- **Quantitative Deconstruction of Enzyme Catalysis**
  - Integrate the concepts of transition state binding, intramolecular versus intermolecular reactions, and the roles of various catalytic strategies (acid-base, metal ion, covalent) to analyze enzyme rate enhancements.
  - Derive and interpret relationships between equilibrium constants for substrate and transition state binding, and relate these to overall catalytic efficiency.
- **Design and Use of Transition State Analogs (Abzymes)**
  - Discuss how stable transition state analogs can be used as inhibitors or to generate antibodies (abzymes) that mimic enzyme catalysis.
  - Evaluate case studies where transition state analogs are synthesized to target specific enzyme active sites.
- **Asymmetric Catalysis and Enantioselectivity in Enzyme Reactions**
  - Describe how enzymes achieve asymmetric catalysis by selectively binding one enantiomer over another, resulting in the formation of a single stereoisomer of product (e.g., in the reaction catalyzed by triose phosphate isomerase).

- Compare enzyme catalysis to small-molecule organocatalysis (such as L-proline catalyzed aldol reactions) to understand the broader implications of chiral catalysis in biological and synthetic chemistry.

By mastering these goals, students will gain a deep understanding of how enzymes and small molecule catalysts lower activation energies through multiple strategies, the quantitative relationships governing these processes, and the practical applications of these concepts in both natural and synthetic systems.

In this section, we will explore chemical and physical factors that speed up reactions and begin to relate these effects to reactions catalyzed by enzymes. We will see that enzymes employ various chemical strategies to increase the rates of reactions, in addition to physical ones like reactant proximity and the introduction of strain. This can result in reactions 10 million (or more!) times faster than the uncatalyzed reaction. To put this 10 million-fold rate enhancement in perspective, if the catalyzed reaction takes one second, the uncatalyzed one would take nearly four months!

This chapter section has been written by Kristen Procko and Henry Jakubowski.

Reactions in solutions that are not catalyzed are slow. Consider the hydrolysis of an ester in water, illustrated in Figure 6.1.1. The ester is stabilized by resonance, and is therefore weakly electrophilic; thus, attack by weakly nucleophilic water is a slow process. Examining the transition state, we can see that charge development and separation occurs in the transition state for the uncatalyzed reaction, resulting in an intermediate (**P**) with both positive and negative charges.

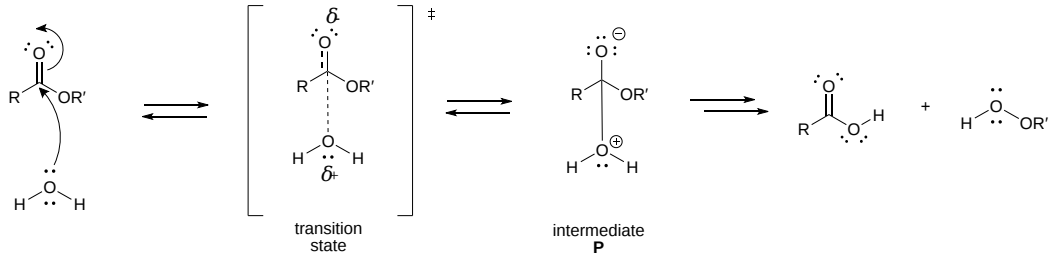


Figure 6.1.1: Charge development in the transition state during ester hydrolysis

When bonds are made or broken, charged intermediates are often formed, which are higher in energy than the reactants. Consider the energy diagram for the first step of a generic endergonic reaction, shown in 6.1.2. The transition state is closer in energy to the intermediate **P** than to the reactant **R.** Therefore, the TS more closely resembles **P** than the starting reactants. Applying this analysis to the ester hydrolysis reaction from Figure 6.1.1, the transition state is closer in energy to the charge-separated intermediate **P** and, therefore, more closely resembles the charge-separated species. In this example, the intermediate is higher in energy than the reactants, thus the transition state is *even higher* in energy than the intermediate.

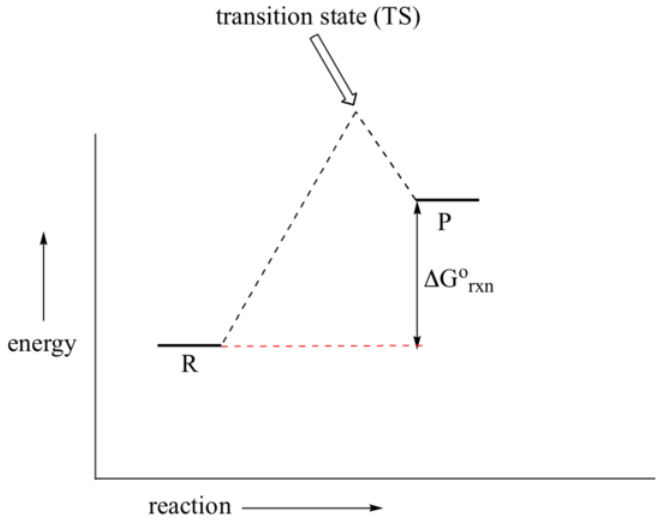


Figure 6.1.2: An energy diagram for an endergonic reaction (Image modified from: Chemistry LibreTexts, 6.9)

Anything that can stabilize the charges on the intermediate will also stabilize the developing charges in the transition state. This lowers the energy of the transition state and catalyzes the reaction. This section will investigate the mechanism underlying the catalysis by small molecules of chemical reactions. Presumably, biological macromolecular catalysts (like protein enzymes) will use similar mechanisms in their catalytic effects (which will be discussed in the next section).

Catalysts, including enzymes, can stabilize transition states in at least five ways.

6.1.1: Chemical Strategies for Rate Enhancement

6.1.1.1: General Acid and Base Catalysis

Considering intermediate **P** in Figure 6.1.1, we can envision two strategies to reduce the charge separation: the negative charge on the anionic oxygen could be protonated, or the positive charge on the cationic oxygen could be removed by deprotonation. If the reaction is pH dependent, and the reaction rate solely depends on hydronium ion concentration, $[H_3O^+]$, then **specific acid catalysis** is operative. **Specific acid catalysis** occurs when the hydronium ion concentration is the sole factor determining the reaction rate, and the concentration of any buffer components present in the solution does not influence the rate. In other words, the reaction rate depends *specifically* on the concentration of hydronium ions. **Specific base catalysis** occurs when the reaction rate depends solely on the hydroxide ion concentration and is again independent of any buffer components in the solution.

By contrast, **general acid catalysis** occurs when the reaction is not solely dependent on the $[H_3O^+]$ concentration; that is, the concentration of a buffer component influences the reaction rate. With general acid catalysis, the charge separation in the transition state is decreased by the donation of a proton to the carbonyl from general acids (e.g., acetic acid or a protonated imidazole ring). Proton donation decreases the developing negative charge in the transition state. In Figure 6.1.3, the first step of the ester hydrolysis mechanism is shown via specific acid catalysis alongside the general acid catalysis mechanism using the weak acid, acetic acid.

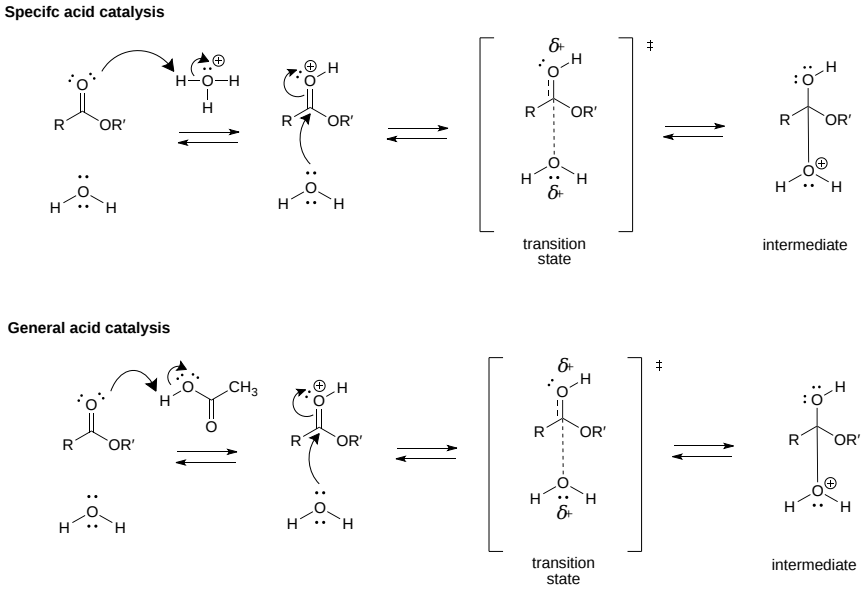


Figure 6.1.3: Specific vs. general acid catalysis

Alternatively, the first step of the ester hydrolysis mechanism can be base-catalyzed, increasing the nucleophile's strength. In Figure 6.1.1, the attacking water molecule develops a partial positive charge in the transition state as it begins to form a bond with the electrophilic carbon of the carbonyl. In the base-catalyzed mechanism shown in Figure 6.1.4, hydroxide becomes the nucleophile in the specific base-catalyzed mechanism. The energy of the transition state can also be lowered by the presence of a general base (e.g., acetate, a deprotonated imidazole ring). Proton abstraction decreases the developing positive charge.

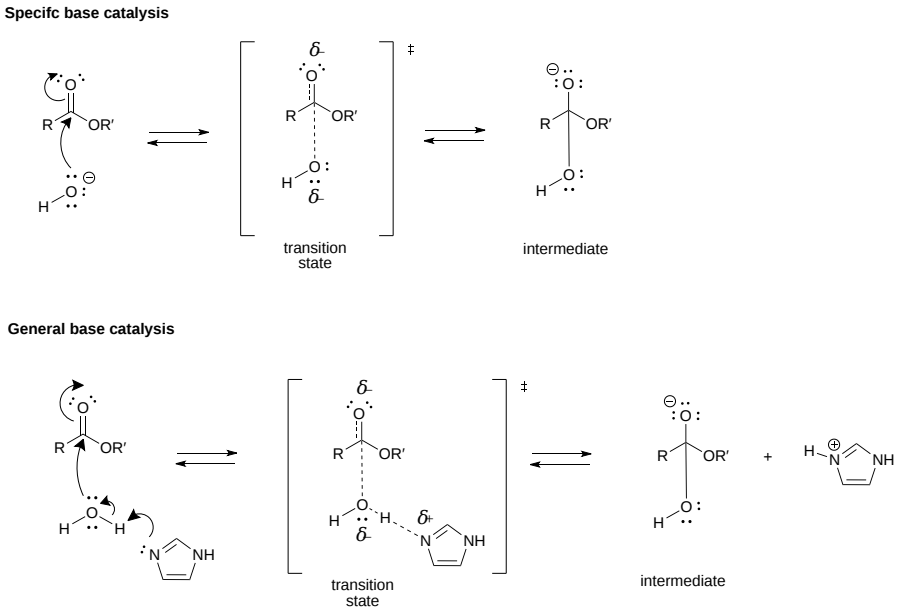


Figure 6.1.4: Specific vs. general base catalysis

General acid/base catalysis is common with enzymes because enzymes often use amino acid side chains to promote acid-base reactions within the **active site**, the enzyme region where the chemical reaction occurs. Acetic acid is similar to glutamic and aspartic acid side chains, and the imidazole ring shown in the general base catalysis reaction in Figure 6.1.4 is present in the side chain of the amino acid histidine.

6.1.1.2: Metal Ion or Electrostatic Catalysis

A metal such as Cu^{2+} or Zn^{2+} can also stabilize the transition state. The metal must be able to bind the charged intermediate and hence the transition state. An oxyanion intermediate formed during the reaction of an electrophilic carbonyl C can interact with a metal, especially when there is an O on an adjacent atom which can help coordinate the metal ion. This charge stabilization of the developing negative in the transition state and the full negative in the intermediate is often called **electrostatic catalysis.** It is illustrated in the decarboxylation of a β-keto carboxylic acid Figure 6.1.5. Coordination of Cu^{2+} to the β-keto carboxylic acid increases the electrophilicity of the carbonyl, making it a superior electron acceptor, which facilitates decarboxylation. The intermediate enolate formed in the decarboxylation is protonated, giving the more stable ketone as the product.

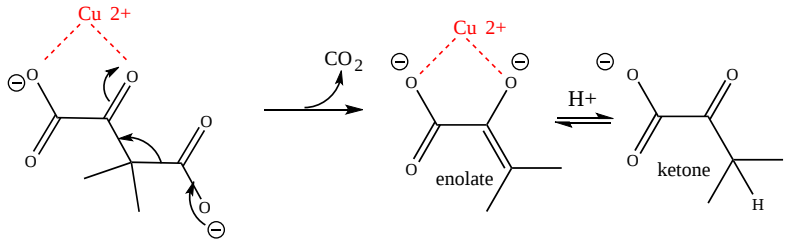


Figure 6.1.5: Metal ion catalysis

Electrostatic catalysis is likely to occur in many enzymes since nearly 1/3 of all enzymes require metal ions. A classic example of an enzyme using metal ion catalysis is carboxypeptidase A. Figure 6.1.6s shows an interactive iCn3D model of Zn and the inhibitor citric acid bound to carboxypeptidase A (3KGQ). Note the histidine and aspartate amino acid side chains of the active site coordinating to the Zn^{2+} ion, along with the carboxylate group of citrate.

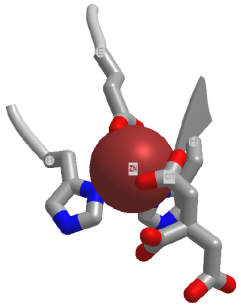


Figure 6.1.6: Zn binding to carboxypeptidase A (3KGQ) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?pH7tydTANZw5ayh9A

Metals can also act differently. They may coordinate a water and, by further polarizing the H-O bond, increase the acidity of the bound water. For instance, a water molecule in the hexaaquairon(III) ion has a pKa of 9.4, compared to pure water, with a pK_a of 14 (6.1.7). The complexed hydroxide is a better nucleophile than bulk water.

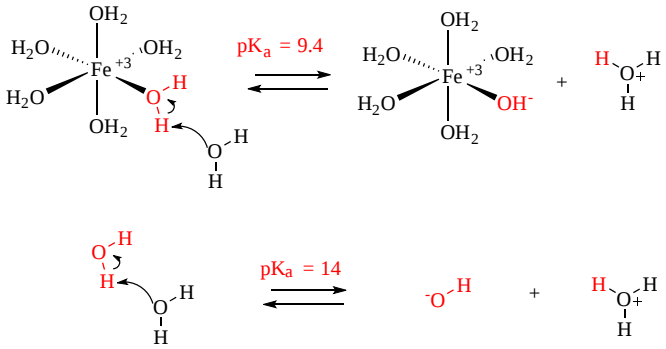


Figure 6.1.7: Metal ion decrease of pK_a of coordinated water

Another enzyme that utilizes Zn^{2+} is carbonic anhydrase. It is among the fastest enzymes, with a k_{cat} of 10^6 s^{-1} and a k_{cat}/K_m of 8.3 x 10^7 $M^{-1}s^{-1}$ (reference). It is diffusion-controlled at low substrate (CO_2) concentration and converts one million bound CO_2 per second to HCO_3^-! The Zn^{2+} appears to bind a water molecule and reduce its pK_a such that the bound form is OH^-. This is illustrated in Figure 6.1.8, which depicts the local environment of the bound Zn^{2+} (coordinated by three histidine side chains and an OH^-) in the absence (left) and presence (right) of CO_2.

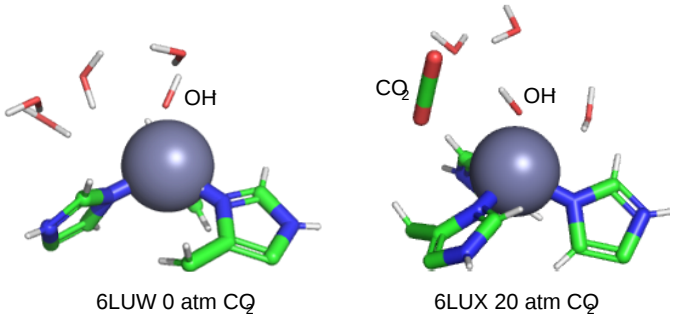


Figure 6.1.8: Coordination of OH^- to Zn^{2+} in carbonic anhydrase in the absence (left) and presence (right) of substrate CO_2.

6.1.1.3: Covalent or Nucleophilic Catalysis

One way to change the activation energy of the reaction is to change the reaction mechanism in ways that introduce new steps with lower activation energy. As shown in Figure 6.1.9, the catalyzed reaction has a new lower energy well, representing the formation of the covalent intermediate, and the activation energy is lowered overall. Formation of the new intermediate results in two transition states, represented by the two high energy points of the blue line in the plot.

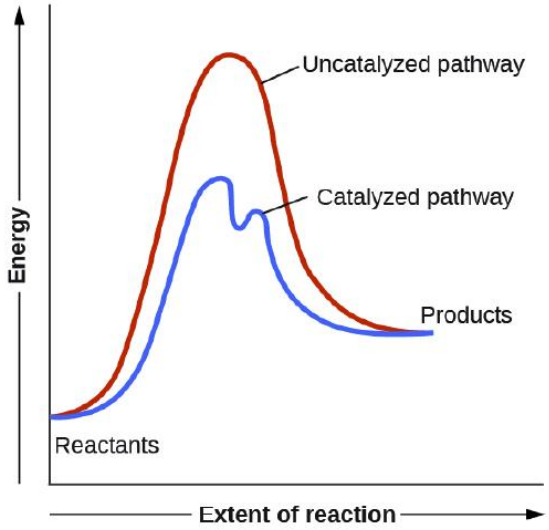


Figure 6.1.9: Energy diagram for an uncatalyzed reaction compared to a catalyzed reaction that utilizes covalent catalysis.

A typical way to achieve covalent catalysis is to add a nucleophilic catalyst, which forms a covalent intermediate with the reactant. Figure 6.1.8 shows how pyridine (red) acts as a nucleophilic or covalent catalyst in the hydrolysis of an anhydride. The anhydride is very reactive initially, and the charged pyridinium ion intermediate contains a very good leaving group. The desired nucleophile, water, can then interact with the intermediate in a nucleophilic substitution reaction. In these reactions, in general, as long as the nucleophilic catalyst is a better nucleophile than the ultimate nucleophile (usually water), the activation energy is lowered, and the reaction is catalyzed. The nucleophilic catalyst and the original nucleophile usually interact with a carbonyl C in a substitution reaction.

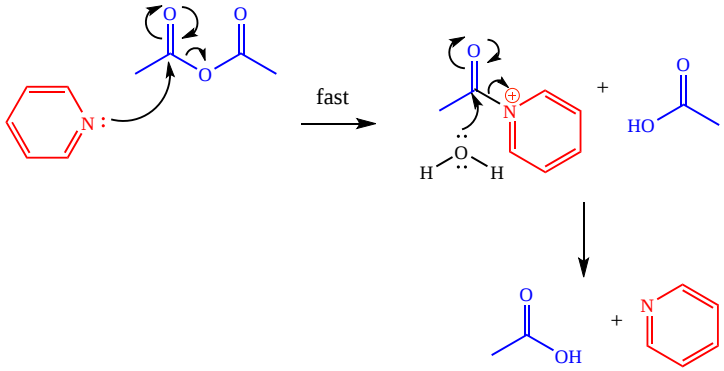


Figure 6.1.10: Nucleophilic covalent catalysis by pyridine

6.1.1.3.1: Reactions involving iminium ions are a recurring theme in biochemistry

Positively charged nitrogen cations (iminium ions) form as intermediates in many biochemical mechanisms. The iminium ion is a powerful electron acceptor and can promote the cleavage of bonds that would otherwise be difficult to break, such as C–C and C–H bonds. To begin our analysis of how iminium ions promote such cleavage reactions, let's revisit a common carbon-carbon bond cleaving reaction, the decarboxylation of a β-keto acid, which we examined briefly above in Figure 6.1.5 with metal ion catalysis.

Under acidic conditions, β-keto acids usually decarboxylate with gentle warming. A cyclic transition state is often invoked, and the presence of the carbonyl of the ketone adjacent to the breaking bond gives the electrons somewhere to go (Figure 6.1.11). The decarboxylation product is an enol, which tautomerizes to the more stable ketone. The equilibrium favors the deprotonated

carboxylate form under the slightly basic conditions that characterize the medium for most biochemical reactions. The adjacent carbonyl again gives the electrons somewhere to go in the decarboxylation reaction, and under these basic conditions, an enolate is formed. Protonation gives the final product, a ketone.

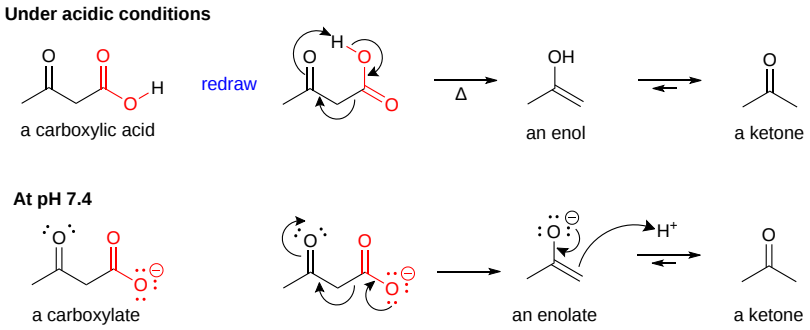


Figure 6.1.11: Decarboxylation of a β-keto acid under acidic conditions, and at pH 7.4

In Figure 6.1.5, we saw that a metal ion can promote the decarboxylation reaction by interacting with the electron-accepting ketone carbonyl, which makes it even more electrophilic. Another strategy to create a better electron acceptor involves forming a full positive charge on the electron-accepting atom, which can be done by converting the ketone to an iminium ion. Amines react with aldehydes or ketones to form iminium ions. Figure 6.1.12 illustrates this strategy involving covalent catalysis. The amine, RNH_2, reacts to form a new intermediate, the iminium ion, with a full positive charge. The protonated nitrogen serves as an excellent electron "sink" for decarboxylation reactions of beta-keto acids.

This **iminium ion** or protonated **Schiff Base** has a pK_a of about 7, so the protonated iminium and deprotonated imine are in equilibrium near pH 7. Figure 6.1.12 also illustrates a simple way to view reaction mechanisms. Electrons in chemical reactions can be viewed as flowing from a source (such as a carboxyl group) to a sink (such as an electrophilic carbonyl O or a positively charged N in a Schiff base).

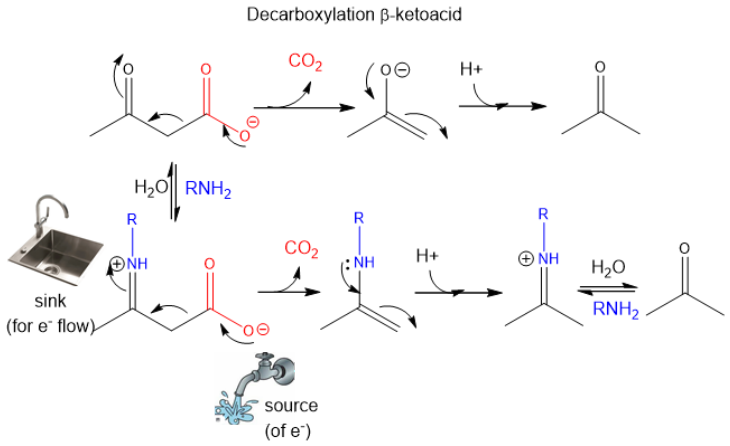


Figure 6.1.12: Schiff-base catalyzed decarboxylation of a β-ketoacid - Source/Sink analogy

Acid- and base-catalyzed reaction mechanisms for Schiff base formation are shown in Figure 6.1.13. An amine is used as the nucleophilic catalyst, forming the initial addition product, a carbinolamine. The carbinolamine dehydrates, since the free pair of electrons on the N are more likely to be shared with the carbon to form a double bond than electrons from the original carbonyl O, which is more electronegative than the N. If catalyzed by a general acid, an iminium ion and the base-catalyzed reaction form an imine. Near pH 7.4, the imine easily protonated to form a positively charged N at the former carbonyl O center.

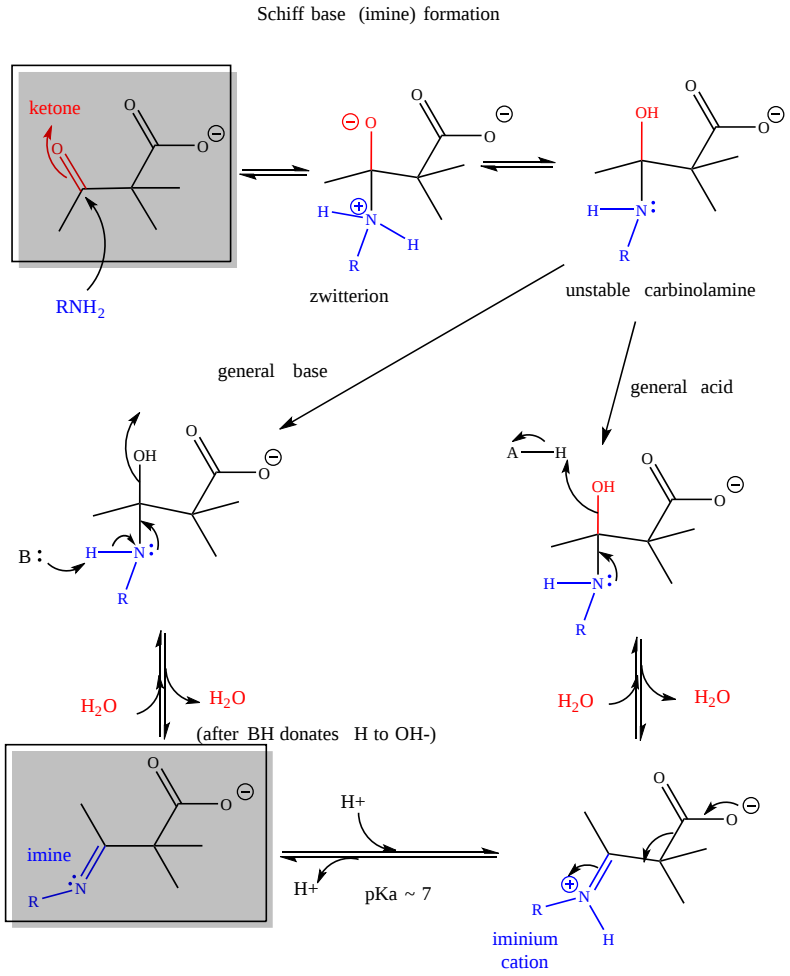


Figure 6.1.13: Schiff Base Formation - reaction mechanism

An actual Schiff base intermediate between fructose-1,6-bisphosphate (2FP-400) and Lys 239 from the enzyme fructose bisphosphate aldolase from Leishmania mexicana is shown in Figure 6.1.14 Only a single bond between the carbon and nitrogen in the Schiff base is shown.

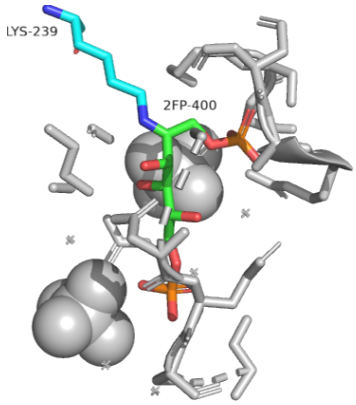


Figure 6.1.14: Schiff base intermediate between fructose-1,6-bisphosphate (2FP-400) and Lys 239 from the enzyme fructose bisphosphate aldolase from Leishmania mexicana(2QDG)

6.1.1.4: Transition State Stabilization

In the middle of the 20th century, Linus Pauling postulated that the only thing that a catalyst must do is bind the transition state more tightly than the substrate. This can be discerned in Figure 6.1.15 and a little math. The diagram shows how the substrate (S) and the transition state (S*) can react with an enzyme (E) to form a complex, which then proceeds to product (following the diagram from "start here" in the right-hand direction), or can go to product in the absence of enzyme (E) (following the diagram from "start here" in the left-hand direction). Note that the diagram is arbitrarily drawn such that the standard Gibbs free energy (G°) of the free product P is higher than that of the free substrate, S.

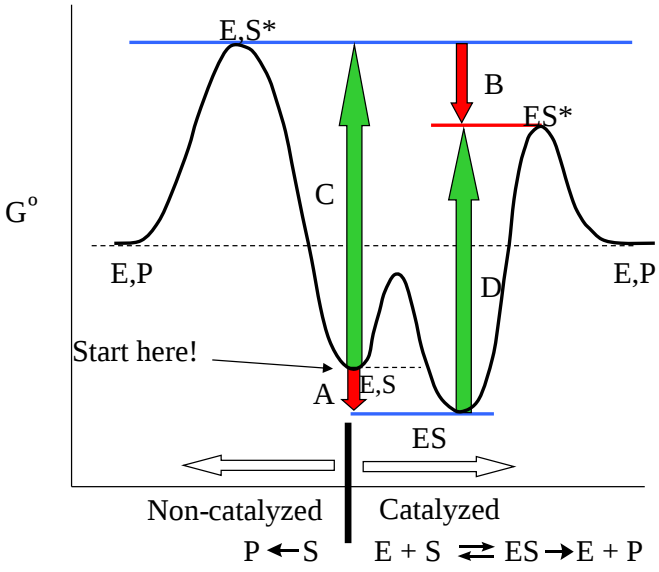


Figure 6.1.15: Enzymes bind the transition state more tightly than the substrate

The large colored vertical arrows represent the ΔG° for the transition shown:

- The red arrows A and B represent the ΔG°s for the binding of E + S (arrow A) and E + S* (arrow B), respectively.
- The green arrows C and D represent the ΔG°s for the activation energy of the free substrate (arrow C) and the enzyme-bound substrate (arrow D).

Now consider the two pairs of arrows (A,C and D,B). Add each set like a vector in elementary physics. Since the distance between the two horizontal blue lines is the same for the left-hand process (uncatalyzed) and the right-hand one (catalyzed), it follows that

$$C - A = D - B \tag{6.1.1}$$

The negative signs for A and B are used since, in the diagram, both A and B have negative ΔG° values.

Now, for an enzyme to be a catalyst, the activation energy D for the reaction in the presence of the enzyme E must be less positive (i.e., smaller) than the activation energy C in the absence of the enzyme. Therefore, after rearranging the equation:

$$C - D = A - B > 0 \tag{6.1.2}$$

and substituting in the ΔG° values for A and B, we can directly compare the free energy of binding the substrate (S) vs. binding the transition state (S*):

$$-RT \ln \mathrm{K_{eqS}} - \left(-\mathrm{RT} \ln \mathrm{K_{eqS^*}}\right) > 0 \tag{6.1.3}$$

Hence, the equilibrium constant for binding the transition state is larger than that for binding free substrate:

$$\mathrm{K_{eq}}_{S^*} > \mathrm{K_{eq\,S}} \tag{6.1.4}$$

Pauling was right. The enzyme just needs to bind the transition state more tightly than the substrate to catalyze the reaction. This is why chemists synthesize stable transition state analogs as potential tight-binding inhibitors of target

proteins.

The stability of the transition state also affects the reaction kinetics (which makes sense, given that the activation energy clearly affects the speed of a reaction). As you probably remember from organic chemistry, biomolecular nucleophilic substitution (S_N2) reactions are slow when the central atom where the substitution will occur is surrounded by bulky substituents (sterics once again). We discussed this in context to nucleophilic substitution on a sp^2 hybridized carbonyl carbon in carboxylic acid derivatives versus on a sp^3 hybridized phosphorous in phosphoesters and diesters. The explanation for this phenomenon has usually been attributed to hindered access to the central atom caused by bulky substituents (intrinsic effects). Is this true? Studies on S_N2 reactions of methylchloroacetonitrile and t-butylchloroacetonitrile (with the reagent labeled with ^{35}Cl) using $^{37}Cl^-$ as the incoming nucleophile in the gas phase showed that the more hindered t-butyl derivative's activation energy was only 1.6 kcal/mol (6.7 kJ/mol) higher than the methyl derivative, but in aqueous solution, the difference is much greater for comparable reactions (Figure 6.1.16).

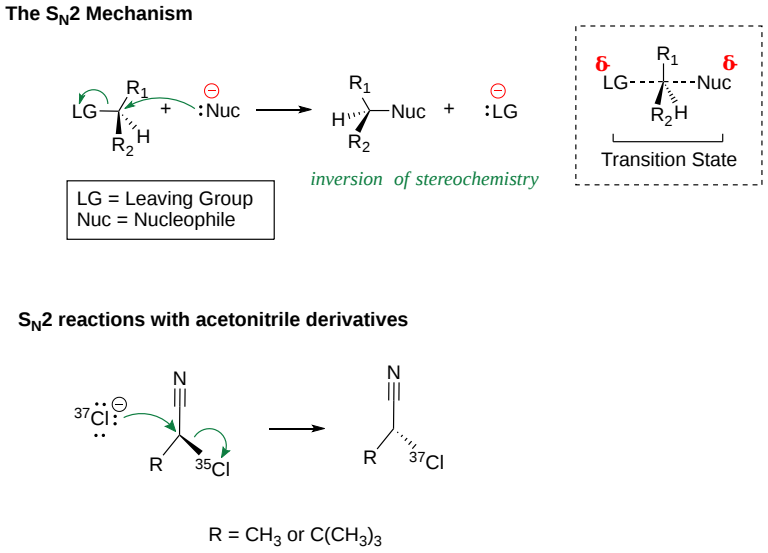


Figure 6.1.16: S_N2 reactions are characterized by a pentavalent transition state

They attributed the differences to solvation effects of the transition state. The bulkier the substituents on the central atom, the more difficult it is to solvate the transition state since water also can't reorient around it as well. In effect, there is steric hindrance for both reactant and solvent.

What does it take for a macromolecule (M) to be a catalyst - an enzyme? It seems the minimum criteria are:

- M binds a reactant
- M binds the transition state more tightly than the substrate

Anything above these is just "icing on the cake." If different functional groups are present in the "active" site of the enzyme that would allow electrostatic, intramolecular, covalent, general acid and/or base catalysis, the better the catalyst.

6.1.1.4.1: A transition state analog case study: Abyzmes (Antibody Catalysis)

Recall that antibodies are immune system proteins that bind foreign molecules (see Chapter 5.4). The usual role of an antibody is to initiate an immune response. When the antigen-binding site, located in the variable region of an antibody, binds to an antigen, it triggers the formation of new antibodies (within B-cells) to optimize the immune response to that antigen. These new antibodies are made with mutations in the antigen-binding region. Those that bind better than the original antibody will be selected to form longer-lived memory B-cells, ready for the next time the body encounters the antigen.

Decades ago, Linus Pauling hypothesized that antibodies could be produced with an atypical role—to act as catalysts! If antibody molecules could be made to bind to a compound resembling the transition state of a chemical reaction, they should also presumably catalyze the chemical reaction. In 1987, his prediction was verified. Lerner et al. made a transition state analog of an ester. When an ester is hydrolyzed, as shown in Figure 6.1.17, the sp^2 hybridized carbonyl carbon is converted to an sp^3 hybridized center in the intermediate, with the carbonyl oxygen becoming an oxyanion.

The transition state presumably resembles this unstable intermediate (sp^3, oxyanion). Thus, Lerner synthesized a phosphonate, an ester mimic with a sp^3 hybridized phosphorous replacing the carbonyl C. It also has a negatively charged oxygen, as does the

intermediate for the ester. This phosphonate ester is very resistant to hydrolysis. When injected into a mouse (after first being covalently attached to a carrier protein so the small molecule becomes "immunogenic"), the mouse makes a protein antibody that binds to the phosphonate. When the corresponding carboxylic acid ester is added to the antibody, it is cleaved with nominal k_{cat} and K_M values. Site-directed mutagenesis can then be done to make it an even better catalyst! The antibody enzymes have been called **abzymes**. The structure shown in Figure 6.1.17 also shows how phosphonamides act as transition state analogs.

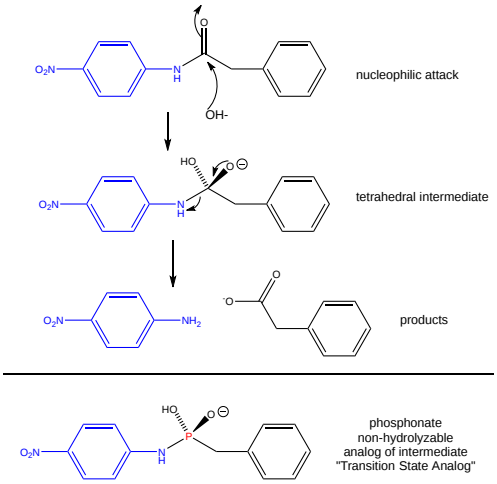


Figure 6.1.17: PHOSPHONAMIDES: TRANSITION STATE ANALOGS

Figure 6.1.18 shows an interactive iCn3D model of the transition state analog 5-(para-nitrophenyl phosphonate)-pentanoic acid bound to a mouse Fab antibody fragment with esterase activity (1aj7)

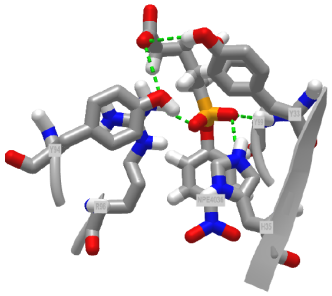


Figure 6.1.18: Transition state analog 5-(Para-Nitrophenyl Phosphonate)-Pentanoic Acid (NPE406) bound to Fab antibody fragment with esterase activity (1aj7) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?zSWAhDag8t4LpJaN7

Transition state theory can be used to quantify the relationships described in the graphical analysis above more clearly. This analysis will use the equilibrium constant (in contrast to the last two chapters, which used dissociation constants to characterize macromolecules, receptors, and enzymes binding to ligands). Let's assume that a substrate S is in equilibrium with its transition state $S^{\ddagger}$. Hence Keq = $[S^{\ddagger}]/[S]$. The following reaction can be written: S → $S^{\ddagger}$ → P. Based on our previous kinetic analysis and experience in writing differential equations, dP/dt = $k1[S^{\ddagger}]$. By analogy, enzyme-bound S (ES) can be converted to ($ES^{\ddagger}$) and then on to product as shown in the following chemical equation:

$$\mathrm{E + S \rightleftharpoons ES \longrightarrow ES^{\dagger} \longrightarrow E + P.}$$

For the non-enzyme catalyzed reaction, transition state theory can be used to show that the first order rate constant $k_1 = kT/h$ where k is the Boltzmann constant, T is the Kelvin temperature, and h is Planck's constant. Hence, using $K_{eq} = [S^{\dagger}]/[S]$, equation 1 can be derived

$$\frac{dP}{dt} = \frac{kT}{h}\left[S^{\dagger}\right] = \frac{kT}{h}K^{\dagger}[S] = k_n[S] \tag{6.1.5}$$

where k_n (hereafter written as k_N) $=(kT/h)K^{\dagger}$ is the effective first-order rate for the non-catalyzed rate. Now, let's create a more complicated linked equilibrium that shows the same reaction in the presence of an enzyme. Figure 6.1.19

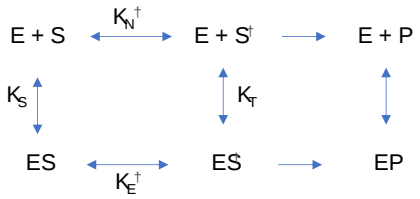


Figure 6.1.19: Linked equilibrium for conversion of free $S^\dagger$ and $ES^\dagger$ to product

Remember that the K values for this analysis are equilibrium constants, not dissociation constants. Note two important equilibrium constants, K_S, the equilibrium constant for the binding of free S to E, and K_T, the equilibrium constant for the binding of free $S^\dagger$ to E (assuming that free $S^\dagger$ could bind to E before it converted to product). As we have seen for linked equilibrium before, since the K_{eq} values are related to the standard free energy changes, which are state functions, the sum of the standard free energies going from E + S to $ES^\dagger$ (by either the top or bottom paths) are path independent, so the products of the K_{eq} for the top path are equal to those for the bottom paths. This gives the following equation:

$$\frac{K_T}{K_S} = \frac{K_{E^\dagger}}{K_{N^\dagger}} = \frac{k_E}{k_N} \tag{6.1.6}$$

The right-hand side is the ratio of the effective first-order rate constant for conversion or $ES^\dagger \rightarrow E + P$, k_E divided by the rate constant for the conversion of $S^\dagger \rightarrow P$ for the noncatalyzed rate, k_N. The final ratio of rate constants can be derived from the simple relationship that $k_x=(kT/h)K^\dagger_x$ where x is either N (non-catalyzed) or E (enzyme-catalyzed). Equation 2 states that the equilibrium constant for the binding of $S^\dagger$ to E, K_T, is greater than that for the binding of S to E, K_S (as $k_E > k_N$). K_T/K_R ranges from 10^8 - 10^{14}. Given common values for the equilibrium constant for binding of S to E (10^3 – 10^5 M^{-1}), which is equivalent to the dissociation constant values K_d = 10 uM – 1 mM, the calculated value of $K_T = 10^{15}$ $M^{-1,}$ which gives a dissociation constant for the enzyme and transition state of $K_d = 10^{-15}$ M (1 femtomolar). This is as tight as one of the highest affinity binding interactions in the biological world, the binding of avidin and biotin. As we noted in Chapter 5.1, assuming that the second order rate constant for avidin/biotin binding and as shown above for $E/S^\dagger$ is diffusion controlled (about 10^8 $M^{-1}s^{-1}$), the off rate for the avidin-biotin or $ES^\dagger$ complex is 10^{-7} s^{-1}, equivalent to a half-life of the complex of 80 days. It doesn't get much tighter than that.

Figure 6.1.20 represents an image of an enzyme and three molecules, 1–3, that could bind to it. Using the analysis above, which molecule do you think represents substrate? Transition state? Product?

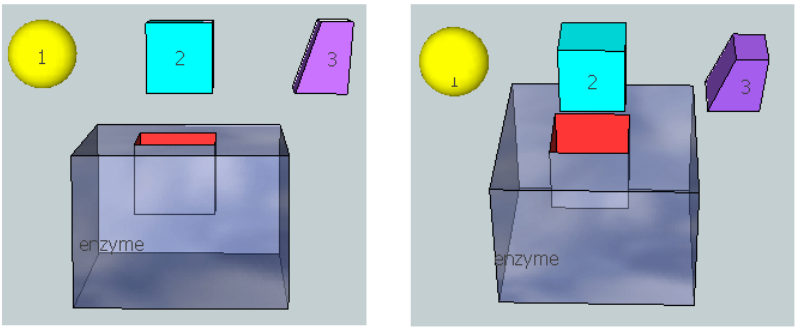


Figure 6.1.20: 3D model for binding substrate, transition state and product to an enzyme

6.1.2: Physical Strategies for Rate Enhancement

6.1.2.1: Intramolecular Catalysis

Consider the hydrolysis of phenylacetate. This reaction, a nucleophilic substitution reaction, could be catalyzed by the addition of the general base acetate to the solution, as described above. Since this reaction rate would double with the doubling of the solution acetate, the reaction is bimolecular (first order in reactant and catalyst). Now consider the same reaction only when the general base part of the catalyst, the carboxyl group, is part of the reactant phenylacetate. Such a case occurs in the acetylated form of salicylic acid—i.e., aspirin. When the carboxy group is ortho compared to the acetylated phenolic OH, it is in the perfect position to accept a proton from water, decreasing the charge development on the O in the transition state. The general base does not have to diffuse to the appropriate site when it is intramolecular with respect to the carbonyl C of the ester link. The rate of this intramolecular base

catalysis is about 100-fold greater than that of an intermolecular base catalyst like acetate. It is as if the effective concentration of the intramolecular carboxyl base catalyst is much higher due to its proximity to the reaction site.

Another type of reaction involving a carboxyl group (in addition to simple proton transfer) is when the negatively charged carboxyl O acts as a nucleophile and attacks an electrophilic carbonyl carbon. When the carbonyl is part of an ester, the carboxyl group engages in a nucleophilic substitution reaction, expelling the alcohol part of the ester as a leaving group. The remaining examples below consider the nucleophilic (carboxyl) substitution on phenylesters, with phenolate as the leaving group. The reactions, in effect, transfer an acyl group to the carboxyl group to create an anhydride.

First, consider acyl transfer with aspirin derivatives. As you know, aspirin contains a carboxyl group ortho to an ester substituent. Hence, the carboxyl group can act as a nucleophile and attack the carbonyl carbon of the ester in a nucleophilic substitution reaction. The net effect is to transfer the acetyl group from the phenolic OH to the carboxyl group, converting it to an anhydride. This is an intramolecular reaction. Compare this reaction to a similar bimolecular reaction shown in Figure 6.1.21.

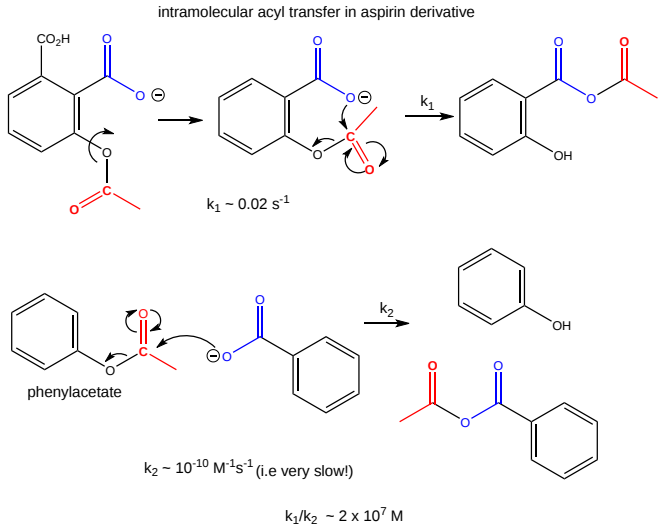


Figure 6.1.21: Acyl transfer in aspirin derivatives.

The first order rate constant of the intramolecular transfer of the acetyl group to the carboxyl group, k_1 = 0.02 s^{-1}. The analogous bimolecular reaction rate constant k_2~ 10^{-10} $M^{-1}s^{-1}$. Dividing k_1/k_2 gives the relative rate enhancement of the intramolecular over the intermolecular reaction. With units of molarity, this ratio can be interpreted as the relative effective concentration of the intramolecular nucleophile. This makes the effective concentration of the carboxylate in the aspirin derivative 2 x 10^7 M.

Now consider the cleavage of phenylacetate using acetate as the nucleophile, as shown in Figure 6.1.22 The products are acetic anhydride and phenolate. This is a bimolecular reaction (a slow one at that), with a bimolecular rate constant, k_2 which I will arbitrarily set to 1 for comparison to some similar reactions.

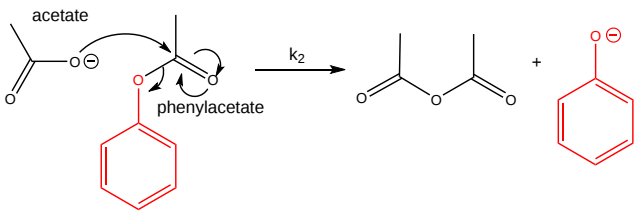


Figure 6.1.22: Reaction of acetate with phenylacetate

Now consider a monoester derivative of succinic acid - phenyl succinate - in which the free carboxyl group of the ester attacks the carbonyl carbon of the ester derivative, as shown in Figure 6.1.23

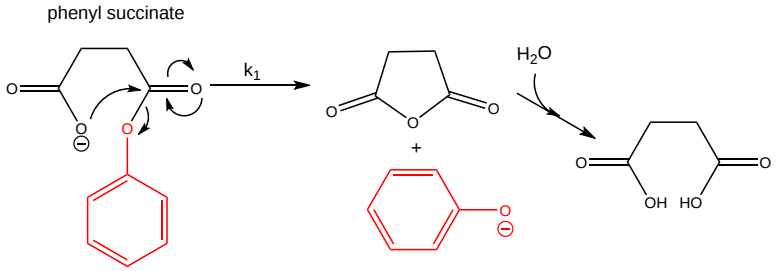


Figure 6.1.23: Intramolecular reaction of phenylsuccinate

Suppose you assign a second-order rate constant k_2 = 1 $M^{-1}s^{-1}$ to the analogous intermolecular reaction of acetate with phenylacetate (as described above). Then, the first-order rate constant for the intramolecular reaction of phenylsuccinate is 10^5 s^{-1}. The ratio of rate constants, $k_1/k_2 = 10^5$ M. That is, it would take 10^5 M concentration of acetate reacting with 1 M phenylacetate in the first bimolecular reaction to get a reaction as fast as the intramolecular reaction of phenylsuccinate. The intramolecular reaction of an even more sterically restricted bicyclic phenylcarboxylate shown in Figure 6.1.24 has a $k_1/k_2 = 10^8$ M.

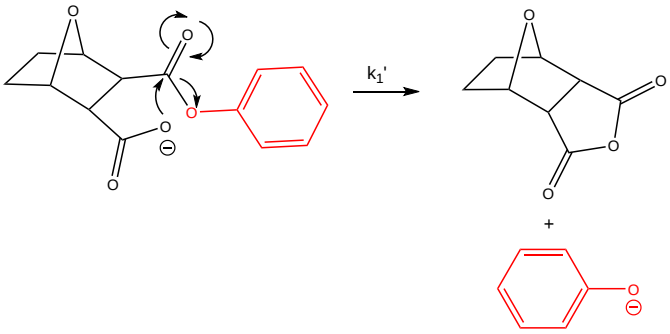


Figure 6.1.24: Intramolecular reaction of bicyclic phenylcarboxylate

Another example is anhydride formation between two carboxyl groups. The ΔG^o for such a reaction is positive, suggesting an unfavorable reaction. Consider two acetic acid molecules condensing to form acetic anhydride. For this intermolecular reaction, $K_{eq} = 3x10^{-12}$ M^{-1}. Now consider the analogous intramolecular reaction of the dicarboxylic acid succinic acid. It condenses in an intramolecular reaction to form succinic anhydride with a $K_{eq} = 8x10^{-7}$ (no units). The ratio $K_{eq\text{-}intra}/K_{eq\ inter} = 3\ x10^5$ M. It is as if the effective concentration of the reacting groups, because they do not have to diffuse together to react, is $3\ x10^5$ M.

How does this apply to enzyme-catalyzed reactions? Enzymes bind substrates in physical steps, which are typically fast. The slow step is often the chemical conversion of the bound substrate, which is effectively intramolecular if the initial binding reaction is fast. These three kinds of reactions, intermolecular, intramolecular, and enzyme-catalyzed, can be broken down into two hypothetical steps: a binding followed by catalysis as shown in Figure 6.1.25.

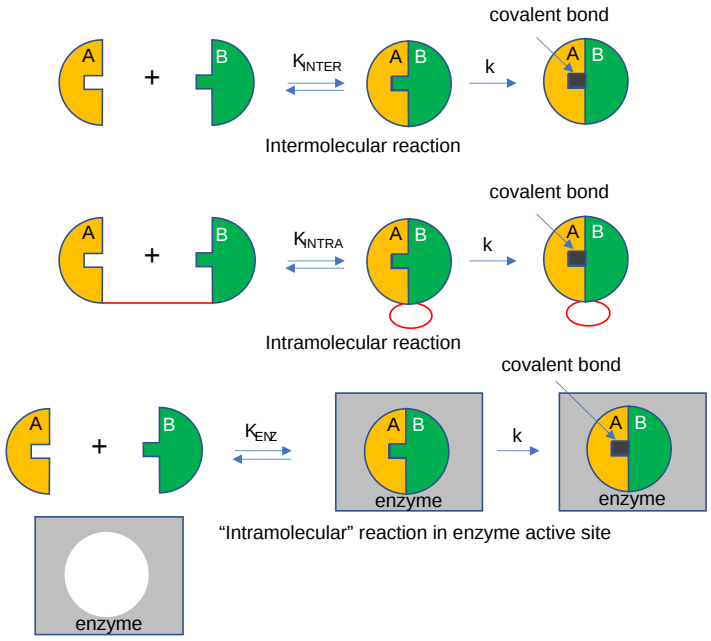


Figure 6.1.25: Intermolecular, intramolecular, and enzyme-bound reactions

Suppose the rate constants for the chemical steps are all identical. Then, the advantages of the intramolecular and enzyme-catalyzed reaction over the intermolecular reaction are K_{INTRA}/K_{INTER} and K_{ENZ}/K_{INTER}, respectively.

The advantage of intramolecular reactions can be seen by studying the Ca-EDTA complex. Calcium in solution exists as an octahedrally coordinated complex with water occupying all the coordination sites. EDTA, a multidentate ligand, first interacts through one of its potential six electron donors to Ca in a reaction entropically disfavored from the Ca-EDTA perspective, although one water is released. Once this first intramolecular complex is formed, the rest of the ligands on the EDTA rapidly coordinate with the Ca and release bound water, as illustrated in Figure 6.1.26. The former is no longer entropically disfavored since it is now an intramolecular process, while the latter is favored by releasing the remaining five water molecules.

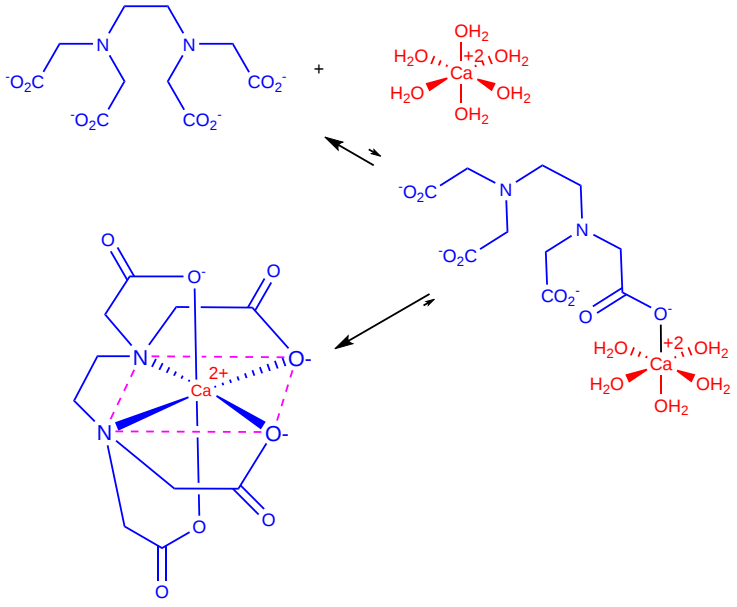


Figure 6.1.26: Binding of Ca^{2+} and EDTA

We've shown above the catalytic advantage offered by intramolecular reaction in terms of a dramatic increase in the effective concentration of reactants, which sometimes reached levels of 10^8 M. Another way is to look at entropy changes associated with

dimer formation. The table below shows that an intramolecular reaction is favored over an intermolecular reaction since, in the latter, significant decreases in translational and rotation entropy results.

Translational, Rotational, and Internal Entropies for Dimer Formation: A + B ↔ A-B (cal/K·mol)

System	A	B	A-B	ΔS
		Gas		
S trans	30	30	30	-30
S rot	20	20	20	-20
S int	5	5	20	+10
Gas → Solution	-10	-10	-15	
S sol	45	46	55	-35 (Correspond to 10^8-10^9 M)

6.1.2.2: Strain Distortion

In organic chemistry, you learned that certain structures, such as three-membered and four-membered ring structures, such as epoxides, were highly reactive due to the strain distortion inherent to the unfavored bond angles inherent to the ring. Enzyme active sites can also utilize strain distortion within a bound substrate to increase the molecule's reactivity and favor the formation of the transition state. Many enzymes that function by the induced fit model also utilize strain distortion within their catalytic mechanism. Within the unbound state they remain in a low catalytic state, however the interaction with the substrate induces the destabilization of the enzyme active site or may induce strain within the substrate causing the initiation of the catalytic activity of the enzyme.

Quantitative Deconstruction of Enzyme Catalysis

New ideas have emerged that quantitatively partition the overall catalytic enhancement by enzymes into discrete features and account for the full rate enhancement conferred by enzymes. We will discuss this in Chapter 6.5, using a class of enzymes called serine proteases as an example.

6.1.3: A Note on Asymmetric Catalysis/Organocatalysis

In a subsequent section, we will discuss how protein enzymes use the catalytic strategies described above. An intriguing question arises: how much of the structure of a large protein is really needed for catalysis? Much work has been directed to developing small molecule catalysis mimetics of large protein enzymes. How small can you reduce the size of a protein and still get catalysis?

One important feature of enzyme catalysis is that they catalyze reactions in which only one enantiomer is produced. That is, the synthesis is asymmetric. This is typically a consequence of the asymmetric enzyme (itself chiral) binding only one enantiomer as a reactant and/or the imposition of steric restrictions on the possible reactions of the bound substrate. L-Pro alone can act as an asymmetric catalyst in an aldol condensation reaction. Figure 6.1.27:

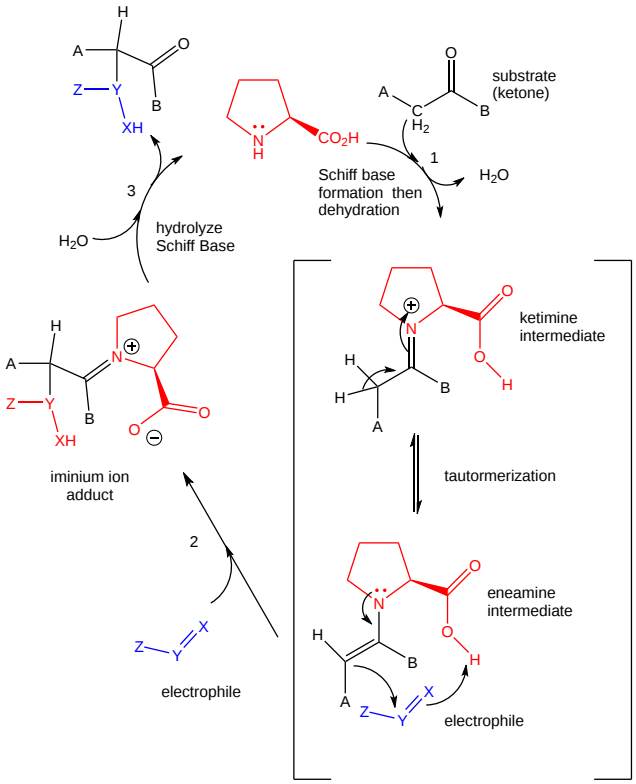


Figure 6.1.27: L-PRO CATALYSIS OF AN ALDOL CONDENSATION: POSSIBLE MECHANISM

Catalysts are vital in biological settings and the laboratory synthesis of molecules that sustain our culture and economy. Transition metal and, increasingly, protein enzymes have been used as industrial catalysts. New **asymmetric catalysts have now joined them** (a subset of **organocatalysts**). The work of Benjamin List and David MacMillan, who were instrumental in developing the ideas of asymmetric catalysts, has been recognized by the Nobel Commission, which awarded them the 2021 Nobel Prize in Chemistry.

The enzyme triose phosphate isomerase catalyzes an asymmetric reaction in which only one enantiomer of glyceraldehyde-3-phosphate is produced from the achiral dihydroxyacetone phosphate. Figure 6.1.28 shows this enantiospecific reaction.

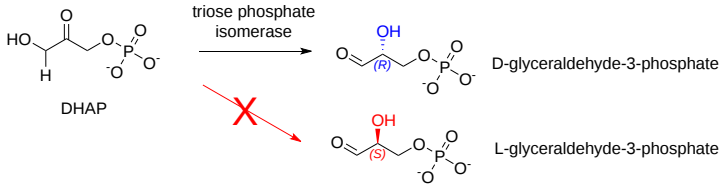


Figure 6.1.28: Reaction catalyzed by triose phosphate isomerase

Figure 6.1.29 contains an interactive iCn3D model of a triose phosphate isomerase from Trypanosoma brucei brucei (1KV5), which shows a conserved active site Pro 168 (spacefill) and amino acid side chains within 4 Å (stick) within the context of one monomer (cartoon) of the dimeric protein.

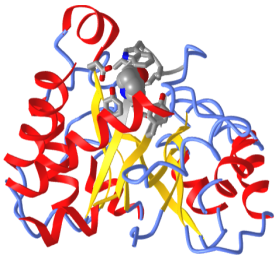


Figure 6.1.29: Active site proline in triose phosphate isomerase from Trypanosoma brucei brucei (1KV5). (Copyright; author via source). Click the image for a popup or use this external link https://structure.ncbi.nlm.nih.gov/icn3d/share.html?1Hnu1xmU5j76Jwrx9

6.1.4: Summary

This chapter provides an in-depth analysis of why uncatalyzed reactions in solution are slow and how catalysts, particularly enzymes, overcome these barriers by lowering the activation energy. The discussion begins with the example of ester hydrolysis, illustrating that the high activation energy of the uncatalyzed reaction stems from the formation of a charge-separated, high-energy transition state that closely resembles an unstable intermediate.

Key Concepts Covered:

1. **Energy Profiles and Transition States:**
   - The chapter explains how the energy diagram of an endergonic reaction features a transition state that is higher in energy than a charge-separated intermediate. This high-energy state, which arises from significant charge development during bond making and breaking, limits the reaction rate in the absence of a catalyst.
2. **Strategies for Transition State Stabilization:**
   - **Acid-Base Catalysis:**
     - *Specific Acid/Base Catalysis:* The reaction rate depends solely on the concentration of hydronium or hydroxide ions.
     - *General Acid/Base Catalysis:* Buffer components (or functional groups within an enzyme) donate or abstract protons to reduce charge separation in the transition state, thereby lowering the activation energy.
   - **Metal Ion (Electrostatic) Catalysis:**
     - Metal ions (e.g., Zn^{2+}, Cu^{2+}) coordinate with reaction intermediates to stabilize developing charges, enhance electrophilicity, and sometimes lower the pKa of bound water molecules, which is critical in enzymes like carbonic anhydrase.
   - **Covalent (Nucleophilic) Catalysis:**
     - Nucleophilic catalysts such as pyridine form covalent intermediates with the substrate. This alternative reaction pathway features lower activation energy due to the formation of a new, more stable intermediate.
   - **Intramolecular Catalysis:**
     - When a catalytic group is part of the same molecule as the reactive site, as in aspirin derivatives or phenylsuccinate, the effective concentration is dramatically increased. This proximity effect accelerates the reaction by minimizing the entropic cost of bringing reactants together.
3. **Quantitative Aspects of Enzyme Catalysis:**
   - Transition state theory is used to show that catalysts, including enzymes, enhance reaction rates by binding the transition state more tightly than the substrate. This principle, first proposed by Linus Pauling, is quantified by comparing equilibrium constants for substrate and transition state binding.
   - The chapter presents mathematical derivations linking free energy differences, rate constants, and equilibrium constants (K_eq) for both catalyzed and uncatalyzed reactions, demonstrating that enzymes can achieve binding affinities for transition states comparable to some of the strongest noncovalent interactions known in biology.
4. **Applications and Implications:**
   - **Transition State Analogs and Abzymes:**

- By synthesizing stable molecules that mimic the transition state, researchers can design potent inhibitors and even generate catalytic antibodies (abzymes) that accelerate chemical reactions.

- **Asymmetric Catalysis:**
  - The inherent chirality of enzymes ensures that only one enantiomer of the product is formed, a principle that is also applied in the development of small-molecule organocatalysts.
- **Effective Concentration and Entropy Considerations:**
  - The chapter discusses how intramolecular reactions overcome entropic barriers that limit bimolecular reactions, with quantitative comparisons illustrating rate enhancements that can approach effective concentrations of up to 10^8 M.

In summary, the chapter unifies various catalytic strategies—acid-base, metal ion, covalent, and intramolecular—under the common theme of transition state stabilization. It highlights how enzymes, through precise structural features and dynamic interactions, achieve extraordinary rate enhancements. This understanding not only elucidates fundamental principles of chemical reactivity but also informs the design of synthetic catalysts and therapeutic inhibitors in modern biochemistry.

6.1.5: References

Amyes TL, Richard JP. Biochemistry. 2013, 52(12), 2021-35. doi: 10.1021/bi301491r

Ferst and Kirby, J. Am. Chem. Soc. 1967, 89, 19, 4857–4863. https://doi-org.ezproxy.csbsju.edu/10.1021/ja00995a007

Komiyama et al., PNAS, **1977,** 74 (7) 2634-2638. https://doi.org/10.1073/pnas.74.7.263

Lerner, R. A., & Tramontano, A. (1987). Antibodies as enzymes. *Trends in Biochemical Sciences*, *12*, 427-430. https://doi.org/10.1016/0968-0004(87)90208-8

OpenStax College. (2022). *Chemistry*. OpenStax. https://openstax.org/details/books/chemistry-2e

Regan, C. K., Craig, S. L., & Brauman, J. I. *Science.* 2002 *295*(5563), 2245-2247. DOI: 10.1126/science.1068849

6.2: Kinetics without Enzymes

Learning Goals (ChatGPT o1, 1/30/25)

- **Fundamentals of Reaction Kinetics and Calculus in Chemistry**
  - Understand the use of calculus in describing reaction kinetics, including the meaning of the derivative dX/dt as the instantaneous rate of change in concentration.
  - Translate between the differential (dX/dt) and integrated (X as a function of time) forms of rate equations.
- **Analyzing First-Order Reactions**
  - Derive and solve the first-order differential equation for an irreversible reaction (e.g., A → P) and explain the significance of the solution, $A = A_o e^{-k_1 t}$.
  - Interpret kinetic graphs, including exponential decay curves for A versus time and the linear relationship in ln[A] versus time plots, and relate these to reaction velocity.
- **Understanding Second-Order and Pseudo First Order Kinetics**
  - Derive and solve the integrated rate equation for second-order reactions (e.g., A + A → P or A + B → P) and explain how the reaction order influences the time course of reactant consumption.
  - Recognize conditions under which a bimolecular reaction can be treated as pseudo first order and calculate the effective rate constant.
- **Initial Rate vs. Integrated Rate Approaches**
  - Compare and contrast initial rate methods (measuring the slope at t ≈ 0) with progress curve analysis (integrated rate equations) and discuss the advantages and limitations of each in kinetic studies.
- **Multi-Step and Reversible Reaction Kinetics**
  - Write differential equations for simple reversible reactions (A ↔ P) and multi-step processes (e.g., consecutive irreversible reactions A → B → C and reversible networks A ↔ B ↔ C).
  - Interpret progress curves and steady-state behavior in reversible systems and understand how changes in rate constants affect the approach to equilibrium.
- **Modeling and Data Analysis in Kinetics**
  - Utilize computational tools (e.g., VCell) to solve ordinary differential equations (ODEs) for complex reaction systems and generate progress curves that illustrate concentration changes over time.
  - Apply statistical methods and curve fitting to experimental kinetic data, linking theoretical models with observed reaction behavior.
- **Developing Chemical and Biological Intuition**
  - Integrate mathematical concepts with chemical intuition by relating the rate equations and progress curves to molecular behavior and reaction mechanisms.
  - Understand how the kinetic analysis of simple chemical reactions provides the foundation for exploring more complex enzyme-catalyzed reactions in biological systems.

By achieving these learning goals, students will be equipped to analyze, model, and interpret the kinetics of single-step and multi-step reactions, forming a critical basis for understanding enzyme-catalyzed processes and broader biochemical reaction mechanisms.

6.2.1: Single Step Reactions

First, we will explore the kinetics of non-catalyzed reactions, which is needed to understand the kinetic of the more complicated enzyme-catalyzed reactions.

Calculus: Derivatives and Integrals

In this book, we will refer to the change of concentration of a chemical species X as a function of time as dX/dt instead of ΔX/Δt, where dX/dt is the derivative of X with respect to time t. This is the language of calculus, which most readers would have studied. The use of calculus will be mainly limited to writing equations of the form dX/dt = f(t), a type of differential equation. We will also use a few integrals, but a working knowledge of calculus is not required. For readers who have not studied calculus, replace dX/dt with ΔX/Δt in your mind, and you will derive the same meaning.

You studied two types of kinetic equations in introductory chemistry to analyze kinetic data:

- Initial Rates: In this method, the initial rate, v_0, is measured as a function of the concentration of reactants. The initial velocity, v_0, is the initial slope of a graph of the concentration of reactants or products as a function of time, taken over a range of times such that only a small fraction of A has reacted. Under this condition, [A] over this short time range is approximately constant and equal to Ao. Initial rate graphs are often based on the measurement of product increase with time, ΔP/Δt, so v_0 vs. A plots have positive slopes. The velocity at time t along the A vs. t curve, dA/dt, constantly changes as [A] decreases since the velocity depends on the [A]. To reiterate, the initial velocity of the reaction is the slope of the initial linear part of the decay curve when the rate is essentially linear over a narrow range of [A].
- Integrated Rates and Progress Curves: In this method, a differential equation that gives the change of A or P with time (dA/dt or dP/dt) is integrated to give an equation that shows the concentration of A or P as a function of time. For any given reaction, it is essential to be able to write the integrated rate equation, often called an ordinary differential equation (ODE). In almost every case, we provide the solutions to the ODEs in either a mathematical equation or a fitted graph of A vs. t. As the reactions get more complicated, we let computers solve them numerically and show the output. The ODEs are also called **progress curves** as they show how the concentration of reactants and products change with time.

In either case, the mathematical equations describing the reaction are used to fit experimental data taken in the lab. Statistics are used to see how closely the experimental data fit the rate equations. The best-fit equation gives the mostly likely chemical reaction equations for the interconversion of reactant and products.

Most biochemistry textbooks focus on initial velocities when explaining enzyme kinetics. However, in many ways, this method is less intuitive than exploring how the concentration of a molecule changes with time. An analogy is the concept of density (mass/volume), which is more complicated than understanding mass or volume separately. We use both methods to develop both chemical and biological intuitions of kinetic properties.

We will first explore simple irreversible reactions (1st and 2nd order), then make them reversible, and then couple them together to form more complex reaction schemes, much as we are compelled to do for enzyme-catalyzed reactions.

6.2.1.1: First Order Reaction

$$\mathrm{A} \xrightarrow{k_1} \mathrm{P}$$

where k_1 is the first-order rate constant. For these reactions, the velocity of the reaction, v, is directly proportional to [A], or

$$v = -\frac{dA}{dt} = +\frac{dP}{dt} = k_1 A \tag{6.2.1}$$

The negative sign in -d[A]/dt indicates that the concentration of A decreases. The equation could also be written as: $v = \frac{dA}{dt} = -k_1 A$

For the rest of the reactions in this book, we follow the convention of writing all velocities expressed as d[x]/dt as positive numbers. A negative sign for a term on the right-hand side of the differential equation will indicate that the concentration dependency of that term will lead to a decrease in [x] with time. Likewise, a positive sign for the term on the right-hand side of the equation will indicate that concentration dependency of that term will lead to an increase in [x] with time.

Examples: dA/dt = **-k**A shows that the A will **decrease** with time. dA/dt = **+k**A shows that A will **increase** with time.

Here is the solution to the differential Equation 6.3.A.1 for [A] as a function of t.

$$A = A_0 e^{-k_1 t} \tag{6.2.2}$$

? A derivation of the first order rate equation

Here it is!

Derivation

$$\begin{aligned} \int_{A_O}^{A} \frac{dA}{A} &= -k_1 \int_0^t dt \\ \ln A - \ln A_0 &= -k_1 t \\ \ln A &= \ln A_0 - k_1 t \\ A &= A_0 e^{-k_1 t} \end{aligned} \tag{6.2.3}$$

Equation 6.3.A.2 is an example of an integrated rate equation. The following graphs show plots o.f A vs t and lnA vs. t for a first-order process. Note that the derivative of the graph of A vs. t (dA/dt) is the velocity of the reaction. The graph of ln[A] vs. t is linear with a slope of $-k_1$. The velocity of the reaction (slope of the A vs t curve) decreases with decreasing A, which is consistent with equation 1. Again, the initial velocity is determined from data taken in the first part of the decay curve when the rate is linear and little A has reacted. That is, [A] is approximately equal to $[A_0]$.

Figure 6.2.1 shows two ways to plot 1st order reaction data. The left graph shows the exponential decay of A with time and the corresponding rise in P when A_0 is 0 and k_1=2. The other shows the linear fall of ln[A] vs time.

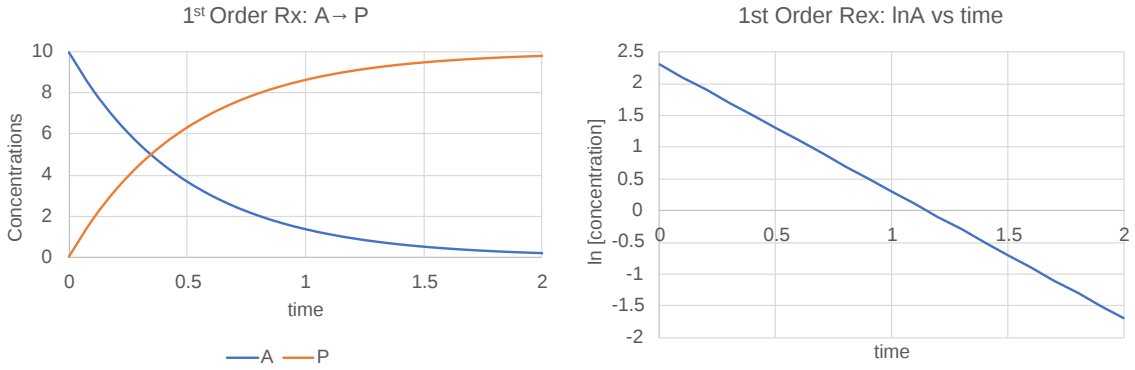


Figure 6.2.1: Two ways to plot 1st order reaction data

Once again, for complete clarification, another way of analyzing the kinetics of a reaction, in addition to following the concentration of a reactant or product as a function of time and fitting the data to an integrated rate equation, is to plot the initial velocity, v_o, of the reaction as a function of the concentration of reactants. The initial velocity is the initial slope of a graph of the concentration of reactants or products as a function of time, taken over a range of times such that only a small fraction of A has reacted, so [A] is approximately constant = A_o. From the first-order graph of A vs. t above, the slope approaches 0 with increasing time as [A] approaches 0, which indicates that the reaction velocity depends on A. For this first-order order process, two equivalent equations can be written showing the

- disappearance of A as v = -d[A]/dt = k_1[A], while
- appearance of A as v = d[A]/dt = -k_1[A],

Both equations show that v is directly proportional to A. As [A] is doubled, the initial velocity is doubled.

Velocity graphs used by biochemists often show the initial velocity of product formation (not reactant decrease) as a function of reactant concentration. Hence, as product concentration increases, the initial velocity are positive. A graph of v (= dP/dt) vs [A] for a first order process would have a positive slope and be interpreted as showing that the rate of appearance of P depends linearly on [A].

6.2.1.2: Second Order/Pseudo First Order Reactions

$$A + B \xrightarrow{k_2} P \text{ or } A + A \xrightarrow{k_2} P \tag{6.2.4}$$

where k_2 is the second-order rate constant. For the first of these irreversible reactions, the velocity of the reaction, v, is directly proportional to [A] and [B], or

$$v = \frac{dA}{dt} = -k_2[A][B] \tag{6.2.5}$$

We will consider two special cases of this reaction type:

1. [B] >> [A]. Under these conditions, [B] never changes, so Equation 5 becomes

$$v = -\left(k_2[B]\right)[A] = -k_1'[A] \tag{6.2.6}$$

where $\mathbf{k_1'}$ is the pseudo first order rate constant (= k2[B]) for the reaction. The reaction appears to be first order, depending only on [A].

2. As illustrated in the second reaction above, the only reactant is A, which must collide with another A to form P.

The following differential equation can be written and solved to find [A] as a function of t.

$$v = \frac{dA}{dt} = 2\frac{dP}{dt} = -k_2A^2 \tag{6.2.7}$$

Solving the differential equation for A gives the following:

$$\frac{1}{A} = \frac{1}{A_0} + k_2t \tag{6.2.8}$$

? A derivation of the second order rate equation

Here it is!

Derivation

$$\begin{gathered}
\frac{dA}{dt} = -k_2A^2 \\
\int_{A_0}^{A}\frac{dA}{A^2} = \int_{A_0}^{A}A^{-2}dA = \int_{A_0}^{A}A^{-2}dA = -k_2\int_0^t dt \\
\frac{A^{n+1}}{n+1}\Bigg]_{A_0}^{A} = \frac{A^{-1}}{-1}\Bigg]_{A_0}^{A} = -k_2t\Bigg]_{A_0}^{A} \\
-\frac{1}{A} - \left(\frac{1}{A_0}\right) = -k_2t
\end{gathered} \tag{6.2.9}$$

Figure 6.2.2 shows plots of A vs. t and 1/A vs. t for a second-order process when A_0 is 0 and k_2=1. The right graph shows the linear rise of 1/[A] with time.

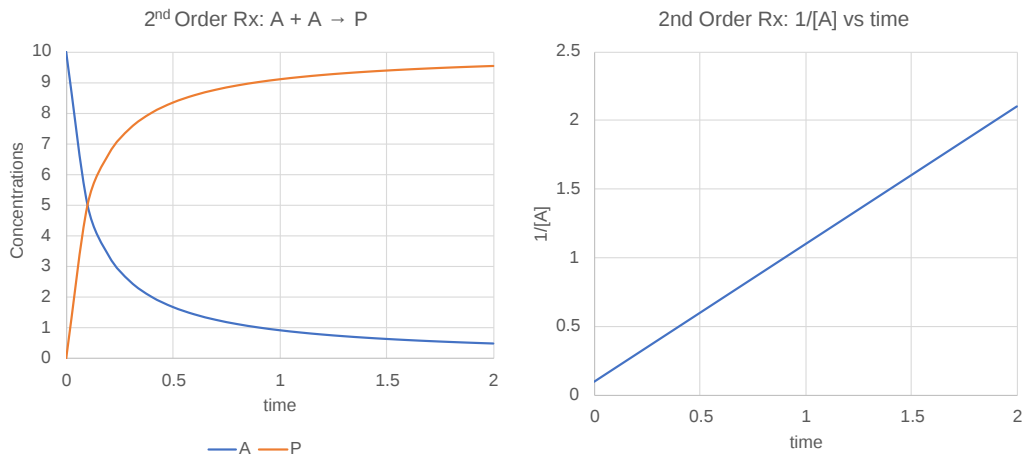
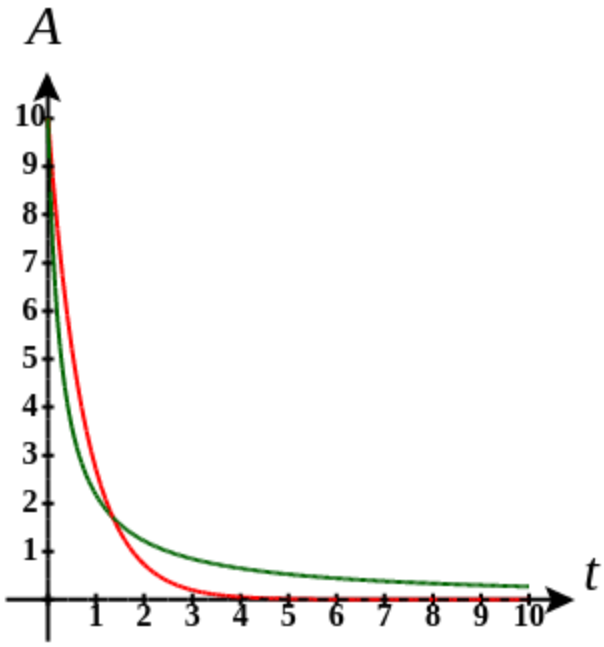


Figure 6.2.2: Plots of A vs t and 1/A vs t for a second order process when A_0 is 0 and k_2=1

Note that just from a plot of A vs. t, it would be difficult to distinguish a first from a second-order reaction. If the plots were superimposed, you would observe that at the same concentration of A (10, for example), the v_o of a first-order reaction would be proportional to 10, but for a second-order reaction, to 10^2 or 100. Therefore, the second-order reaction is faster (assuming similarity in the relative magnitude of the rate constants), as indicated by the steeper negative slope of the curve. However, at low A (0.1 example), the v_o of a first-order reaction would be proportional to 0.1 but second-order order reaction to 0.1^2 or 0.01. Therefore, at low A, the second-order reaction is slower.

The interactive graphs below show the first and second-order conversion of reactant A to product. Change the sliders to see how the curves are different.



By comparing these curves, you should see how difficult it is to differentiate between a 1^{st} and 2^{nd} order process unless the reaction progresses to almost completion.

6.2.2: Multi-Step Reactions

6.2.2.1: Reversible First Order Reactions

$$\mathrm{A} \underset{k_2}{\overset{k_1}{\rightleftarrows}} \mathrm{P} \tag{6.2.10}$$

Here is the differential equation for dA/dt..

$$\mathrm{v} = \frac{\mathrm{dA}}{\mathrm{dt}} = -\mathrm{k}_1\ \mathrm{A} + \mathrm{k}_2\mathrm{P} \tag{6.2.11}$$

Here are the solution for both A and P as a function of time

$$A = \frac{A_0\left(k_2 + k_1\left[e^{-(k_1+k_2)t}\right]\right.}{k_1 + k_2}$$
$$P = A_0 - \left(\frac{A_0\left(k_2 + k_1\left[e^{-(k_1+k_2)t}\right]\right.}{k_1 + k_2}\right) \tag{6.2.12}$$

Figure 6.2.3 shows graphs of A and P vs t for this reaction at two different sets of values of k_1 and k_2.

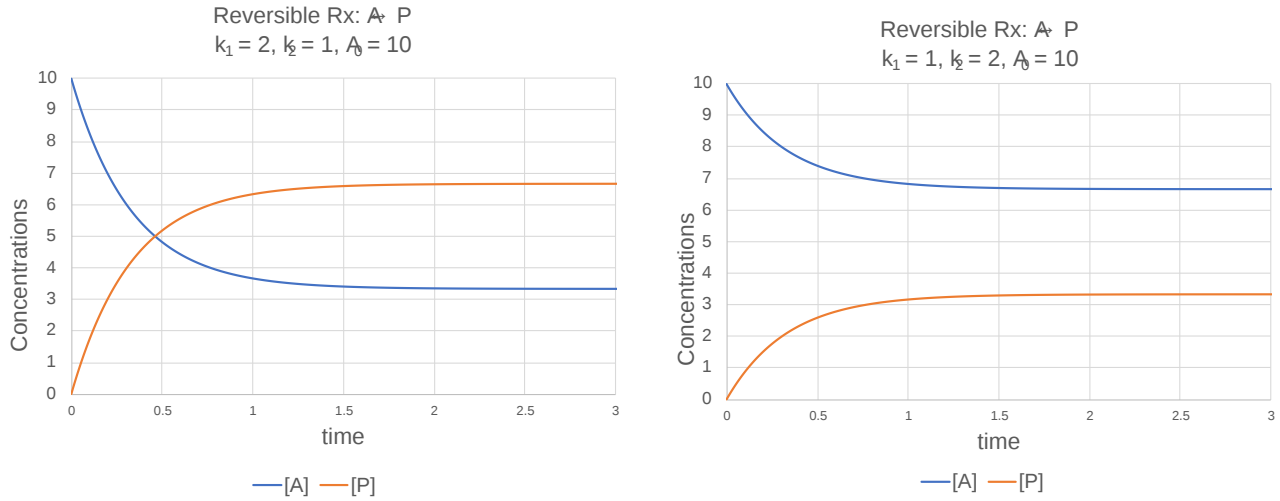
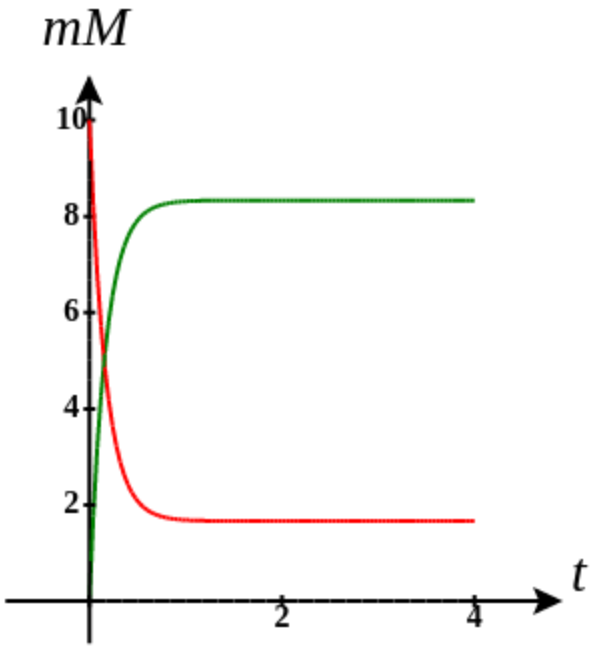


Figure 6.2.3: Graphs of A and P vs t for the reversible reaction $A \leftrightarrow P$ at two different sets of values of k_1 and k_2

Change the sliders on the interactive graph below of a reversible reaction to see how changing the relative values of the forward and reverse rate constants affects the concentrations at which the concentration plateaus are reached.



We all grew up on mathematical graphs that give you valuable insight into textual descriptions and data tables from which the graphs were made. These graphs are enhanced when you can use sliders to change constants as for the reversible reaction A ↔ P above. Even then, you might not infer that when the reaction has reached equilibrium, product is still being made from reactant, and reactant from product, since the equilibrium is dynamic. To add insight into simple and complex reactions, animations showing the continual disappearance of reactants and products are valuable.

This book will incorporate many animations to visually show the changes in the reactant and product concentrations. Hui Liu and Shraddha Nayak (Animation Lab, University of Utah) made all the animations in this book using mathematical solutions to the progress curves for the reactions. Multiple modes of presentation are useful as learners with different backgrounds and preferred ways of learning attempt to understand complex materials.

It is relatively simple to write the differential equations (differential) to show how the rate of disappearance of a reactant A (for example), dA/dt, depends on the concentration of its immediate participants in the reaction. It is not so easy to solve the equations (as we did above) for the progress curve, which shows how [A] changes with time t (i.e. [A] = f(t)). Luckily, many programs have been developed that produce numerical solutions to the differential equations and give progress curve graphs like [A] = f(t). Two interrelated, freely available programs, Copasi and Virtual cell (Vcell), can solve all the equations for hundreds of cellular reactions simultaneously. They use a format called **Systems Biology Markup Language** (**SBML**) for describing and storing computational models. We will use Vcell models in this book as they are straightforward to create. All the coding to describe the reactions is built into Vcell and this book and hidden from you. All you will see are the output results. You can change the progress curves by moving sliders to change constants and see the resulting changes in graphical outputs.

The VCell models use a reaction diagram that shows all of the interconnected species. The first Vcell model we will run is for the reversible conversion of A to P (A ↔P), which we just discussed and displayed in the graph above. Here is the Vcell reaction diagram and a description of its components.

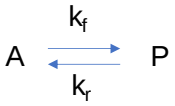
- The reactant A and product P are called species and are shown as green spheres.
- The yellow square indicates a reaction node connecting A to P.
- Lines connect the species that participate in the reaction. The arrows appear unidirectional, **BUT** the equations describing the concentrations of A and P are derived assuming a reversible reaction with rate constants k_f (forward reaction) and k_r (reverse reaction).

The program calculates A and P as a function of time (i.e. it solves the differential equations for both species). The graphs of concentration vs time are called progress curves. It also can calculate **fluxes (J)** (velocities) for each species. The flux at any given time is the slope of the concentration vs time curve at any given time. When we get to metabolism, we will talk about fluxes of metabolites through pathways. Also, fluxes are used to describe the rate of movement of solute through membranes. Here is the result of the simulation run in Vcell, exported as a sbml file, and displayed in the book using a program called MiniSideWinder.

VCell MODEL

Reversible reaction A ↔ P

Vcell reaction diagram (1-way arrows defined as reversible in actual mathematical model) and chemical equation



Initial parameter values: $k_f = 2$, $k_r = 4$

Select Load [model name] below

Load AtoP_Rev_Uncatalyzed

Select **Start** to begin the simulation.

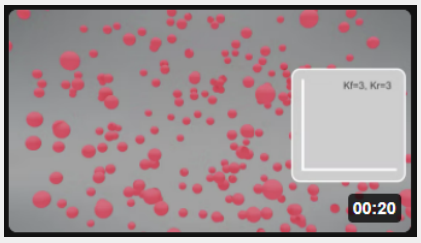
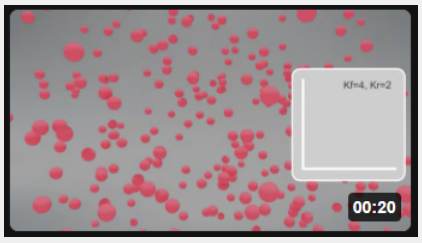
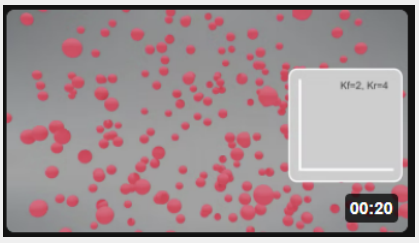
Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

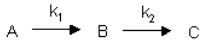
Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Animations

The video animations show particles representing A (red) and P (cyan) interconverting in a reversible process with embedded progress curves showing A (red) and P (cyan) vs. time.

Reversible Rx A (red) ↔ B (cyan): $k_F = k_R = 3$	Reversible Rx A (red) ↔ B (cyan): $k_F=4$, $k_R=2$	Reversible Rx A (red) ↔ B (cyan): $k_F=2$, $k_R=4$
		

6.2.2.2: Consecutive Irreversible First Order Reactions



The following differential equations can be written for these reactions:

$$\begin{aligned} \frac{dA}{dt} &= -k_1\ A \\ \frac{dB}{dt} &= k_1\ A - k_2\ B \\ \frac{dC}{dt} &= k_2\ B \end{aligned} \tag{6.2.13}$$

Here are the solutions to the differential equations:

$$\begin{aligned} A &= A_0 e^{-k_1 t} \\ B &= \frac{k_1 A_0}{k_2 - k_1}\left(e^{-k_1 t} - e^{-k_2 t}\right) \\ C &= A_0 - A - B = A_0\left[1 + \frac{1}{k_1 - k_2}\left(k_2 e^{-k_1 t} - k_1 e^{-k_2 t}\right]\right. \end{aligned} \tag{6.2.14}$$

Figure 6.2.4 shows graphs of A, B, and C vs. t for these reactions for a fixed value of k_1 and k_2.

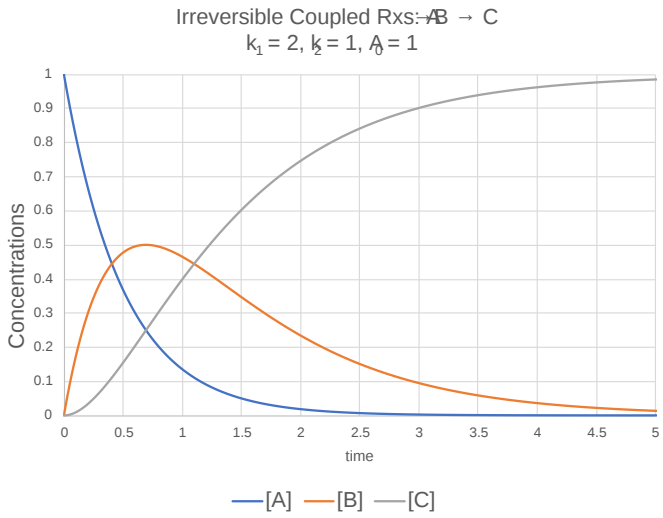
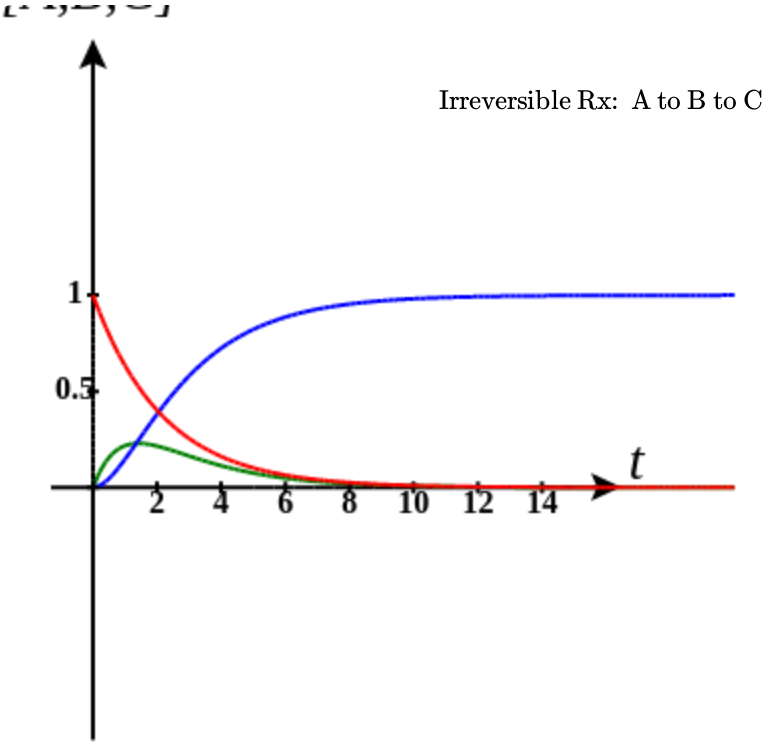
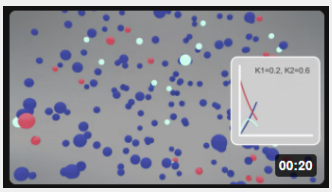
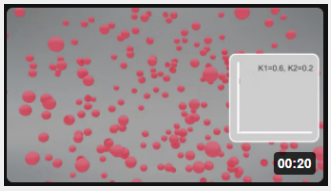


Figure 6.2.4: Graphs of A, B, and C vs. t for the irreversible reactions A → B → C for a fixed value of k_1 and k_2.

Change the sliders on the interactive graph below to see how changing the forward and reverse rate constants affect the curves.



Here are two different animations for the **irreversible** reaction using different rate constants.

Irreversible Rx A (red) → B (cyan) → C (blue) k_1 = 0.2, k_2 = 0.6	Irreversible Rx A (red) → B (cyan)→ C (blue) k_1= 0.6, k_2 = 0.2
	

6.2.2.3: Consecutive Reversible First Order Reactions

You can imagine that solving the equations for the completely reversible reactions of A ↔ B ↔ C would be very difficult. However, writing the differential equations for each step is straight-forward and can be done easily in Vcell by choosing the built-in equations for each separate reaction based on mass action. The program can then solve the equations numerically to produce progress curve graphs.

Now let's look at the simulation for the fully **reversible** reactions A ↔ B ↔ C. Again, the model was built and solved in VCell, and then exported in the system's biology markup language (sbml) format. The interactive graphs are made using a program called miniSideWinder.

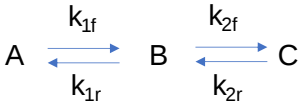
A note: Arrows in VCell Diagrams - In the reaction diagram for the reversible reaction A ↔ B ↔ C below, the arrows go in only one direction, left to right, and simply show that the species are connected. However, in the Vcell program, the equations for the

reversible reaction were used to produce the graphs below. To run the simulation of the irreversible reaction, the rate constants for the reverse reaction would be set to 0.

VCell MODEL

Reversible reaction A ↔ B ↔ C.

Vcell reaction diagram (1-way arrows defined as reversible in actual mathematical model)



Initial parameter values: k_{1f} = 0.2, k_{1r} = 0.1, k_{2f} = 0.6, k2r = 0.3 A_0 = 1

Select Load [model name] below

Load A to B to C_RevNoEnz

Select **Start** to begin the simulation.

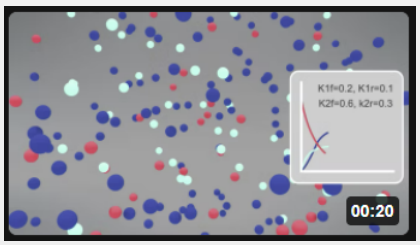
Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Here is the corresponding animation for the fully reversible reaction A ↔ B ↔ C.

Reversible Rx A (red) ↔ B (cyan) ↔ C Blue)
k_{1f} = 0.2, k_{1r} = 0.1; k_{2f} =0.6, k_{2r} = 0.3



6.2.3: Summary

This chapter introduces the fundamental principles of chemical kinetics using single-step and multi-step reactions as models. The focus is on understanding how reaction rates are described mathematically, both for simple, uncatalyzed reactions and for more complex reversible and consecutive processes. Key topics include:

1. **Calculus in Kinetics**
   The chapter begins by establishing the language of calculus for describing reaction rates. Students learn to interpret the derivative (dX/dt) as the instantaneous rate of change in concentration and to use integrals to derive expressions that relate concentration to time. This mathematical framework provides the basis for writing and solving differential equations that describe reaction progress.

2. **First-Order Reactions**
   Using the simple irreversible reaction A → P as a model, the integrated rate equation

   $[A]=[A]_0e^{-k_1t}$

   is derived and explained. Graphical representations, such as exponential decay plots of [A] versus time and linear plots of ln[A] versus time, are used to illustrate how the reaction rate decreases as the concentration of A diminishes.

3. **Second-Order and Pseudo First-Order Reactions**
   The chapter then extends the discussion to second-order reactions, exemplified by reactions involving two molecules (A + B → P) or a self-reaction (A + A → P). The corresponding integrated rate equation,

   $1[A]=1[A]_0+k_2t$

   is derived. Special cases, such as pseudo first-order kinetics (when one reactant is in large excess), are introduced to show how the complexity of bimolecular reactions can be simplified.

4. **Initial Rate vs. Integrated Rate Methods**
   Two approaches for kinetic analysis are compared. The initial rate method involves measuring the slope of the concentration versus time curve at the very beginning of the reaction, while progress curve analysis relies on integrated rate equations to describe the entire time course of the reaction. Each method provides unique insights into the reaction mechanism and rate constants.

5. **Multi-Step and Reversible Reactions**
   The concepts are further developed by considering reversible reactions (A ↔ P) and consecutive reactions (A → B → C). Differential equations for these systems are formulated and solved to yield progress curves that depict how reactant and product concentrations change over time. The dynamic nature of equilibrium in reversible reactions is emphasized, illustrating that even when concentrations plateau, the system remains dynamic.

6. **Computational Modeling and Visualization**
   Finally, the chapter highlights the role of computational tools such as VCell in solving complex kinetic equations. Interactive simulations and animations help visualize progress curves and the impact of varying rate constants on reaction dynamics, thereby linking mathematical models with experimental observations.

In summary, this chapter equips students with a strong quantitative foundation in reaction kinetics, bridging theoretical mathematical concepts with practical applications in chemical and enzymatic processes. Understanding these principles is essential for analyzing more complex biochemical reactions and for developing an intuition about how molecular processes unfold over time.

6.3: Kinetics with Enzymes

Learning Goals (ChatGPT o1, 1/30/25)

Below is a series of learning goals designed to help junior and senior biochemistry majors master the principles of enzyme-catalyzed reactions, kinetic derivations, and data analysis:

- **Understanding Enzyme Catalysis and Reaction Thermodynamics**
  - Explain how enzymes accelerate the rate at which equilibrium is achieved without altering the overall equilibrium constant (Keq) of a reaction.
  - Describe the concept that enzymes lower the activation energy by stabilizing bound transition states and, in doing so, alter the reaction mechanism.
- **Mechanistic Foundations of Enzyme-Catalyzed Reactions**
  - Analyze the basic enzyme-catalyzed reaction scheme, $E+S \leftrightarrow ES \rightarrow E+P$, and discuss the significance of each step in the pathway.
  - Understand the physical meaning of binding substrate to an enzyme and how this enables intramolecular conversion of substrate to product before release.
- **Rapid Equilibrium vs. Steady State Assumptions**
  - Differentiate between the rapid equilibrium and steady state assumptions in enzyme kinetics and identify conditions under which each is applicable.
  - Derive the expression for the enzyme-substrate complex concentration [ES][ES] using the rapid equilibrium assumption and mass balance for enzyme, leading to $[ES]=E_0[S]/(K_S+[S])$.
- **Derivation and Interpretation of the Michaelis-Menten Equation**
  - Derive the Michaelis-Menten equation ($v_0=V_{max}[S]/(K_M+[S])$ under both rapid equilibrium and steady state conditions, understanding the assumptions and limitations of each derivation.
  - Define and interpret the kinetic parameters K_M (Michaelis constant), V_{max} (maximum velocity), and k_{cat} (turnover number), including their units and physical significance.
- **Graphical Analysis and Data Interpretation**
  - Explain how initial rate (v_0) measurements are obtained from progress curves and how these data are used to generate hyperbolic vv vs. [S] plots.
  - Describe methods for linearizing Michaelis-Menten kinetics (e.g., Lineweaver-Burk, Eadie-Hofstee, and Scatchard plots) and discuss their advantages and potential pitfalls.
- **Integration of Reaction Progress Curves**
  - Differentiate between analyses based on initial velocities and those based on integrated rate equations (progress curves) for enzyme-catalyzed reactions.
  - Understand how integrated rate equations for enzyme kinetics can be derived and used to extract kinetic parameters from time-course data.
- **Modeling Reversible and Complex Enzyme-Catalyzed Reactions**
  - Analyze reversible enzyme-catalyzed reactions and understand how both forward and reverse reaction rates contribute to overall kinetics.
  - Describe kinetic mechanisms that involve covalent intermediates (such as acyl-enzyme complexes) and incorporate these into an extended Michaelis-Menten framework.
- **Application of Computational Tools in Kinetic Modeling**
  - Utilize computer modeling platforms (e.g., VCell) to simulate enzyme-catalyzed reactions and visualize progress curves under varying conditions.
  - Interpret simulation outputs to understand the effects of changing kinetic constants (e.g., k_1, k_2, k_3) on reaction profiles and the dynamic equilibrium between substrate, enzyme–substrate complex, and product.

By achieving these learning goals, students will develop a comprehensive understanding of the quantitative and mechanistic aspects of enzyme kinetics—from deriving fundamental equations to applying them in experimental and computational contexts.

An enzyme alters the pathways for converting a reactant to a product by binding to the reactant and facilitating the intramolecular conversion of bound substrate to bound product before it releases the product. Enzymes do not affect the thermodynamics of reactions. For reversible reactions (as an example), the equilibrium constant, $K_{eq,}$ is unchanged. What is charged is the rate at which equilibrium is achieved. Enzymes lower the activation energy for bound transition states and change the reaction mechanism.

Figure 6.3.1 shows the simplest chemical reaction that can be written to show how an enzyme catalyzes a reaction.

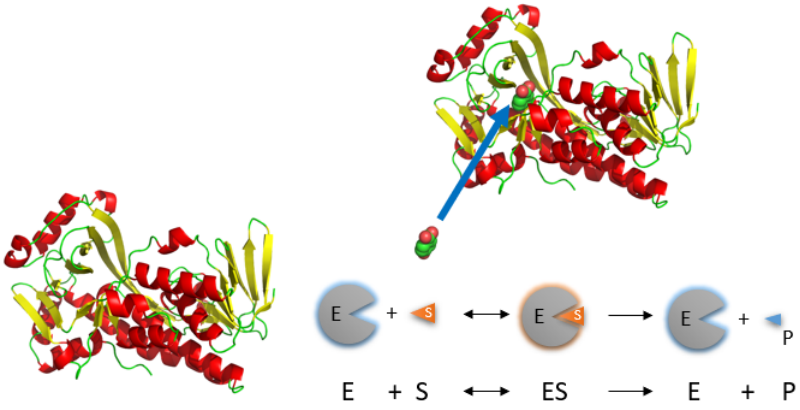


Figure 6.3.1: Chemical equation for a simple enzyme-catalyzed reaction

6.3.1: Rapid Equilibrium Enzyme-Catalyzed Reactions

We have previously derived equations for the reversible binding of a ligand to a macromolecule. Next, we derived equations for the receptor-mediated facilitated transport of a molecule through a semipermeable membrane. This latter case extended the former case by adding a physical transport step. Now, in what hopefully will seem like *deja vu*, we will derive almost identical equations for the chemical transformation of a ligand, commonly referred to as a substrate, into a product by an enzyme. We will study two scenarios based on two different assumptions, each enabling a straightforward mathematical derivation of kinetic equations:

1. **Rapid Equilibrium Assumption** - enzyme E (macromolecule) and substrate S (ligand) concentrations can be determined using the dissociation constant since E, S, and ES are in rapid equilibrium, as we previously used in deriving the equations for facilitated transport. Sorry about the switch from A to S in the designation of the substrate. Biochemists use S to represent the substrate (ligand) and A, B, P, and Q to represent reactants and products in the case of multi-substrate and multi-product reactions.
2. **Steady State Assumption** (more general) - enzyme and substrate concentrations are not those determined using the dissociation constant.

Enzyme kinetics experiments, as we will see in the following chapters, must be used to determine the detailed mechanism of the catalyzed reaction. Using kinetic analysis, you can determine the order of binding/dissociation of substrates and products, the rate constants for individual steps, and clues to the mechanism used by the enzyme in catalysis.

Consider the following reaction mechanism for the enzyme-catalyzed conversion of substrate S into product P. (We will assume that the catalyzed rate is much greater than the noncatalyzed rate.)

$$\mathrm{E} + \mathrm{S} \overset{K_s}{\rightleftharpoons} \mathrm{ES} \overset{k_3}{\longrightarrow} \mathrm{E} + \mathrm{P}$$

As we did for the derivation of the equations for the facilitated transport reactions under rapid equilibrium conditions, this derivation is based on the assumption that the relative concentrations of S, E, and ES can be determined by the dissociation constant, K_S, for the interactions and the concentrations of each species during the early part of the reaction (i.e., under initial rate conditions). Assume also the S >> E_0. Remember that under these conditions, S does not change much with time. Is this a valid

assumption? Examine the mechanism shown above. S binds to E with a second-order rate constant k_1. ES has two fates. It can dissociate with a first-order rate constant k_2 to S + E, or it can be converted to the product with a first-order rate constant of k_3 to give P + E. If we assume that $k_2 >> k_3$ (i.e. that the complex falls apart much more quickly than S is converted to P), then the relative ratios of S, E, and ES can be described by Ks. Alternatively, you can think about it this way. If S binds to E, most of S will dissociate, and a small amount will be converted to P. If it does, then E is now free and will quickly bind S and reequilibrate since the most likely fate of bound S is to dissociate, not be converted to P (since $k_3 \ll k_2$). This also makes sense if you consider that the physical step, characterized by k_2, is likely to be quicker than the chemical step, characterized by k_3. Hence the following assumptions have been used:

- $S \gg E_0$
- $P_0 = 0$
- k_3 is rate limiting (i.e. the slow step)

We will derive equations showing the initial velocity v as a function of the initial substrate concentration, S_0, assuming that P is negligible over the time period used to measure the initial velocity. Also assume that $v_{catalyzed} \gg v_{noncatalyzed}$. In contrast to the first-order reaction of S to P in the absence of E, v is not proportional to S_0 but to S_{bound}. Therefore, $v \propto [ES]$, or

$$v = const[ES] = k_3[ES] \tag{6.3.1}$$

where v is the velocity (i.e., reaction rate).

Now, let's get ES from the dissociation constant K_S (assuming rapid equilibrium of E, S, and ES) and mass balance for E (E_0 = E + ES , so E = E_0 - ES). We will use mass balance for S when we derive the equation of the steady state. In many cases, S is approximately equal to S_0 or the total amount of substrate. This makes sense if you consider that the enzyme is a catalyst that acts repeatedly to produce the product, so you don't need significant amounts of the enzyme. The resulting equation of [ES] is shown below.

$$(ES) = \frac{(E_0)\,(S)}{K_S + (S)} \tag{6.3.2}$$

? A derivation of [ES] under rapid equilibrium conditions

Here it is!

Derivation

$$\begin{aligned} K_S = \frac{[E][S]}{[ES]} &= \left(\frac{([E_0] - [ES])\,[S])}{[ES]} \right. \\ (ES)K_S &= (E_0)\,(S) - (ES)(S) \\ (ES)K_S + (ES)(S) &= (E_0)\,(S) \\ (ES)\,(K_S + (S)) &= (E_0)\,(S) \end{aligned} \tag{6.3.3}$$

This derivation assumes that we know S (equal to S_0) and E_{tot} (which is E_0).

Let us assume that S is much greater than E, as is the likely biological case. We can calculate ES using the following equations and the same procedure we used for the derivation of the binding equation, which gives the equation below:

$$[ES] = \frac{[E_0]\,[S]}{K_S + S} \tag{6.3.4}$$

which is analogous to

$$[ML] = \frac{[M_0]\,[S]}{K_D + L} \tag{6.3.5}$$

This leads to

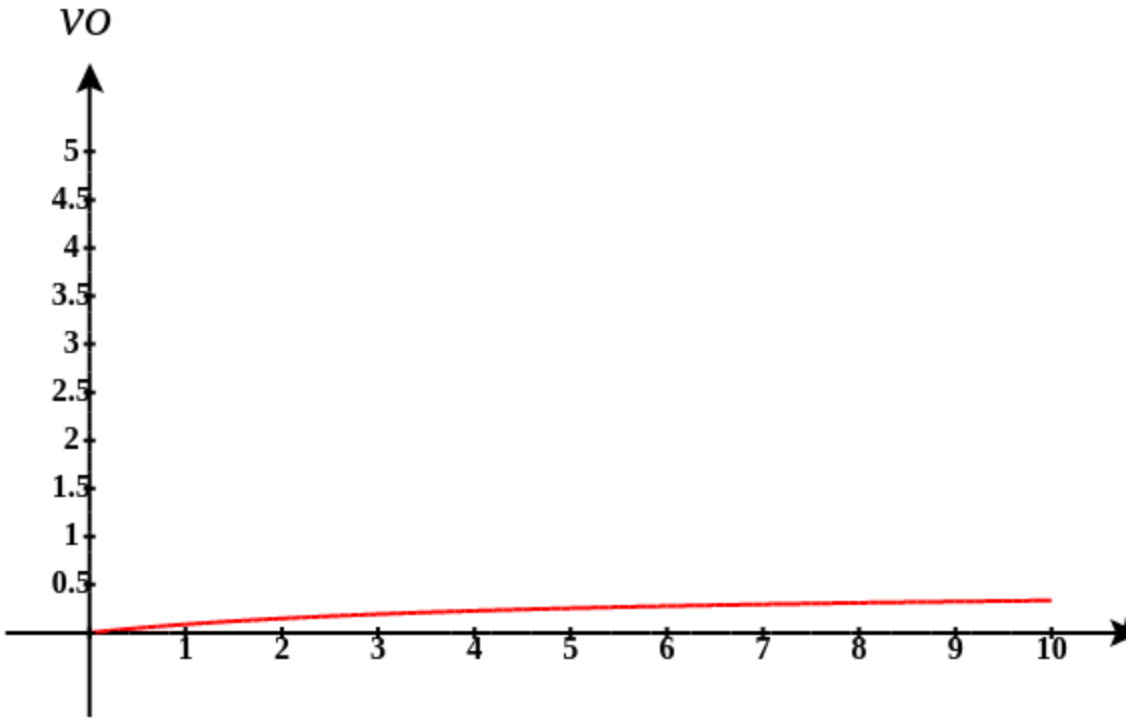
$$v_0 = \frac{(k_3)\,(E_0)\,(S)}{K_S + (S)} = \frac{(V_M)\,(S)}{K_S + (S)} \tag{6.3.6}$$

where

$$V_M = k_3 E_0 \tag{6.3.7}$$

This is the world-famous Henri-Michaelis-Menten Equation. It is a hyperbola just like the graph for binding of a ligand to a macromolecule with a given dissociation constant, K_D.

Move the sliders in the interactive graph below to see how changing Km and Vm alters the graph. Note that this graph is identical to the graph for M + L ↔ ML.



Just as in the case with noncatalyzed first-order decay, it is easiest to measure the initial velocity of the reaction when [S] does not change much with time, and the velocity is constant (i.e., the slope of the dP/dt curve is constant). A plot of [P] vs t (called a **progress curve**) is made for each studied substrate concentration. From these curves, the initial rates at each [S] is determined.

Alternatively, one reaction time that gives a linear rise in [P] with time is determined for all the different substrate concentrations. At that specified time, the reaction can be stopped (quenched) with a reagent that does not cause any change in S or P. Then, initial rates can be easily calculated for each [S] from a single data point.

Under these conditions:

- a plot of v vs. S is hyperbolic
- v = 0 when S = 0
- v is a linear function of S when S<<Ks.
- v = V_{max} (or V_M) when S is much greater than Ks
- S = K_S when v = V_M/2.

These are the same conditions we detailed for our understanding of the binding equation

$$(ML) = \frac{(M_0)\,L}{K_D + L} \tag{6.3.8}$$

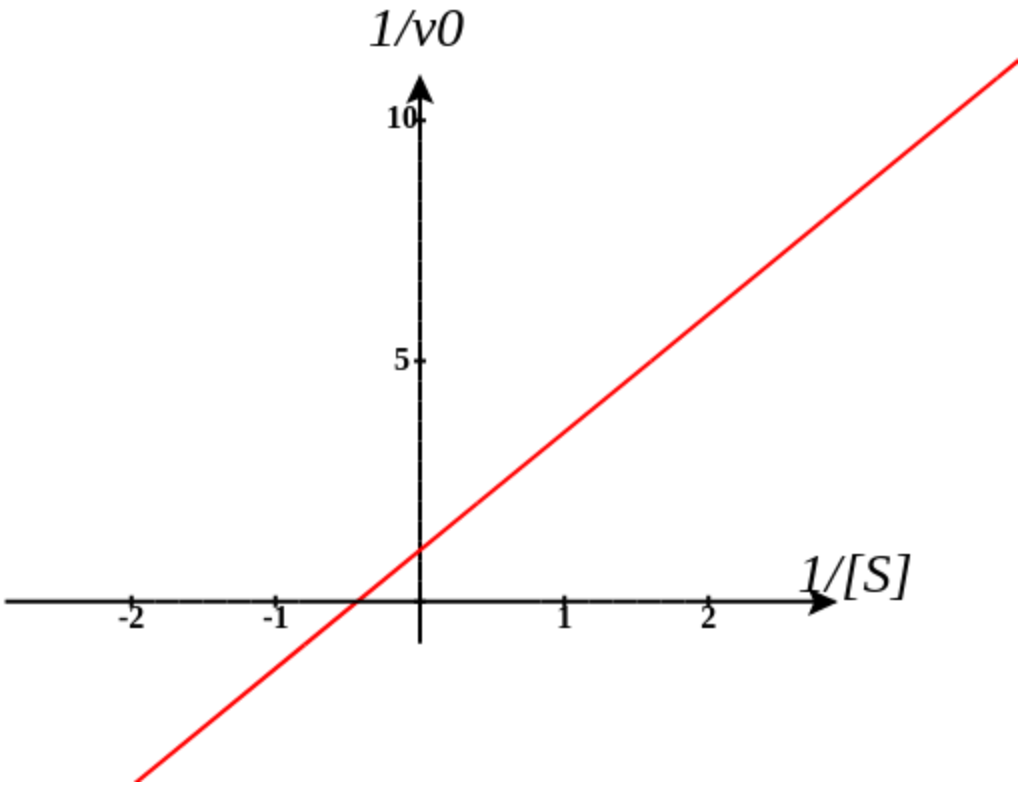
Note that when S is not >> K_S, the graph does not reach saturation or look hyperbolic. It should be apparent from the graph that only if S >> K_S (or when S is approximately 100 x > K_s) will saturation be achieved.

The K_S constant is usually called the **Michaelis constant, K_M**. We will see in a bit that the K_M for most enzyme-catalyzed reactions is **not** equal to the dissociation constant for ES, which we call K_S.

These graphs are often transformed into **double reciprocal or Lineweaver-Burk** plots, as shown below.

$$\frac{1}{v_0} = \frac{K_M + S}{V_M S} = \left(\frac{K_M}{V_M}\right)\frac{1}{S} + \frac{1}{V_M} \tag{6.3.9}$$

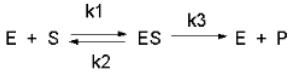
These plots are used to estimate V_M from the 1/v intercept ($1/V_M$) and K_M from the 1/S axis ($-1/K_M$). These values should be used as "seed" values for a nonlinear fit to the hyperbola that models the actual v vs. S curve. An interactive graph of $1/v_0$ vs 1/[S] is shown below. Change the sliders for K_M and V_M and note the change in the slope and intercepts of the plot



Help

As we saw in the graph of A or P vs t for a noncatalyzed, first-order reaction, the velocity of the reaction, given as the slope of those curves, is always changing. Which velocity should we use in Equation 5? The answer invariably is the initial velocity, v_0, measured in the early part of the reaction when little substrate is depleted. Hence, v vs S curves for enzyme-catalyzed reactions invariably are really v_0 vs [S] curves.

6.3.2: Steady State Enzyme-Catalyzed Reactions



In this derivation, we will consider the following equations and all the rate constants and will not arbitrarily assume that $k_2 >> k_3$. We will still assume that S >> E_0 and that P_0 = 0. An added assumption, however, is that d[ES]/dt is approximately 0. Look at this assumption this way. When an excess of S is added to E, ES is formed. In the rapid equilibrium assumption, we assumed that it would fall back to E + S (a physical step) faster than it would go onto product (a chemical step). In the steady state case, we will assume that ES might go on to product either less or more quickly than it will fall back to E + S. In either case, a steady state concentration of ES arises within a few milliseconds, and its concentration does not change significantly during the initial part of the reaction under which the initial rates are measured. Therefore, d[ES]/dt is about 0. For the rapid equilibrium derivation, v = k_3[ES]. We then solved for ES using K_S and mass balance of E. In the steady state assumption, the equation v= k_3[ES] still holds, but now we will solve for [ES] using the steady state assumption that d[ES]/dt =0.

$$\frac{d[ES]}{\mathrm{dt}} = \mathrm{k_1}[E][S] - \mathrm{k_2}[ES] - \mathrm{k_3}[ES] = 0 \tag{6.3.10}$$

We can solve this and obtain the Michaelis-Menten equation for reaction.

$$\mathrm{v} = \mathrm{k_3}[\mathrm{ES}] = \frac{\mathrm{k_3}\,[\mathrm{E_0}]\,[\mathrm{S}]}{\frac{\mathrm{k_2}+\mathrm{k_3}}{\mathrm{k_1}} + \mathrm{S}} = \frac{\mathrm{V_M}[\mathrm{S}]}{\mathrm{K_M} + \mathrm{S}} \tag{6.3.11}$$

? Derivation of the Michaelis-Menten Equation for the steady state

To see how to derive this, click below.

Answer

Applying mass balance for E (i.e E = E_0-ES) in the appropriate term below gives

$$\begin{aligned}
&\mathrm{k_1}[E][S] = (\mathrm{k_2} + \mathrm{k_3})\,[ES] \\
&\mathrm{k_1}\,[\mathrm{E_0} - \mathrm{ES}]\,[S] = (\mathrm{k_2} + \mathrm{k_3})\,[ES] \\
&\mathrm{k_1}\,[\mathrm{E_0}]\,[S] - \mathrm{k_1}[\mathrm{ES}][S] = (\mathrm{k_2} + \mathrm{k_3})\,[ES] \\
&\mathrm{k_1}\,[\mathrm{E_0}]\,[S] = (\mathrm{k_2} + \mathrm{k_3})\,[ES] + \mathrm{k_1}[\mathrm{ES}][S] \\
&\mathrm{k_1}\,[\mathrm{E_0}]\,[S] = [ES]\,(\mathrm{k_2} + \mathrm{k_3} + \mathrm{k_1 S})
\end{aligned} \tag{6.3.12}$$

Solving for [ES] gives

$$[\mathrm{ES}] = \left(\frac{[\mathrm{E_0}]\,[\mathrm{S}]}{\frac{\mathrm{k_2}+\mathrm{k_3}}{\mathrm{k_1}} + \mathrm{S}} \right) \tag{6.3.13}$$

Substituting into the Henri Michaelis Menten equation gives

$$\mathrm{v} = \mathrm{k_3}[\mathrm{ES}] = \frac{\mathrm{k_3}\,[\mathrm{E_0}]\,[\mathrm{S}]}{\frac{\mathrm{k_2}+\mathrm{k_3}}{\mathrm{k_1}} + \mathrm{S}} = \frac{\mathrm{V_M}[\mathrm{S}]}{\mathrm{K_M} + \mathrm{S}} \tag{6.3.14}$$

Note that

$$V_M = k_3 E_0 \tag{6.3.15}$$

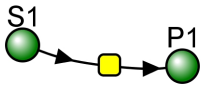
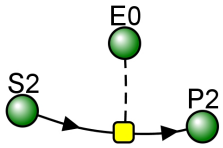
and

$$K_M = \frac{k_2 + k_3}{k_1} \tag{6.3.16}$$

Now let's look at a progress curve simulation that compares the Michaelis-Menten equation derived from rapid equilibrium assumptions (when k_2>>k_3) and in which K_M is the actual dissociation constant to the steady state approximation when k_2 is not >>k_3. Reaction 1 is the rapid equilibrium, and Reaction 2 is the steady state.

VCell MODEL

A comparison of the rapid equilibrium (left) and steady state (right)Michaels-Menten reaction

	
$v_0 = \frac{k_3\,[E_O]\,[S]}{K_M + S} = \frac{V_M[S]}{K_M + S}$ (6.3.17)	$v_0 = \frac{k_3\,[E_O]\,[S]}{\frac{k_2+k_3}{k_1} + S} = \frac{V_M[S]}{K_M + S}$ (6.3.18)
Initial Condition: V_M = 10 ; K_M = 10; S = 50	Initial Conditions: k_1 = 10 ; k_2 = 90 ; k_3 = 10 ; E_0 = 1 uM; S = 50 $V_M = k_3E_0$ =10; $K_M = (k_2 + k_3)/k_2$ = 10)

Select Load [model name] below

Load MMvsSteadyState

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

The initial conditions in the graph are set so the graphs of the rapid equilibrium and steady state are identical.

Recommendations:

- Scale the y-axis to 50
- change the slider k_{2_r2} for the steady state graph to other values. Watch the curves separate. Although these plots are only for one substrate concentration (50 uM), the effects of changing k_2 for the steady state plot are very dramatic. This should convince you that in general, unless $k_3 << k_2$, the calculated value of K_M is **not** equal to the thermodynamic dissociation constant, K_D.

6.3.3: Analysis of the General Michaelis-Menten Equation

This equation can be simplified and studied under different conditions. First, notice that $(k_2 + k_3)/k_1$ is a constant, which is a function of relevant rate constants. This term is usually replaced by K_M, which is called the **Michaelis constant**. Likewise, when S approaches infinity (i.e. S >> K_M, equation 5 becomes v = $k_3(E_0)$ which is also a constant called V_M for maximal velocity. Substituting V_M and K_M into equation 5 gives the simplified equation:

$$\mathrm{v} = \frac{\mathrm{V_M[S]}}{\mathrm{K_M + S}} \tag{6.3.19}$$

It is extremely important to note that K_M in the general equation does **n**ot equal the K_S, the dissociation constant used in the rapid equilibrium assumption! K_M and K_S have the same units of molarity, however. A closer examination of K_M shows that under the limiting case when k_2 >> k_3 (the rapid equilibrium assumption) then,

$$\mathrm{K_M} = \frac{\mathrm{k_2 + k_3}}{\mathrm{k_1}} = \frac{\mathrm{k_2}}{\mathrm{k_1}} = \mathrm{K_D} = \mathrm{K_S} \tag{6.3.20}$$

If we examine these equations under several different scenarios, we can better understand the equation and the kinetic parameters:

- when S = 0, v = 0.

- when S >> K_M, v = V_M = k_3E_0. (i.e. v is zero order with respect to S and first order in E. Remember, k_3 has units of $s^{-1>}$ since it is a first-order rate constant. k_3 is often called the **turnover number,** because it describes how many molecules of S "turn over" to product per second.
- v = V_M2, when S = K_M.
- when S << K_M, v = V_MS/K_M = k_3E_0S/K_M (i.e. the reaction is bimolecular, dependent on both on S and E. k3/K_M has units of $M^{-1}s^{-1}$, the same as a second order rate constant.

6.3.4: More Complicated Enzyme-catalyzed Reactions

6.3.4.1: A reversibly-catalyzed reaction

You have learned previously that enzymes don't change the equilibrium constant for a reaction but rather decrease the activation energy barrier to move from reactants to products. This implies that the activation energy to move in the reverse direction, from products to reactants, is also lowered. Hence, the enzyme speeds up both the forward and reverse reaction. We haven't accounted for that yet in our kinetic equations. Many reactions in metabolic reactions that do not have large negative values of ΔG (ie. they are not significantly favored) are reversible, allowing the enzyme to be used in the reverse direction. Take, for example, the pathway to break down glucose to pyruvate (glycolysis). It has nine steps, of which five are reversible, allowing them to be used in the reverse pathway to take pyruvate to glucose (gluconeogenesis).

Let's set up the equations for the reversible reaction of substrate S to product P catalyzed by enzyme E. Assume that the K_M for the forward reaction is K_{MS} (or K_S) and for the reverse reaction is K_{MP} (or K_P) as shown in the reaction scheme in Figure 6.3.2. The rate constant k_2 is the k_{cat} (forward rate constant) for conversion of ES to EP and k_{-2} is the k_{cat} (reverse rate constant) for conversion of EP to ES

$$\mathrm{E + S \underset{K_S}{\overset{K_{MS}}{\leftrightarrow}} ES \underset{k_{-2}}{\overset{k_2}{\leftrightarrow}} EP \underset{K_P}{\overset{K_{MP}}{\leftrightarrow}} E + P}$$

Figure 6.3.2: Reaction scheme for a simple enzyme-catalyzed reaction

The following simple Michaelis-Menten equations can be written for just the forward reaction and for the reverse reaction:

$$\begin{aligned} v_f &= \frac{V_f S}{K_{MS} + S} = \frac{\frac{V_f S}{K_{MS}}}{1 + \frac{S}{K_{MS}}} \\ v_r &= \frac{V_r P}{K_{MP} + P} = \frac{\frac{V_r P}{K_{MP}}}{1 + \frac{P}{K_{MP}}} \end{aligned} \tag{6.3.21}$$

You might think that simply subtracting the two would give the net velocity in the forward direction, but that is **NOT** the case.

$$v \neq \left[\frac{\frac{V_f S}{K_{MS}}}{1 + \frac{S}{K_{MS}}} - \frac{\frac{V_r P}{K_{MP}}}{1 + \frac{P}{K_{MP}}} \right] \tag{6.3.22}$$

The reason is that in the derivation, the equations for both the forward and reverse rates must have terms for the reverse and forward reactions, respectively.

A simple derivation shows that this is the equation for the reversible conversion of substrate to product.

$$v = k_2[ES] - k_{-2}[EP] = \frac{V_f \frac{[S]}{K_S}}{\left[1 + \frac{[S]}{K_S} + \frac{[P]}{K_p}\right]} - \frac{V_r \frac{[P]}{K_P}}{\left[1 + \frac{[S]}{K_S} + \frac{[P]}{K_P}\right]} = \frac{V_f \frac{[S]}{K_S} - V_r \frac{[P]}{K_P}}{\left[1 + \frac{[S]}{K_S} + \frac{[P]}{K_P}\right]} \tag{6.3.23}$$

This reversible form of the Michaelis-Menten equation and other equations, written in the format shown in equation 6.24, are commonly used in programs such as VCell and Copasi to model the kinetics of whole pathways of biological interactions and reactions. The figures below show a reaction diagram, graphical results showing S and P vs time for the selected K_M and V_M values, and animations for the reaction. The chemicals (S and P) are shown as green spheres connected by a line. The red dot again

represents the enzyme (shown as a node of connection between S and P). Equation 6.24 was used to model the reversible reaction (even though the arrows shown between S and P are unidirectional.

VCell MODEL

Reversible Enzyme-Catalyzed Reaction: E + S ↔ ES ↔ EP ↔ E + P .

Vcell reaction diagram (1-way arrows defined as reversible in the actual mathematical model) and chemical equation

J	reaction rate	$\frac{\left(\frac{S \cdot VmaxFwd}{KmFwd} - \frac{P \cdot VmaxRev}{KmRev}\right)}{\left(1.0 + \frac{S}{KmFwd} + \frac{P}{KmRev}\right)}$
KmFwd	Km forward	10.0
VmaxFwd	max forward rate	10.0
KmRev	Km reverese	20.0
VmaxRev	max reverse rate	5.0
Size	size	1.0
initConc	initial concentration for S	5.0
initConc	initial concentration for P	0.0

Initial parameter values: as shown in above

Select Load [Enz Rev] below

Load Enz Rev

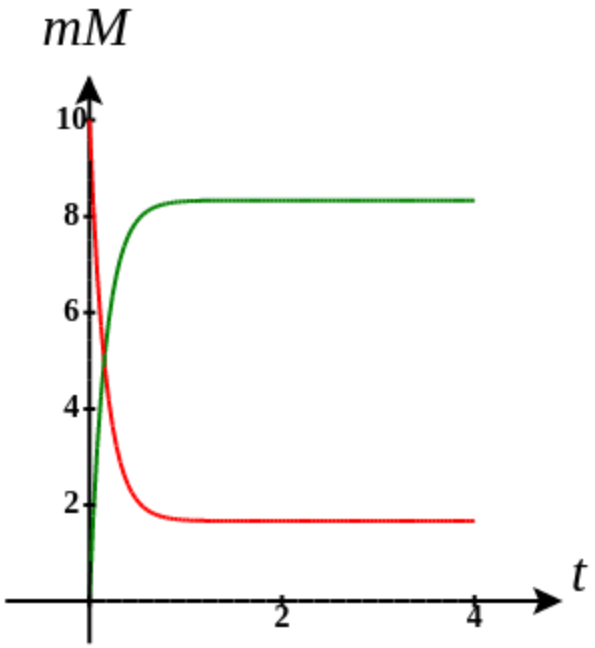
Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

If you reflect on it, this reaction is very similar to the reversible reaction of A ↔ P in the **absence** of an enzyme, which we explored in Chapter section 6.2. Just for comparison, the graph for that reaction is shown below. Change the sliders to produce a curve similar to the enzyme-catalyzed reaction shown in the figure above.



6.3.4.2: Reaction with intermediates

Not all reactions can be characterized simply as a simple substrate interacting with an enzyme to form an ES complex, which then turns over to form the product. Sometimes, intermediates form. For example, a substrate S might interact with E to form a complex, which then is cleaved to products P and Q. Q is released from the enzyme, but P might stay covalently attached. This often happens in the hydrolytic cleavage of a peptide bond by a protease when an activated nucleophile like Ser reacts with the sessile peptide bond in a nucleophilic substitution reaction, releasing the amine end of the former peptide bond as the leaving group. The carboxy end of the peptide bond remains bonded to the Ser as a Ser-acyl intermediate. Water then enters and cleaves the acyl intermediate, freeing the carboxyl end of the original peptide bond. This is shown in the written reaction in Figure 6.3.3:

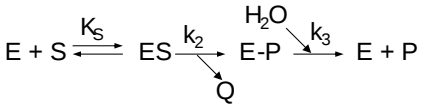


Figure 6.3.3: Reaction scheme for enzyme-catalyzed reaction with an intermediate

Even for this seemingly complicated reaction, you get the standard Michaelis-Menten equation.

To simplify the derivation of the kinetic equation, let's assume that E, S, and ES are in rapid equilibrium defined by the dissociation constant, Ks. Assume Q has a visible absorbance, so it is easy to monitor. Assume from the steady state assumption that:

$$\frac{d[E-P]}{dt} = k_2[ES] - k_3[E-P] = 0 \tag{6.3.24}$$

assuming that k_3 is a pseudo first-order rate constant and that $[H_2O]$ doesn't change.

The velocity depends on which step is rate-limiting. If $k_3 << k_2$, then the k_3 step is rate-limiting. Then

$$v = k_3[E-P] \tag{6.3.25}$$

If $k_2 << k_3$, then the k_2 step is rate-limiting. Then

$$v = k_2[ES] \tag{6.3.26}$$

The following kinetic equation for this reaction can be derived, assuming $v = k_2[ES]$.

$$v = \frac{\frac{k_2 k_3}{k_2+k_3}[E_0][S]}{\left(\frac{k_3}{k_2+k_3}\right)K_S + S} = \frac{k_{cat}[E_0][S]}{K_M + S} = \frac{V_M S}{K_M + S} \tag{6.3.27}$$

You can verify that you get the same equation if you assume that $v = k_3[E\text{-}P]$.

This equation looks quite complicated, especially if you substitute for Ks, k_{-1}/k_1. All the kinetic constants can be expressed as functions of the individual rate constants. However, this equation can be simplified by realizing the following:

- When $S >> \frac{K_S k_3}{k_2+k_3}$, $v = \left(\frac{k_2 k_3}{k_2+k_3}\right)E_0 = V_M$
- $\frac{K_S k_3}{k_2+k_3} = \text{constant} = K_M$

Substituting these into equation 7 gives:

$$v = \frac{V_M S}{K_M + S}$$

This, again is the general form of the Michaelis-Menten equation

The expression for V_M in the first bulleted expression above is more complicated than our earlier definition of $V_M = k_3E_0$. They are similar in that the term E_0 is multiplied by a constant which is itself a function of rate constant(s). The rate constants are generally lumped together into a generic constant called k_{cat}.

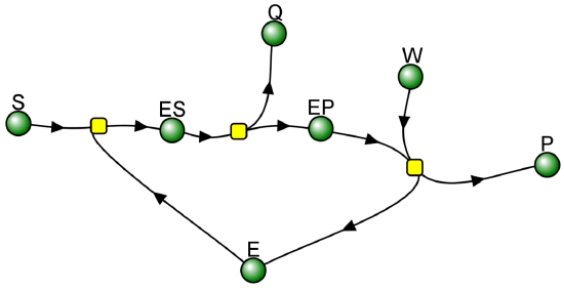
- For the simple reaction $\mathbf{k_{cat} = k3}$
- For the more complicated reaction above with a covalent intermediate, $k_{cat} = \frac{k_2 k_3}{k_2+k_3}$
- For all reactions, $V_M = k_{cat}E_0$.

Recent Updates (2/7/24): Here is a VCell computational model for a bisubsrate and biproduct (BiBi) reaction with a covalent intermediate (ping pong reaction).

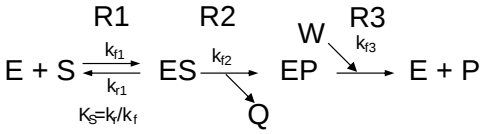
VCell MODEL

BiBi-Ping Pong_Covalent Intermediate Irreversible reaction

Vcell reaction diagram (1-way arrows defined as reversible in actual mathematical model) and chemical equation



Yellow dots: Reaction Nodes (R1, R2 and R3 left to right)



Reaction made irreversible since $k_{r2} = 0$, $k_{r3} = 0$.

Initial parameter values:

- S_0 = 100, E_0 = 1, W (water in a hydrolysis reaction) = 50 and fixed throughout
- k_{1f} = 5, k_{1r} = 1, k_{2f} = 0.6,
- k_{2f} = 50, k_{2r} = 0
- k_{3f} = 0.05, k_{3f} = 0

Select Load [model name] below

Load BiBiPingPongChymo

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

To see the burst phase for reaction, change the time and parameters to these values:

- set Run time to 0.3
- Select **Plot** then Update Y axis max to 2
- Click Edit Plot Species and check just P and Q
- reset

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

When k_2 >> k_3, Q forms quickly in a **burst phase** due to a high k_2. This phase is followed by the slower conversion of the accumulated EP to E + P in a **steady state phase**, as shown below in Figure 6.3.4. Make sure to change the parameters in the model above to see the burst phase for the production of product Q. This reaction is an example of a BiBi (two substrates, two products) Ping Pong reaction, since one reactant binds, followed by one product departing before the second react binds and the second product departs.

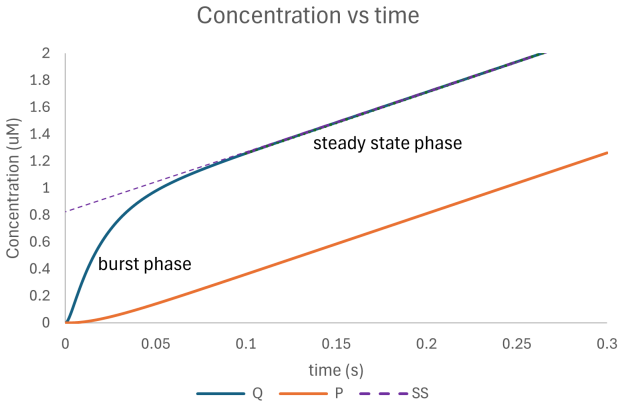


Figure 6.3.4: Burst and steady-state phase in an enzyme-catalyzed reaction with a covalent intermediate and k_2 >> k_3.

Reverse the rate constants in the model to k_2 = 5 and k_3 = 0.5 so that k_2 is **not** >> k_3, and note that the burst phase for Q disappears!

When plotting v_0 vs S Michaelis-Menten plots of this type of reaction, the state state v_0 should be used and **NOT** $(v_0)_{t=0}$ = $(dQ/dt)_{t=0}$, since that rate changes very quickly after the initial burst formation of Q.

Summary to this point

Figure 6.3.5 compares the Michaelis-Menten kinetic equations for the rapid equilibrium, steady state assumptions, and covalent intermediate cases.

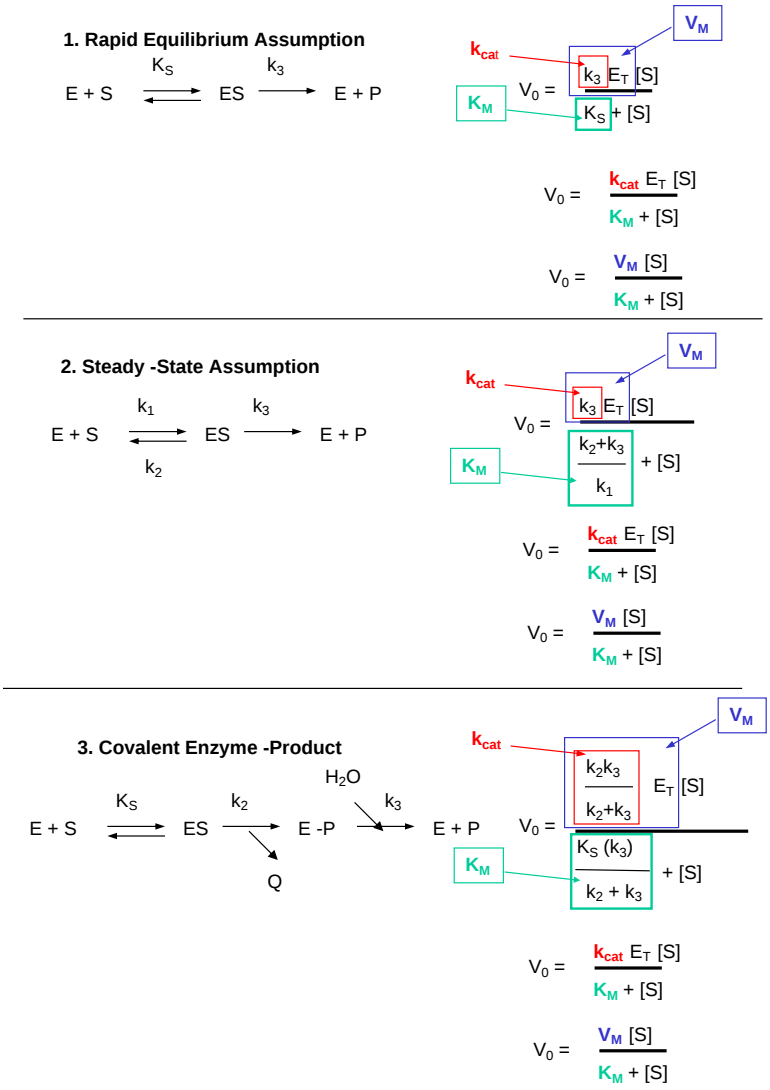


Figure 6.3.5: Michaelis-Menten kinetic equations for the rapid equilibrium, steady state assumptions, and covalent intermediate cases

6.3.5: Meaning of Kinetic Constants

Getting a "gut-level" understanding of the significance of the rate constants is important. Here they are:

- K_M: The Michaelis constant, with units of molarity (M), is operationally defined as the substrate concentration at which the initial velocity is half of V_M. It is equal to the dissociation constant of E and S only if if E, S and ES are in rapid equilibrium. It can be thought of as an "**effective**" (but not actual) K_D in other cases.
- k_{cat}: The catalytic rate constants, with units of s^{-1} is often called the **turnover number**. It is a measure of how many bound substrate molecules "turnover" or form product in 1 second. This is evident from equation $v_0 = k_{cat}[ES]$
- k_{cat}/K_M: Under condition when [S] << K_M, the Michaelis-Menten equation becomes $v_0 = (k_{cat}/K_M)[E_0][S]$. This really describes a biomolecular rate constant (k_{cat}/K_M), with units of $M^{-1}s^{-1}$, for conversion of free substrate to product. Some enzymes have k_{cat}/Km values around 10^8, indicating they are diffusion-controlled. That implies that the reaction is essentially done as soon as the enzyme and substrate collide. The constant k_{cat}/Km is also referred to as the **specificity constant** in that it describes how well an enzyme can differentiate between two different and competing substrates. (We will show this mathematically in the next chapter.)

Table 6.3.1 below shows K_M and k_{cat} values for various enzymes

K_M values		
enzyme	substrate	Km (mM)
catalase	H_2O_2	25
hexokinase (brain)	ATP	0.4
	D-Glucose	0.05
	D-Fructose	1.5
carbonic anhydrase	HCO_3^-	9
chymotrypsin	glycyltyrosinylglycine	108
	N-benzoyltyrosinamide	2.5
b-galactosidase	D-lactose	4.0
threonine dehydratase	L-Thr	5.0
k_{cat} values		
enzyme	substrate	k_{cat} (s^{-1})
catalase	H_2O_2	40,000,000
carbonic anhydrase	HCO_3^-	400,000
acetylcholinesterase	acetylcholine	140,000
b-lactamase	benzylpenicillin	2,000
fumarase	fumarate	800
RecA protein (ATPase)	ATP	0.4

Table 6.3.1: K_M and k_{cat} values for various enzymes

Table 6.3.2 below shows k_{cat}, K_M, and k_{cat}/K_M values for diffusion-controlled enzymes

Enzymes with k_{cat}/K_M values close to diffusion controlled (10^8 - 10^9 $M^{-1}s^{-1}$)					
enzyme	substrate		kcat (s^{-1})	Km (M)	kcat/Km ($M^{-1}s^{-1}$)
acetylcholinesterase	acetylcholine		1.4×10^4	9×10^{-5}	1.6×10^8
carbonic anhydrase	CO_2		1×10^6	1.2×10^{-2}	8.3×10^7
	HCO_3^-		4×10^5	2.6 x 10-2	1.5×10^7
catalase	H_2O_2		4×10^7	1.1	4×10^7
crotonase	crotonyl-CoA		5.7×10^3	2×10^{-5}	2.8×10^8
fumarase	fumarate		8×10^2	5×10^{-6}	1.6×10^8
	malate		9×10^2	2.5×10^{-5}	3.6×10^7
triose phosphate isomerase	glyceraldehyde-3-P		4.3×10^3	4.7×10^{-4}	2.4×10^8
b-lactamase	benzylpenicillin		2.0×10^3	2×10^{-4}	1×10^8

Table 6.3.2 below show k_{cat}, K_M and k_{cat}/K_M values for diffusion controlled enzymes

6.3.6: Experimental Determination of V_M and K_M

How can V_M and K_M be determined from experimental data?

6.3.6.1: From the initial rate data

The most common way to determine V_M and K_M is through initial rates, v_0, obtained from P or S vs time curves. Hyperbolic graphs of v_0 vs [S] can be fitted or transformed as we explored the different mathematical transformations of the hyperbolic binding equation to determine K_D. These included:

- Michaelis-Menten plot: nonlinear hyperbolic fit
- Lineweaver-Burk double reciprocal plot
- Scatchard plot
- Eadie-Hofstee plot

We discussed all of these plots, except for the **Eadie-Hofstee** plot, in the chapter on binding. The Eadie-Hofstee plot is another linearized version of the Michaelis-Menten equation

Here is a derivation of that equation, which starts with each side of the double-reciprocal plot being multiplied by v_0V_M.

$$\begin{aligned} (v_0 V_M)\frac{1}{v_0} &= (v_0 V_M)\left(\frac{K_M}{V_M}\right)\frac{1}{S} + (v_0 V_M)\frac{1}{V_M} \\ V_M &= (v_0)\,K_M\frac{1}{S} + (v_0) \\ v_0 &= -K_M\left(\frac{v_0}{S}\right) + V_M \end{aligned} \tag{6.3.28}$$

Note that a graph of v_0 vs v_0/S is linear, so slopes and intercepts can be used to obtain values for V_M and K_M.

The double-reciprocal plot is commonly used to analyze initial velocity vs substrate concentration data. When used for such purposes, the graphs are referred to as **Lineweaver-Burk** plots, where plots of 1/v vs 1/S are straight lines with slope m = K_M/V_M, and y-intercept b = $1/V_M$. Figure 6.3.6 common graphs used to display initial rate enzyme kinetic data.

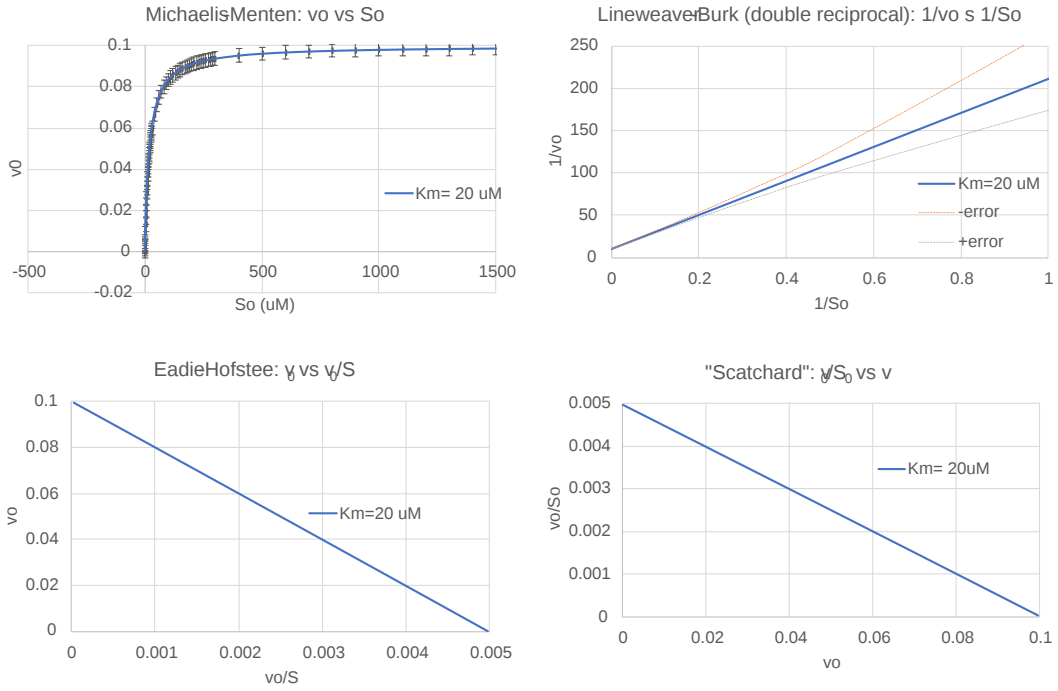


Figure 6.3.6: Common graphs used to display initial rate enzyme kinetic data

The straight-line plots shown above should **not** be analyzed using linear regression since simple linear regression assumes constant error in v_0 values. A weighted linear regression or even better, a nonlinear fit to a hyperbolic equation should be used. (Common

Error in Biochemistry Textbooks: The Shape of the Hyperbola). A rearrangement of the corresponding Scatchard equations in the Eadie-Hofstee plot is also commonly used, especially to visualize enzyme inhibition data as we will see in the next chapter.

6.3.7: An Extension: k_{cat} and V_M from integrated rate equations

K_M and V_M could be theoretically extracted from progress curves of A or P as a function of t at one single A concentration by deriving an integrated rate equation for A or P as a function of t, as we did in equation 2 (the integrated rate equation for the conversion of A → P in the absence of enzyme). In principle, this method would be better than the initial rates methods. Why? It is not easy to be certain you are measuring the initial rate for each and every [S] which should vary over a wide range. It's also time-intensive. In addition, think how much data is discarded if you take an entire progress curve at each substrate concentration, especially if you quench the reaction at a given time point, which effectively limits the data to one time point per substrate.

In practice, the mathematics is complicated, and it is impossible to get a simple explicit function of [P] or [S] as a function of time. A slight variant of a progress curve can be derived. Let us consider the simple case of a single substrate S (or A) being converted to product P in an enzyme-catalyzed reaction. The analogous equations for first-order, noncatalyzed rates were $A=A_0e^{-k_1t}$ or $P = A_0(1-e^{-k_1t})$.

We can derive the equation for the enzyme-catalyzed reaction shown below.

$$E + S \underset{k_2}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_3} E + P$$

Here it is!

$$\frac{P}{t} = \frac{K_M \ln\left(\frac{S_0 - P}{S_0}\right)}{t} + V_M \tag{6.3.29}$$

? Derivation: k_{cat} and V_M from integrated rate equation

Click below to see the derivation

Derivation

$$v = -\frac{dS}{dt} = +\frac{dP}{dt} = \frac{V_M S}{K_M + S}$$
$$\int_{S_0}^{S} \frac{K_M + S}{V_M S} dS = -\int_0^t t \tag{6.3.30}$$

$$-t = \frac{S + K_M \ln S - S_0 - K_M \ln S_0}{V_M} \tag{6.3.31}$$

On rearrangement, this gives:

$$S_0 - S + K_M \ln \frac{S_0}{S} = V_M t \tag{6.3.32}$$

This equation is an **implicit** equation, not an explicit one, as it does NOT give S(t) explicitly as a function of t.

Equation yy can be written with respect to product P as follows:

$$P = S_0 - S'' \text{ or } ''S = S_0 - P$$
$$S_0 - (S_0 - P) + K_M \ln \frac{S_0}{S_0 - P} = V_M t$$
$$P - K_M \ln\left(\frac{S_0 - P}{S_0}\right) = V_M t \tag{6.3.33}$$

Rearranging this gives

$$\frac{P}{t} = \frac{K_M \ln\left(\frac{S_0 - P}{S_0}\right)}{t} + V_M \tag{6.3.34}$$

S(t) explicitly as a function of t.

This equation does not give P(t) explicitly as a function of time. Rather, one can get a graph of P/t vs [ln (1-P/S_0)]/t (shown below) from the derived equation, which *does* give a straight line with a slope of Km and a y-intercept of Vm. Note that the calculated values of V_M and K_M are derived from only one substrate concentration, and the values may be affected by product inhibition.

Figure 6.3.7 compares a first-order noncatalyzed conversion of A → P to the enzyme-catalyzed rate. The V_M for the enzyme-catalyzed reaction was chosen to be small to make the two graphs comparable.

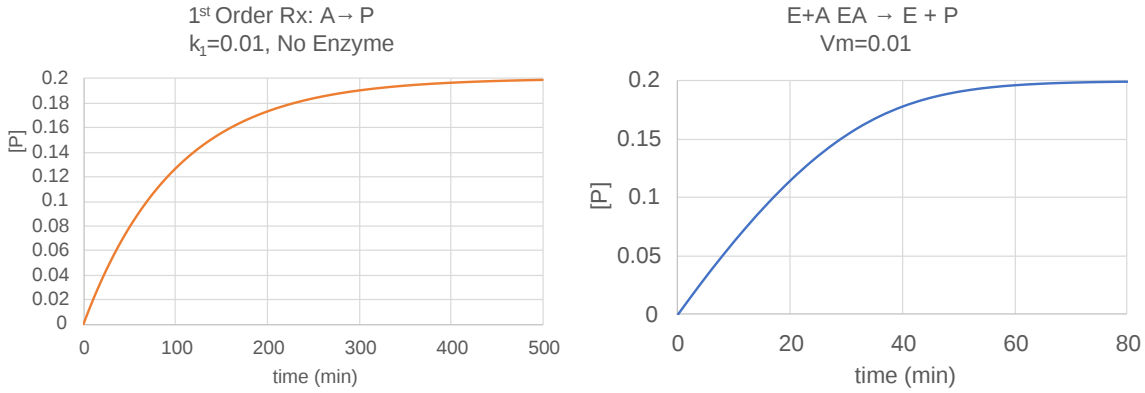


Figure 6.3.7: Comparison of a first order noncatalyzed conversion of A → P to the enzyme catalyzed rate

Note that the curves are similar but not identical. If you didn't know an enzyme was present, you could fit the data to a first-order rise in [P] with time, but it would not be the optimal fit. The progress curves are a lot more complicated to analyze if the product, which shares structural similarities with the substrate, binds the enzyme tightly and inhibits it (called product inhibition).

6.3.7.1: Comparison of progress Curves for uncatalyzed and catalyzed reactions.

Let's explore progress curves for enzyme-catalyzed reactions a bit further. Students usually see v_0 vs. [S] Michaelis-Menten plots in textbooks. These plots are, in some ways, less intuitive than seeing P vs. t curves, which are more in line with how we might contemplate how a reaction proceeds. Hence, it would be illuminating to compare progress curve graphs of A → P (irreversible) for the uncatalyzed and S → P for the enzyme-catalyzed reactions. What might you expect? We saw one example in Figure 6.3.6.

In each case, P should increase with time. In the uncatalyzed reaction, S exponentially decreases to 0 and P concomitantly rises to a value of P = S_0. You also get a rise in P vs t for the enzyme-catalyzed rate, but you would think it would be a faster rise with time since the enzyme catalyzes the reaction. Figure 6.3.87 shows a comparison of the progress curves for the uncatalyzed first-order reaction of A → P_1 (red) and S → P_2 (blue) for enzyme- catalyzed reaction (blue, right) for these conditions: uncatalyzed reaction A → P_1, k_1 = 0.1; Catalyzed reaction: S → P_2, V_M=10, K_M=5. Note that the rate at which bound S (i.e ES) goes to P for the catalyzed rate is 100x faster than the rate constant for the catalyzed rate.

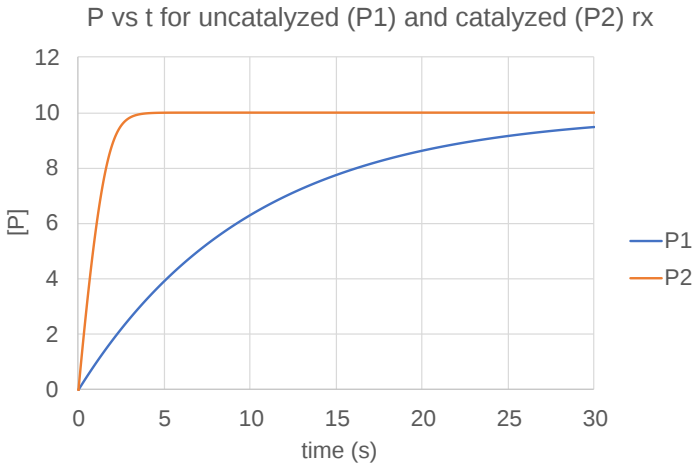


Figure 6.3.8: Another comparison of progress curves (P vs t) for a catalyzed and uncatalyzed conversion of A → P.

Note that the curves are somewhat similar in shape but also clearly different in comparison to the one shown in Figure 6.3.6.

Now, let's use Vcell to compare the reactions for different values for the kinetic constants for the uncatalyzed and the enzyme-catalyzed reactions. Change the constants and try to find a set of conditions so that the catalyzed and catalyzed rates for conversion of reactions to products are superimposable. How can that be?

VCell MODEL

Irreversible reactions: A → P1 and E + S ↔ ES → E + P

Vcell reaction diagram (1-way arrows defined as reversible in actual mathematical model) and chemical equation

Initial conditions

J	reaction rate	$(Kf \cdot A - Kr \cdot P1)$
Kf	forward rate constant	0.1
Kr	reverse rate constant	0.0
J	reaction rate	$\frac{Vmax \cdot S}{(Km + S)}$
Km	Km (1/2 max)	5.0
Vmax	max reaction rate	10.0
Size	size	1.0
initConc	initial concentration for A	10.0
initConc	initial concentration for P1	0.0
initConc	initial concentration for S	10.0
initConc	initial concentration for P2	0.0

Select Load [model name] below

Load Compare Irrev Rx No EnzW ithEnz

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), The Center for Cell Analysis & Modeling, at UConn Health. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Animations

Now let's look at an animation of the same irreversible reactions in which enzyme-catalyzed reaction is no faster than the noncatalyzed rate (a worthless enzyme!). Here are the reactions:

$A \rightarrow P$, k_1 =0.1; $S \rightarrow P$, V_M=?, K_M = ?.

The animations show just the accumulation of products. Animations are by Shraddha Nakak and Hui Liu.

? Exercise 6.3.1

Use the Vcell model above to find a set of values for K_M and V_M that would make the two graphs superimposable - i.e. when the graphs in the absence and presence of E are identical. (Hint: that would be a really bad enzyme if it didn't increase the reaction over the uncatalyzed rate!)

Answer

K_M = 96, V_M = 10

6.3.8: Summary

This chapter delves into the fundamental principles of enzyme catalysis and the quantitative analysis of reaction kinetics. It explains how enzymes accelerate chemical reactions by binding substrates, stabilizing transition states, and facilitating intramolecular conversions, all without changing the overall thermodynamics or equilibrium constant (Keq) of the reaction. Instead, enzymes dramatically increase the rate at which equilibrium is reached by lowering the activation energy.

Key topics covered include:

1. **Mechanistic Overview of Enzyme Catalysis**
   - **Reaction Pathway:** The enzyme binds to the substrate to form an enzyme–substrate complex (ES), converts the bound substrate to a bound product via a lower-energy transition state, and finally releases the product.
   - **Thermodynamics vs. Kinetics:** Although enzymes do not alter the equilibrium constant, they modify the reaction pathway by lowering activation barriers and changing the mechanism.
2. **Derivation of the Michaelis-Menten Equation**
   - **Rapid Equilibrium Assumption:** Under conditions where the substrate concentration far exceeds that of the enzyme, the rapid equilibrium between enzyme, substrate, and the ES complex allows the derivation of $[ES]=E_0[S]/(K_S+[S])$.
   - **Steady State Assumption:** A more general approach assumes that the formation and breakdown of ES reach a quasi-steady state (i.e., $d[ES]/dt \approx 0$, leading to the same functional form of the Michaelis-Menten equation: $v_0=V_{max}[S]/(K_M+[S])$ where $V_{max}=k_{cat}E_0$ and $K_M=(k_2+k_3)/k_1$.
3. **Graphical and Mathematical Analysis of Kinetic Data**

- **Progress Curves and Initial Rates:** The chapter explains how to measure reaction velocities from progress curves (concentration vs. time plots) and emphasizes the importance of using initial rate data when substrate depletion is minimal.
- **Linearization Techniques:** Various plots (e.g., Lineweaver-Burk, Eadie-Hofstee) are discussed as tools to extract KMK_M and VmaxV_{max} from experimental data, while noting the potential pitfalls and the value of nonlinear curve fitting.

4. **Complex Enzyme-Catalyzed Reactions**

- **Reversible Reactions:** The extension of the Michaelis-Menten framework to reversible reactions is explored, showing how both forward and reverse reactions are incorporated into kinetic models.
- **Reactions with Intermediates:** The chapter addresses mechanisms involving covalent intermediates (e.g., acyl-enzyme intermediates in proteases) and discusses how these multi-step processes still conform to a generalized Michaelis-Menten equation under appropriate assumptions.

5. **Computational Modeling and Simulation**

- The use of computational tools, such as VCell, is highlighted to simulate reaction progress curves and model complex kinetic systems. Interactive graphs and animations illustrate how changes in kinetic parameters (e.g., rate constants, substrate concentration) influence the reaction trajectory, thereby reinforcing the concepts developed through the analytical derivations.

In summary, this chapter provides a comprehensive treatment of enzyme kinetics, from basic principles and derivations to experimental data analysis and computational modeling. The insights gained here lay the groundwork for understanding more complex biochemical pathways and for the rational design of experiments to probe enzyme mechanisms.

6.4: Enzyme Inhibition

Learning Goals (ChatGPT o1, 1/30/25)

- **Understanding Irreversible Covalent Inhibition:**
  - Explain how irreversible inhibitors modify key amino acid side chains (e.g., cysteine modification by iodoacetamide) to permanently inactivate an enzyme.
  - Recognize the difference between enzyme denaturation (via extremes of pH and temperature) and irreversible covalent modification.
- **Mechanisms of Reversible Enzyme Inhibition:**
  - Define competitive inhibition and describe how substrate and inhibitor bind mutually exclusively to the active site.
  - Interpret the competitive inhibition kinetic equation, $v_0=V_MS/(K_M(1+IKis)+S)$ and explain how this alters the apparent $K_{M\ aoo}$ without affecting V_M.
- **Graphical Analysis of Competitive Inhibition:**
  - Interpret Lineweaver-Burk plots for competitive inhibition, identifying that the y-intercept ($1/V_M$) remains constant while the x-intercept ($–1/K_M$) shifts as inhibitor concentration increases.
  - Explain how plotting v_0 vs. log [S] can reveal the sigmoidal shifts associated with competitive inhibition.
- **Exploring Uncompetitive Inhibition:**
  - Describe uncompetitive inhibition where the inhibitor binds only to the enzyme–substrate (ES) complex, leading to a decrease in both V_M and K_M.
  - Understand that Lineweaver-Burk plots for uncompetitive inhibition display parallel lines due to proportional decreases in both V_M and K_M.
- **Noncompetitive and Mixed Inhibition:**
  - Distinguish between noncompetitive inhibition (where the inhibitor binds to both free enzyme and ES with equal affinity) and mixed inhibition (with differing affinities).
  - Interpret kinetic expressions for noncompetitive/mixed inhibition and explain how they affect V_M and K_M (e.g., K_M remains unchanged in pure noncompetitive inhibition, while both parameters change in mixed inhibition).
  - Analyze interactive graphs and Lineweaver-Burk plots to visualize these effects.
- **Product Inhibition and Competing Substrate Effects:**
  - Explain how product inhibition occurs when the reaction product, due to its structural similarity to the substrate, binds to the enzyme and decreases the overall reaction rate.
  - Understand the concept of specificity constants (k_{cat}/K_M) and how competing substrates can effectively act as inhibitors.
- **In Vivo vs. In Vitro Inhibition:**
  - Contrast the experimental conditions for enzyme inhibition studies in vitro (fixed enzyme, varying substrate) versus in vivo (dynamic flux through metabolic pathways).
  - Interpret how metabolic flux and substrate accumulation can influence the effectiveness of competitive versus uncompetitive inhibitors in a cellular context.
- **Environmental Effects on Enzyme Activity:**
  - Describe how temperature and pH changes affect enzyme activity, including the role of pH in altering the protonation states of key catalytic residues.
  - Predict how variations in temperature or pH might lead to reversible changes in enzyme kinetics or irreversible enzyme denaturation.

By achieving these learning goals, students will gain a thorough understanding of both the mechanistic basis of enzyme inhibition and the practical techniques used to measure and analyze inhibitor effects. This knowledge forms a critical foundation for exploring enzyme regulation in metabolic pathways and designing effective pharmaceutical inhibitors.

6.4.1: Irreversible Covalent Inhibition

Given what you already know about protein structure, it should be easy to determine how to inhibit an enzyme. Since structure mediates function, anything that would significantly alter an enzyme's structure would inhibit the enzyme's activity. Hence, extremes of pH and high temperature, all of which can denature the enzyme, would irreversibly inhibit the enzyme unless it could refold properly. Alternatively, we could add a small molecule, which interacts noncovalently with the enzyme to either change its conformation or directly prevent substrate binding. Finally, we could covalently modify certain side chains, that if they are essential to enzymatic activity, would irreversibly inhibit the enzyme.

We discussed previously the types of reagents that would chemically modify specific side chains that might be critical for enzymatic activity. For example, iodoacetamide might abolish enzyme activity if a cysteine side chain is required for activity. However, these reagents will usually modify several side chains. Determining which is critical for binding or catalytic conversion of the substrate can be difficult. One way would be to protect the active site with saturating concentrations of a ligand that binds reversibly at the active site. Then, the chemical modification can be performed at varying reaction times. The critical side chain would be protected from the chemical modification with the extent of protection depending on the K_D, the concentration of the protecting ligand., and the length of the reaction.

The rest of the chapter will deal with **reversible, noncovalent inhibition**

6.4.2: Competitive Inhibition

Reversible Competitive inhibition occurs when substrate (S) and inhibitor (I) both bind to the same site on the enzyme. In effect, they compete for the active site and bind in a mutually exclusive fashion. This is illustrated in the chemical equations and molecular cartoons shown in Figure 6.4.1.

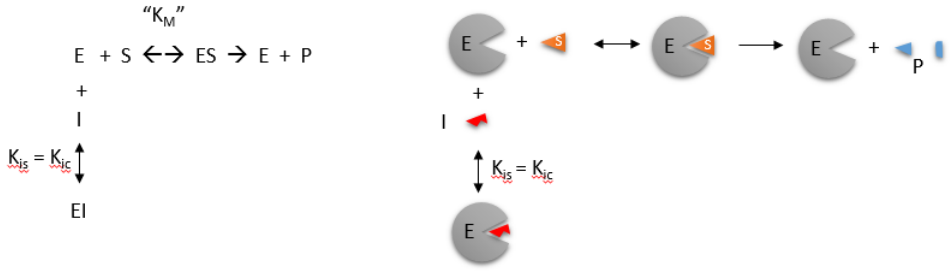


Figure 6.4.1: Competitive Inhibition

$$v_0 = \frac{V_M S}{K_M \left(1 + \frac{I}{Kis}\right) + S} \tag{6.4.1}$$

There is another type of inhibition that would give the same kinetic data. If S and I were bound to different sites, and S was bound to E and produced a conformational change in E such that I could not bind (and vice versa), then the binding of S and I would be mutually exclusive. This is called allosteric competitive inhibition. Inhibition studies are usually done at several fixed and non-saturating concentrations of I and varying S concentrations.

The key kinetic parameters to understand are V_M and K_M. Let us assume for ease of equation derivation that I binds reversibly and with rapid equilibrium to E, with a dissociation constant K_{IS}. The "s" in the subscript "is" indicates that the slope of the 1/v vs. 1/S Lineweaver-Burk plot changes while the y-intercept stays constant. K_{IS} is also named K_{IC}, where the subscript "c" stands for competitive inhibition constant.

A look at the top mechanism shows that even in the presence of I, as S increases to infinity, all E is converted to ES. That is, there is no free E to which I could bind. Now, remember that $V_M = k_{cat}E_0$. Under these conditions, ES = E_0; hence v = V_M. V_M is not changed. However, the apparent K_M, K_{Mapp}, will change. We can use LaChatelier's principle to understand this. If I binds to E alone and not ES, it will shift the equilibrium of E + S → ES to the left. This would increase the K_{Mapp} (i.e. it would appear that the affinity of E and S has decreased.). The double reciprocal plot (Lineweaver-Burk plot) offers a great way to visualize the inhibition as shown in Figure 6.4.2.

$$v_0 = \frac{V_M S}{K_M\left(1 + \frac{I}{Kis}\right) + S}$$

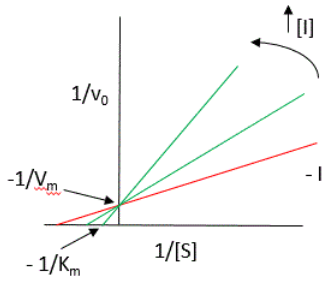
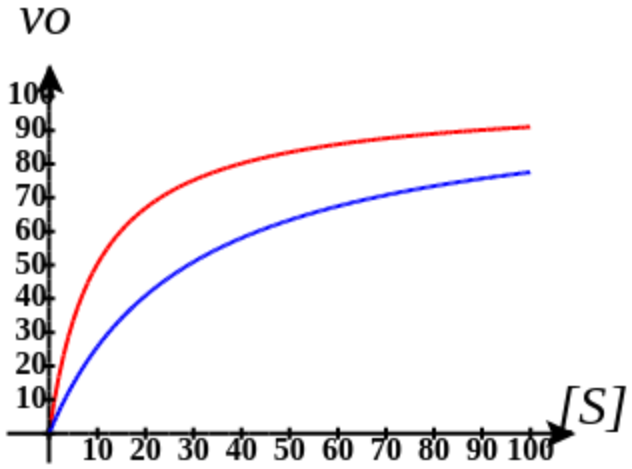


Figure 6.4.2: Competitive Inhibition: Lineweaver-Burk plots

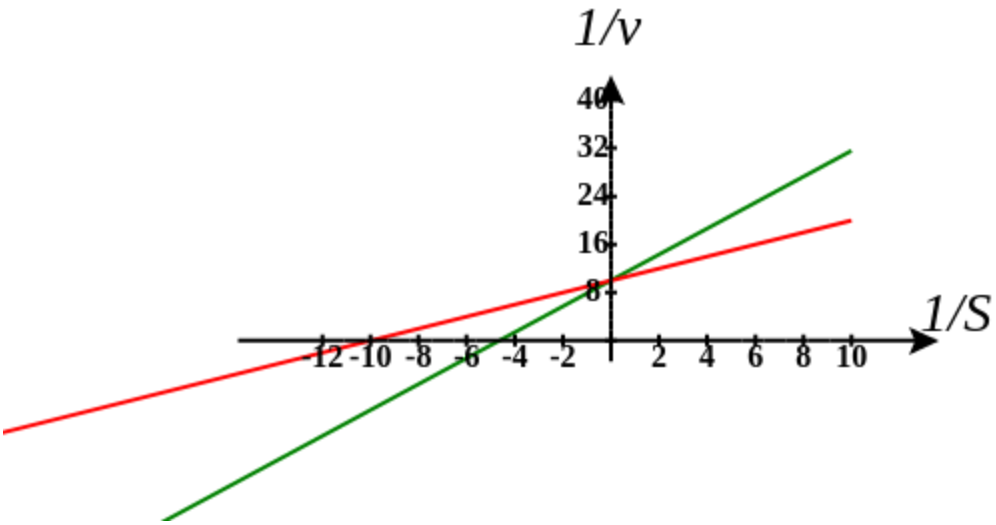
In the presence of I, V_M does not change, but K_M appears to increase. Therefore, $1/K_M$, the x-intercept on the plot will get smaller, and closer to 0. Therefore the plots will consist of a series of lines, with the same y-intercept ($1/V_M$), and the x-intercepts ($-1/K_M$) closer and closer to 0 as I increases. These intersecting plots are the hallmarks of competitive inhibition.

Here is an interactive graph showing v_0 vs [S] for competitive inhibition with Vm and Km both set to 100. Change the sliders for [I] and Kis and see the effect on the graph.



Here is the interactive graphs showing $1/v_0$ vs 1/[S] for competitive inhibition, with Vm and Km both set to 10.

Competitive Inhibition



Note that in the first three inhibition models discussed in this section, the Lineweaver-Burk plots are linear in the presence and absence of an inhibitor. This suggests that plots of v vs. S in each case would be hyperbolic and conform to the usual form of the Michaelis Menton equation, each with potentially different apparent V_M and K_M values.

An equation for v_0 in the presence of a competitive inhibitor is shown in the above figure. The only change compared to the equation for the initial velocity in the absence of the inhibitor is that the K_M term is multiplied by the factor $1+I/K_{is}$. Hence $K_{Mapp} = K_M(1+I/Kis)$. This shows that the apparent K_M does increase as we predicted. K_{IS} is the inhibitor dissociation constant in which the inhibitor affects the slope of the double reciprocal plot.

v_0 vs log [Inhibitor] plots

If the data were plotted as v_0 vs. log S, the plots would be sigmoidal, as we saw for plots of ML vs. log L in Chapter 5B. In the case of a competitive inhibitor, the plot of v_0 vs log S in the presence of different fixed concentrations of inhibitor would consist of a series of sigmoidal curves, each with the same V_M, but with different apparent K_M values (where $K_{Mapp} = K_M(1+I/K_{is})$, progressively shifted to the right. Enzyme kinetic data is rarely plotted this way. These plots are mostly used for simple binding data for the M + L ↔ ML equilibrium in the presence of different inhibitor concentrations.

Reconsider our discussion of the simple binding equilibrium, M + L ↔ ML. For fractional saturation Y vs a log L graphs, we considered three examples:

1. L = 0.01 K_D (i.e., L << K_D), which implies that K_D = 100L. Then Y = L/[K_D+L] = L/[100L + L] ≈1/100. This implies that irrespective of the actual [L], if L = 0.01 K_D, then Y ≈0.01.
2. L = 100 K_D (i.e,. L >> K_D), which implies that K_D = L/100. Then Y = L/[K_D+L] = L/[(L/100) + L] = 100L/101L ≈ 1. This implies that irrespective of the actual [L], if L = 100 Kd, then Y ≈1.
3. L = K_D, then Y = 0.5

These scenarios show that if L varies over four orders of magnitude ($0.01K_D < K_D < 100K_D$), or, in log terms, from $-2 + \log K_D < \log K_D < 2 + \log Kd$), irrespective of the magnitude of the K_D, that Y varies from approximately 0 - 1.

In other words, Y varies from 0-1 when L varies from log K_D by +2. Hence, plots of Y vs log L for a series of binding reactions of increasingly higher K_D (lower affinity) would reveal a series of identical sigmoidal curves shifted progressively to the right, as shown below in Figure 6.4.3.

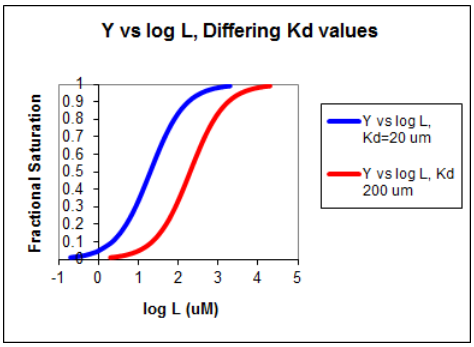


Figure 6.4.3: Plot of fractional saturation Y vs log L for two different K_D values.

The same would be true of v_0 vs. S in the presence of different concentrations of a competitive inhibitor, for initial flux, Jo vs ligand outside, in the presence of a competitive inhibitor, or ML vs. L (or Y vs L) in the presence of a competitive inhibitor.

In many ways, plots of v_0 vs. lnS are easier to visually interpret than plots of v_0 vs. S . As noted for simple binding plots, textbook illustrations of hyperbolas are often misdrawn, showing curves that level off too quickly as a function of [S] as compared to plots of v_0 vs lnS, in which it is easy to see if saturation has been achieved. In addition, as the curves above show, multiple complete plots of v_0 vs lnS at varying fixed inhibitor concentrations or for variant enzyme forms (different isoforms, site-specific mutants) over a broad range of lnS can be made, which facilitates comparisons of the experimental kinetics under these different conditions. This is especially true if Km values differ widely.

Now that you are more familiar with binding and enzyme kinetics curves, in the presence and absence of inhibitors, you should be able to apply the above analysis to inhibition curves where the binding or the initial velocity is plotted at varying competitive inhibitor concentrations at different fixed nonsaturating concentrations of ligand or substrate. Consider the activity of an enzyme. Let's say that at some reasonable substrate concentration (not infinite), the enzyme is approximately 100% active. If a competitive inhibitor is added, the activity of the enzyme decreases until at saturating (infinite) I, no activity would remain. Graphs showing this are shown below in Figure 6.4.4.

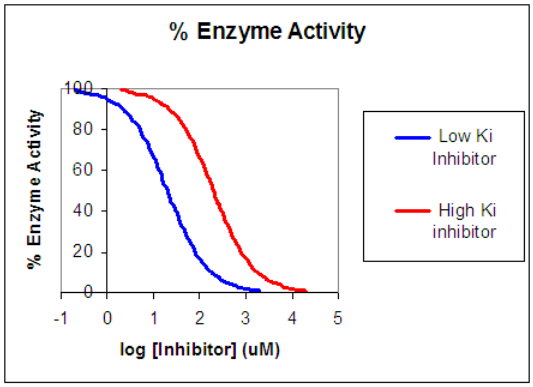


Figure 6.4.4: Inhibition of Enzyme Activity - % Activity vs log [Inhibitor]

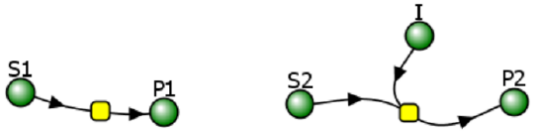
6.4.2.1: Progress Curves for Competitive Inhibition

In the previous section, we explored how important progress curve (Product vs time) analyses are in understanding both uncatalyzed and enzyme-catalyzed reactions. We are aware of no textbooks that cover progress curves for enzyme inhibition. Yet progress curves are what most investigators record and analyze to determine initial rates v_0 and to calculate V_M, K_M, and inhibition constants, as described above. We will use Vcell to produce progress curves for reversibly inhibited enzyme-catalyzed reactions.

VCell MODEL

Competitive Inhibition with constant [I]:

No inhibition (left) and competitive inhibition (right)



Initial conditions for no inhibition

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S1}{(Km + S1)}$
Km	Km (1/2 max)	☐	5.0
Vmax	max reaction rate	☐	10.0
S1	Species Concentration	☑	Variable

Initial conditions for competitive inhibition

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S2}{(Km + S2 + \frac{Km \cdot I}{Ki})}$
Km	user defined	☐	5.0
Vmax	user defined	☐	10.0
Ki	user defined	☐	5.0
S2	Species Concentration	☑	Variable
I	Species Concentration	☑	Variable

I is fixed for each simulation (as it is not converted to a product) but can be changed in the simulation below.

Select Load [model name] below

Load NoInhib_CompInhib_MM_I

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed. For this model, select Vm, Km, Ki and I | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

The graphs from your initial run show the concentrations of S, P, and I as a function of time for just the initial conditions shown above. In typical initial rate analyses of competitive inhibition, at least three sets of reactions are run with the same varying substrate concentrations and different fixed inhibitor concentrations. In the analyses above, [I] is fixed at 5 uM.

Conduct a series of runs at different values of I. Vary the K_I, the dissociation constant for the EI complex, as follows:

- $I << K_I$, the dissociation constant for the EI complex
- $I >> K_I$, the dissociation constant for the EI complex. Then, download the data and determine the initial rate for each of the initial conditions.

Figure 6.4.5 shows an interactive iCn3D model of human low molecular weight phosphotyrosyl phosphatase bound to a competitive inhibitor (5PNT)

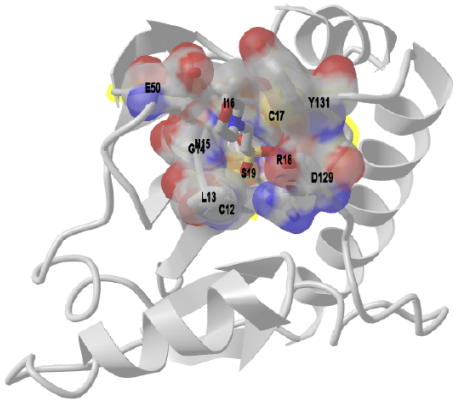


Figure 6.4.5: Human low molecular weight phosphotyrosyl phosphatase bound to a competitive inhibitor (5PNT). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...XsEacG2tixDDi9

The competitive inhibitor, the deprotonated form of 2-(N-morpholino)-ethanesulfonic acid (MES), is actually the conjugate base of the weak acid (pKa = 6.15) of a commonly used component of a buffered solution. It is shown in color sticks with the negatively charged sulfonate at the bottom of the active site pocket. The amino acids comprising the active site binding pocket are shown as color sticks underneath the transparent colored surface of the binding pocket. The normal substrates for the enzyme are proteins phosphorylated on tyrosine side chains, so the sulfonate mimics the negatively charged phosphate group of the phosphoprotein target.

6.4.2.2: Two special cases of competition inhibition

Product Inhibition

Let's look at an enzyme that converts reactant S to product P. Since P arises from S, they may have structural similarities. For example, what if GTP was the reactant and GDP was a product? If so, then P might also bind in the active site and inhibit the conversion of S to P. This is called **product inhibition**. It probably occurs in most enzymes, and when it does occur, it will start bending downward at the beginning of the progress curve for P formation. If the product binds very tightly, it might cause a significant underestimation of the initial velocity (v_0) or flux (J_0) of the enzyme. Let's use Vcell to explore product inhibition. The model will explore two reactions:

- E + R ↔ ER → E + Q (no product inhibition)
- E + S ↔ ES → E + P (with product inhibition)

The chemical equation above does not explicitly show the product P binding the enzyme to form an EP complex. An actual reaction diagram showing the inhibition of an enzyme by an inhibitor I and by the product P is shown in Figure 6.4.6 below.

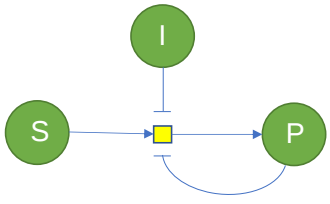


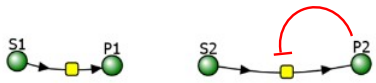
Figure 6.4.6: reaction diagram showing inhibition of an enzyme by an inhibitor I and by the product P

Vcell uses much simpler diagrams since it is most often used for modeling whole pathways or even entire cells. In the simpler Vcell reaction diagrams, the inhibitor is typically not shown since the inhibition is built into the equation for the enzyme, represented by the node or yellow square in the figure above.

Let's now explore product inhibition in Vcell. In the reaction without product inhibition, R and Q are the reactant and product, respectively. S and P are used for the reaction *with* product P inhibition.

VCell MODEL

Irreversible MM Kinetics - Without (left rx 1) and With (right, rx 2) Product Inhibition



Initial Conditions: No product inhibition

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S1}{(Km + S1)}$
Km	Km (1/2 max)	☐	5.0
Vmax	max reaction rate	☐	10.0
S1	Species Concentration	☑	Variable

Initial Conditions: With product inhibition

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S2}{(Km + S2 + \frac{Km \cdot P2}{Ki})}$
Km	user defined	☐	5.0
Vmax	user defined	☐	10.0
Ki	user defined	☐	5.0
S2	Species Concentration	☑	Variable
P2	Species Concentration	☑	Variable

Select Load [model name] below

Load ProdInhib_NoProdInhib

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Inhibition by a competing substrate - the specificity constant

In the previous chapter, the specificity constant was defined as k_{cat}/K_M, which we also described as the second-order rate constant associated with the bimolecular reaction of E and S when S << K_M. It also describes how good an enzyme is in differentiating between different substrates. If an enzyme encounters two different substrates, one can be considered a competitive inhibitor of the other. The following equation gives the ratio of initial velocities for two competing substrates at the same concentration equal to the ratio of their k_{cat}/K_M values.

$$\frac{v_A}{v_B} = \frac{\frac{k_{cat\,A}}{K_A} A}{\frac{k_{cat\,B}}{K_B} B} \tag{6.4.2}$$

? A derivation of the specificity constant for an enzyme with competing substrates

Here it is!

Derivation

$$v_A = \frac{V_A A}{K_A\left(1+\frac{B}{K_B}\right)+A} \quad v_B = \frac{V_B B}{K_B\left(1+\frac{A}{K_A}\right)+B} \tag{6.4.3}$$

$$\frac{v_A}{v_B} = \frac{\frac{V_A A}{K_A\left(1+\frac{B}{K_B}\right)+A}}{\frac{V_B B}{K_B\left(1+\frac{A}{K_A}\right)+B}} = \frac{\frac{V_A A}{K_A+\frac{K_A B}{K_B}+A}}{\frac{V_B B}{K_B+\frac{K_B A}{K_A}+B}} \tag{6.4.4}$$

Now, in the above equation:

multiple the top half of the right-hand expression by

$$\frac{\frac{1}{K_A}}{\frac{1}{K_A}} \tag{6.4.5}$$

multiple the bottom half of the right-hand expression by

$$\frac{\frac{1}{K_B}}{\frac{1}{K_B}} \tag{6.4.6}$$

replace V_A with $k_{catA}E_0$ and V_B with $k_{catB}E_0$

This gives the following expression for v_A/v_B:

$$\frac{v_A}{v_B} = \frac{\frac{k_{catA}}{K_A}A}{\frac{k_{catB}}{K_B}B} \tag{6.4.7}$$

6.4.3: Uncompetitive Inhibition

Reversible uncompetitive inhibition occurs when I binds only to ES and not free E. One can hypothesize that on binding S, a conformational change in E occurs, which presents a binding site for I. Inhibition occurs since ESI can not form the product. It is a dead-end complex that has only one fate: to return to ES. This is illustrated in the chemical equations and molecular cartoon shown in Figure 6.4.7.

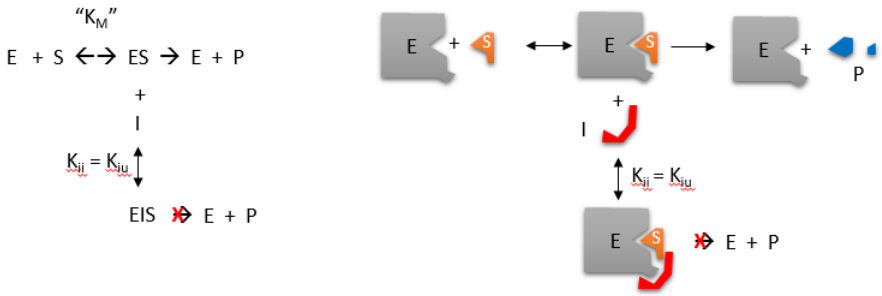


Figure 6.4.7: Uncompetitive Inhibition

$$v_0 = \frac{V_M S}{K_M + S\left(1+\frac{I}{Kii}\right)} = \frac{\left(\frac{V_M}{1+\frac{I}{Kii}}\right)S}{\left(\frac{K_M}{1+\frac{I}{Kii}}\right)+S} \tag{6.4.8}$$

Let us assume that I binds reversibly to ES with a dissociation constant K_{ii}. The second "i" in the subscript "ii" indicates that the intercept of the 1/v vs 1/S Lineweaver-Burk plot changes while the slope stays constant. K_{ii} is also named Kiu, where the subscript "u" stands for the uncompetitive inhibition constant.

A look at the top mechanism shows that in the presence of I, as S increases to infinity, not all of E is converted to ES. There is a finite amount of ESI, even at an infinite S. Now remember that $V_m = k_{cat}E_0$ if and only if all E is in the form ES. Under these conditions, the apparent V_m, V_{mapp} is less than the real Vm without the inhibitor. In addition, the apparent K_m, K_{mapp}, will change. We can use LaChatelier's principle to understand this. If I binds to ES alone, and not E, it will shift the equilibrium of E + S ↔ ES to the right, which would have the effect of decreasing the K_{mapp} (i.e. it would appear that the affinity of E and S has increased.). The double reciprocal plot (Lineweaver Burk plot) offers a great way to visualize the inhibition. In the presence of I, both Vm and Km decrease. Therefore, $-1/K_m$, the x-intercept on the plot, will get more negative, and $1/V_m$ will get more positive. It turns out that they change to the same extent. Therefore, the plots will consist of a series of parallel lines, which is the hallmark of uncompetitive inhibition, as shown in Figure 6.4.8.

$$v_0 = \frac{V_M S}{K_M + S\left(1 + \frac{I}{Kii}\right)} = \frac{\left(\frac{V_M}{1 + \frac{I}{Kii}}\right) S}{\left(\frac{K_M}{1 + \frac{I}{Kii}}\right) + S}$$

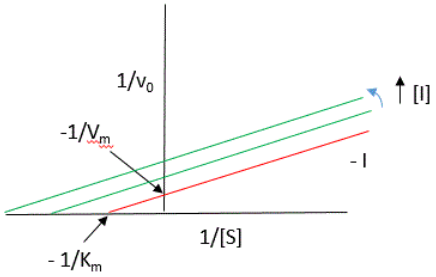
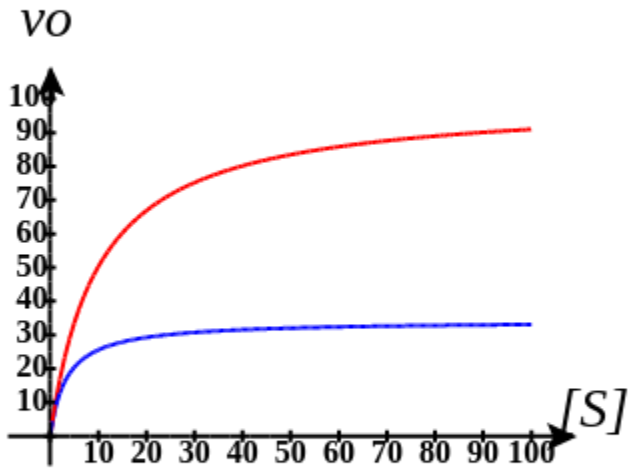


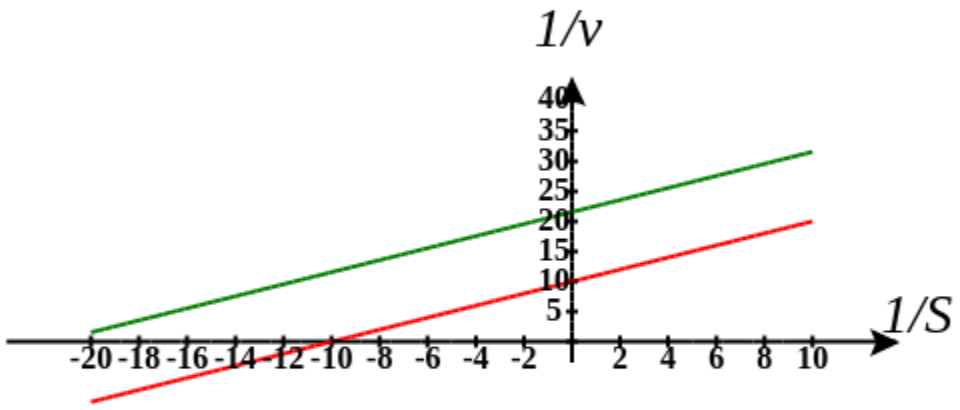
Figure 6.4.8: Uncompetitive Inhibition: Lineweaver-Burk plots

Here is an interactive graph showing v_0 vs [S] for uncompetitive inhibition with Vm and Km both set to 100. Change the sliders for [I] and Kis and see the effect on the graph.



Here is an interactive graph showing uncompetitive inhibition with Vm and Km both set to 10. Change the sliders for [I] and Kii and see the effect on the graph

Uncompetitive Inhibition



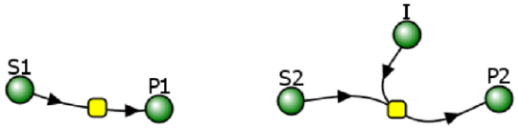
An equation, shown in the diagram above, can be derived to show the effect of the uncompetitive inhibitor on the velocity of the reaction. The only change is that the S term in the denominator is multiplied by the factor 1+I/Kii. We want to rearrange this equation to show how Km and Vm are affected by the inhibitor, not S. Rearranging the equation above shows that Kmapp = Km/(1+I/Kii) and Vmapp = Vm/(1+I/Kii). This shows that the apparent Km and Vm do decrease as we predicted. Kii is the inhibitor dissociation constant in which the inhibitor affects the intercept of the double reciprocal plot. Note that if I is zero, Km and Vm are unchanged.

6.4.3.1: Progress Curves: Uncompetitive Inhibition

Now, let's compare the progress curves for an enzyme-catalyzed reaction in the absence and presence of an uncompetitive inhibitor.

VCell MODEL

No inhibition (left) and Uncompetitive Inhibition (right)



(Note the the Vcell reaction diagram is the same as for competitive inhibition. However, the mathematical equations differ as shown below.

Initial values No Inhibition

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S1}{(Km + S1)}$
Km	Km (1/2 max)	☐	5.0
Vmax	max reaction rate	☐	10.0
S1	Species Concentration	☑	Variable

Initial values With Uncompetitive Inhibitor

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S2}{(Km + S2 \cdot (1.0 + \frac{I}{Ki}))}$
Km	user defined	☐	5.0
Vmax	user defined	☐	10.0
Ki	user defined	☐	5.0
S2	Species Concentration	☑	Variable
I	Species Concentration	☑	Variable

I is fixed for each simulation (as it is not converted to a product) but can be changed in the simulation below.

Select Load [model name] below

Load NoInhib_UncompInhibMM

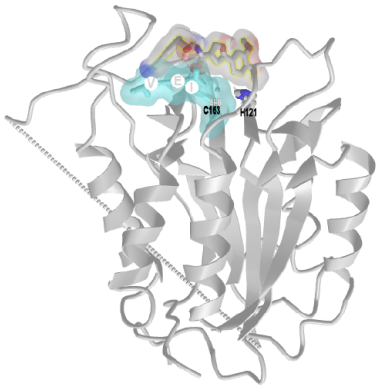
Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Figure 6.4.9 shows an interactive iCn3D model of an uncompetitive inhibitor of the cysteine protease caspase-6 (4HVA)



NCBI iCn3D Figure 6.4.9: Uncompetitive inhibitor of the cysteine protease caspase-6 (4HVA) . (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...TVfKio7NxcpX28

The "substrate" in this model (cyan, transparent surface, with labels on V, E, and I) is a substrate analog, VEI-CHO, in which the tripeptide substrate VEI ends not in a free carboxyl or amide group but an aldehyde, which causes this "substrate" to become covalently attached to the enzyme and act as an inhibitor. The uncompetitive inhibitor (gray transparent surface) binds externally

to the blue surface. Hence, it binds to the ES complex. Two active site residues, Cys 163, the active site nucleophile, and His 121, a catalytic acid/base, are shown in colored sticks and labeled.

Figure 6.4.10 below shows a Lineweaver-Burk plot showing $1/v_o$ vs 1/[S], where the substrate is the divalent compound $(VEID)_2R110$. Its N-terminus is capped with a benzyloxy (Z) group. R110 is a rhodamine-type fluorophore, which, on cleavage, gives a strong fluorescent signal that the benzyloxy group initially quenched in the uncleaved substrate.

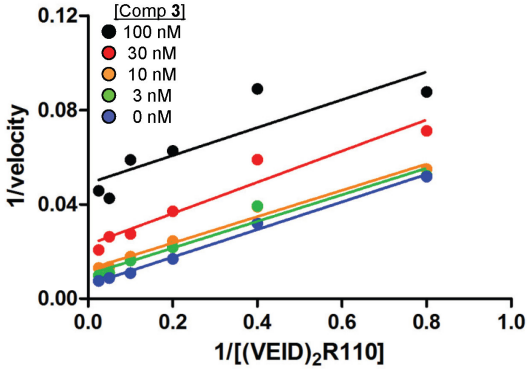


Figure 6.4.10: Double-reciprocal Lineweaver-Burk plot of compound 3 with $(VEID)_2R110$ substrate showing uncompetitive inhibition. Heise CE et al. (2012) Mechanistic and Structural Understanding of Uncompetitive Inhibitors of Caspase-6. PLoS ONE 7(12): e50864. https://doi.org/10.1371/journal.pone.0050864. Creative Commons Attribution License.

6.4.4: Noncompetitive and Mixed Inhibition

Reversible noncompetitive inhibition occurs when I binds to both E and ES. We will look at only the special case in which the dissociation constants of I for E and ES are the same. This is called noncompetitive inhibition. It is quite rare as it would be difficult to imagine a large inhibitor that inhibits the turnover of a bound substrate and does not affect the binding of S to E. However, the covalent interaction of protons with both E and ES can lead to noncompetitive inhibition. In the more general case, the Kd's are different, and the inhibition is called mixed. Since inhibition occurs, we will hypothesize that ESI can not form the product. It is a dead-end complex with only one fate: to return to ES or EI. This is illustrated in the chemical equations and the molecular cartoon in Figure 6.4.11.

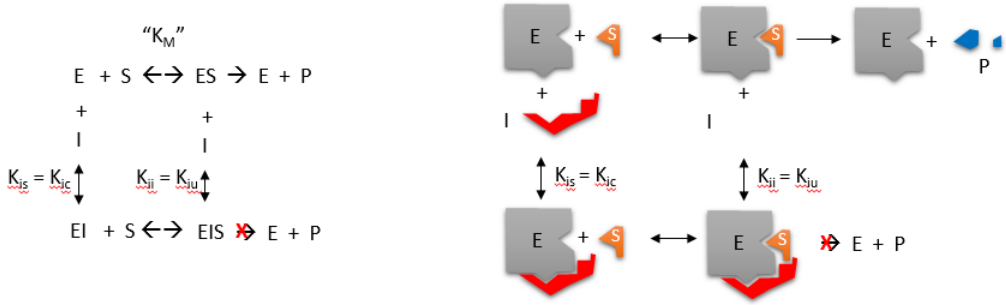


Figure 6.4.11: Noncompetitive and Mixed Inhibition

Let us assume for ease of equation derivation that I binds reversibly to E with a dissociation constant of K_{is} (as we denoted for competitive inhibition) and to ES with a dissociation constant K_{ii} (as we noted for uncompetitive inhibition). Assume that for noncompetitive inhibition, $K_{is} = K_{ii}$. A look at the top mechanism shows that in the presence of I, as S increases to infinity, not all of E is converted to ES. There is a finite amount of ESI, even at infinite S. Now remember that $V_m = k_{cat}E_0$ if and only if all E is in the form ES. Under these conditions, the apparent V_m, V_{mapp}, is less than the real Vm without an inhibitor. In contrast, the apparent K_m, K_{mapp}, will not change since I binds to both E and ES with the same affinity and hence will not perturb that equilibrium, as deduced from LaChatelier's principle. The double reciprocal plot (Lineweaver Burk plot) offers a great way to visualize the inhibition. In the presence of I, just V_m will decrease. Therefore, $-1/K_m$, the x-intercept will stay the same, and $1/V_m$ will get more positive. Therefore, the plots will consist of a series of lines intersecting on the x-axis, the hallmark of noncompetitive inhibition. You should be able to figure out how the plots would appear if K_{is} is different from K_{ii} (mixed inhibition).

$$v_0 = \frac{V_M S}{K_M\left(1+\frac{I}{Kis}\right)+S\left(1+\frac{I}{Kii}\right)} \tag{6.4.9}$$

An equation, shown in the diagram above, can be derived to show the effect of the noncompetitive inhibitor on the velocity of the reaction. Km is multiplied by $1+I/K_{is}$, and S by $1+I/K_{ii}$ in the denominator. We want to rearrange this equation to show how the inhibitor affects Km and Vm, not S. Rearranging the equation as shown above shows that $K_{mapp} = K_m(1+I/K_{is})/(1+I/K_{ii}) = K_m$ when $K_{is}=K_{ii}$, and $V_{mapp} = V_m/(1+I/K_{ii})$. This shows that the Km is unchanged, and Vm decreases as we predicted. The plot shows a series of lines intersecting on the x-axis as shown in Figure 6.4.12 Both the slope and the y-intercept are changed, which are reflected in the names of the two dissociation constants, K_{is} and K_{ii}. Note that if I is zero, $K_{mapp} = K_m$ and $V_{mapp} = V_m$. Sometimes, the K_{is} and K_{ii} inhibition dissociation constants are called $\mathbf{K_c}$ and $\mathbf{K_u}$ (competitive and uncompetitive inhibition dissociation constants.

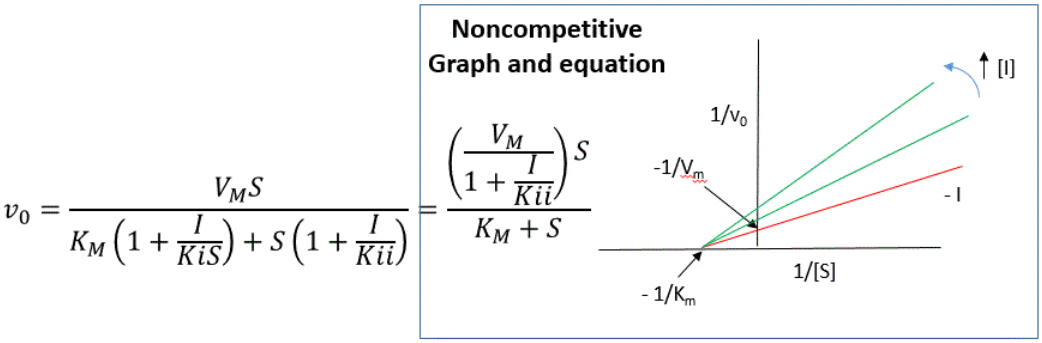
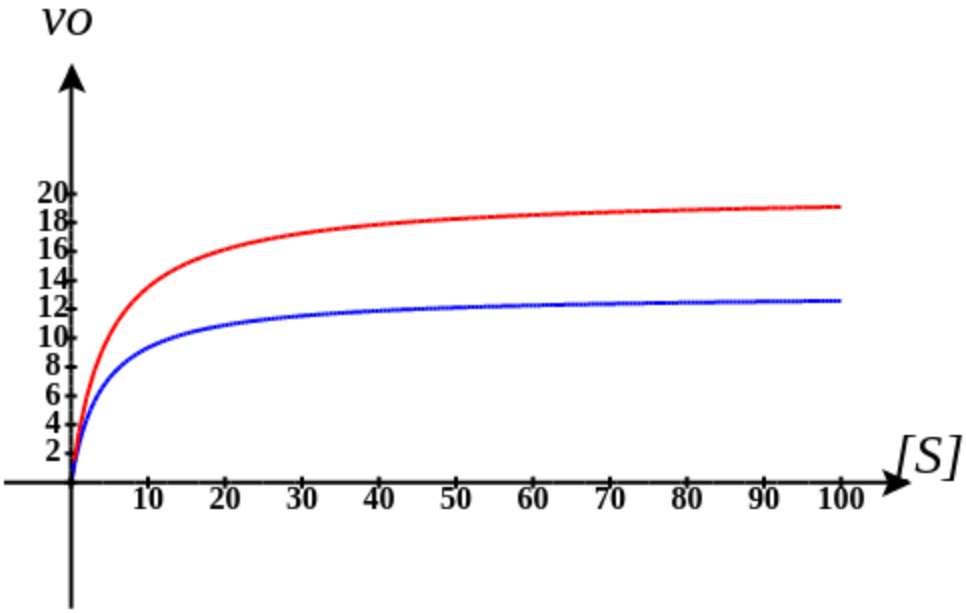


Figure 6.4.12: Noncompetitive Inhibition: Lineweaver-Burk plots

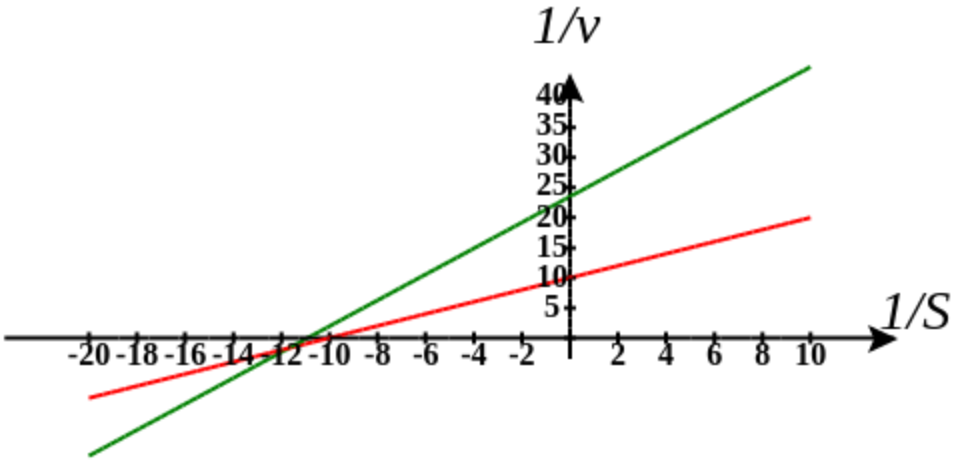
Move the sliders on this interactive graph below to show changes in K_{is} and K_{ii} affect position on the graph where the lines intersect. Try to change their values to move the intersections of the graphs from the left top quadrant to the x-axis to the left bottom quadrant.

Here is an interactive graph showing v_0 vs [S] for mixed inhibition with Vm and Km both set to 100. Change the sliders for [I] and Kis and see the effect on the graph



Here is an interactive graph showing mixed inhibition with Vm and Km both set to 10. Change the sliders for [I] and Kii and see the effect on the graph

Mixed Inhibition

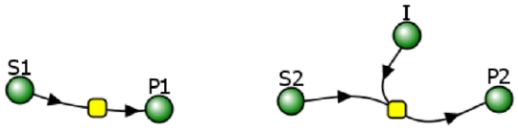


6.4.4.1: Progress Curves: Mixed/Noncompetitive Inhibition

Now let's compare the progress curves for an enzyme-catalyzed reaction in the absence and presence of a mixed inhibitor.

VCell MODEL

No inhibition (left) and Uncompetitive Inhibition (right)



Note that the Vcell reaction diagram is the same as for competitive and uncompetive inhibition. It doesn't explicitly show that the mixed inhibitor binds to both free and substrate-bound enzymes. Those interactions are addressed in the mathematical equations for mixed inhibition.

Initial values No Inhibition

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S1}{(Km + S1)}$
Km	Km (1/2 max)	☐	5.0
Vmax	max reaction rate	☐	10.0
S1	Species Concentration	☑	Variable

Initial values With Uncompetitive Inhibitor

Name	Description	Global	
J	reaction rate	☐	$\frac{Vmax \cdot S2}{(Km \cdot (1.0 + \frac{I}{Kc}) + S2 \cdot (1.0 + \frac{I}{Ki}))}$
Km	user defined	☐	5.0
Vmax	user defined	☐	10.0
Kc	user defined	☐	1.0
Ki	user defined	☐	1.0
S2	Species Concentration	☑	Variable
I	Species Concentration	☑	Variable

I is fixed for each simulation (as it is not converted to a product) but can be changed in the simulation below.

Select Load [model name] below

Load NoInhib_MIXEDInhib_MM

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

Mixed and noncompetitive inhibition (as shown by the mechanism above) differ from competitive and uncompetitive inhibition. The inhibitor binding is not simply a dead-end reaction in which the inhibitor can only dissociate in a single reverse step. In the above equilibrium, S can dissociate from ESI to form EI, so the system may not be at equilibrium. With dead-end steps, no flux of reactants occurs through the dead-end complex, so the equilibrium for the dead-end step is not perturbed.

Other mechanisms can commonly give mixed inhibition. For example, the product released in a ping pong mechanism (discussed in the next chapter) can give mixed inhibition. A ping pong reaction mechanism is shown and superficially explained in Figure 6.4.13

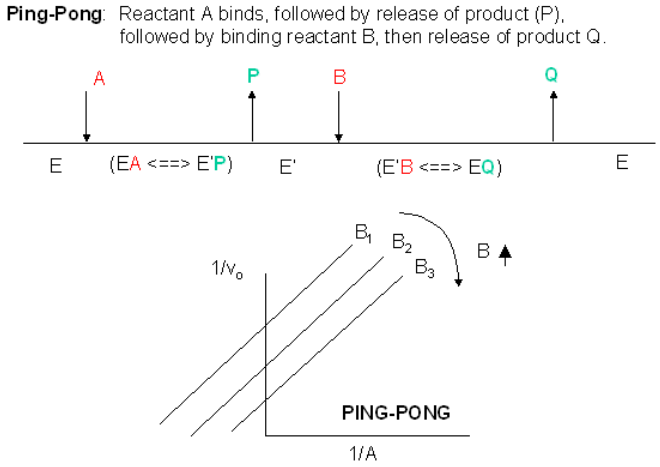


Figure 6.4.13: Ping pong reaction mechanism

If P, acting as a product inhibitor, can bind to two different forms of the enzyme (E' and also E), it will act as a mixed inhibitor.

Figure 6.4.14 shows an interactive iCn3D model of a noncompetitive inhibitor of M. tuberculosis's class IIa fructose 1,6-bisphosphate aldolase (4LV4).

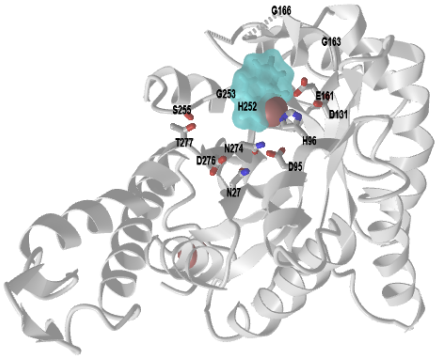


Figure 6.4.14: Noncompetitive inhibitor of M. tuberculosis's class IIa fructose 1,6-bisphosphate aldolase (4LV4). (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...wZDtkvtwqnrVy7

The non-competitive inhibitor, 8-hydroxyquinoline carboxylic acid (HCA), is shown with a cyan surface. Active site side chains of the Zn^{2+}- containing enzyme are shown as colored sticks and labeled.

The noncompetitive inhibitor does not enter into the pocket where the substrate (not shown in the model above) binds. The binding of the noncompetitive inhibitor causes a conformational change that moves the Zn-binding loop (Z-loop), which contains His 212, which coordinates the Zn^{2+} ion. Hence, the Zn-loop is part of the active site, and when it moves on binding of the inhibitor, the access to the empty pocket where the substrate binds is hindered. When the Zn^{2+} moves away from the active site, it can no longer engage in catalytic activity.

6.4.5: Enzyme Inhibition *in Vivo*

The pharmaceutical industry is devoted to finding drug molecules that affect biological processes. Typically, this means the development of small molecule inhibitors for target proteins. Recent work has expanded to developing inhibitory RNA molecules that affect DNA transcription and mRNA translation. Using combinatorial synthetic techniques and computational modeling, developing small molecule inhibitors (especially competitive ones) that inhibit proteins in vitro using purified enzymes, substrates, and inhibitors in lab testing has gotten easier. Assuming that the inhibitor could pass through the membrane and accumulate to a sufficient concentration, would it have the same inhibitory properties in the cell as in the test tube? The answer turns out to be maybe. Remember that a cell is tightly packed with a multitude of other small molecules and macromolecules. In addition, the enzyme targeted for inhibition is most likely part of a pathway of enzymes that feeds reactant into the enzyme and removes the product. Hence, the flux of substrate and product is controlled by the entire pathway and not just the single target enzyme. The target enzyme's product concentration is determined by kinetic parameters for the enzyme and available substrate concentration.

The conditions under which the enzymes are studied (in vitro) and operate (in vivo) are very different.

- In vitro (in the lab), the enzyme is held at a constant concentration while the substrate is varied (i.e., the substrate concentration is the independent variable). The velocity is determined by the substrate concentration. When inhibition is studied, the substrate is varied while the inhibitor is held constant at several different fixed concentrations.
- In vivo (in the cell), the velocity might be held at a relatively fixed level, with the substrate determined by the velocity. The substrate can't build up at the enzyme to avoid a bottleneck in flux, so the enzyme processes it in a steady state fashion to produce the product as determined by the Michael-Menten equation.

What happens when an inhibitor is added *in vivo*? Let's assume that the enzyme is running at $v = V_m/2$. How might in vivo inhibition plots look at constant velocity (for example, $v=V_m/2$) when both I and S can vary and in which S for an enzyme in the middle of a pathway is determined by v?

The equations and graph below show the ratio of S/K_m vs I/K_{ix} for inhibition at constant v, a condition encountered when an enzyme in a metabolic pathway is subject to flux controls imposed by the entire pathway. The x-axis reflects the ratio of inhibitor concentration to its inhibition constant. Likewise, the y-axis reflects the relative amount of substrate compared to its K_m. The graph for *in vivo* competitive inhibition is linear, but it "blows up" for uncompetitive inhibition, as shown in Figure 6.4.15.

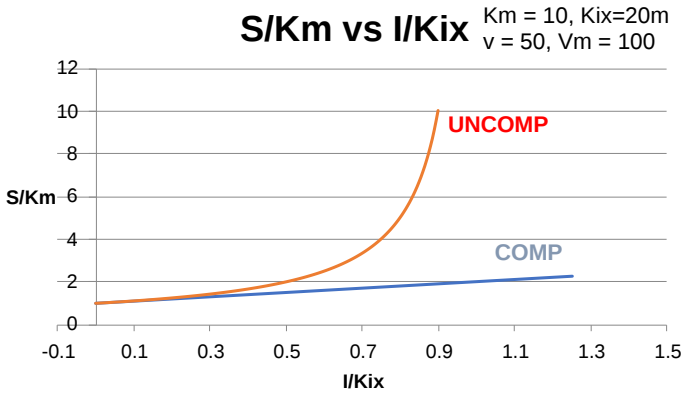


Figure 6.4.15: S/K_m vs $1/K_{ix}$ plots for competitive and uncompetitive inhibition

Here are derivations used to produce the graphs in Figure 6.4.13

? A derivation of the competitive and uncompetitive inhibition in vivo when v = constant

Here it is!

Derivation

Competitive Inhibition at constant velocity v:

Let's start with the equation of competitive inhibition.

$$\mathrm{v} = \frac{\mathrm{V_M S}}{\mathrm{K_M}\left(1+\frac{\mathrm{I}}{\mathrm{K_{is}}}\right)+\mathrm{S}} \tag{6.4.10}$$

Let

$$\left(1+\frac{\mathrm{I}}{\mathrm{K_{is}}}\right) = \mathrm{y} \tag{6.4.11}$$

Then,

$$\begin{gathered} v = \frac{V_M S}{K_M y + S} \\ v\left(K_M y + S\right) = V_M S \\ v K_M y + v S = V_M S \\ v K_M y = V_M S - v S = S\left(V_M - v\right) \\ K_M = \frac{S\left(V_M - v\right)}{v y} \end{gathered} \tag{6.4.12}$$

which gives:

$$\begin{gathered} \frac{S}{K_M} = \frac{v y}{\left(V_M - v\right)} = \frac{y}{\frac{V_M}{v}-1} = \frac{1+\frac{I}{K_{is}}}{\frac{V_M}{v}-1} = \frac{1}{\frac{V_M}{v}-1} + \frac{\frac{I}{K_{is}}}{\frac{V_M}{v}-1} \\ \frac{S}{K_M} = \frac{1}{\frac{V_M}{v}-1} + \left(\frac{1}{\frac{V_M}{v}-1}\right)\frac{I}{K_{is}} \end{gathered} \tag{6.4.13}$$

Note from the last equation that the graph of S/K_M vs I/K_{is} is linear (at a fixed v), as shown in the above figure.

Uncompetitive Inhibition at constant velocity v:

Let's start with the equation of uncompetitive inhibition.

$$\mathrm{v} = \frac{\mathrm{V_M S}}{\mathrm{K_M + S\left(1 + \frac{I}{K_{ii}}\right)}} \tag{6.4.14}$$

Let

$$\left(1 + \frac{1}{\mathrm{K_{ii}}}\right) = \mathrm{y} \tag{6.4.15}$$

then:

$$\begin{aligned} v &= \frac{V_M S}{K_M + Sy} \\ v\left(K_M + Sy\right) &= V_M S \\ vK_M = V_M S - vSy &= S\left(V_M - vy\right) \end{aligned} \tag{6.4.16}$$

which gives:

$$\frac{S}{K_M} = \frac{v}{V_M - vy}\left(\frac{\frac{1}{v}}{\frac{1}{v}}\right) = \frac{1}{\frac{V_M}{v} - y} = \frac{1}{v - 1 - \frac{I}{K_{ii}}} \tag{6.4.17}$$

This graph is not a linear function of I/K_{ii} as it was for in vivo competitive inhibition.

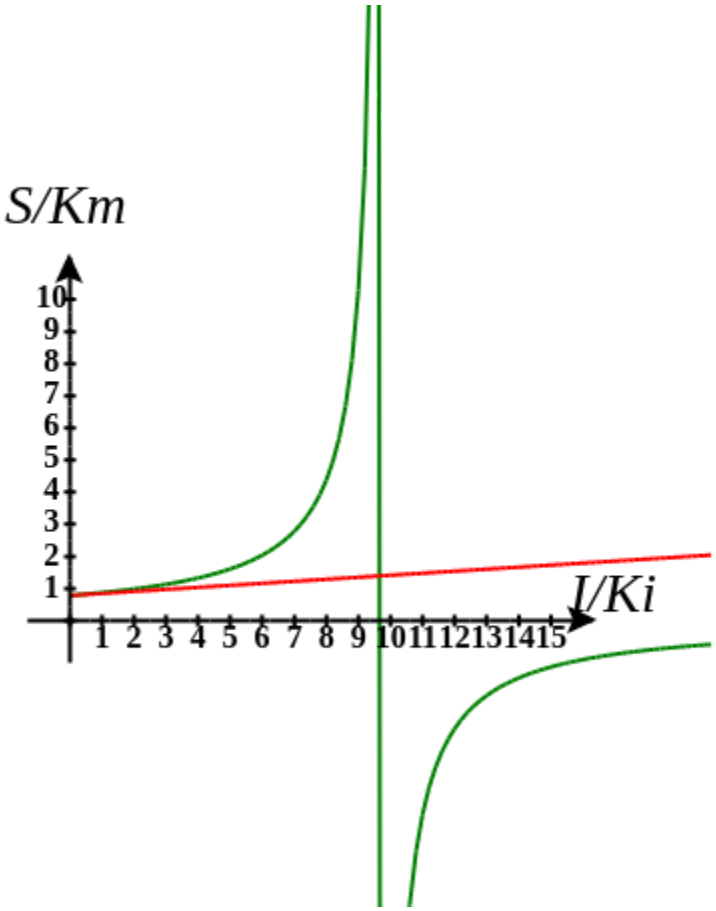
- for competitive inhibition, the graph of S/K_M is a linear function of I/K_{is}
- for uncompetitive inhibition, the graph of S/K_M is **NOT** a linear function of I/K_{ii} but rather "blows up" to infinity.

These graphs and associated equations are **dramatically** different from the very similar forms of inhibition equations and curves for in vitro inhibition at varying S and different fixed values of inhibitor. Consider the uncompetitive graph and equation. In the absence of an inhibitor, if S=Km, then Vm/v = 2, so the calculated value from the equation above of S/Km = 1. The y-intercept of the graph above is 1 for uncompetitive (and competitive) inhibition. If I is allowed to increase to a value of Kii (so I/Kii = 1), again at constant v=Vm/2, then the right-hand side goes to infinity.

In a linked series of reactions, if the middle reaction is inhibited, the substrate for that enzyme builds, whether the inhibition is competitive or uncompetitive. With competitive inhibition, the substrate concentration can be raised to meet the requirements of the enzyme. However, as the above figure shows, this can't happen for uncompetitive inhibition since as more substrate accumulates, the reaction reaches a point where the steady state is lost.

Obviously, this limiting case can't be realistically reached, but it does suggest that uncompetitive inhibitors would be more effective *in vivo* in controlling a metabolic pathway than competitive inhibitors. Cornish-Bowden argues that purely uncompetitive inhibitors are rare because of the degree of inhibition they can hypothetically produce (1986). Likewise, he suggests that medicinal chemists should synthesize uncompetitive inhibitors if their goal is to maximally inhibit a metabolic pathway under the kind of flux control described above. Although it is more difficult to synthesize a purely uncompetitive inhibitor (as it can't be easily modeled after the structure of a natural ligand that binds to the active site and are competitive inhibitors), he notes that synthesizing mixed (and noncompetitive) inhibitors whose Kii values are of reasonable size compared to their Kis values, would be one approach.

Move the sliders on the interactive graph below to show how the graphs change. Disregard the graph in the right lower quadrant as that location would require negative values of either S or K.



6.4.6: Inhibition by Temperature and pH Changes

From 0 to about 40-50 ^{o}C, enzyme activity usually increases, as do the rates of most reactions in the absence of catalysts. (Remember the general rule of thumb that reaction velocities double for each increase of 10 ^{o}C.). At higher temperatures, the activity decreases dramatically as the enzyme denatures. These features are illustrated in Figure 6.4.16.

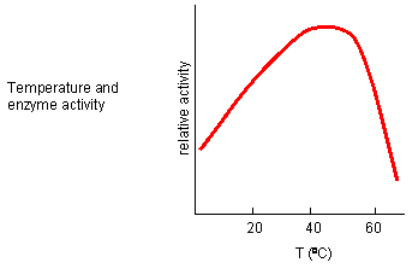


Figure 6.4.16: Temperature effects on enzyme activity

Likewise, pH has a marked effect on the velocity of enzyme-catalyzed reactions as illustrated in Figure 6.4.17.

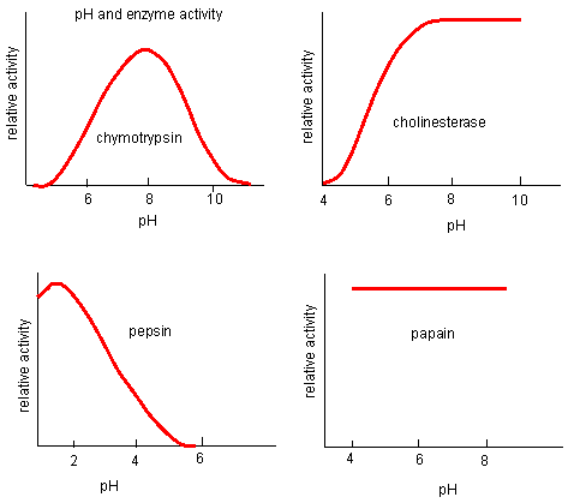


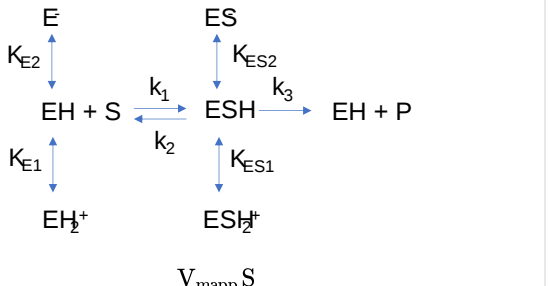
Figure 6.4.17: pH effects on enzyme activity

Think of all the things that pH changes might affect. It might

- affect E in ways to alter the binding of S to E, which would affect Km
- affect E in ways to alter the actual catalysis of bound S, which would affect kcat
- affect E by globally changing the conformation of the protein
- affect S by altering the protonation state of the substrate

The easiest assumption is that key side chains necessary for catalysis must be in the correct protonation state. Thus, a sidechain, with an apparent pKa of around 6, must be deprotonated for optimal activity of trypsin, which shows increases in activity with pH centered at pH 6. Which amino acid side chain would be a likely candidate?

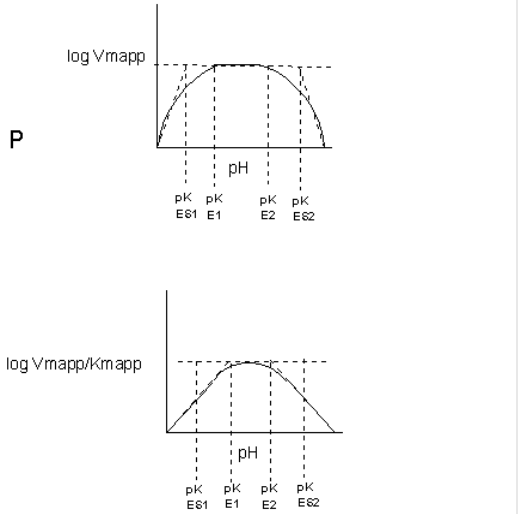
Table 6.4.1 below shows how pH effects on enzyme kinetics can be modeled at the chemical and mathematical level.



$$v_0 = \frac{V_{mapp}\,S}{K_{m\ app}} + S$$

$$V_{m\,app} = \frac{V_m}{1 + \frac{H^+}{K_{ES1}} + \frac{K_{ES2}}{H^+}} \qquad (6.4.18)$$

$$K_{m\,app} = \frac{K_m\left(1 + \frac{H^+}{K_{ES1}} + \frac{K_{ES2}}{H^+}\right)}{1 + \frac{H^+}{K_{ES1}} + \frac{K_{ES2}}{H^+}}$$



6.4.7: Summary

Chapter Summary: Enzyme Inhibition – From Irreversible Covalent Modifications to Reversible Inhibition Mechanisms

This chapter explores the diverse strategies by which enzymes can be inhibited, emphasizing the molecular basis of both irreversible and reversible inhibition. Understanding these mechanisms is crucial for appreciating how enzyme activity is regulated

in biological systems and for the design of therapeutic inhibitors.

1. **Irreversible Covalent Inhibition**
   The chapter begins by discussing how enzyme activity can be permanently abolished through irreversible modifications. These modifications, whether induced by extremes of pH, high temperature, or covalent binding of small molecules (e.g., iodoacetamide modifying critical cysteine residues), result in structural changes that render the enzyme inactive. Techniques to identify essential catalytic side chains using protecting groups are also introduced.
2. **Reversible Inhibition Mechanisms**
   The discussion then shifts to reversible inhibition, which is subdivided into several mechanistic types:
   - **Competitive Inhibition:**
     Substrate and inhibitor compete for the same active site on the enzyme. Kinetic analysis shows that while VMV_M remains unchanged, the apparent K_M increases (i.e., $K_{M\ app}=K_M(1+I/K_{is})$. Graphical representations (e.g., Lineweaver-Burk plots) illustrate that the lines intersect on the y-axis.
   - **Uncompetitive Inhibition:**
     The inhibitor binds exclusively to the enzyme–substrate complex, reducing both V_M and K_M proportionally, which is reflected as parallel lines in Lineweaver-Burk plots.
   - **Noncompetitive and Mixed Inhibition:**
     When an inhibitor binds both to free enzyme and the ES complex (with equal or different affinities), the kinetic effects differ. In pure noncompetitive inhibition, V_M decreases without changing K_M, while mixed inhibition shows changes in both parameters. These distinctions are crucial for understanding enzyme regulation and designing inhibitors.
3. **Graphical and Mathematical Analysis of Inhibition**
   The chapter emphasizes the use of various kinetic plots (e.g., Michaelis-Menten, Lineweaver-Burk, and progress curves) to analyze how inhibitors affect enzyme kinetics. Interactive graphs and simulations illustrate how changing inhibitor concentrations and their dissociation constants (e.g., K_{is} and K_{ii}) influence parameters like K_M and V_M, as well as the overall reaction velocity.
4. **In Vivo versus In Vitro Considerations**
   The text contrasts inhibition studies conducted in vitro—where enzyme and substrate concentrations are controlled—with the complex, flux-driven environment in vivo. It discusses how metabolic pathways influence enzyme kinetics and how different inhibition mechanisms may have distinct effects under cellular conditions. For example, while competitive inhibitors can be overcome by increasing substrate concentration in vitro, uncompetitive inhibitors may exert more pronounced effects in vivo by preventing the accumulation of substrate.
5. **Environmental Effects on Enzyme Activity**
   Finally, the chapter reviews how external factors such as temperature and pH influence enzyme activity. Moderate increases in temperature typically enhance reaction rates, whereas excessive heat leads to irreversible denaturation. Similarly, pH changes can alter the protonation state of key catalytic residues, thereby affecting both substrate binding and catalytic turnover.

In summary, this chapter provides a comprehensive understanding of enzyme inhibition by detailing both irreversible covalent modifications and reversible inhibition mechanisms. It integrates mathematical derivations with graphical analyses and computational models to illustrate how inhibitors modulate enzyme kinetics. These insights are essential for the rational design of inhibitors in drug discovery and for a deeper understanding of metabolic regulation in living systems.

6.05A: Enzyme Reaction Mechanisms - Arrow Pushing

Learning Goals (ChatGPT o1, 1/30/25)

1. **Understand the Fundamentals of Arrow Pushing in Enzyme Mechanisms:**
   - Define arrow pushing and its significance in illustrating electron flow during enzyme-catalyzed reactions.
   - Explain how arrow pushing provides insight into reaction mechanisms and intermediate states.
2. **Analyze Mechanistic Steps in Enzymatic Reactions:**
   - Identify and interpret key mechanistic steps in enzyme-catalyzed reactions using arrow pushing diagrams.
   - Evaluate the role of catalytic residues and cofactors in facilitating chemical transformations.
3. **Apply Arrow Pushing Techniques to Reaction Mechanisms:**
   - Develop the ability to construct and annotate arrow pushing diagrams for complex enzymatic reactions.
   - Practice using arrow pushing to predict reaction outcomes and intermediate species.
4. **Critically Evaluate Mechanism Diagrams:**
   - Assess the clarity and accuracy of enzyme reaction mechanisms depicted through arrow pushing.
   - Compare different mechanistic representations to determine the most plausible reaction pathway.
5. **Integrate Biochemical Concepts with Mechanistic Analysis:**
   - Connect the principles of enzyme structure and function with the detailed steps illustrated in arrow pushing diagrams.
   - Apply theoretical knowledge to real-world biochemical problems by analyzing enzyme reaction mechanisms.
6. **Develop Problem-Solving Skills in Enzymology:**
   - Enhance critical thinking by troubleshooting common mistakes and misconceptions in arrow pushing.
   - Use mechanistic insights to propose modifications or improvements in enzyme function based on structural changes.

prompt: Write a series of learning goals for the following web page. The page is designed for junior and senior biochemistry majors.

KristenProcko and HenryJakubowski have written this chapter section.

We can apply what we learned about catalysis by small molecules (e.g., acids and bases) to enzyme-catalyzed reactions. To understand the mechanism of an enzyme-catalyzed reaction, we try to alter as many variables, one at a time, and ascertain the effects of the changes on the activity of the enzyme. Kinetic methods can be used to obtain data, from which inferences about the mechanism can be made. Crystal structures of the enzyme in the presence and absence of a competitive inhibitor give abundant information about possible mechanisms. It is amazing, however, how much information about enzyme mechanism can be gained even if all you have is a blender, a stopwatch, an impure enzyme, and a few substrates and inhibiting reagents.

6.05A.1: Introduction

Almost every chemical reaction in the biological world is catalyzed by protein enzymes. The human genome encodes over 20,000 proteins, thousands of which are enzymes. The total number of different enzymes in the biosphere must be staggering. Yet simultaneously, all of these enzymes catalyze different sets of similar reactions. To bring order to the world of enzyme catalysis, the IUBMB has classified enzymes based on the type of chemical reactions they catalyze. There are seven main categories, as shown in the expandable Table 6.5.1 below. Each reaction type is given a four-digit **Enzyme Commission** number. For example, alcohol dehydrogenase, the enzyme that catalyzes the oxidation of ethanol, a primary alcohol, to acetaldehyde using an oxidizing agent called NAD^+, is given the enzyme commission number EC 1.1.1.1. Other enzymes that oxidize primary alcohols to aldehydes or secondary alcohols to ketones are also given the same EC number.

Subclass	Description
Oxidoreductases	Redox reactions
Transferases	Transfer/exchange of group from one molecule to another
Hydrolases	Hydrolysis reactions

Lyases	elimination forming a double bond
Isomerases	conversions of geometric, stereo- or constitutive isomers
Ligases	condensation of two molecules into one
Translocases	movement of species across a semipermeable membrane

Table 6.5.1: ExplorEnz database for the curation and dissemination of the International Union of Biochemistry and Molecular Biology (IUBMB) Enzyme Nomenclature. (Source: https://academic.oup.com/nar/article...1/D593/1000297)

Enzymes with the same or similar EC numbers probably have similar reaction mechanisms. Throughout this book, we will explore the reaction mechanisms of many enzymes, but we can't and shouldn't explore all of them. You can take your acquired understanding of the reaction mechanism for key representative enzymes and apply it to others. Of course, experimental evidence is needed to validate a given mechanism. This chapter section will focus on the mechanisms of a few transferases (EC2) and hydrolases (EC3) as prototypical examples.

The rules for electron pushing in biochemical mechanisms mirror those from organic chemistry. However, because of the length of some biochemical mechanisms, abbreviated mechanisms are often accepted and will be presented in the literature and textbooks. This section aims to review arrow pushing by presenting some simple biochemical mechanisms and to familiarize you with acceptable alternative ways to show arrow pushing.

6.05A.2: Class I Methyltransferases

Coenzymes are organic molecules that participate in some enzyme-catalyzed reactions (see Section 6.8 for a detailed discussion). These "enzyme helpers" often impart reactivity that an enzyme would not have. We'll begin our investigation with a mechanism catalyzed by class I methyltransferases using the coenzyme S-adenosylmethionine (SAM). SAM is formed from the reaction of methionine with ATP, resulting in the positively charged sulfur shown in Figure 6.5.1. The blue methyl group attached to the sulfur is very electrophilic due to the sulfur cation and is transferred to a nucleophilic substrate.

The reaction in Figure 6.5.1 shows the SAM-promoted conversion of norepinephrine to epinephrine. It is catalyzed by the enzyme phenylethanolamine N-methyltransferase (EC: 2.1.1.28). The nucleophilic nitrogen atom in norepinephrine is brought closer to the electrophilic methyl group when it binds to the enzyme. In an S_N2 reaction, nitrogen attacks, and we show a second arrow to keep track of the electrons from the carbon-sulfur bond, becoming a lone pair on the sulfur atom.

An amine has a pK_a close to 30, but a protonated amine has a pK_a around 8–10. Specifically, the conjugate acid of epinephrine has a pK_a near 8. Therefore, a protonated amine is a reasonable product to show at physiological pH (near 7.4) following the attack step. However, biochemical products are often shown uncharged when depicting an overall reaction. Therefore, a general base can be shown deprotonating the ammonium ion. The general base could be an amino acid side chain or a general base within the buffer components contained in a cell; we will represent the general base as B: here.

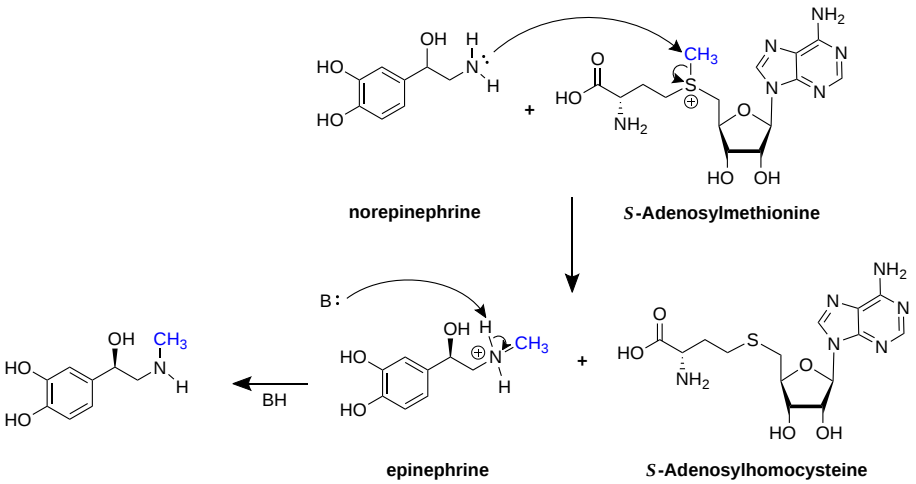


Figure 6.5.1: Class I methyltransferase mechanism with substrate norepinephrine

To simplify the mechanism in Figure 6.5.1, biochemists may abbreviate the arrow-pushing steps. The one-step deprotonation in Figure 6.5.2 shows that the general base deprotates the amine. The electrons from that bond are used to attack, and deprotonated epinephrine is produced in a single step with that arrow pushing. It is important to remember that anionic nitrogen is not made in the reaction mechanism. With pK_a values near 30, most amines are only deprotonated in practice using very strong organic bases, such as butyllithiums.

Because of this, the authors prefer the alternative arrow-pushing mechanism shown at the bottom of Figure 6.5.2. Partial bond formation between the nitrogen and the methyl group must occur before a biological general base can deprotonate norepinephrine, so showing the arrow coming from the lone pair on nitrogen is perhaps a better depiction of the biological reality. The arrow pushing shown in each methyltransferase mechanism is acceptable, and you may see each in different contexts.

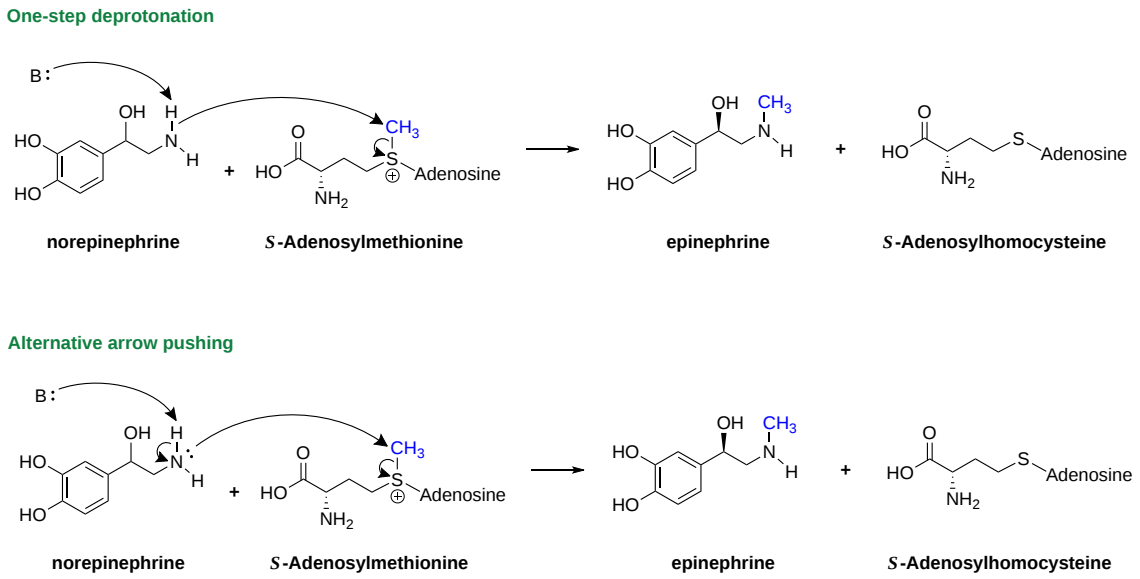


Figure 6.5.2: Class I methyltransferase mechanism alternative arrow pushing

6.05A.3: Kinases

Transfer of a phosphate group is common in biochemistry. The phosphate group is often transferred via kinase enzymes from adenosine triphosphate (ATP). The following excerpt from Chemistry LibreTexts describes the phosphate transfer mechanism:

In a phosphate transfer reaction, a phosphate group is transferred from a **phosphate group** donor molecule to a phosphate group acceptor molecule, as shown in Figure 6.5.3.

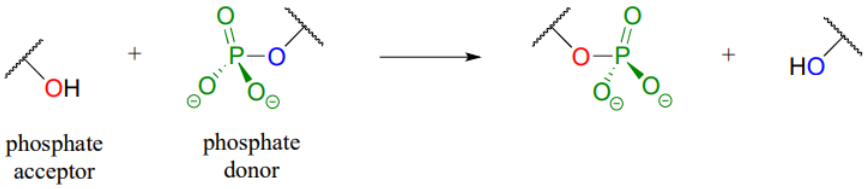


Figure 6.5.3: Phosphate transfer to an acceptor

A very important aspect of biological phosphate transfer reactions is that the electrophilicity of the phosphorus atom is usually enhanced by the Lewis acid (electron-accepting) effect of one or more magnesium ions. Phosphate transfer enzymes generally contain a Mg^{2+} ion bound in the active site in a position where it can interact with non-bridging phosphate oxygens on the substrate (Figure 6.5.4). The magnesium ion pulls electron density away from the phosphorus atom, making it more electrophilic.

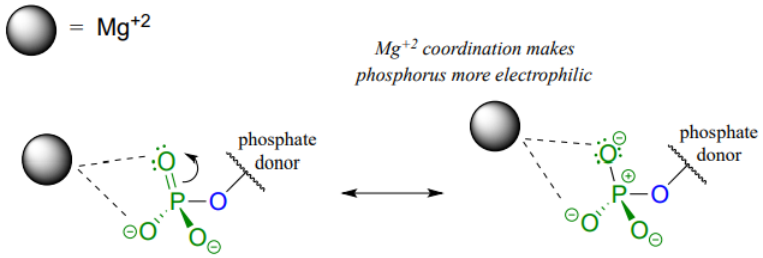


Figure 6.5.4: Phosphate interaction with magnesium ion

Without this metal ion interaction, phosphate is a poor electrophile, as the negatively charged oxygens shield the phosphorus center from attack by a nucleophile.

Note: For simplicity and clarity, we may omit the magnesium ion or other active site groups interacting with phosphate oxygens in some of the following figures. However, these interactions play an integral role in phosphate transfer reactions.

Mechanistically speaking, a phosphate transfer reaction at a phosphorus center can be thought of as much like a S_N2 reaction at a carbon center. Like in an SN2 reaction, the nucleophile in a phosphoryl transfer approaches the electrophilic center opposite the leaving group from the back side, as shown in Figure 6.5.5.

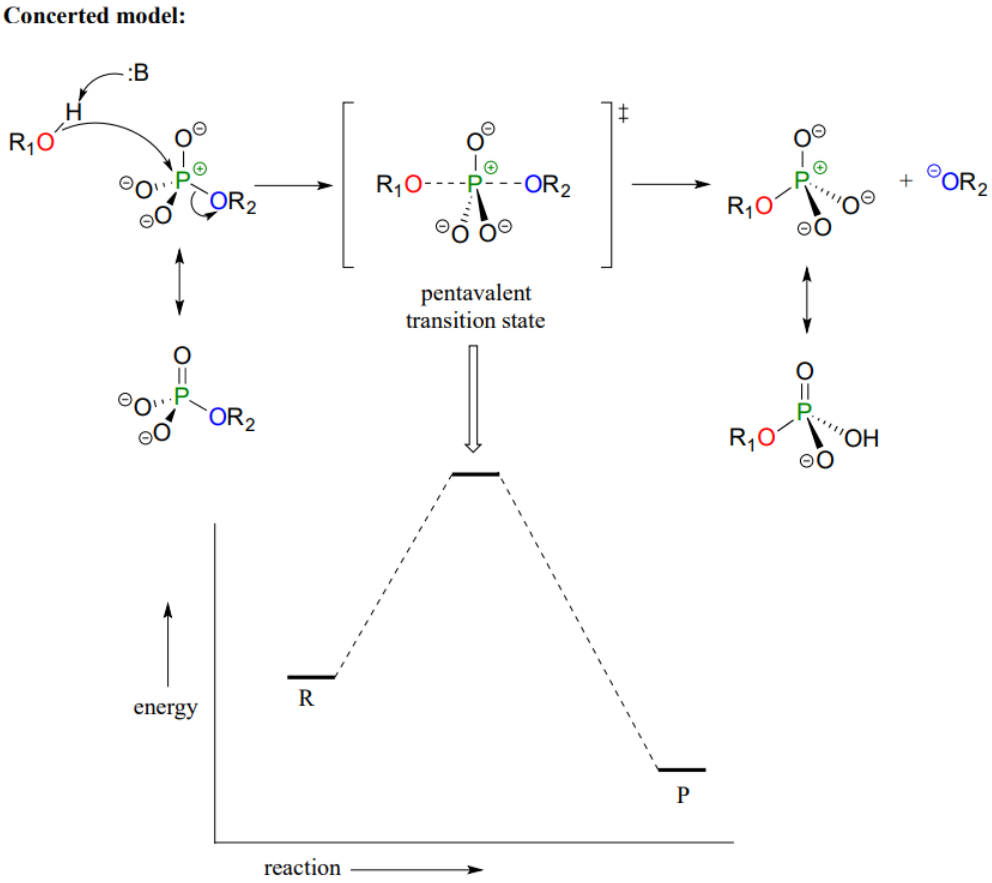


Figure 6.5.5: Phosphate transfer mechanism and energy diagram

As the nucleophile gets closer and the leaving group begins its departure, the bonding geometry at the phosphorus atom changes from tetrahedral to trigonal bipyramidal at the pentavalent (5-bond) transition state (Figure 6.5.6). As the phosphorus-nucleophile bond gets shorter and the phosphorus-leaving group bond gets longer, the bonding picture around the phosphorus atom returns to its original tetrahedral state. However, the stereochemical configuration has been 'flipped' or inverted.

In the trigonal bipyramidal transition state, the five substituents are not equivalent: the three non-bridging oxygens are said to be equatorial (forming the base of a trigonal bipyramid). In contrast, the nucleophile and the leaving group are said to be apical (occupying the tips of the two pyramids).

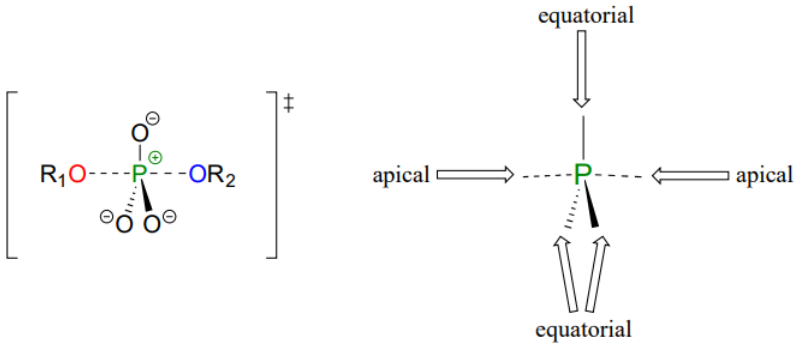


Figure 6.5.6: Transition state of the phosphate transfer reaction

Although stereochemical inversion in phosphoryl transfer reactions is predicted by theory, the fact that phosphoryl groups are achiral made it impossible to observe the phenomenon directly until 1978, when a group of researchers was able to synthesize organic phosphate esters in which stable oxygen isotopes ^{17}O and ^{18}O were specifically incorporated (Figure 6.5.7). This created a chiral phosphate center.

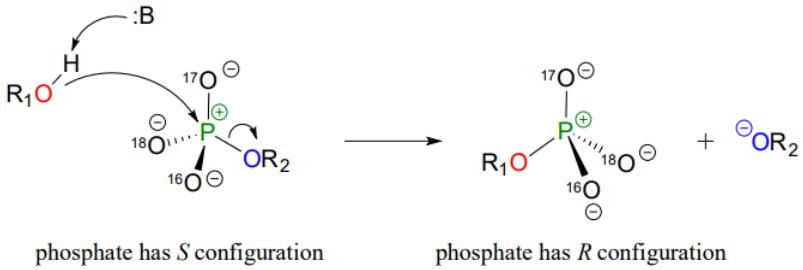


Figure 6.5.7: Inversion of stereochemistry in the phosphate transfer reaction

Subsequent experiments with phosphoryl transfer-catalyzing enzymes confirmed that these reactions proceed with stereochemical inversion. (Nature 1978 275, 564; Ann Rev Biochem 1980 49, 877). (The previous excerpt has been adapted from Chemistry LibreTexts.)

We should note that although the charge-separated resonance form shown above contributes to the structure and, therefore, contributes to the resonance hybrid, the phosphate group is almost always shown with a double-bonded oxygen atom to the phosphorous. Note that we have also depicted the mechanism below with our preferred biochemical arrow pushing, where oxygen attack at phosphorous facilitates the deprotonation of the alcohol (R_1OH) by a general base.

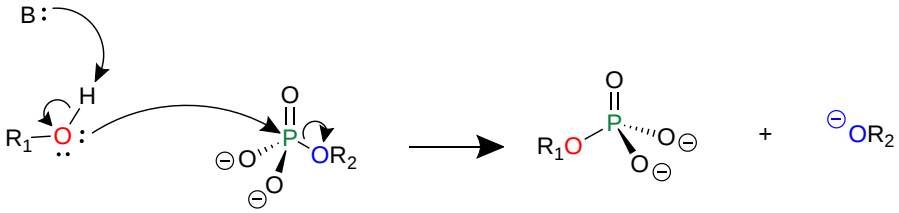


Figure 6.5.8: Common arrow pushing for phosphate transfer

ATP would be an extremely poor electrophile with a formal charge of negative four. Therefore, the phosphoryl transfer from ATP, shown for hexokinase (EC 2.7.1.1) in Figure 6.5.9, requires the binding of a magnesium ion with ATP. The magnesium ion partially neutralizes the negative charge, allowing for nucleophilic attack by the oxygen atom of glucose.

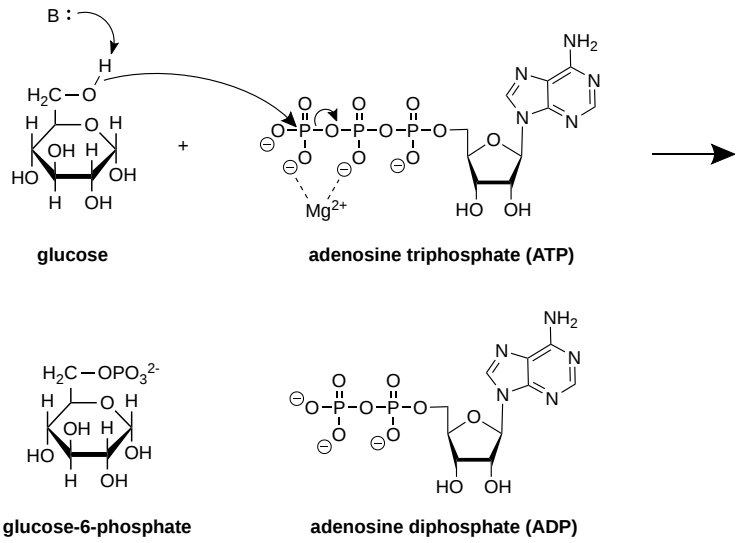


Figure 6.5.9: Hexokinase mechanism with phosphate transfer from ATP

A crystal structure of the enzyme was solved with the substrate, glucose-6-phosphate, bound. The magnesium cofactor and a non-hydrolyzable analog of ATP also occupy the active site. In the substrate analog, ANP, the oxygen of the terminal phosphoanhydride bond of ATP is replaced with a nitrogen atom. It allows us to view probable interactions between the enzyme and the substrate analog of the catalytic complex. In the iCn3D image in Figure 6.5.10, ANP and glucose are bound in the active site. The magnesium ion is shown in green.

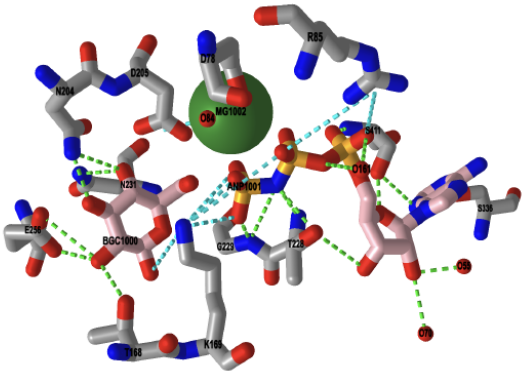


Figure 6.5.10: Interactive iCn3D image of the catalytic complex of human glucokinase (a hexokinase isoform, 3FGU). (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?EajeSCpF8GM16HYq5

6.05A.4: Serine and other Endoproteases

The mechanism for chymotrypsin, a prototypical serine protease, will be presented to explore the different types of acceptable arrow pushing one can show for this nucleophilic acyl substitution mechanism. Only the relevant peptide bond will be shown in the mechanism. For arrow pushing, it is only necessary to focus on this small portion of the molecule shown in Figure 6.5.15.

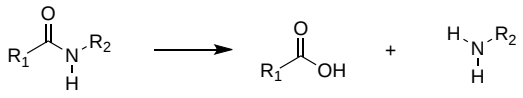


Figure 6.5.15: Abbreviated chymotrypsin peptide cleavage reaction

The active site of chymotrypsin contains a catalytic triad, three amino acids working together to carry out the reaction that cleaves the peptide bond. The amino acids involved are the aspartate, histidine, and serine residues mentioned earlier (Figure 6.5.16).

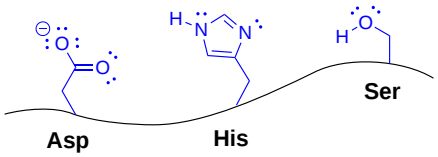


Figure 6.5.16: Serine protease catalytic triad

The deprotonated aspartate side chain increases histidine's basicity, allowing it to accept a proton from serine in the catalytic mechanism. In some mechanisms, Asp is shown accepting a proton from histidine; however, simplified arrow pushing can be shown without Asp acting as a proton acceptor, which is how the mechanism will be represented here.

In the first stage of the mechanism, histidine deprotonates serine, which acts as a nucleophile and attacks the partially electropositive carbon atom of the carbonyl functional group (Figure 6.5.17). In the simplest form of arrow pushing shown, histidine deprotonates the nucleophilic using the lone pair on nitrogen, and the electrons from the hydrogen-oxygen bond are shown attacking. The carbonyl double bond breaks, shifting the electrons onto oxygen.

This forms a tetrahedral sp^3 hybridized carbon atom from the sp^2 hybridized carbonyl group, which is why it is called a tetrahedral intermediate.

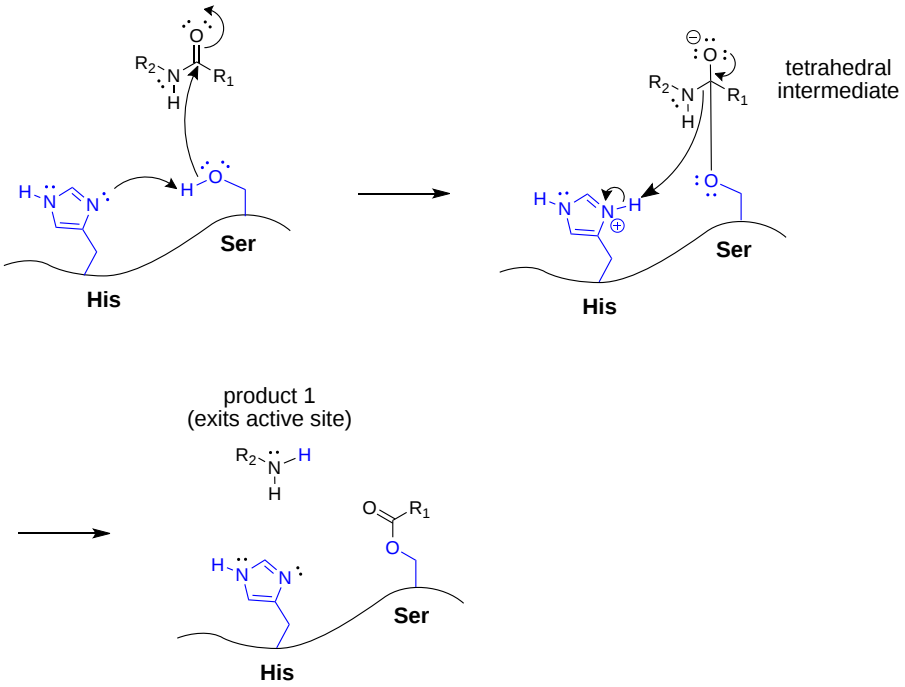


Figure 6.5.17: First stage of the chymotrypsin mechanism

Using the simple arrow-pushing model again, the electrons from the negatively charged tetrahedral intermediate reform a double bond, kicking out the amine leaving group, which accepts an H+ from protonated histidine, neutralizing the charge on histidine. The arrow from the N-H bond neutralizes the positive charge on histidine; please note that this arrow is essential to show. In all deprotonations, it is conventionally to show electrons from a breaking bond becoming a lone pair on the atom receiving them.

Another acceptable form of arrow pushing for this stage of the reaction requires more arrows but better depicts how the enzyme is interacting with the substrate in the active site (Figure 6.5.18. Because an active site often uses entropy reduction, bringing substrates close together in a reactive orientation, the lone pairs on heteroatoms already interact favorably to form the new bond. Additionally, in the first mechanism, it almost appears that a serine alkoxide attacks the carbon of the peptide bond and that R-NH^- (a poor leaving group) departs the molecule, picking up a proton after the bond breaks.

Therefore, these subtleties are considered in this second mechanism option. In the first step, the proton on serine is deprotonated by histidine, and those electrons are pushed toward the serine oxygen atom. The serine oxygen is positioned close to the carbonyl carbon of the amide bond, and an arrow originating from the serine lone pair depicts the attack. This type of arrow pushing implies that the attack and deprotonation steps are in concert.

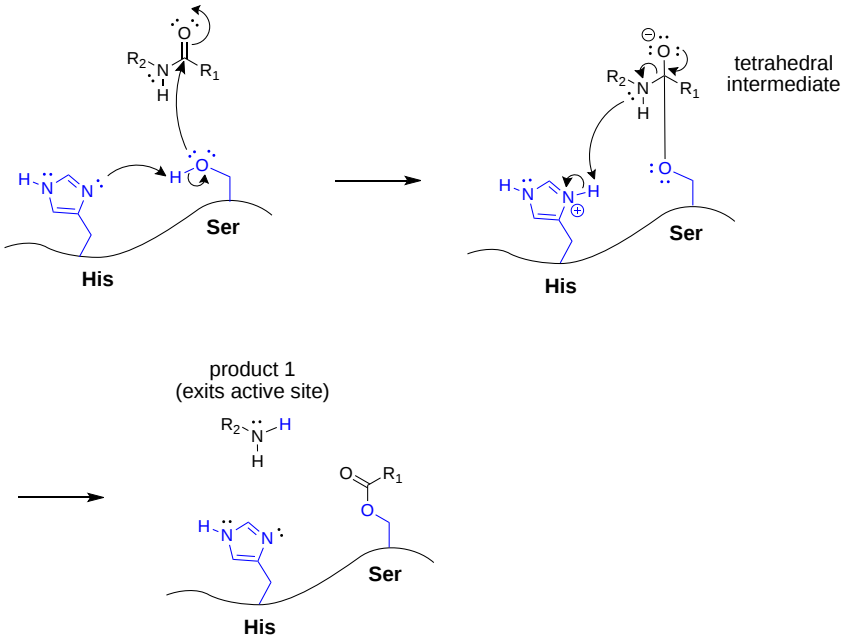


Figure 6.5.18: Alternate arrow pushing for the first stage of the chymotrypsin mechanism

In the second step, the electrons that push down from the negatively charged oxygen atom break the N-H bond, becoming a lone pair on oxygen. Simultaneously, the lone pair already present on nitrogen is shown deprotonating histidine.

It is important to note that because biochemical mechanisms are often lengthy, they may be shown with the tetrahedral intermediate omitted. This arrow pushing for the first step, which may be shown using either convention described above, shows the serine oxygen attacking and the amine leaving group departure in one step, as shown in Figure 6.5.19 below. Note that this abbreviated style of arrow pushing, which does not show the tetrahedral intermediate, often uses generic acids and bases, so it does not keep track of protonation states or account for the amino acid residues performing protonation and deprotonation steps (which is quite essential in the chymotrypsin mechanism).

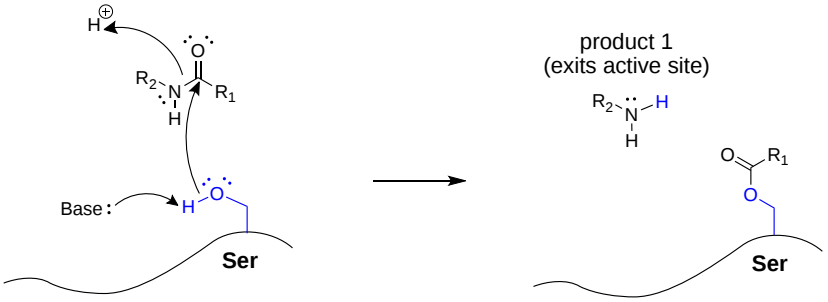


Figure 6.5.19: Abbreviated arrow pushing for the first stage of the chymotrypsin mechanism

The covalent intermediate must be released from the enzyme for chymotrypsin to catalyze another reaction. This second nucleophilic acyl substitution also proceeds through a tetrahedral intermediate; this time, water is the nucleophile, as shown in Figure 6.5.20.

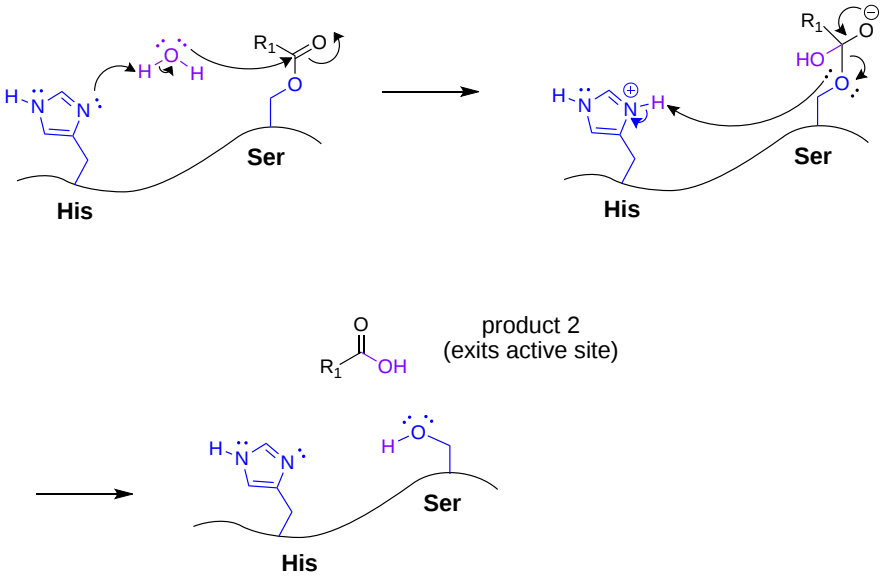


Figure 6.5.20: Second stage of the chymotrypsin mechanism

Now that we've seen the steps in detail let's put all this together to show the full mechanism for serine protease cleavage of protein, shown in Figure 6.5.21.

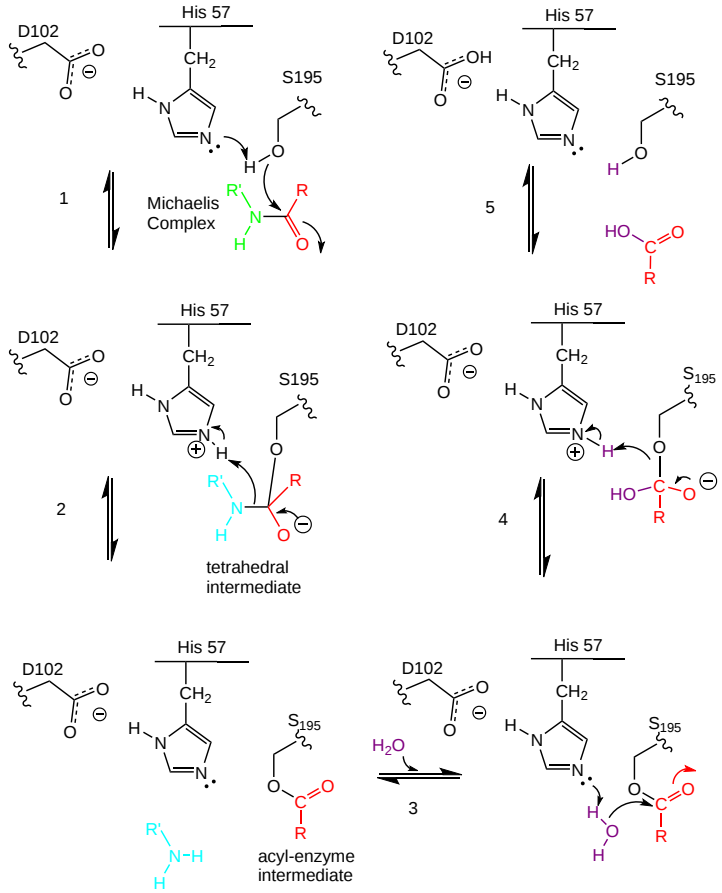


Figure 6.5.21: General mechanism of peptide bond cleavage by chymotrypsin and other serine proteases

Here a summary of what the Figure 6.5.21 mechanism shows:

- The deprotonated His 57 acts as a general base to abstract a proton from Ser 195, enhancing its nucleophilicity as it attacks the electrophilic C of the amide or ester link, creating the oxyanion tetrahedral intermediate. Asp 102 acts electrostatically to

stabilize the positive charge on the His.

- The oxyanion collapses to form a double bond between the O and the original carbonyl C, with the amine product as the leaving group. The protonated His 57 acts as a general acid, donating a proton to the amine leaving group and regenerating the unprotonated His 57.
- The mechanism repeats itself, with water as the nucleophile, which attacks the acyl-enzyme intermediate to form the tetrahedral intermediate.
- The intermediate collapses again, releasing the E-SerO—as the leaving group, which is deprotonated by His 57. This regenerates both His 57 and Ser 195 in the normal protonation state, and the enzyme is now ready for another catalytic round of activity.
- The mechanism for the first nucleophilic attack (by Ser) is the same as for the second (by water). The reverse mechanism of condensation of two peptides would be the reverse of the above mechanism and is an example of the **principle of microscopic reversibility.**

In short, many of the previously encountered catalytic mechanisms are deployed in chymotrypsin catalysis. These include nucleophilic catalysis (with the Ser 195 forming a covalent intermediate with the substrates), general acid/base catalysis with His 57, and loosely, electrostatic catalysis with Asp 102 stabilizing not the transition state or intermediate, but the protonated form of His 57. An important point to note is that His, as a general acid and base catalyst, stabilizes developing charges in the transition state and provides a path for proton transfer, without which reactions would have difficulty proceeding.

One final mechanism is at work. The enzyme does indeed bind the transition state more tightly than the substrate. Crystal structures with poor "pseudo"-substrates that get trapped as partial tetrahedrally-distorted substrates of the enzyme and with inhibitors show that the oxyanion intermediate, and hence presumably the TS, can form H-bonds with the amide H (from the main chain) of Gly 193 and Ser 195. These **cannot be made** to the trigonal, sp^2 hybridized substrate. In the enzyme alone, the hole into which the oxyanion intermediate and TS would be placed is not occupied. This oxyanion hole is occupied in the tetrahedral intermediate.

A crystal structure of a relative of chymotrypsin, **trypsin**, which cleaves after positively charged lysine and arginine side chains, has been determined with a bound transition state analog inhibitor. The transition state inhibitor is t-butoxy-Ala-Val-boro-Lys methyl ester, as shown in Figure 6.5.22

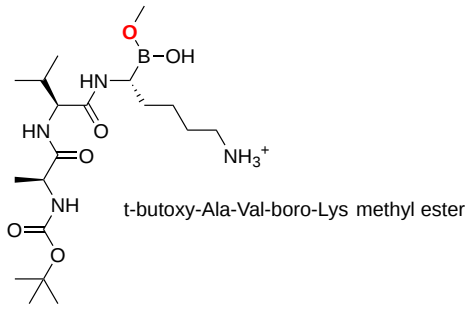


Figure 6.5.22: The transition state inhibitor of trypsin, t-butoxy-Ala-Val-boro-Lys methyl ester

Recall from introductory chemistry that neutral boron compounds like BH_3 and BF_3 are trigonal planar (sp^2) and electron deficient. Although the boron is not charged, it has a significant partial positive charge (δ^+) so it is electrophilic. The nucleophilic oxygen of Ser 195 can then attack the boron to form a tetrahedral intermediate. This intermediate is not an oxyanion; one of the attached oxygens with a δ^- charge occupies the oxyanion hole.

Figure 6.5.23 shows the active site group in trypsin interacting with a part of the transition state analog (1BTZ). The serine 195 side chain O is covalently attached to the boron, so the boron is now tetrahedral (sp^3).

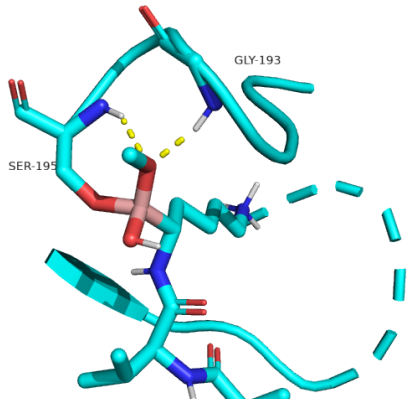


Figure 6.5.23: Stabilization of a transition state analog by the oxyanion hole

The yellow dotted lines show hydrogen bonding between the **backbone** amide hydrogens of Ser 195 and Gly 193 with the methoxy oxygen of the now tetrahedral borate transition state analog inhibitor. The boron atom is the yellow/orange sp^3 atom connected to 3 oxygen (red) atoms and one carbon (cyan) atom. Normally, the oxyanion O^- from the tetrahedral intermediate in amide or ester cleavage would occupy the oxyanion hole.

Figure 6.5.24 shows an interactive iCn3D model of the active site of the phenylethane boronic acid (PBA) complex of alpha-chymotrypsin (6cha).

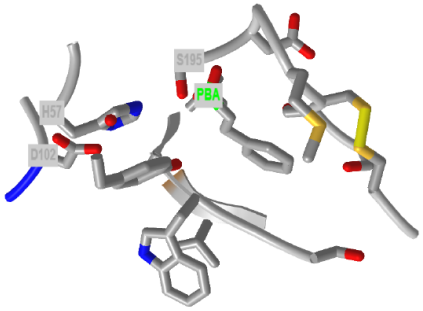


Figure 6.5.24: Active site of the phenylethane boronic acid (PBA) complex of alpha-chymotrypsin (6cha). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...qXUxbrakurYmx6

Many enzymes have active site serines which act as nucleophilic catalysts in nucleophilic substitution reactions (usually hydrolysis). One such enzyme is acetylcholinesterase, which cleaves the neurotransmitter acetylcholine in the synapse of the neuromuscular junction (Figure 6.5.25).

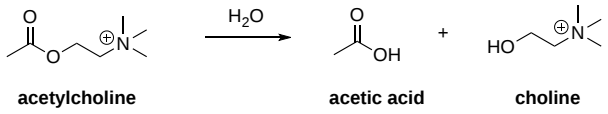


Figure 6.5.25: Reaction catalyzed by acetylcholinesterase

The neurotransmitter leads to muscle contraction when it binds its receptor on the muscle cell surface. The transmitter must not reside too long in the synapse; otherwise, muscle contraction will continue uncontrolled. To prevent this, a hydrolytic enzyme, acetylcholinesterase, a serine esterase found in the synapse, cleaves the transmitter at rates close to diffusion-controlled. Diisopropylphosphofluoridate (DIPF) also inhibits this enzyme, effectively making it a potent chemical warfare agent. Another fluoride-based inhibitor of this enzyme, sarin (Figure 6.5.26), is the most potent lethal chemical agent of this class known. Only 1 mg is necessary to kill a human being.

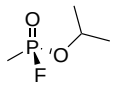


Figure 6.5.26: Sarin

Serine Protease Specificity

Serine proteases have unique specificities that allow the cleavage of target proteins after a different subset of side chains. They cleave the peptide bond on the carboxylic acid side of specific amino acids, and the specificity is determined by the size/shape/charge of the amino acid side chain that fits into the enzyme's S1 binding pocket (Figure 6.5.27). The pancreatic digestive enzymes, trypsin, chymotrypsin, and elastase are three chymotrypsin-like family members with high sequence homology. The protein cleavage sites of these enzymes vary. Trypsin cleaves proteins on the carboxylic side of basic residues, such as lysine and arginine, while chymotrypsin cleaves after aromatic hydrophobic amino acids, such as phenylalanine, tyrosine, and tryptophan. Elastase cleaves after small, hydrophobic residues like glycine, alanine, and valine. As shown in Figure 6.5.27, variations in the amino acid residues within the binding pocket of these proteases enable electrostatic interactions with the substrate and determine sequence specificity.

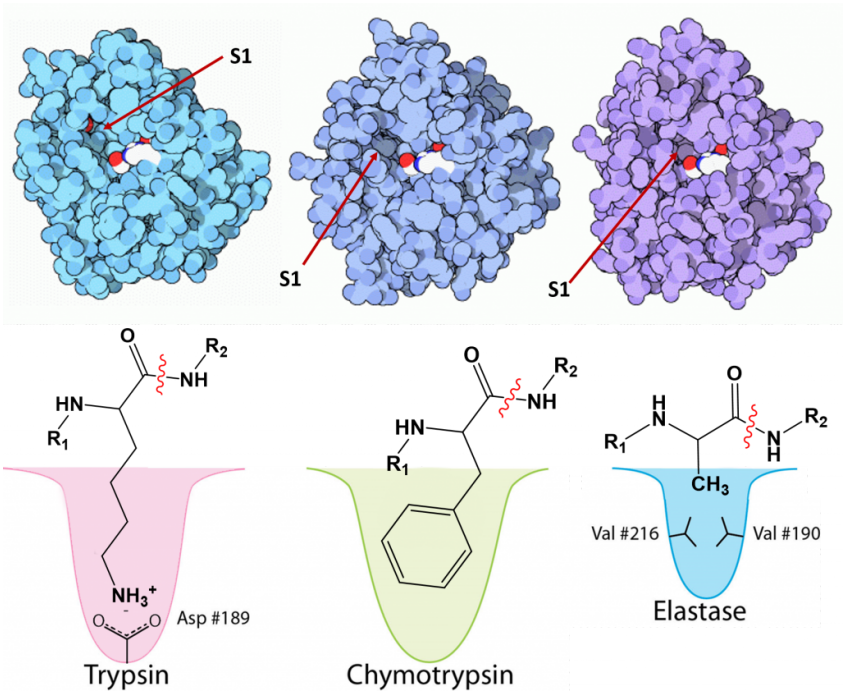


Figure 6.5.27: Substrate Specificity of Trypsin, Chymotrypsin, and Elastase. The *upper panel* shows the space-filling crystal structures of Trypsin, Chymotrypsin, and Elastase, respectively, with the S1 substrate binding pocket indicated. The *lower panel* depicts the S1 binding domains of each protease in more detail, with the indicated important amino acid R-groups. For Trypsin, an aspartate residue in the lower portion of the S1 pocket aids in electrostatic interactions with basic residues of the substrate. The Chymotrypsin S1 binding pocket is large and hydrophobic, accommodating the substrate's aromatic residues. In contrast, the Elastase S1 binding pocket is small and hydrophobic, only allowing other small and hydrophobic R-groups to dock in this location. Image modified from: Goodsell, D. (2012) Molecule of the Month, Protein Database and Aleia Kim

A schematic nomenclature developed by Berger and Schechter is often used to show the sites on the substrate (labeled P3, P2, P1, P1', P2' and P3') referring to the products made after cleavage of the peptide/protein that is cleaved between P1 and P' (the scissile bond) and the corresponding sites on the protease (S3, S2, S1, S1', S2' and S3'). This is illustrated in Figure 6.5.28

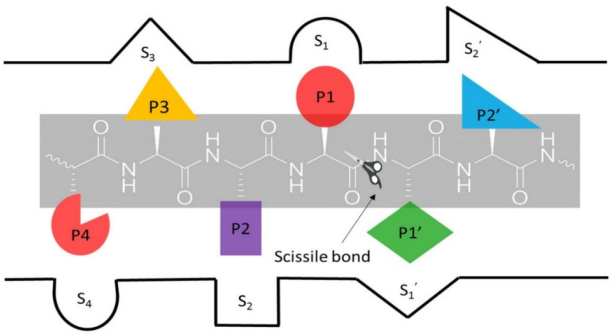


Figure 6.5.28: Nomenclature of denote key interacting groups in the enzyme (S_n) and in the substrate (P_n) around the scissile bond for protein cleavage

Serine proteases are just one type of endoprotease. However, they are highly abundant in both prokaryotes and eukaryotes. Protease A, a chymotrypsin-like protease from Streptomyces griseus, has a very different primary sequence than chymotrypsin, but its overall tertiary structure is similar to chymotrypsin. The positions of the catalytic triad amino acids in the primary sequences of the protein are very similar, indicating that the genes for the proteins diverged from a common precursor gene. In contrast, subtilisin, a serine protease from B. Subtilis, has both limited sequence and tertiary structure homology to chymotrypsin. However, when folded, it also has a catalytic triad (Ser 221 - His 64 - Asp 32) similar to chymotrypsin (Ser 195 - His 57 - Asp 102). The alignment of the core structures of chymotrypsin (5cha, magenta) and subtilisin (1sbc, cyan) are shown in Figure 6.5.29.

Figure 6.5.29: Convergent Evolution of the Serine Proteases, Chymotrypsin and Subtilisin.

The list of serine proteases is quite long. They are grouped into two broad categories - 1) those that are chymotrypsin-like and 2) those that are subtilisin-like. Though subtilisin-type and chymotrypsin-like enzymes use the exact mechanism of action, including the catalytic triad, the enzymes are otherwise unrelated to each other by sequence and appear to have evolved independently. They are, thus, an example of ***convergent evolution*** - a process where the evolution of different forms converges on a structure to provide a common function.

Proteases have multiple functions besides digestion, including degrading old or misfolded proteins and activating precursor proteins (such as clotting proteases and proteases involved in programmed cell death). Four different protease classes have been found based on residues in their active sites. Proteases can also be integral membrane proteins and carry out their activities in the membrane's hydrophobic environment. For example, aberrant cleavage of the amyloid precursor protein by the membrane protease presenilin can lead to the development of Alzheimer's Disease.

Table 6.5.4 below shows a classification of proteases based on their active site nucleophiles.

Class (active site)	Active Site Nucleophile	Location	Examples
Serine/Threonine Hydrolases	Ser/Thr	soluble	trypsin, chymotrypsin, subtilisin, elastase, clotting enzymes, proteasome
		membrane	Rhomboid family

Aspartic Hydrolases	H_2O activated by 2 Asps	soluble	pepsin, cathepsin, renin, HIV protease
		membrane	β-secretase (BACE), presenilin I, signal peptide peptidase
Cysteinyl Hydrolases	Cys	soluble	bromelain, papain, cathepsins, caspases
		membrane	?
Metallo Hydrolases	H_2O activated by 1 or 2 metal ions	soluble	thermolysin, angiotensin-converting enzyme
		membrane	S2P family
Glutamate Hydrolases	Glu	.	endolysins (fungal)
Asparagine Lysases (EC4) (elimination rx which are self-cleavage and hence not catalytic)	Asn	.	Tsh autotransporter E. Coli

Table 6.5.4: Protease classification

How do integral membrane proteases catalyze the hydrolysis (using water) of transmembrane domains in proteins, given the hydrophobic environment of the bilayer? The **rhomboid** class of membrane proteases, found in prokaryotic and eukaryotic cells, is one of the most conserved membrane proteins. Instead of using a catalytic triad, these serine proteases use a dyad of Ser 201 as a nucleophile and His 254 as a general acid/base.

The chief requirement for rhomboids' protein substrates is a transmembrane domain in the target protein. No specific amino acid sequence seems required for the specificity of one particular substrate, the Drosophila transmembrane protein Spitz, found in Golgi membranes. When this protein is cleaved, the remaining part is released as a water-soluble protein to the lumen of the Golgi, where it can eventually be released from the cell. The soluble protein fragment released from the cell contains an epidermal growth factor domain.

The structure of a rhomboid protease, GlpG (EC:3.4.21.105), from E. Coli, was determined. It is a serine protease with a catalytic dyad (Ser 201 and His 254) instead of a triad, as in most serine proteases. This transmembrane protein has six transmembrane helices. The enzyme has a polar active site at the bottom of a V-shape opening situated laterally in the membrane. The active site His and Ser residues are deep in this V-shaped cleft, well below the surface of the membrane. Access to the transmembrane strand of the protein substrate is blocked by a loop, which must be gated open to allow substrate access between the V-shaped gap between helices S1 and S3. Ser 201 (nucleophile) and His 254 (general base/acid) are essential for activity. The active site His 254 can be covalently modified with different chloromethylketone peptide derivatives. Figure 6.5.30 shows an interactive iCn3D model of the Rhomboid intramembrane protease GlpG 4QO2.

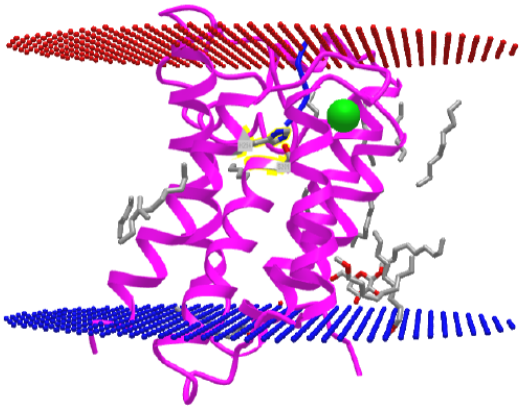


Figure 6.5.30: Rhomboid intramembrane protease GlpG (4QO2). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...ft9vFn7WqcujM7

Proteolytic enzymes (peptidases, proteases, and proteinases) are found in all living organisms, from viruses to animals and humans. Proteolytic enzymes have great medical and pharmaceutical importance due to their key role in biological processes and the life cycle of many pathogens. Proteases are extensively applied enzymes in several sectors of industry and biotechnology. Furthermore, numerous research applications require their use, including the production of Klenow fragments, peptide synthesis, digestion of unwanted proteins during nucleic acid purification, cell culturing, and tissue dissociation, preparation of recombinant antibody fragments for research, diagnostics, and therapy, and the exploration of the structure-function relationships.

Proteolytic enzymes belong to the hydrolase class of enzymes and are grouped into the subclass of peptide hydrolases or peptidases. Depending on the site of enzyme action, the proteases can also be subdivided into exopeptidases (like chymotrypsin) or endopeptidases (like carboxypeptidase A), as we will discuss next. Exopeptidases, such as aminopeptidases and carboxypeptidases, catalyze the hydrolysis of the peptide bonds near the substrate's N- or C-terminal ends, respectively. Endopeptidases cleave peptide bonds at internal locations within the peptide sequence. These differences are illustrated in Figure 6.5.31. Proteases may also be nonspecific and cleave all peptide bonds equally, or they may be highly sequence-specific and only cleave peptides after certain residues or within specific localized sequences.

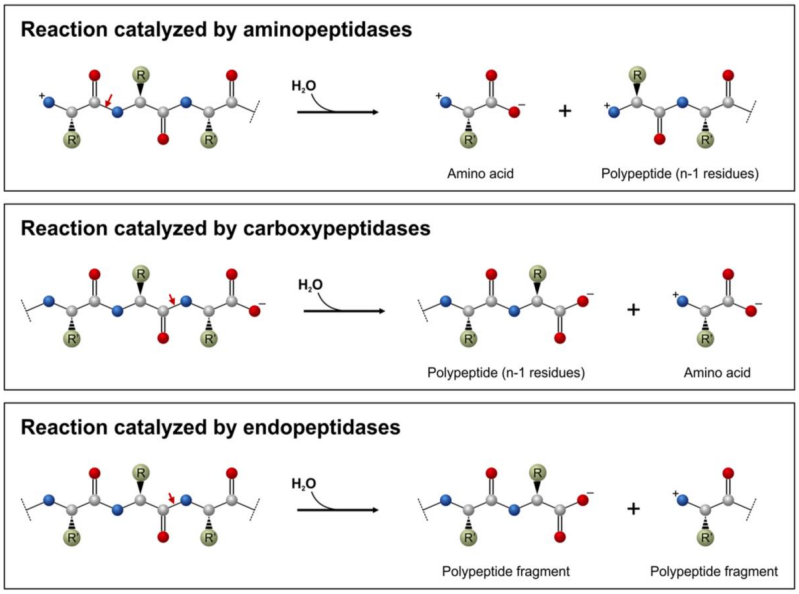


Figure 6.5.31: Peptide bond cleavage by various proteases (Figure from Mótyán, J.A., et al. (2013) *Biomolecules 3*(4), 923-942).

The action of proteolytic enzymes is essential in many physiological processes. For example, proteases function in the digestion of food proteins, protein turnover, cell division, the blood-clotting cascade, signal transduction, processing of polypeptide hormones, apoptosis, and the life-cycle of several disease-causing organisms, including the replication of retroviruses such as the human immunodeficiency virus (HIV). Due to their key role in the life cycle of many hosts and pathogens, they have significant medical, pharmaceutical, and academic importance.

About 2% of human genes encode proteolytic enzymes necessary in many biological processes. Because proteases have become important therapeutic targets, they are intensively studied to explore their structure-function relationships, investigate their interactions with substrates and inhibitors, develop therapeutic agents for antiviral therapies, improve their thermostability and efficiency, and change their specificity by protein engineering for industrial or therapeutic purposes.

6.05A.5: Carboxypeptidase A

This enzyme (EC 3.4.17.1) cleaves the C-terminal amino acid from a protein through a hydrolysis reaction. It is an **exoprotease** (not an endoprotease, which cleaves proteins internally within the sequence). Regarding selectivity toward C-terminal amino acids, its activity is increased if the C-terminal side chain group is aromatic or branched aliphatic (Phe, Tyr, Trp, Leu, or Ile).

X-ray structures of the enzyme with and without a competitive inhibitor show a sizeable conformational change at the active site when an inhibitor or substrate is bound. Without inhibitors, several waters occupy the active site. When an inhibitor (and presumably, by extension, a substrate) is bound, the water leaves (which is entropically favored), and Tyr 248 swings around from near the surface of the protein into the active site to interact with the carboxyl group of the bound molecule, a distance of motion equal to about 1/4 the diameter of the protein. This effectively closes off the active site and expels the water.

A Zn^{2+} ion is present at the active site. It is bound by His 69, His 196, Glu 72, and finally, a water molecule as the fourth ligand. A hydrophobic pocket interacting with the substrate's phenolic group accounts for the protein's specificity. In the catalytic mechanism, Zn^{2+} might have several roles. In one, it may help coordinate water and make it more nucleophilic by either polarizing the water or converting it to a more potent nucleophile OH^-. It might also stabilize developing negative charges in the transition state and in an intermediate. Two possible mechanisms have been offered.

The Water Pathway: In this proposed mechanism, water acts as a nucleophile and is deprotonated by Glu 270, acting as a general base. Glu 270, along with Zn_{2+}, helps to promote the dissociation of a proton from the bound water, making it a better nucleophile. Water attacks the electrophilic carbon of the sessile bond, forming a tetrahedral intermediate. The tetrahedral intermediate then collapses, expelling the alkoxy leaving group, which picks up a proton from Glu 270, now acting as a general acid catalyst. People used to believe that Tyr 248 acted as a general acid, but mutagenesis showed that Tyr 248 could be replaced with Phe 248 without significantly affecting the reaction rate. Figure 6.5.33 shows a simplified reaction scheme.

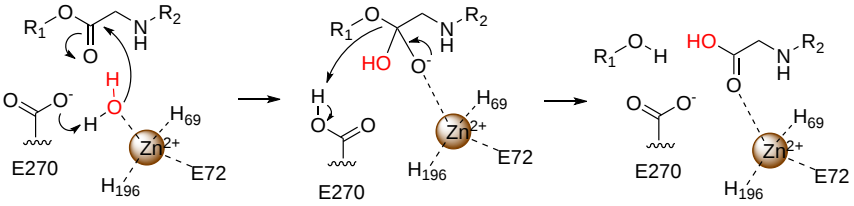


Figure 6.5.33: Water pathway mechanism for carboxypeptide A. After Wu et al. J Phys Chem B. 2010 July 22; 114(28): 9259–9267. doi:10.1021/jp101448j

Nucleophilic Pathway: In this pathway, Glu 270 is the primary initial nucleophile that forms the initial tetrahedral intermediate. The role of Zn^{2+} is in charge stabilization. This pathway is illustrated in Figure 6.5.34

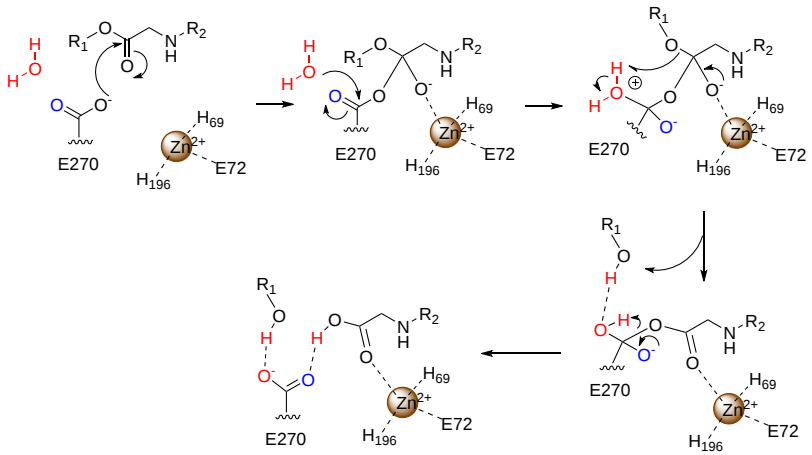


Figure 6.5.34: Nucleophilic Pathway after Wu et al. J Phys Chem B. 2010 July 22; 114(28): 9259–9267. doi:10.1021/jp101448j

Figure 6.5.35 shows an interactive iCn3D model of the active site of bovine carboxypeptidase in the absence of a substrate or inhibitor (1M4L). The Zn^{2+} ion is shown as a red sphere.

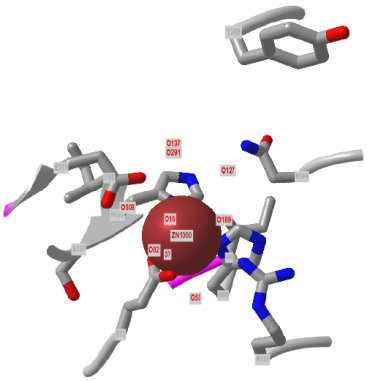


Figure 6.5.35: Bovine carboxypeptidase A (1M4L) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...sX1NYbGav1CFY6

Note how far Tyr248 is away from the active site in the model. Glu72 and Glu270 are negatively charged in the resting state of the enzyme at pH 7.5. The values are much higher (weaker acid) than the solution pK_a of the side chain of glutamic acid. Also, the water bound to the Zn^{2+} is long enough to suggest that the water is neutral and not in the form of OH^- in this form of the enzyme. If OH^- were present, the distance between it and the Zn^{2+} would be shorter due to the great electrostatic force.

Figure 6.5.36 shows an interactive iCn3D model of the active site of bovine carboxypeptidase bound to the inhibitor aminocarbonylphenylalanine (1HDU). The Zn^{2+} ion is shown as a red sphere.

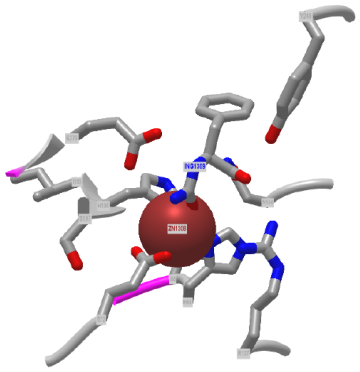


Figure 6.5.36: Bovine carboxypeptidase A bound to the inhibitor aminocarbonylphenylalanine (1HDU). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...vWt2qeZtUrwj78

Note the closer proximity of tyrosine 248 to the active site.

6.05A.6: Lysozyme

Lysozyme (EC 3.2.1.17), found in cells and secretions of vertebrates but also in viruses that infect bacteria, cleaves peptidoglycan GlcNAc (β-1,4) MurNAc repeat linkages (NAG-NAM) in the cell walls of bacteria and the GlcNAc(β-1,4) GlcNAc (poly-NAG) in chitin, found in the cells walls of certain fungi. Since these polymers are hydrophilic, the enzyme's active site would be expected to contain a solvent-accessible channel into which the polymer could bind. The crystal structures of lysozyme and complexes of lysozyme and NAG have been solved to high resolution. The inhibitors and substrates form strong H bonds and some hydrophobic interactions with the enzyme cleft. Kinetic studies using $(NAG)_n$ polymers show a sharp increase in k_{cat} as n increases from 4 to 5. The k_{cat} for $(NAG)_6$ and $(NAG\text{-}NAM)_3$ are similar. Model studies have shown that for catalysis to occur $(NAG\text{-}NAM)_3$ binds to the active site with each sugar in the chair conformation, except the fourth, which is distorted to a half chair form. This labilizes the glycosidic link between the 4^{th} and 5^{th} sugars. Additional studies show that if the sugars that fit into the binding site are labeled A-F, then because of the bulky lactyl substituent on the NAM, residues C and E cannot be NAM, which suggests that B, D, and F must be NAM residues. Cleavage occurs between residues D and E.

A review of the chemistry of glycosidic bond (an acetal) formation and cleavage shows that acids catalyze the acetal cleavage and proceed through an oxonium ion, which exists in resonance form as a carbocation. A reaction mechanism of hemiacetal/acetal formation and cleavage is illustrated in Figure 6.5.37.

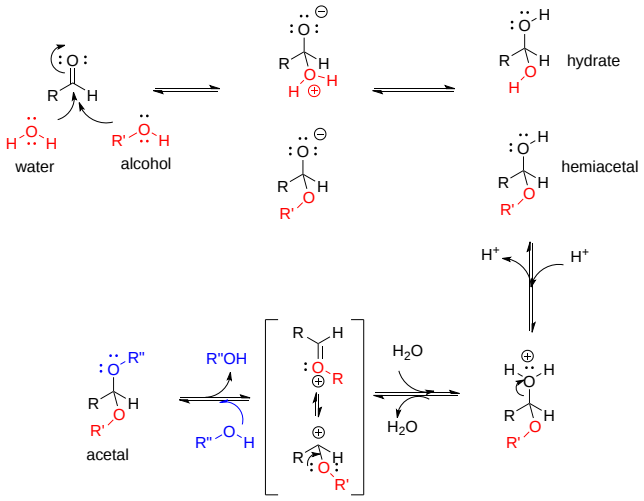


Figure 6.5.37: A reaction mechanism of hemiacetal/acetal formation and cleavage

Catalysis by the enzyme involves Glu 35 and Asp 52, which are in the active site. Polar groups surround Asp 52, but Glu 35 is in a hydrophobic environment. This should increase the apparent pK_a of Glu 35, making it less likely to donate a proton and acquire a negative charge at low pH values, making it a better general acid at higher pH values. Here is a possible general mechanism:

- binding of a hexasaccharide unit of the peptidoglycan with concomitant distortion of the NAM.
- protonation of the sessile acetal O by the general acid Glu 35 (with the elevated pK_a) facilitates cleavage of the glycosidic link and the resonance stabilized oxonium ion formation.
- Asp 52 stabilizes the positive oxonium through electrostatic catalysis. The distorted half-chair form of the NAM stabilizes the oxonium, which requires co-planarity of the substituents attached to the sp^2 hybridized carbon of the carbocation resonant form (much like we saw with the planar peptide bond).
- water attacks the stabilized carbocation, forming the hemiacetal by releasing the extra proton from water to the deprotonated Glu 35, reforming the general acid catalysis.

Part of a mechanism illustrating the roles of Glu 35 and Asp 52 is shown below in Figure 6.5.38.

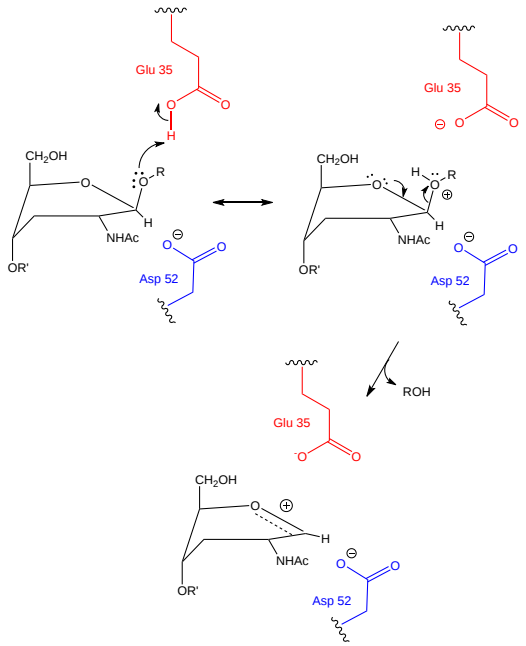


Figure 6.5.38: Mechanism of acetal cleavage by lysozyme

Binding and distortion of the D substituent of the substrate (to the half chair form as shown above) occurs before catalysis. Since this distortion helps stabilize the oxonium ion intermediate, it presumably stabilizes the transition state. Hence, this enzyme appears to bind the transition state more tightly than the free, undistorted substrate, yet another catalysis method.

pH studies show that side chains with pKa's of 3.5 and 6.3 are required for activity. These presumably correspond to Asp 52 and Glu 35, respectively. If the carboxy groups of lysozyme are chemically modified in the presence of a competitive inhibitor of the enzyme, the only protected carboxy groups are Asp 52 and Glu 35.

In an alternative mechanism, Asp 52 acts as a nucleophilic catalyst and forms a covalent bond with NAM, expelling a NAG leaving group with Glu 35 acting as a general acid as shown in Figure 6.5.39. This alternative mechanism is also consistent with other β-glycosidic bond cleavage enzymes. Substrate distortion is also important in this alternative mechanism.

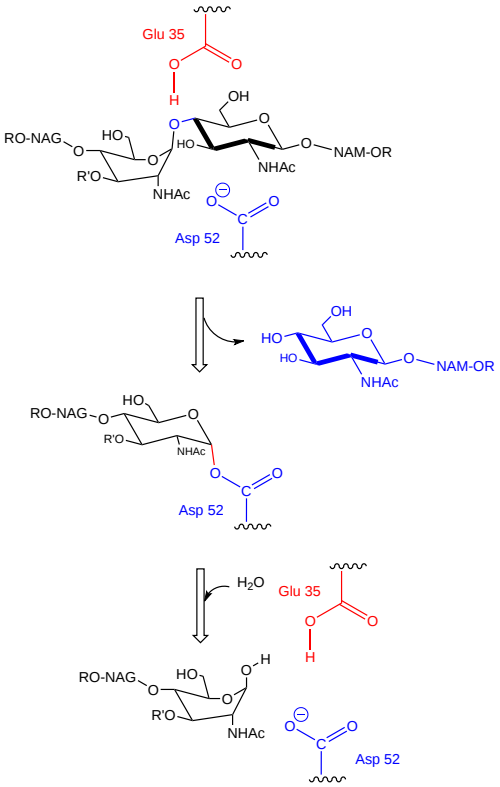


Figure 6.5.39: Alternative mechanism for lysozyme catalysis employing Asp 52 as a nucleophilic catalyst (after Vocadlo et al., Nature, **412** (2001), https://www.nature.com/articles/35090602.

Recent structural work shows that Asp 52 is involved in a strong hydrogen bond network that might preclude its ability to form a covalent bond with the glycan substrate. An earlier structure (1H6M) did show a covalent bond.

Figure 6.5.40 shows an interactive iCn3D model of the active site of hen egg white lysozyme bound to a $(NAG)_4$ glycan (7BR5). Note the positions of E35 and D52.

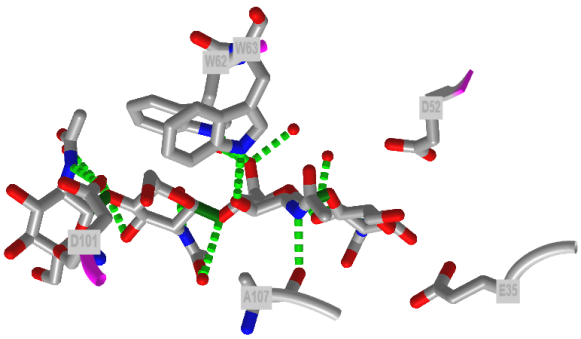


Figure 6.5.40: Interactions of hen egg white lysozyme with bound $(NAG)_4$ glycan (7BR5) (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?mnZuP4W4pfTxKRmX7

6.05A.7: Summary

The chapter introduces the essential concept of arrow pushing as a tool to depict enzyme reaction mechanisms, emphasizing its importance in visualizing the flow of electrons during catalysis. It explains that arrow pushing is not merely a drawing exercise but a method to understand how specific chemical transformations occur within enzyme active sites. The text outlines the conventions of curved arrow notation, detailing how arrows indicate the movement of electron pairs from nucleophiles to electrophiles and illustrate the formation and breaking of chemical bonds.

Key concepts include:

- **Electron Flow and Mechanistic Steps:** The chapter discusses how the movement of electrons is tracked through each step of an enzymatic reaction, from substrate binding to the formation of intermediates and eventual product release. This approach helps in rationalizing the role of catalytic residues and cofactors in lowering activation energies.
- **Interpretation of Reaction Mechanisms:** By breaking down complex reactions into individual, understandable steps, the chapter guides students in dissecting enzyme mechanisms. It shows how different reaction intermediates and transition states can be identified and how their stability is crucial for efficient catalysis.
- **Application in Biochemical Problem-Solving:** The chapter emphasizes that mastering arrow pushing enhances one's ability to predict reaction outcomes, propose alternative pathways, and understand the impact of mutations or inhibitors on enzyme activity. This skill is fundamental for advanced studies in biochemistry and for research involving enzyme kinetics and molecular enzymology.

Overall, the chapter equips junior and senior biochemistry majors with a robust framework for analyzing and depicting enzyme-catalyzed reactions, fostering a deeper understanding of the molecular underpinnings of catalysis through a detailed and systematic approach to arrow pushing.

6.05B: Enzyme Reaction Mechanisms - Quantiative Analyses of Serine Protease Catalysis

Learning Goals (ChatGPT o1, 1/30/25)

Mechanistic and Structural Insights

- **Integrate Multiple Data Types:**
  - Explain how structural (PDB data), kinetic, and thermodynamic data combine to elucidate enzyme reaction mechanisms.
- **Identify Key Catalytic Residues:**
  - Describe the roles of Ser 195, His 57 (or His 95 in some texts), and Asp 102 in the catalytic triad of serine proteases.
- **Interpret Enzyme–Substrate Interactions:**
  - Analyze how changes in the substrate (e.g., leaving group modifications) provide evidence for the formation of a covalent acyl-enzyme intermediate.

Kinetic Analysis and Reaction Mechanisms

- **Understand Kinetic Parameters:**
  - Interpret kinetic constants (kcat, Km, and kcat/Km) and relate them to specific steps in the enzyme mechanism (acylation vs. deacylation).
- **Distinguish Reaction Mechanisms:**
  - Explain the concept of a BiBi Ping Pong reaction and how kinetic bursts (rapid formation of an intermediate followed by a slower reaction phase) support this mechanism.
- **Apply Mathematical Models:**
  - Simplify complex kinetic equations (e.g., under the rapid equilibrium assumption) to describe enzyme kinetics for ester versus amide hydrolysis.

Effects of Experimental Variables

- **Impact of pH and Ionic Strength:**
  - Use pH-dependent kinetic data to identify the roles of ionizable groups (e.g., histidine as a general base) in enzyme activity.
- **Enzyme Modification Strategies:**
  - Evaluate how site-specific mutagenesis and chemical modification (e.g., using DIPF or chloromethyl ketones) confirm the roles of individual amino acids in catalysis.
- **Solvent Effects on Enzyme Activity:**
  - Compare enzyme function in aqueous versus nonaqueous (organic) solvents, including the kinetic and thermodynamic reasons for retained or altered activity in nonpolar environments.

Dynamics and Thermodynamic Contributions

- **Quantitative Analysis of Catalysis:**
  - Calculate how changes in free energy ($\Delta G‡$) reflect the stabilization of the transition state versus the destabilization of the ground state, incorporating concepts like ground-state destabilization and pre-orientation.
- **Assess the Role of Molecular Dynamics:**
  - Describe how changes in torsion angles, distances, and dihedral angles (e.g., $\chi 1$ for Ser 195) are correlated with catalysis and substrate positioning.
- **Relate Physical Chemistry to Biochemical Function:**
  - Explain how fundamental physical and chemical principles (hydrogen bonding, entropy, van der Waals interactions, and torsional strain) contribute quantitatively to the overall catalytic rate enhancement.

Broader Applications and Synthetic Utility

- **Enzyme Specificity and Selectivity:**
  - Discuss how the specificity of chymotrypsin (and other enzymes) can change depending on the solvent environment, with implications for substrate binding and competitive inhibition.
- **Practical Implications in Organic Synthesis:**
  - Evaluate how enzymes used in organic solvents serve as stereoselective, regioselective, and chemoselective catalysts, and outline potential applications in novel organic transformations.

These learning goals provide a structured framework for understanding the multifaceted nature of enzyme catalysis, integrating detailed mechanistic insights with experimental evidence and quantitative analysis.

KristenProcko and HenryJakubowski have written this chapter section.

6.05B.1: Introduction

To this point, we've presented mechanisms with the support of PDB structures alone. However, much was known about enzyme mechanisms before ready access to crystal structures in the Protein Data Bank. Systematically, the kineticists, medicinal chemists, and molecular biologists (i.e., well-trained chemists) can change:

A. the substrate - for example, changing the leaving group or substituents of a hydrolyzable substrate;
B. the pH or ionic strength - which can give data about general acids/bases in the active site;
C. the enzyme - by chemical modification of specific amino acids or through site-specific mutagenesis;
D. the solvent - an odd idea on the surface but it leads to new insights into enzyme catalysis.

We will concentrate on reaction mechanisms based on a mix of structural, kinetic, and thermodynamic data for the following enzymes to hypothesize a reaction mechanism consistent with the findings. Even with lots of data, there are often different proposed mechanisms for a given reaction. Kinetic data is vital as it can help to determine:

- the order of binding/dissociation of substrates and products;
- the rate constants for individual steps;
- and clues to the nature of catalytic groups found in the enzyme.

6.05B.2: Chymotrypsin and Other Serine Proteases

Chymotrypsin (EC 3.4.21.1), an endoprotease, cleaves an internal peptide bond after aromatic side chains by hydrolysis. It also cleaves small ester and amide substrates after aromatic residues. For example, in Figure 6.5.1, cleavage occurs on the C-terminal side of the tyrosine residue, giving two peptide fragments.

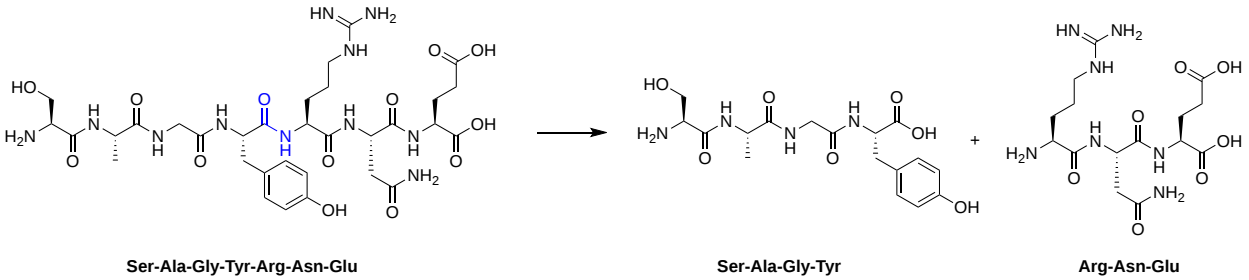


Figure 6.5.1: Chymotrypsin cleavage of an example peptide substrate

Chymotrypsin has a similar mechanism to many other **serine proteases** that use the same catalytic triad, **Ser** 195, Asp 102, and His 95, so we'll study it in significant detail. In determining the mechanism of an enzyme, you have to change an experimental variable and see how catalytic activity changes. What can be changed? It turns out everything, including the solvent! Let's explore these changes and how they affect chymotrypsin activity.

A. **Changing the substrate** (for example, changing the leaving group or acyl substituents of a hydrolyzable substrate):

In the lab, studying the enzyme using small substrate mimics of a protein is easier than using a full protein substrate. The mimics include both esters and amides. Data from the small amide and ester substrates cleavage shown in Figure 6.5.2 suggest that a

covalent intermediate is formed during chymotrypsin-catalyzed cleavage.

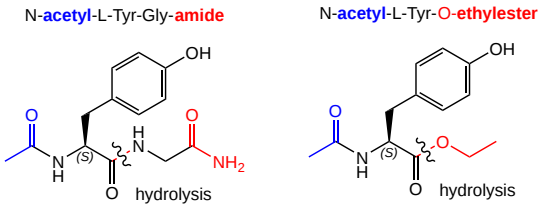


Figure 6.5.2: Small amide and ester substrates of chymotrypsin

Table 6.5.1 below shows kinetic data for the cleavage of these substrates.

Chymotrypsin substrate cleavage, 25 °C, pH 7.9			
kinetic constants	Acetyl-Tyr-Gly-amide	Acetyl-Tyr-O Ethylester	Ester/Amide
k_{cat} (s^{-1})	0.50	193	390
K_m (M)	0.023	0.0007	0.03
k_{cat}/K_m ($M^{-1}s^{-1}$)	22	280,000	12,700

Kinetic constants for chymotrypsin cleavage of N-acetyl-L-Trp Derivatives - N-acetyl-L-Trp-X		
X	k_{cat} (s^{-1})	K_m x 103 (M)
$-OCH_2CH_3$	27	0.097
$-OCH_3$	28	0.095
-p-nitrophenol	31	0.002
$-NH_2$	0.026	7.3

Table 6.5.1: Cleavage of peptides and ethylester substrate analogs by chymotrypsin

Here's how these data can be interpreted.

1. The k_{cat} and k_{cat}/K_m are larger and the K_m smaller for ester substrates compared to amide substrates, suggesting that amides are more challenging to hydrolyze (Table 2 above). This is expected given the poorer leaving group of the amide.
2. The k_{cat} for the hydrolysis of ester substrates doesn't depend on the nature of the leaving group (i.e., whether it is a poorer leaving group such as methoxy or a better leaving group such as p-nitrophenolate), suggesting that this step is not the rate-limiting step for ester cleavage. Without the enzyme, p-nitrophenyl esters are cleaved much more rapidly than methyl esters. Therefore, deacylation must be rate-limiting. But deacylation of what? If water were the nucleophile, releasing the leaving group would result in both products, the free carboxyl group and the amine, being formed simultaneously. Since they are not released simultaneously, this suggests an acyl-enzyme covalent intermediate.

A covalent intermediate can be trapped when the acyl end of the ester substrate is changed without changing the leaving group (a p-nitrophenyl group). Specifically, the deacylation of a trimethylacetyl group is much slower than that of an acetyl group. It is so slow that a ^{14}C-labeled trimethylacetyl-labeled chymotrypsin intermediate can be isolated after incubation of chymotrypsin with ^{14}C-labeled p-nitrophenyltrimethylacetate using gel filtration chromatography.

We have seen a kinetic mechanism previously consistent with these ideas before. The data suggest a mechanism based on the chemical equations shown in Figure 6.5.3:

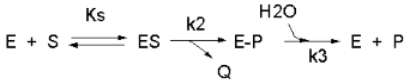


Figure 6.5.3: Chemical equations for chymotrypsin hydrolysis of a substrate involved a covalent intermediate with ping-pong kinetics.

In this reaction, a substrate S might interact with E to form a complex, which then is cleaved to products P and Q. Q is released from the enzyme, but P might stay covalently attached until it is expelled. This conforms exactly to the mechanism described above. For chymotrypsin-catalyzed cleavage, the step characterized by $\mathbf{k}_2$ is the acylation step. The step characterized by $\mathbf{k}_3$ is the deacylation step in which water attacks the acyl-enzyme to release product P (free phosphate in Lab 5). The mathematical equation for this reaction is shown below (without derivation)

$$v_0 = \frac{\left(\frac{k_2 k_3}{k_2+k_3}\right) E_0\, S}{K_S\left(\frac{k_3}{k_2+k_3}\right) + S} \tag{6.5.1}$$

For hydrolysis of ester substrates, which have better leaving groups than amides, deacylation is rate limiting ($k_3 << k_2$). For amide hydrolysis, as mentioned above, acylation can be rate-limiting ($k_2 << k_3$). From this, equation 6.5.1 can be simplified as shown in Table 6.5.2 below for ester and amide hydrolysis.

Ester hydrolysis (deacylation rate limiting, $k_3 << k_2$)	Amide hydrolysis (deacylation rate limiting, $k_2 << k_3$)
$v_0 = \frac{k_3 E_0\, S}{K_S\left(\frac{k_3}{k_2}\right) + S}$ (6.5.2)	$v_0 = \frac{k_2 E_0\, S}{K_S + S}$ (6.5.3)
$V_M = k_3 E_0$ (6.5.4)	$V_M = k_2 E_0$ (6.5.5)
$K_M = K_S\left(\frac{k_3}{k_2}\right)$ (6.5.6)	$K_M = K_S$ (6.5.7)

Table 6.5.2: Simplification of equation 6.5.1

As we saw before, for the rapid equilibrium assumption (when ES falls apart to E + S more quickly than it goes to the product, Chapter 6.3), K_M = Ks in the case of amide hydrolysis.

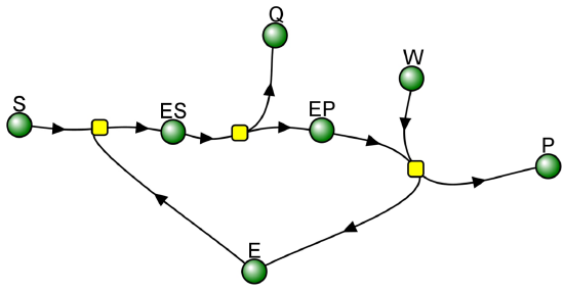
This reaction with two reactants (bi) and two products (bi), a covalent enzyme intermediate with the second reactant binding after the first product Q is released, is called a **BiBi Ping Pong** reaction. If $k_2 >> k_3$, an immediate and fast **burst** or release of product Q happens, followed by a slow release of P since the covalent E-P complex reacts with the second reactant with a small rate constant k_3. When doing initial rate Michaelis-Menten kinetics, the initial velocity of Q formation, v_0, is not $(dQ/dt)_{t=0}$, but the slower constant rate after the burst phase, which is determined by k_3, the rate of cleavage of the E-P intermediate.

The VCell computational model below shows the reaction BiBi Ping Pong reaction for the reaction involving an enzyme-P covalent intermediate. The burst phases in Q is clearly seen if you rescale the graph as described below.

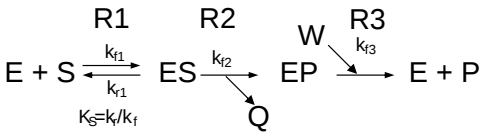
VCell MODEL

BiBi-Ping Pong_Covalent Intermediate Irreversible reaction

Vcell reaction diagram (1-way arrows defined as reversible in actual mathematical model) and chemical equation



Yellow dots: Reaction Nodes (R1, R2 and R3 left to right)



Reaction made irreversible since $k_{r2} = 0$, $k_{r3} = 0$.

Initial parameter values:

- $S_0 = 100$, $E_0 = 1$, W (water in a hydrolysis reaction) = 50 and fixed throughout
- $k_{1f} = 5$, $k_{1r} = 1$, $k_{2f} = 0.6$,
- $k_{2f} = 50$, $k_{2r} = 0$
- $k_{3f} = 0.05$, $k_{3f} = 0$

Select Load [model name] below

Load BiBiPingPongChymo

Select **Start** to begin the simulation.

Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

To see the burst phase for reaction, change the time and parameters to these values:

- set Run time to 0.3
- Select **Plot** then Update Y axis max to 2
- Click Edit Plot Species and check just P and Q
- reset

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

The burst phase is seen with ester hydrolysis as described above when $k_2 >> k_3$.

B. **Changing the pH or ionic strength** - which can give data about general acids/bases in the active site:

- a graph of k_{cat} as a function of pH indicates that a group of pK_a of approximately 6 must be deprotonated to express activity (i.e., $V_{max}/2$ is at about pH 6). This suggests that an active site histidine is necessary, which, if it must be deprotonated to express activity, must be acting as a general base.
- a graph of k_{cat}/K_m shows a bell-shaped curve indicating the necessity of a deprotonated side chain with a pK_a of about 6 (i.e., the same His above) and a group that must be protonated with a pK_a of about 10. This is an N terminal Ile in chymotrypsin, which must be protonated to form a stabilizing salt bridge in the protein. *Note*: This N-terminal Ile is actually at the 16 position in the inactive precursor of chymotrypsin (called chymotrypsinogen); upon activation of chymotrypsinogen, it loses the first 15 amino acids by selective proteolysis.

(*Note*: The PKAD is a database of experimentally measured pK_a values of protein ionizable groups. It is searchable by the PDB ID.)

C. **Changing the enzyme** - by chemical modification of specific amino acids, or through site-specific mutagenesis:

Here are some specific examples.

1. Modification of chymotrypsin (and many other proteases) with diisopropylphosphofluoridate (DIPF) modifies only one (Ser 195) of many serines in the protein, suggesting that it is hypernucleophilic and probably the amino acid that attacks the carbonyl C in the substrate, forming the acyl-intermediate. This reaction is illustrated in Figure 6.5.3. The figure also shows analogous molecules used in common insecticides, which act through a similar mechanism.

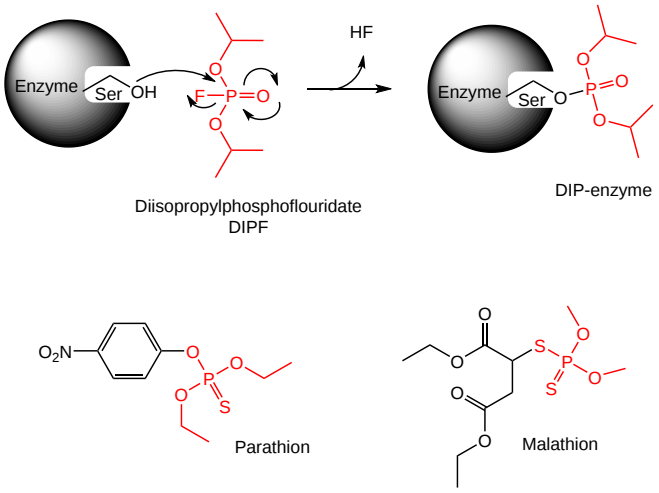


Figure 6.5.3: Mechanism of chymotrypsin inhibition by covalent modification by diisopropylphosphoflouridate.

2. Modification of the enzyme with tos-L-Phe-chloromethyl ketone inactivates the enzyme with a 1:1 stoichiometry, which results in a modified His, as shown in Figure 6.5.4.

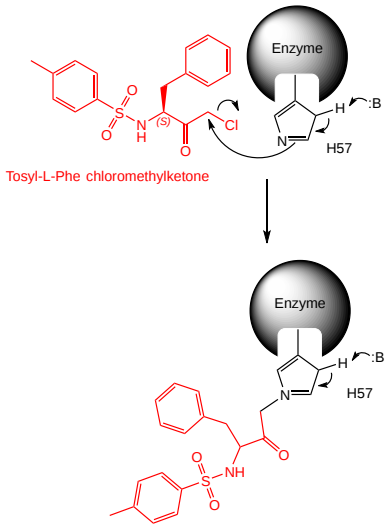


Figure 6.5.4: Reaction of chymotrypsin and other serine protease with chloromethyketones.

3. A comparison of the primary sequences of many proteases shows that three residues are invariant: a Ser, a His, and an Asp.
4. Site-specific mutagenesis shows that if Ser 195 is changed to Ala 195, the enzymatic activity is almost reduced to background levels. This strongly suggests that Ser 195 is an active site nucleophile.

D. **Changing the solvent**. Yes, indeed, you can take chymotrypsin and show that it is active in anhydrous organic solvents. Surely this is impossible, you say! It is true, and we will explore it at the end of the chapter since it's challenging enough to understand chymotrypsin activity in an aqueous solution. No new chemistry is needed; it is just a change in what you can conceptualize.

6.05B.3: Quantitative Analysis of Catalysis by Serine Proteases

Based on Siyuan Du *et al.* Conformational ensembles reveal the origins of serine protease catalysis. *Science* **387**, eado5068 (2025). DOI:10.1126/science.ado5068.

Our understanding of enzyme catalysis is incomplete. Other factors, in addition to those described in Chapter 6.1: How Enzymes Work, likely contribute to the 10^{12} increase in the rate of serine protease-catalyzed hydrolysis of peptide bonds compared to solution hydrolysis. A quantitative analysis of each factor would help define its contribution and help find new ones. Siyuan et al. have done just that.

So far, we've seen that the active site serine side chain (pKa ≈ 13, not considering its environment) acts as a nucleophile, made better by a nearby general base/acid histidine (pKa ≈ 7), which, after proton abstraction to form His^+, is stabilized by a nearby

negatively-charged aspartate. The developing δ^- on the carbonyl O in the transition state and the intermediate's full negative oxyanion are stabilized in the "oxyanion hole" through hydrogen bonds from N-H main chain donors, Ser 193 and Gly 195 (pKa ≈ 15). The reaction is intramolecular (entropically favored over bimolecular reactions) after substrate binding, except for the reaction with solvent water (55 M) needed to complete the reaction.

Now compare this to the uncatalyzed hydrolysis in water. H_2O (pKa ≈ 16) is the nucleophile/substrate, and other waters (pKa ≈ 16) stabilize the developing δ^-O in the transition state and the oxyanion intermediate. By analogy to the enzyme-catalyzed reaction, a stabilizing water molecule, which becomes, on accepting a proton, H_3O^+ (pKa ≈ -2) increases the nucleophilicity of the "substrate" water by proton abstraction.

Except for the adjacent His (pKa ≈ 7), which can more readily act as a general acid/base), there is nothing extraordinary about the groups involved in catalysis. The constrained intramolecular reaction involving the positioning and likely distortion of substrates in the active site is a critical feature to quantify. Another factor we haven't discussed is the dynamics of the reaction. The reaction has to proceed along reaction coordinates involving bond-making and breaking. In addition, substrates, waters, and side chains move to position themselves for catalysis.

Siyuan Du et al. show that fundamental concepts in chemistry and physics (torsion angles and strain, van der Waals interactions, hydrogen bonds, and entropy) can explain catalysis. They state that this "simplicity may inspire new ways of teaching enzyme catalysis, allowing instructors to reinforce the value of fundamental physical and chemical concepts and students to appreciate the tight connection between these concepts and the emergent, complex functions of biomolecules." This is precisely what Fundamentals of Biochemistry attempts to do!

They used molecular dynamics simulations and protein structures to analyze contributions to catalysis quantitatively. Although we'll focus on serine proteases here, they extended their analyses to other enzymes. They compared the PDB structures of 1231 wild-type serine proteases from 4 clans (structural superfamilies). These structures included enzymes:

- without substrate or other ligands (i.e. the apo form of the enzyme);
- with bound ligand when the ligand structure is unperturbed (i.e., the ligand is a **ground-state analog** or **GSA)** which is typically a noncovalently-bound peptide or peptide analog);
- with covalently bound ligands resembling the transition state for a substrate (i.e., the ligand is a **transition state analog or TSA**). The enzyme's active site would have a slightly altered structure to accommodate the tetrahedral TSA. The TSAs are covalently attached to Ser 195 through an sp^3 tetrahedral bond. These TSAs include fluoromethylketones, peptide aldehydes, and boronic acids, which underwent nucleophilic addition at the analog's sp^2 carbonyl carbon or boron center as they form adducts. The attached TSA stays covalently attached and does not react further. The active site differs from the apo and GSA-bound form from the movement required to form the covalent adduct. The active site is trapped in an "active" conformation.

These 1231 structures form a group or "pseudo-ensemble," which collectively provides data that have been hidden until the present analysis. These data describe subtle factors favoring catalysis and hint at the dynamical motion required within the active site, where groups move from the apo state to the transition/intermediate state.

Quantum mechanical analyses of reaction paths and dynamics for the hydrolysis reaction of N-methylacetamide (NMA) in water (without enzyme) were studied for a meaningful comparison to the solution-phase reaction. (We saw NMA in Chapter 4.9: Protein Stability - Thermodynamics.) NMA mimics the peptide bond as shown below in Figure 6.5.5 below. It likewise forms a tetrahedral intermediate during hydrolysis.

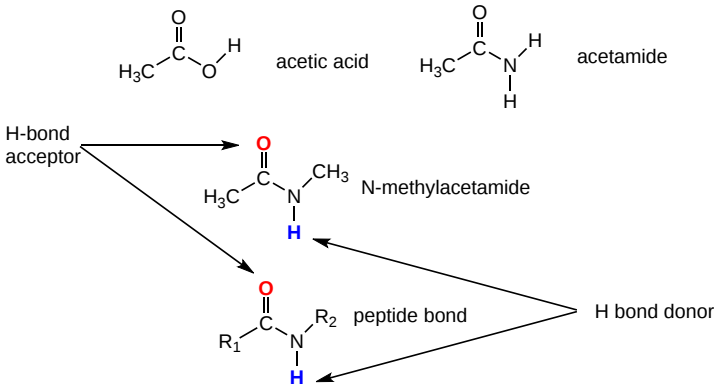


Figure 6.5.5: Comparison of N-methylacetamide to a peptide bond.

Molecular dynamics simulations were used to obtain a distribution or ensemble of reacting molecules in solution to parallel the ensemble of protein structures.

They then compared the distribution of molecules (solution) and sidechains (enzyme) to compare the positioning of the serine side chain (compared to water), as catalysis has to involve its movement. Three geometric variables were used: the distance (d_{attack}), angle (α_{attack}) and dihedral (Φ_{attack}). These parameters are described in Figure 6.5.6 below.

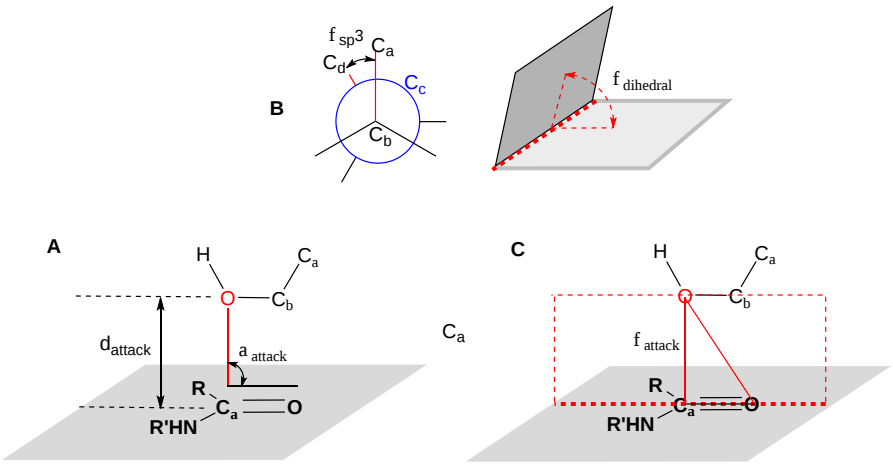


Figure 6.5.6: Visual representations of 3 key geometric parameters for hydrolysis of amide bonds.

Panel **A** shown $\mathbf{d_{attack}}$, the distance between the nucleophile O on serine (or water in the uncatalyzed case) and the electrophilic C on the amide substrate, and $\boldsymbol{\alpha_{attack}}$, the angle between three atoms (the nucleophilic O, the electrophilic C, and the carbonyl O. The other key angle is a dihedral. Panel **B** reviews examples of a dihedral angle. The left Newman project shows the Φ dihedral, which is the angle of rotation around the C_b-C_c bond in the four atoms, three bonds C_a-C_b-C_c-C_d structure, where C_b is the front carbon, and C_c is the back carbon represented by the blue circle. The right structure shows that the $\Phi_{dihedral}$ is also the angle of rotation around the red dotted line (---) connecting two gray planes. Panel **C** shows the $\Phi_{attack.}$ In this case, the gray plane contains the sp^2 hybridized amide substrate, and the red-dotted plane/triangle contains the plane defined by the nucleophilic O, the electrophilic C, and the carbonyl O.

The values for these parameters were very similar for all serine proteases studied with mode (not median) values as follows: d_{attack} = 2.68 Å (SD = 0.14 Å), α_{attack} = 93° (SD = 7°), and Φ_{attack} = 84° (SD = 8°). The distribution in water for the nucleophilic attack of water on DMA was much broader, and the distance was larger. So, it seems that the serine in proteases has a shorter path and higher efficiency for attack. A caveat is that the active site can't be too rigid, which could hinder the motions required for catalysis. Also, the serine must be pointed in the right direction.

To look for catalytic-specific motion, they compared torsion angles (backbone, side chain) for every residue in the enzyme with GSA and TSA bound. In addition, they measured changes between the apo and GSA-bound enzyme to see changes just on the binding substrate. Differences among these were interpreted as movement along the reaction pathway. For trypsin, 32 torsion angle changes were found when substrate bound, and 23 were found for nucleophilic attack (comparing TSA and GSA-bound enzymes). These changes were distributed throughout the enzyme. Changes in the torsion angle χ^1 for the Cα–Cβ in the side chain of Ser 195 were large and occurred during substrate binding (−14°, leading to torsional strain on partial eclipsing) and the actual nucleophilic attack (+14°, relieving the torsional strain). This large change was also seen for chymotrypsin and elastase (in the same clan). Indeed, this common single change was the single largest torsional change for the 3 enzymes.

Figure 6.5.7 shows an interactive iCn3D model of the alignment of bovine pancreatic trypsin with a Kunitz Type serine protease Inhibitor-1 (3M7Q), a ground state noncovalent inhibitor and trypsin with APA (1TPP), a covalent transition state analog (TSA). The gray backbone is the GS inhibitor complex, and the cyan backbone is the TSA analog covalent complex. The inhibitors are not shown for clarity and simplicity. Toggle the "A" key to switch between each form.

NCBI iCn3D Figure 6.5.7: Kunitz Type serine protease Inhibitor-1 (3M7Q), a ground state noncovalent inhibitor and trypsin with APA (1TPP), a covalent transition state analog. (Copyright; author via source). Click the image for a popup or use this external link: https://www.ncbi.nlm.nih.gov/Structu...pUFy1femfcy4Y7

Look at the change in orientation of the serine O nucleophile. Rotate the models to see distance and χ^1 since the side chain dihedral angle changes.

In the GSA-Trypsin, the nucleophilic Ser-O to the electrophilic carbonyl carbon in the planar sp^2 peptide bond is 2.7 Å (open this iCn3D link to see the distance). The van der Waals radius for a carbonyl carbon (C) ≈ 1.7 Å, and ≈ 1.52 Å for an O, so the optimal distance is 3.2 Å. Hence, the ground state structure is partially destabilized before bond-making actually occurs. The destabilization in the GSA contributes to catalysis.

A distance change occurs in the TSA complex as the serine oxygen moved ≈ 0.3 Å closer to the electrophilic carbon (now sp^3), which is now raised 0.8 Å above the plane from its previous position in the sp^2 substrate since it is attached to the C. This move is expected during the dynamic progress of the reaction as it proceeds along the reaction path. Now, the distance from the Ser-O to the electrophilic carbon in the covalent tetrahedral sp^3 (nonplanar) TSA adduct is 1.6 Å (open this iCn3D link to see the distance). (Note: the paper gives values of 1.6 in the text and 1.7 in a figure.) The net movement is ≈ 1.1 Å movement.

How does this compare to the solution hydrolysis path? Quantum mechanical analyses show that the distance between the water O and the NMA electrophilic C in the substrate in solution is 3.6 Å, greater the the vdW distance. This decreases ≈ 2 Å, not ≈ 1.1 Å, as we saw for the enzyme reaction. No bond rotations are noted. Most of the changes arise from a simple translation of the nucleophilic O on the water, which is farther from the electrophilic carbon in the ground state. Hence, the enzyme has a shorter reaction path with part contributed by 1D bond dihedral rotation, faster than positioning the reactants with a 3D translation.

Let's translate some of these changes into the thermodynamics parameter ΔG. Rates of catalysis can be related to the equilibrium constant, which is related to the ΔG^0 for the reaction. For example, suppose that breaking a strong bond is rate limiting in a pathway, which often occurs in a thermodynamically stable molecule that reacts. In that case, the reaction rate also depends on the bond strength (i.e., a measure of its thermodynamic stability). For acids and bases, this is reflected in a linear relationship between log(k), the rate constant for the reaction, and the $\log(K_a)$, the equilibrium acid constant, as shown by the linear Bronsted catalysis equation below:

$$\log k = a\log(K_a) + C = -apK_a + C \tag{6.5.8}$$

ΔG for individual catalytic steps can be calculated using this linear free energy relationship example. For example, the observed 10^6 increase in rate for trypsin from the histidine general base catalysis and its high effective intramolecular concentration gives a value of ΔG around 8.2 kcal/mol with an extra 0.8 kcal/mol from the stabilizing catalytic triad aspartate. The overall 10^{12} increase in catalytic rate (net 17.1 kcal/mol) suggests another 8.1 kcal/mol remains unexplained. The authors attribute the remaining catalysis to several factors.

One feature is based on the ground state of the enzyme. When the substrate binds, and since the reaction proceeds in an intramolecular fashion, there is reduced conformational entropy and some unfavorable interactions, as described above. These are decreased (relieved) in the transition site. The net effect is another 7.6 kcal/mol of catalysis. They refer to these combined effects as ground-state destabilization.

Let's consider all of the factors they explored. The first two deal with the entropy of positioning both the reactant and the active site serine:

- **pre-orientation of the reactant in the active site**: This makes it entropically easier to reach the "single" transition state than the reaction in water.
- **pre-orientation of the enzyme serine nucleophile**: Similarly, pre-orientation of the serine nucleophile in the vicinity of the substrate decreases the number of possible microstates needed to reach the transition state, entropically favoring the reaction.
- **ground state destabilization of the complex through O-C bond shortening:** We mentioned above that the serine O to carbonyl C distance is shorter than the van der Waals interaction distance, which raises its energy as it moves into the direction of the transition state.
- **ground state destabilization of the complex through partial eclipsing from the torsion angle of the catalytic serine:** As mentioned above, this rotation is reversed on conversion to the transition state.
- **ground state destabilization of the main oxyanion hole by hydrogen bonds:** Solution phase H bonds between amide Hs and carbonyl Os are more stable when planar. In the GSA, they are not planar and hence destabilized.

The authors calculated ΔG for each of these effects. For an enzymatic reaction, the free energy of the transition state $\Delta G‡_{solution} > \Delta G‡_{enzyme}$. The following equation describes ΔΔG‡:

$$\Delta\Delta G^{‡} = \Delta G^{‡}_{\text{solution}} - \Delta G^{‡}_{\text{enzyme}} \tag{6.5.9}$$

Hence, the ΔΔG‡ value for the enzyme-catalyzed reaction and each contributing factor are all positive (>0). Figure 6.5.8 shows the individual contributions (combined to 16.6 kcal/mol) compared to the experimentally observed value of 17 kcal/mol.

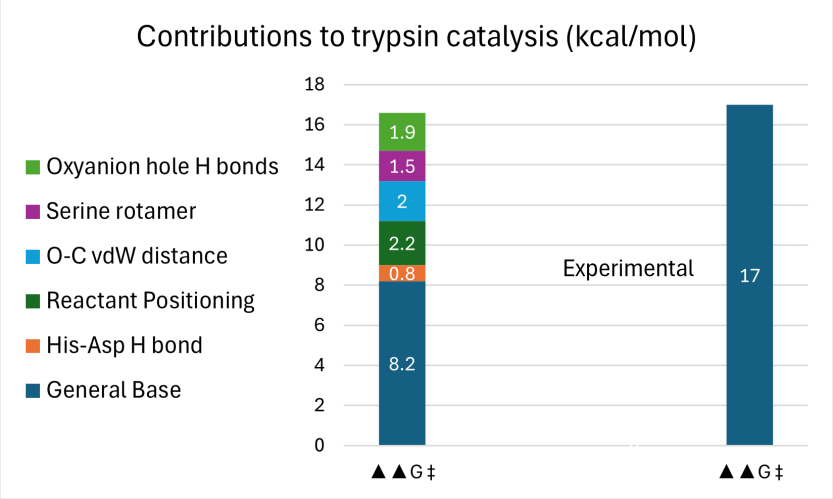


Figure 6.5.8: Data from Siyuan Du et al., ibid. Error bars not included.

Within error, catalysis can be described by destabilizing the ground state substrate and stabilizing the transition state based on principles you learned in introductory and organic chemistry: acid/base reactions, hydrogen bonds, entropy, torsional strain, bond angle strain, and van der Waals interactions. The authors extended their analyses and found the same factors account for catalysis by other enzymes of distinct folds using nucleophilic attack on carbonyl carbons. These include lactamases, caspases, transcylases, acylases, peptidases, and other proteases. They found similar structures and catalytic features (listed above) when aligning the PDB structures on active site nucleophiles, oxyanion holes, oxyanion, and the substrates electrophile carbon.

6.05B.4: Enzyme catalysis in organic solvents

In our earlier lists, we mentioned changing the solvent and exploring its effect on enzyme catalysis. It might seem a bit wild, but as we saw with the rhomboid protease, some enzymes work in hydrophobic environments. Also, lipases work at the boundary between the aqueous and hydrophobic worlds. For those interested, let's see what happens when we change solvents. These include putting the enzyme in various solvents or mixtures of solvents, as described below:

- Water-miscible solvents like ethanol and acetone were added. If the water concentration was high enough, activity remained.
- Biphasic mixtures in which an aqueous solution of an enzyme was emulsified in a water-immiscible solvent like chloroform or ethyl acetate. The substrate would partition into both phases, while the product hopefully would end up in the organic phase.

- Nearly nonaqueous solvents, with a few % water at less than the solubility limits of water.
- Anhydrous organic solvents (0.01% water). This case is most astonishing since enzymatic activity is often retained.

It is important to realize that in this last case, the enzyme is **not in solution**. Instead, it is in **suspension** and acts as a **heterogeneous catalyst**, much like palladium, which acts as a heterogeneous catalyst in the hydrogenation of alkenes. The suspension must be mixed vigorously and then sonicated to produce small suspended particles, so diffusion of reactants into the enzyme and out is not rate limiting. Let's explore the activity of chymotrypsin in a nonpolar solvent.

Why aren't the enzymes inactive? Indeed, it must seem ridiculous that they aren't since, as we learned earlier, proteins are marginally stable. On average, a 100 amino acid protein is stabilized only about 10 kcal/mol (41 kJ/mol) over the denatured state, or the equivalent of a few H bonds. Indeed, the hydrophobic effect, one of the dominant contributors to protein folding and stability, would not stabilize the native structure of enzymes in nonpolar organic solvents, and the protein would denature. It doesn't, however! Maybe the real question should be not whether water is necessary but how much water is essential. The enzyme can't "see" more than a monolayer or so of water around it. The data suggests that the nature of the organic solvent is very important. The most hydrophobic solvents are best in terms of their ability to maintain active enzymes! Chymotrypsin retains 10^4 more activity in octane than pyridine (see $\mathbf{k_{cat}/K_m}$ below), which is more hydrophilic than octane. The more polar the solvent, the more it can strip bound water away from the protein. If you add 1.5% water to acetone, the bound water increases from 1.2 to 2.4%, and the activity of chymotrypsin increases 1000-fold.

Table 6.5.3 below shows chymotrypsin activity in organic solvents.

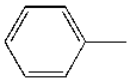
Solvent	Structure	$\mathbf{k_{cat}}/K_m$ ($M^{-1}min^{-1}$)	relative ratio k_{cat}/K_m	$\mathbf{H_2O}$ bound to enzyme (%, w/w)
Octane		63	15000x	2.5
Toluene		4.4	1000x	2.3
Tetrahydrofuran		0.27	175x	1.6
Acetone		0.022	5.5x	1.2
Pyridine		<0.004	1x (.004)	1.0

Table 6.5.3: Chymotrypsin activity in organic solvents

Consider the following questions:

- *How much water do the enzymes need?* One chymotrypsin molecule in octane has less than 50 molecules of water associated with it and can demonstrate activity. About 500 water molecules are required to form a monolayer. Water can be added, presumably leading to more water binding and higher activity.
- *How stable are the enzymes?* Denaturation requires conformational flexibility, which requires water. The half-life of chymotrypsin in water at 60 ºC is minutes, but in octane at 100 ºC, it is hours. At 20 ºC, the half-life in water is a few days, but in octane, it is greater than 6 months. Remember two factors contribute to stability: 1. The protein can denature at high temperatures. 2. Chymotrypsin is a protease. It can cleave itself in an autoproteolytic reaction.

Table 6.5.4 below shows the half-life of chymotrypsin activity in water and octane

Solvent	60ºC	100ºC	20ºC
water	minutes	-	few days
octane	-	hours	> 6 months

Table 6.5.4: Half-life of chymotrypsin activity in water and octane at different temperatures

- *Has the enzyme specificity changed?* The net binding energy is a function of the substrate's binding energy minus the water's binding energy since water must be displaced from the active site on binding. In an anhydrous solvent, specificity changes must be expected. For chymotrypsin, the driving force for binding substrates in water is primarily hydrophobic. In water, the k_{cat}/K_M for the reaction of N-acetyl-L-Ser-esters is reduced 50,000 times compared to the Phe ester. However, chymotrypsin is three times more active toward Ser esters in octane than Phe esters.

Table 6.5.5 shows specificity changes in chymotrypsin in water and octane

Substrate	k_{cat}/K_m	
	solvent: H2O	solvent: Octane
N-acetyl-L-Ser-ester	1x	3x
N-acetyl-L-Phe-ester	50,000x	1x

Table 6.5.5: Specificity changes in chymotrypsin in water and octane

Now, consider competitive inhibitors. Naphthalene binds 18 times more tightly than 1-naphthoic acid, but chymotrypsin binds naphthoic acid 310 times as tightly in octane. Likewise, the ratio of [k_{cat}/K_m (L isomer)]/[k_{cat}/K_m (D isomer)] of N-acetyl-D- or N-acetyl-L-Ala-chloroethyl esters is 1000-10,000 in water, but less than 10 in octane.

Table 6.5.6 shows chymotrypsin inhibition constants in water and octane.

Inhibitor	Inhibition Constant Ki (nM)	
	In water	In Octane
Benzene	21	1000
Benzoic acid	140	40
Toluene	12	1200
Phenylacetic acid	160	25

Naphthalene	0.4	1100
1-Naphthoic acid	7.2	3

Table 6.5.6: Chymotrypsin inhibition constants in water and octane

Can new reactions be carried out in nonpolar solvents? The quick answer is yes since reactions in aqueous solutions can be unfavorable due to low K_{eq} values, side reactions, or insolubility of reactants. Consider lipases, which cleave fatty acid esters by hydrolysis in aqueous solutions. In nonaqueous solutions, reactions such as transesterification or ammonolysis can be performed.

Enzymes are active in organic solvents, which contradicts our central concepts of protein stability. Two reasons could could explain this stability:

1. From a thermodynamic perspective, the enzyme may be stable in organic solvents. However, as discussed above, this is **inconceivable** given the delicate balance of noncovalent and hydrophobic interactions required for protein stability.
2. The second reason must prevail: the protein cannot unfold from a **kinetic** point of view. Conformational flexibility is required for denaturation, which requires water as the solvent. Denaturation in organic solvents is kinetically, not thermodynamically, controlled.

A specific example helps illustrate the effects of different solvents on chymotrypsin activity. Dry chymotrypsin can be dissolved in DMSO, a water-miscible solvent. In this solvent, chymotrypsin is entirely and irreversibly denatured. No activity is observed if it is now diluted 50X with acetone with 3% water. (In the final dilution, the concentrations of solvents are 98% acetone, 2.9% water, and 2% DMSO.) However, the enzyme is very active if dry chymotrypsin is added to a mixture of 98% acetone, 2.9% water, and 2% DMSO. We end up with the same final solvent state, but the enzyme has no activity in the first case, while in the second case, it retains activity. These ideas are illustrated in Figure 6.5.9.

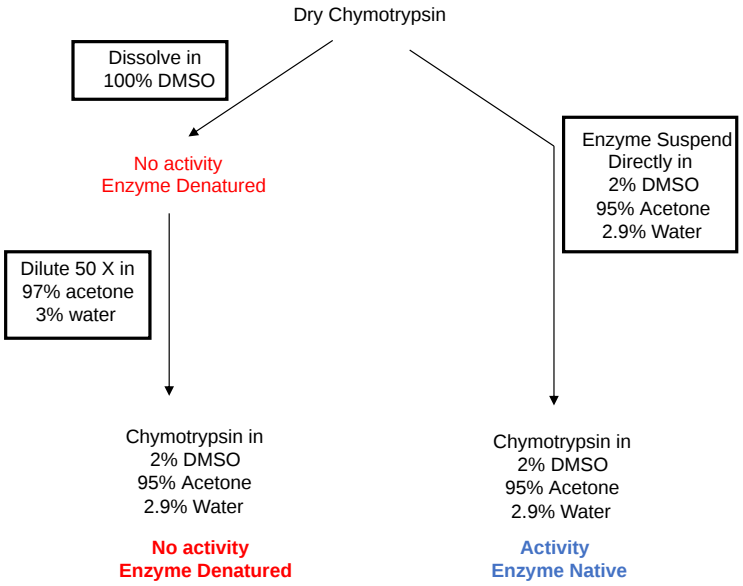


Figure 6.5.9: Chymotrypsin activity in acetone depends on the order of solvent addition

Dry enzymes added to a concentrated water-miscible organic solvent (like DMSO) will dissolve and surely denature but will retain activity when added to a concentrated water-immiscible solvent (like octane), in which the enzyme will not dissolve but stay in suspension.

It appears the enzymes have very restricted conformational mobility in nonpolar solvents. By lyophilizing (freeze-drying) the enzyme against a specific ligand, a given conformation of a protein can be trapped or imprinted onto the enzyme. For example, if the enzyme is dialyzed against a competitive inhibitor (which can be extracted by the organic solvent), freeze-dried to remove water, and then added to a nonpolar solvent, the enzyme activity of the "imprinted" enzyme in nonpolar solvents is as much as 100x as great as when no inhibitor was present during the dialysis. Suppose chymotrypsin is lyophilized from solutions of different pHs.

In that case, the resulting curve of V/Km for ester hydrolysis in octane is bell-shaped, with the initial rise in activity reaching half-maximum activity at a pH of around 6.0 and a fall in activity reaching half-maximum at a pH of approximately 9.

The use of enzymes in organic solvent allows new routes to organic synthesis. Enzymes, which are so helpful in synthetic reactions, are:

- stereoselective - can differentiate between enantiomers and between prochiral substrates
- regioselective - can differentiate between identical functional groups in a single substrate
- chemoselective - can differentiate between different functional groups in a substrate (such as between a hydroxyl group and an amine for an acylation reaction)

Enzymes in anhydrous organic solvents are useful (from a synthetic point of view) not only because they can catalyze new types of reactions (such as transesterification, ammonolysis, and thiolysis) but also because the stereoselectivity, regioselectivity, and chemoselectivity of the enzyme often differ from those of the enzyme in water.

Organic reactions are usually conducted in organic solvents since many organic molecules react with water, and the reagents and products are generally not soluble in water. In a manner analogous to using an enzyme as a heterogeneous catalyst in a nonpolar solvent, Sharpless is pioneering a technique to conduct organic reactions in water. They (Narayan et al.) have shown that many unimolecular and bimolecular reactions occur faster in water than in organic solvents. As in enzyme catalysis in nonpolar solvent, the reactions must be mixed vigorously to disperse reactants in micro-drops (a suspension) in water, significantly increasing the surface area that might allow water to act on transition states or intermediates to stabilize them through hydrogen bonding. They called these "on water" reactions since reactants usually float on water. Using this process, they have performed cycloadditions, alkene reactions, Claisen rearrangements, and nucleophilic substitution reactions. One cycloaddition reaction went to completion in ten minutes at room temperature, compared to 18 hours in methanol and 120 in toluene. Adding nonpolar solvents at certain times significantly increased the rate of the reaction.

6.05B.5: Summary

This chapter explores the mechanistic and thermodynamic foundations of enzyme catalysis with an emphasis on serine proteases—particularly chymotrypsin. It demonstrates how a combination of structural, kinetic, and thermodynamic data can be employed to develop detailed reaction mechanisms, even before the advent of high-resolution crystal structures.

Enzyme Mechanisms and the Catalytic Triad

- **Catalytic Strategy:**
  The chapter begins by describing how early enzyme mechanism studies relied on kinetic experiments and chemical modifications to elucidate catalytic strategies. In serine proteases, the catalytic triad—comprising Ser 195, His (57/95), and Asp 102—is critical. Ser 195 acts as a nucleophile, His functions as a general base, and Asp stabilizes the positive charge on His.
- **Evidence for Covalent Intermediates:**
  Kinetic analyses using small ester and amide substrates reveal that chymotrypsin forms a covalent acyl-enzyme intermediate. Differences in k_cat and k_cat/K_m values between ester and amide substrates, along with substrate modifications (e.g., changing leaving groups), support this mechanistic hypothesis.

Kinetic Analysis and the Ping Pong Mechanism

- **BiBi Ping Pong Kinetics:**
  The reaction is characterized by a burst phase—rapid formation of an initial product followed by a slower phase corresponding to deacylation. This two-step process is described by a simplified kinetic model that distinguishes between acylation (k_2) and deacylation (k_3) steps. Depending on which step is rate-limiting, the kinetic equations can be simplified for ester or amide hydrolysis.
- **Mathematical Modelling:**
  The chapter outlines the derivation of kinetic equations that link substrate concentration, enzyme concentration, and rate constants, allowing for a quantitative dissection of each catalytic step.

Modulation of Catalysis: Substrate, pH, and Enzyme Modifications

- **Substrate Effects:**
  Modifying the substrate's leaving group or acyl substituents alters the kinetic parameters, thereby providing insight into the

chemical steps of the reaction, such as the formation and breakdown of the acyl-enzyme intermediate.

- **pH and Ionic Strength:**
  pH-dependent studies reveal the involvement of ionizable groups. For instance, the activity profiles indicate that a deprotonated histidine is necessary for optimal catalysis, while other groups (e.g., the N-terminal Ile) contribute to structural stabilization through salt bridge formation.

- **Enzyme Modifications:**
  Chemical modification (using reagents like diisopropylphosphofluoridate) and site-specific mutagenesis confirm the essential roles of specific amino acids in the active site. Inactivating modifications that target Ser 195 or His validate their roles in nucleophilic attack and proton transfer.

The Role of Solvent in Enzyme Catalysis

- **Aqueous vs. Organic Environments:**
  A striking portion of the chapter is devoted to examining how chymotrypsin operates in nonaqueous solvents. In water-miscible and nearly nonaqueous solvents, the enzyme retains activity, whereas in anhydrous organic solvents it acts as a heterogeneous catalyst. The enzyme's minimal hydration shell is sufficient for catalysis, and its kinetic stability can even increase due to the restricted conformational flexibility in nonpolar environments.

- **Effects on Specificity and Inhibition:**
  Solvent changes impact substrate specificity and inhibitor binding. For example, substrate preference shifts in octane compared to water, and competitive inhibitors display different binding affinities, underscoring how the solvent environment can influence enzyme–substrate interactions.

Quantitative Contributions to Catalysis

- **Thermodynamic Perspective:**
  The chapter concludes with a quantitative analysis that breaks down the contributions to the overall rate enhancement of serine proteases. Key factors include the stabilization of the transition state, pre-orientation of the reactants, and ground state destabilization upon substrate binding. Using linear free energy relationships (e.g., Bronsted relationships), the authors correlate changes in catalytic rate with specific free energy contributions, ultimately accounting for nearly the entire observed rate acceleration.

- **Molecular Dynamics and Structural Ensembles:**
  Analyses of PDB-derived structural ensembles and molecular dynamics simulations illustrate how subtle conformational changes (such as torsion angle adjustments in the catalytic serine) and geometric parameters (attack distances and angles) are finely tuned for catalysis, setting the enzyme apart from the uncatalyzed reaction in solution.

By integrating experimental data with computational modeling, the chapter not only deepens our understanding of enzyme catalysis at a molecular level but also highlights the interplay between structure, dynamics, and energetics. This comprehensive approach reinforces the idea that fundamental chemical principles—such as acid/base chemistry, hydrogen bonding, entropy, and van der Waals forces—are central to the sophisticated function of enzymes, inspiring new ways of teaching and exploring biomolecular catalysis.

6.6: Enzymes and Protein Regulation

Learning Goals (ChatGPT o1, 1/30/25)

- **Proteome Complexity Beyond the Genome**
  - Explain how a relatively modest number of genes can give rise to a proteome that is two to three orders of magnitude more complex through alternative splicing and post-translational modifications (PTMs).
- **Diversity and Function of Post-Translational Modifications (PTMs)**
  - Identify and describe common enzymatic and nonenzymatic PTMs—including phosphorylation, acetylation, glycosylation, ubiquitination, sumoylation, oxidation, and methylation—and their effects on protein structure, stability, activity, and localization.
  - Explain how PTMs create proteoforms and modulate protein–protein interactions and signal transduction pathways.
- **Detection and Analysis of PTMs**
  - Evaluate the various methods for detecting PTMs (e.g., antibody-based techniques, Western blotting, mass spectrometry) and discuss the advantages and limitations of each approach for both targeted and global PTM analysis.
- **Allosteric Regulation and Conformational Control**
  - Describe the principles of allosteric regulation, including how binding of small molecules, ions, or proteins at sites distinct from the active site can induce conformational changes that modulate enzymatic activity.
  - Differentiate between allosteric changes induced by small molecules, phosphorylation, protein–protein interactions, and alterations in disulfide bonds.
- **Zymogen Activation and Enzyme Maturation**
  - Explain the concept of zymogens (proenzymes) and the mechanisms by which they are activated, with examples such as the activation of pancreatic proteases and caspases during apoptosis.
- **Isoenzymes (Isozymes) and Differential Regulation**
  - Define isozymes and distinguish them from allozymes, highlighting their origins via gene duplication or alternative splicing and their roles in fine-tuning metabolic pathways.
  - Illustrate how isozymes, such as cyclooxygenases (COX-1 and COX-2), differ in kinetic properties, regulatory mechanisms, and tissue-specific expression, and discuss their relevance to therapeutic interventions.
- **Cyclooxygenases and Prostaglandin Biosynthesis**
  - Describe the structural and functional characteristics of COX-1 and COX-2, including their bifunctional enzyme activities (cyclooxygenase and peroxidase), and explain how their regulation influences the synthesis of prostaglandins and thromboxanes.
  - Discuss how selective and non-selective NSAIDs target COX isozymes and the physiological implications of disrupting the balance between prostacyclin and thromboxane in cardiovascular and gastrointestinal contexts.

By mastering these goals, students will be able to integrate molecular, structural, and regulatory aspects of protein control and appreciate the intricate layers that modulate protein function in response to cellular and environmental cues.

Many different mechanisms control protein activity within a cell. The primary sequence of a protein is a main determinant of protein folding and final conformation, as well as biochemical activity, stability, and half-life. However, at any given moment, the proteome, the full complement of proteins within a cell, is up to two or three orders of magnitude more complex than the encoding genomes would predict. This chapter will give an overview of the major mechanisms biological systems utilize to regulate protein functions after the protein has been synthesized. Note that these mechanisms seldom work in isolation. Multiple levels of protein control function at any given time and in response to many different environmental cues and signals. We will discuss many ways to regulate protein activity. At the end, we will discuss regulation through the use of different **isozymes** of an enzyme, which are

variants of an enzyme arising from differential splicing of a gene or from slightly different genes that arose from a common precursor gene. We will focus on cyclooxygenases, the target of so many medicinal drugs.

6.6.1: Post-translational modifications (PTMs)

The human genome contains approximately 20,000 genes. When analyzing the transcriptome, it becomes apparent that the genome becomes amplified by the wide array of splice variants that can occur when primary RNA transcripts are spliced to form mature messenger RNAs. There are about 200,000 expressed transcripts, of which about 146,00 encode proteins, with about 7.4 transcripts per gene. A summary of human gene transcripts is shown in Table 6.6.1 below.

Type	GTEx dataset	Expressed all genes	Expressed protein-coding genes
Transcripts per Gene	3.42	3.63	7.43
Gene counts	58,219	53,539	19,591
Transcript counts	199,324	194,146	145,571

Table 6.6.1: **Numbers of transcript and gene in the GTEx dataset.** Tung, KF., Pan, CY., Chen, CH. *et al.* Top-ranked expressed gene transcripts of human protein-coding genes investigated with GTEx dataset. *Sci Rep* **10**, 16245 (2020). https://doi.org/10.1038/s41598-020-73081-5. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

The number of distinct primary structures of proteins (**proteoforms**) from 20,000 human genes is amplified again through posttranslational modifications (PTMs) of proteins, which produce up to a million different variants. PTMs are present in both eukaryotes and prokaryotes. Still, they are more common in eukaryotic cells, in which about 5% of the genome is dedicated to enzymes that carry out posttranslational modifications of proteins. We discussed the chemical modification of specific amino acids in Chapter 3.1: Amino Acids and Peptides.

Protein PTM results from the enzymatic or nonenzymatic attachment of specific chemical groups to amino acid side chains. Such modifications occur either following protein translation or concomitant with translation. PTM influences both protein structure and physiological and cellular functions. Examples of enzymatic PTMs include phosphorylation, glycosylation, acetylation, methylation, sumoylation, palmitoylation, biotinylation, ubiquitinylation, nitration, chlorination, and oxidation/reduction. Nonenzymatic PTMs include glycation, nitrosylation, oxidation/reduction, acetylation, and succinylation. Some rare and unconventional PTMs, such as glypiation, neddylation, siderophorylation, AMPylation, and cholesteroylation, are also known to influence protein structure and function. Note that many of these modifications are not made in isolation. It is common for proteins to have several different types of modifications and that these modifications can differ depending on the tissue type and environmental circumstances present. The major PTMs in eukaryotes, their target amino acid residue(s), and the types of enzyme(s) or protein(s) involved are shown in Table 6.6.2 below.

Posttranslational modification	Target amino acid residue(s)	Enzyme(s) or proteins involved
Phosphorylation	Tyrosine, serine, threonine	Kinases, phosphatases
Glycosylation N-linked	Asparagine	Glycosyltransferases, deglycosylases
Glycosylation O-linked	Serine/threonine	Glycosyltransferases, deglycosylases
Acetylation	Lysine	Acetyltransferases, deacetylases
Methylation	Lysine, arginine	Methyltransferases, demethylases
Ubiquitination	Lysine	Ubiquitin-activating enzymes, ubiquitin-conjugating enzymes, ubiquitin ligases, deubiquitinases
Sumoylation	Lysine	Ubiquitin-activating enzymes, ubiquitin-conjugating enzymes, ubiquitin ligases, deubiquitinases
Myristoylation	Glycine	N-Myristoyltransferases
Prenylation	Cysteine	Farnesyltransferases, geranyl geranyltransferases
Palmitoylation	Cysteine	DHHC protein acyltransferases, acyl-protein thioesterases
Sulfation	Tyrosine	Sulfatases, desulfatases
S-Nitrosylation	Cysteine, methionine	
Glycation	Lysine	
Nitration	Tyrosine	Denitrases
Chlorination	Tyrosine	Myeloperoxidases
Oxidation/reduction	Cysteine	Peroxidases, oxidases, glutathione, thioredoxin
Carbonylation	Lysine, proline, arginine, threonine	

Table 6.6.2: **Common Protein Post-Translational Modifications, Their Target Amino Acid Residues, and the Enzyme(s) or Proteins Involved.** *Santos, A.L, and Lindner, A.B. (2017) Oxidatve Medicine and Longevity, Article ID: 5716409*

PTMs can exert their effect through many different structural changes, including opening and closing the active site, changing the conformation and electrostatic properties of binding sites, changing the chain's flexibility, increasing or decreasing intrinsically disordered regions, altering protein-protein interactions, etc. These effects can also arise from the binding of small molecules.

6.6.1.1: Methods to Detect Protein Posttranslational Modifications

Specific amino acid residues are subjected to PTMs depending on the chemistry of the reaction and the sequence specificity of the enzyme involved. Initially, PTMs were detected by various analytical methods, such as radiolabeling of the proteins, thin-layer chromatography, column chromatography, and/or polyacrylamide gel electrophoresis. Other methods have since been developed, such as protein sequencing by Edman degradation and Western blotting using protein-specific antibodies. Antibody-based detection methods and mass spectrometry-based proteomic analysis are the predominant methods used to detect and analyze PTMs. However, mass spectrometric methods are the only available tools for global or large-scale PTM analysis.

Antibody-based methods mostly rely on the availability of antibodies that can specifically recognize a modified amino acid residue within a protein or peptide. Such antibodies can be polyclonal or monoclonal and are developed against either the modified peptide/protein or against the modified amino acid. Moreover, antibody-based detection and quantification of PTMs on protein/peptide samples can be performed by chemiluminescence-based Western blotting and absorbance/fluorescence-based ELISA. However, the detection of PTMs depends entirely on the recognition site of the antibody used. If the antibody detects only the modified amino acid, additional analysis—for instance, protein/peptide isolation and sequencing—should be performed to detect the sequence context of the modification. However, if the antibody detects the PTM within a specific sequence context, the presence of PTM at other sites will remain undetected (ie the antibody will be specific for only that single modification).

Mass spectrometric detection of specific PTMs is based on mass changes. Depending on the type of modification, a specific change in the mass of the modified amino acid or peptide occurs. Subsequently, the change in mass is detected by the mass spectrometer to identify the presence of a PTM in a peptide sample. Using tandem mass spectrometric methods, identification of the specific site of PTM can be achieved by subsequent fragmentation and sequencing of the relevant peptide. Yet, technical challenges hamper MS-based investigation of biologically important PTMs, such as ADP-ribosylation, one of the key signaling molecules that regulate DNA repair, a critical process in maintaining genome stability that is compromised in cancer and aging.

Data increasingly implicate PTMs during aging and/or under pathological conditions and also in the cell's normal functioning. PTMs are increasingly studied for their role in health and disease. For example, the precise and accurate measurement of distinct PTM-containing moieties offers potential biomarker utility to aid early diagnosis, prognosis, monitoring response to therapy, and decisions regarding inclusion in clinical trials as new medicines are developed. However, technical difficulties limit these studies, leaving many unanswered questions. Identifying unknown/unexpected PTMs by proteomic data reanalysis, is an emerging subfield of proteomics recently boosted by the increased availability of raw data shared in public repositories. Notably, a sampling of the proteome in a given organism or cell provides only a snapshot of a highly dynamic process, confounding the analytical problem and ultimately arguing for time-resolved inventories. Thus, while many tools are currently available for studying PTMs, new methods are needed to advance the study of these modifications further.

6.6.1.2: Examples of PTMs

Protein PTMs involve the covalent addition of some chemical groups by enzymatic catalysis. Typically, an electrophilic fragment of a co-substrate is added to an electron-rich protein side chain, which acts as a nucleophile in the transfer. Common covalent protein PTMs include phosphorylation, acylation, methylation, sumoylation, ubiquitination, glycosylation, lipidation, oxidation and disulfide bond formation (internal within a single protein or linking two protein/peptide chains together). Examples of common PTMs are provided in Figure \(\PageIndex{1}\ below.

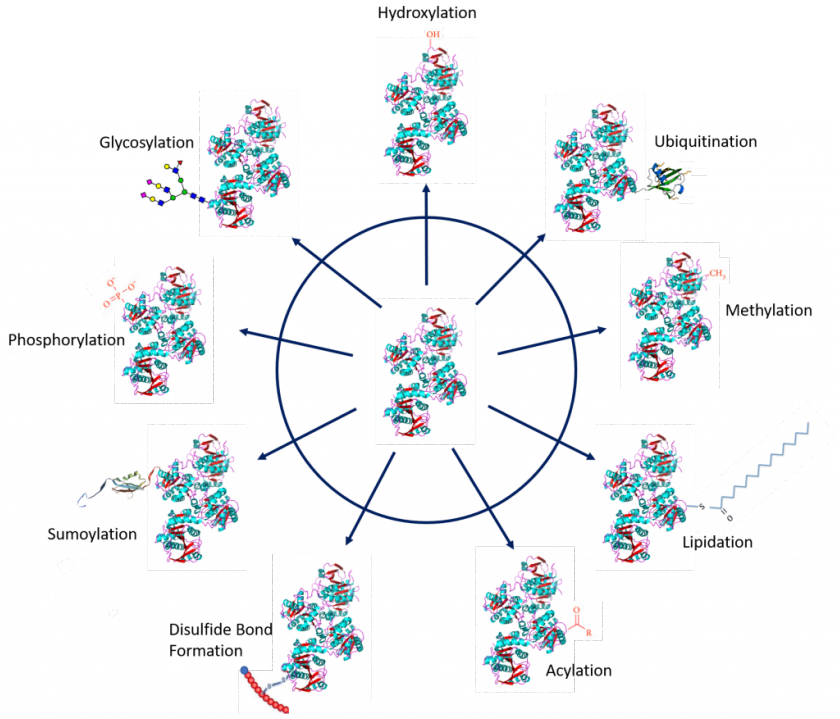


Figure 6.6.1: **Major types of covalent additions to protein side chains.** Common modifications include oxidation, ubiquitination, methylation, lipidation, acylation, disulfide bond formation, sumoylation, phosphorylation, and glycosylation. *Figure modified from: Santos, A.L, and Lindner, A.B. (2017) Oxidative Medicine and Longevity, Article ID: 5716409, Goettig, P. (2016) Int. J. Mol. Sci. 17(12), 1969, Rogerdodd, Jakob Suckale, Akinlade, A., et al (2014) Int. Archives of Med 7(50):28, WilsonNR, and Sivart13*

Protein Phosphorylation

One of the most common posttranslational modifications, protein phosphorylation, is the reversible addition of a phosphoryl group from adenosine triphosphate (ATP) to amino acid side chains such as serine, threonine, and tyrosine residues as shown in Figure

6.6.2. This modification causes conformational changes that either (1) affect the catalytic activity to activate or inactivate the protein and/or cause the tendency of a protein to misfold and aggregate or (2) recruit other proteins to bind; both responses result in altered protein function and cell signaling. Phosphorylated proteins have critical and well-known functions in diverse cellular processes across eukaryotes, but phosphorylation also occurs in prokaryotic cells. In humans, about one-third of proteins are estimated to be substrates for phosphorylation. Indeed, phosphorylated proteins are now identified and characterized by high-throughput phosphoproteomics studies. The reversibility of protein phosphorylation is attributed to the actions of ***kinases*** and ***phosphatases***, which phosphorylate and dephosphorylate substrates, respectively. The temporal and spatial balance of kinase and phosphatase concentrations within a cell mediates the size of its phosphoproteome.

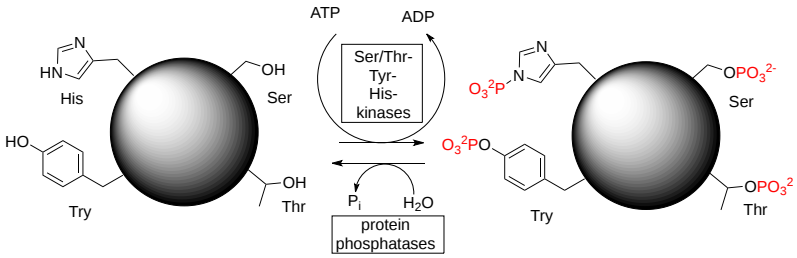


Figure 6.6.2: Amino acid phosphorylation in proteins

Protein Acylation

The simplest form of acylation is protein N-acetylation. This occurs at the amino terminus amine and the ε-amino group of the lysine side chains through the action of acetylases. The acetylation of lysine side chains can be reversed through deacetylases (similar to the combined actions of protein kinases and phosphatases). Figure 6.6.3 shows the general reactions of protein N-acetylation. Interestingly, 80–90% of eukaryotic proteins are acetylated, yet the underlying biological significance remains unclear.

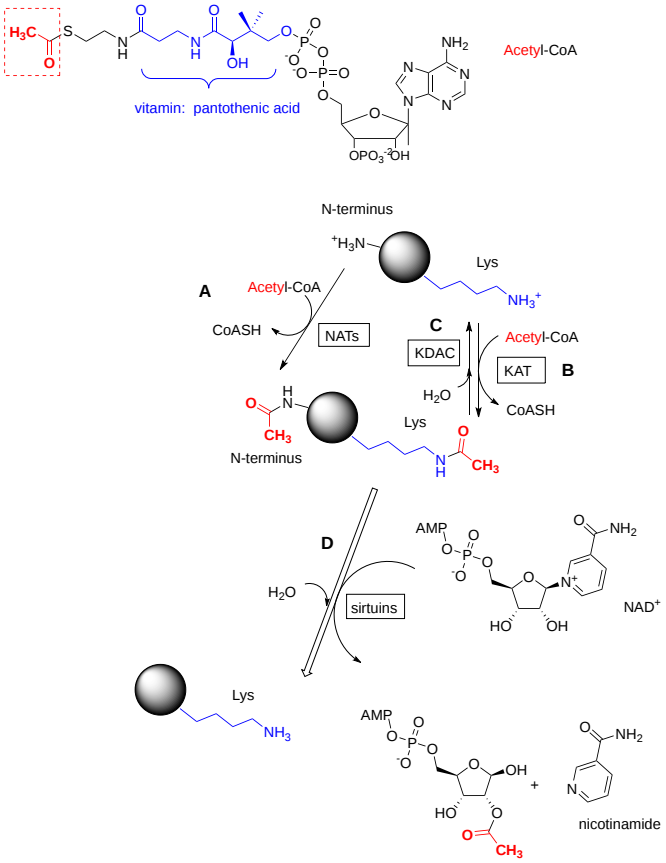


Figure 6.6.3: Protein N-acetylation and deacetylation reactions

The acetyl donor is acetyl-CoA. The N-terminus (reaction A) is acetylated by **N-terminal acetyltransferases (NATs)**. The ε-amino group of lysine (reaction B) is acetylated by **lysine acetyltransferases (KATs).** The latter modification is reversed through two different enzymes. One is a simple hydrolysis (reaction C) catalyzed by lysine deacetylases (KDACs). The other (reaction D)

requires NAD^+ and is catalyzed by enzymes called **sirtuins** (silent information regulators). Since sirtuins use NAD^+, one of the main oxidizing agents in the biological world, they link the protein **acetylome** with metabolic status and promote metabolic health.

Histones, positively charged proteins found in eukaryotic cell nuclei, pack and order the DNA into structural units called nucleosomes, as shown in Figure 6.6.4.

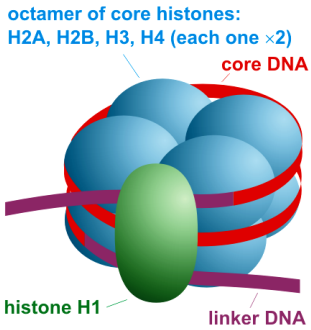


Figure 6.6.4: Nucleosome organization. https://commons.wikimedia.org/wiki/F...ganization.png. Creative Commons Attribution-Share Alike 3.0 Unported license.

The histones have multiple lysine and arginine residues that interact with the negatively charged phosphate groups of the DNA backbone. The nucleosome core comprises two H2A-H2B dimers and an H3-H4 tetramer, forming two nearly symmetrical halves by tertiary structure (C2 symmetry; one macromolecule is the mirror image of the other). The 4 'core' histones (H2A, H2B, H3 and H4) are relatively similar in structure and highly conserved through evolution, all featuring a 'helix-turn-helix' motif (a DNA-binding protein motif that recognizes specific DNA sequence). They also share the feature of long 'tails' on one end of the amino acid structure, which is the location of post-translational modification, specifically N-acetylation.

Histone acetylation typically results in transcriptional activation; deacetylation typically results in transcriptional suppression. Acetylation occurs via ***histone acetyltransferases (HATs)*** and is reversible via the action of ***histone deacetylases (HDACs)***. One group of histone deacetylases is the **sirtuins** (silent information regulator), which maintain gene silencing via hypoacetylation. Sirtuins have been reported to aid in maintaining genomic stability. Figure 6.6.5 shows the effect of acetylation on histone:DNA packing.

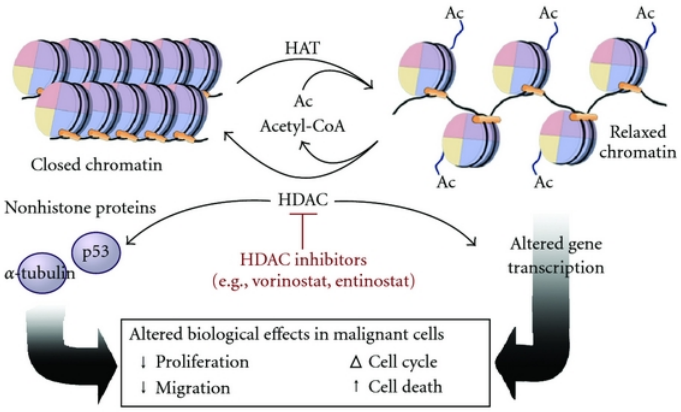


Figure 6.6.5: **Histone octamers form the backbone for chromosomal structure.** Acetylation and methylation affect histone-DNA interactions. DNA methylation increases histone-DNA affinity and blocks transcriptional activation. Acetylation of the histone tails disrupts histone-DNA interactions and facilitates gene expression. https://commons.wikimedia.org/wiki/F...cetylation.jpg. Creative Commons Attribution 3.0 Unported

Although first described in histones, acetylation is also observed in cytoplasmic proteins. Acetylated proteins can also be modulated by cross-talk with other posttranslational modifications, including phosphorylation, ubiquitination, and methylation. Therefore, acetylation may contribute to cell biology beyond transcriptional regulation.

Protein Glycosylation

Protein glycosylation involves adding diverse sugar moieties to the protein core. Glycosylation significantly affects protein folding, conformation, distribution, stability, and activity. Glycosylated proteins can have additions of simple monosaccharides. For

example, many nuclear transcription factors are modified in this way. Alternatively, some proteins are modified with highly complex branched polysaccharides, such as those on cell surface protein receptors.

More than half of all mammalian proteins are believed to be glycosylated. While proteins exhibit improved stability and trafficking after glycosylation *in vivo*, glycan structures can alter protein functions or activities. Glycan structures are often modified by glycan-processing enzymes working within a cell at any given time. However, the structures are sometimes protein-specific, depending on protein trafficking properties and interactions with other cellular factors.

There are three types of protein glycosylation in higher eukaryotes: N-linked, O-linked, and C-linked. These types reflect their glycosidic linkages to amino acid side chains. In N-linked glycosylation, β-N-acetylglucosamine (GlcNAc) is attached through an amide linkage to the side chain of Asn within a consensus sequence of AsnXaaSer/Thr (Figure 8.8). N-linked glycans have multiple functions. While they act as ligands for glycan-binding proteins in cell-cell communication, they also can regulate glycoprotein aggregation in the plasma membrane and affect the half-life of antibodies, cytokines, and hormones in serum.

O-linked glycosylation in higher eukaryotes occurs through several different mechanisms. The most abundant type of O-linked glycosylation is the ***mucin-type***, which involves the attachment of an α-N-acetylgalactosamine (GalNAc) to the hydroxyl group of Ser/Thr side chains. ***Mucins*** are a family of high molecular weight, heavily glycosylated proteins (glycoconjugates) produced by epithelial tissues in most animals. A key characteristic of mucin proteins is their ability to form gels; therefore, they are a key component in most gel-like secretions, serving functions from lubrication to cell signaling to forming chemical barriers. Aberrant expression of mucin-type O-linked glycans occurs in cancer cells and may provide targets for anticancer vaccines.

O-linked glycosylation occurring with the addition of α-O-mannose is the only form of O-linked glycosylation in yeast. Still, it also occurs in the brains of higher eukaryotes. Higher eukaryotes also have an α-O-fucose modification of Ser/Thr residues. This type of glycosylation modulates signaling pathways during eukaryotic development. Another modification, β-O-galactosylation, may contribute to rheumatoid arthritis.

Finally, C-linked glycosylation involves the addition of α-mannose (Man) to the 2-position of the indole side chain of tryptophan residues. While first identified on ribonuclease 2, it also occurs on other proteins, including some mucin proteins, thrombospondin, and the Ebola virus soluble glycoprotein. Figure 6.6.6 shows examples of glycan adducts.

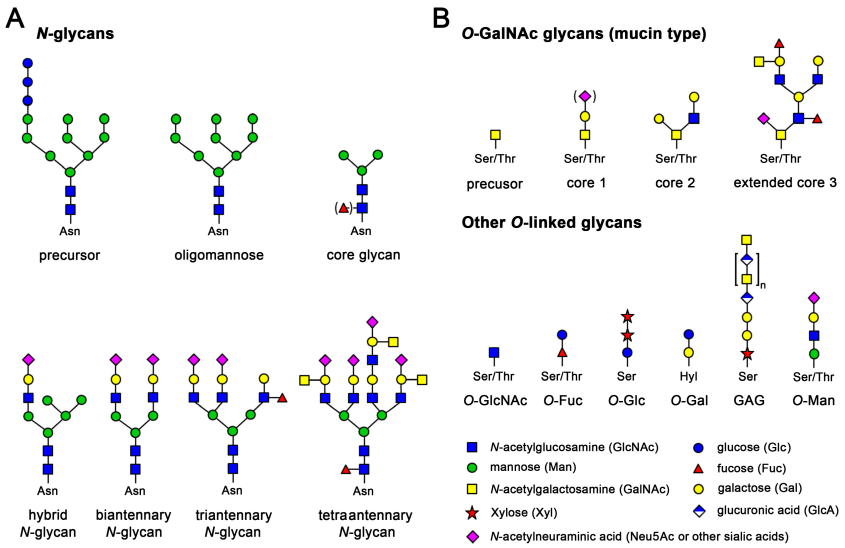


Figure 6.6.6: Glycan adducts to proteins Goettig, P. (2016) Int. J. Mol. Sci. 17(12), 1969

We will discuss glycoprotein in great detail in the next chapter.

Protein Ubiquitination and Sumoylation

Ubiquitination (or ubiquinitylation) is the addition of an 8 kDa polypeptide, called ubiquitin, to lysine residues of target proteins via the C-terminal glycine residues of ubiquitin. Adding one ubiquitin can lead to adding more to form ubiquitin chains on the target protein. The addition of one ubiquitin can modify the activity of the target protein. This can regulate the activity of a protein, its location within the cell, and its interaction with other proteins. Adding a single ubiquitin to lysine 120 (K120) in humans and to K123 in yeast alters gene transcription through the remodeling of chromatin. However, if multiple ubiquitins are added to the target

protein, forming a polyubiquitinated protein, the protein is tagged for degradation by the 26S proteasome. Figure 6.6.7 shows different patterns of ubiquitylation for proteins.

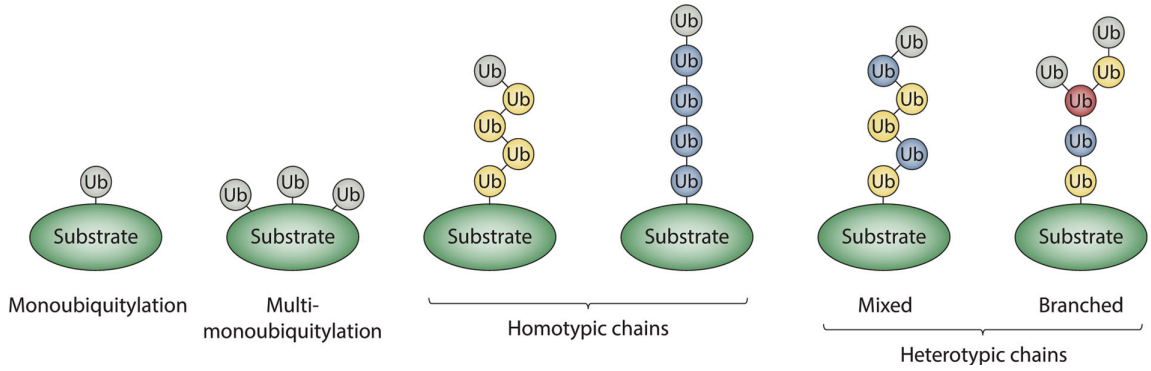


Figure 6.6.7: Classification of ubiquitin modifications. Ubiquitins modified on one acceptor site are colored in blue or yellow; the branch point ubiquitin is colored in red; unmodified or "terminal" ubiquitins are colored gray. French et al. Cell Discovery (2021) 7:6 Cell Discovery. https://doi.org/10.1038/s41421-020-00237-y. Creative Commons Attribution 4.0 International License. http://creativecommons.org/licenses/by/4.0/.

An elaborate system of enzymes attaches a ubiquitin through its C-terminal Gly 76 to an internal lysine (or cysteine) in a target protein. Then it adds more ubiquitins to form the structures shown above. Figure 6.6.8 shows the first step in the process, which requires three enzymes: E1, E2, and E3.

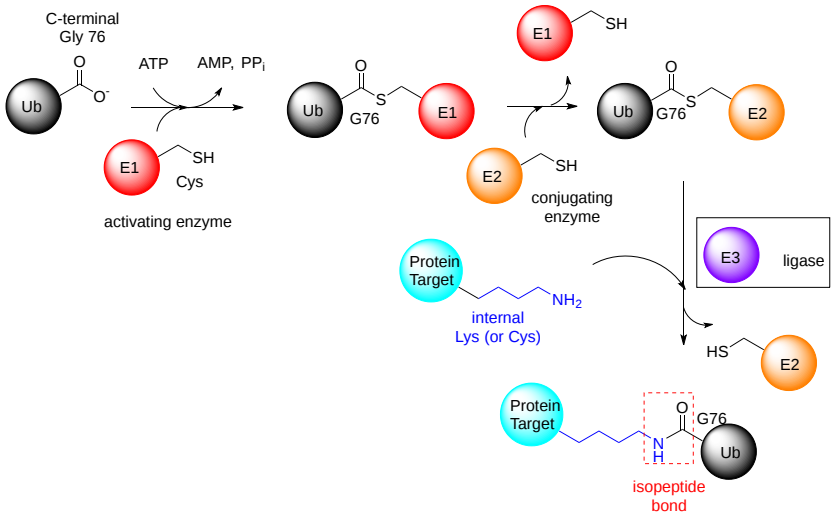


Figure 6.6.8: The first step in the process, which requires three enzymes, E1 (Ub-activating), E2 (Ub-conjugating), and E3 (Ub-ligating) enzymes

Once the first (or proximal) ubiquitin is added, more can be added to form chains. The steps involved in adding a second ubiquitin to the proximal one are shown in Figure 6.6.9.

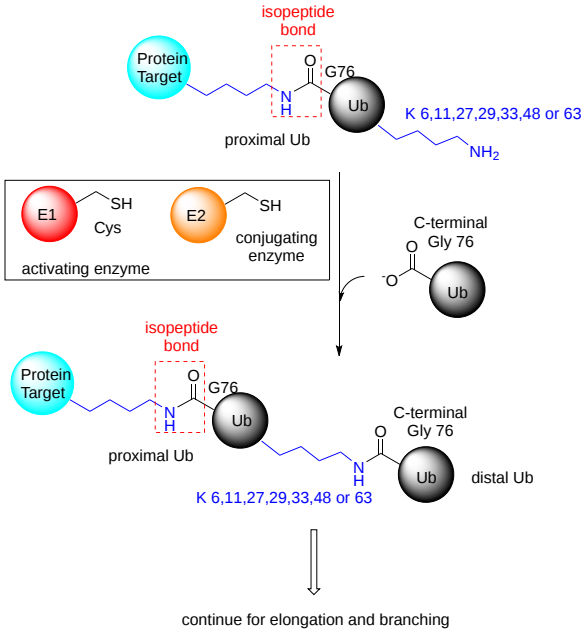


Figure 6.6.19: Addition of additional ubiquitin to form chains

Similarly, protein sumoylation is a reversible posttranslational modification whereby a **s**mall **u**biquitin-like **mo**difier (**SUMO**) protein is covalently attached to a target. Like ubiquitin, SUMO is linked to a lysine side chain of target proteins and is removed by SUMO-specific isopeptidases. Sumoylation controls many aspects of nuclear function. Protein sumoylation is involved in many extranuclear neuronal processes and potentially in many neuropathological conditions.

Figure 6.6.10 shows ubiquitin (1UBQ, magenta) and SUMO-1 (1A5R, cyan) superimposed. Note the similarity in their structures even though they share only 18% sequence identity. Key amino acids are shown in spacefill.

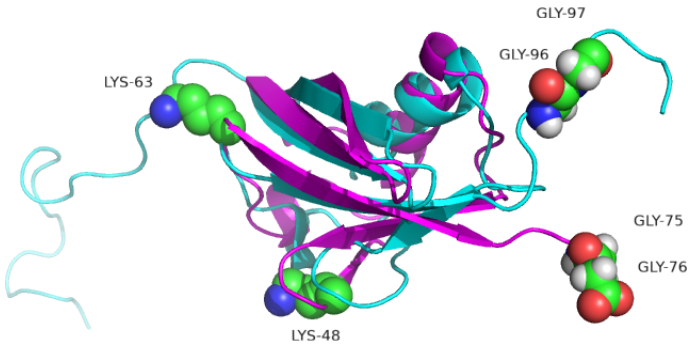


Figure 6.6.10: Ubiquitin (1UBQ) and SUMO-1 (1A5R) superimposed

SUMO-1 doesn't target proteins for degradation. One of its roles is in the transport of proteins between the nucleus and cytoplasm, and it is involved in the modulation of protein-protein interactions. Both ubiquitin and SUMO-1 have C-terminal glycine involved in isopeptide bond formation between ubiquitin or SUMO and the target protein. A significant difference between ubiquitin and SUMO-1 is along the disordered N-terminal end in SUMO. Lysines 48 and 63, found in ubiquitin, are also replaced by other amino acids. This may account for the observation that SUMO-1 does not form polymers.

The interplay between ubiquitylation and SUMOlaytion is complex. If the ubiquitylation of proteins is inhibited, SUMOylation of newly synthesized proteins increases. These can end up in phase-separated condensates called Promyelocytic Leukemia Nuclear Bodies (PML nuclear bodies), 0.1–1 μm condensates without a membrane found in most mammalian cell nuclei. These are somehow involved in gene expression in the nucleus.

Protein Oxidation

The reaction of proteins with a variety of free radicals and reactive oxygen species (ROS) leads to oxidative protein modifications such as the formation of protein hydroperoxides, hydroxylation of aromatic groups and aliphatic amino acid side chains, oxidation

of sulfhydryl groups, oxidation of methionine residues, conversion of some amino acid residues into carbonyl groups, cleavage of the polypeptide chain, and formation of cross-linking bonds. Aromatic and sulfur-containing amino acid residues are particularly susceptible to oxidative modification.

Unless repaired or removed from cells, oxidized proteins are often toxic and can impair cellular viability since oxidatively modified proteins can form large aggregates. Oxidatively damaged proteins undergo selective proteolysis, primarily by the 26S proteasome, in a ubiquitin- and ATP-independent way. Upon extensive protein oxidation, these aggregates can become progressively resistant to proteolytic digestion, bind the 20S proteasome, and irreversibly inhibit its activity.

Protein carbonylation is defined as an irreversible posttranslational modification (PTM) whereby a reactive carbonyl moiety, such as an aldehyde, ketone, or lactam, is introduced into a protein. The first identified source of protein-bound carbonyls was metal-catalyzed oxidation (MCO). MCO results from the Fenton reaction when transition metal ions are reduced in the presence of hydrogen peroxide, generating highly reactive hydroxyl radicals. These hydroxyl radicals can oxidize amino acid side chains or cleave the protein backbone, leading to numerous modifications, including reactive carbonyls. For example, the oxidation of proline and arginine results in the production of glutamic semialdehyde, while lysine is oxidized to aminoadipic semialdehyde and threonine to 2-amino-3-ketobutyric acid. Direct oxidation of other amino acid residues can also lead to protein-bound carbonyls. Tryptophan oxidation by ROS produces at least seven oxidation products. Among them are kynurenine and N-formyl kynurenine and their hydroxylated analogs, which contain aldehyde or keto groups formed by oxidative cleavage of the indole ring.

Another important source of protein-bound carbonyls is reactive lipid peroxidation products, which are produced during the oxidation of polyunsaturated fatty acids. Protein carbonylation can also occur via glycoxidation. Reactive α-carbonyls formed during glycoxidation, such as glyoxal, methylglyoxal, and 3-deoxyglucosone, can then modify the basic residues Lys and Arg to generate, for example, pyrralines and imidazolones. Glycation (i.e., the reaction of reducing sugars such as glucose or fructose with the side chains of lysine and arginine residues) forms Amadori and/or Hynes products. ROS can further change these glycated residues into advanced glycation end products (AGE) carrying carbonylated moieties.

Some oxidative modifications are made enzymatically and have key regulatory or structural functions within the modified proteins. For example, proline can be converted to hydroxyproline and lysine to hydroxylysine, as shown in Figure 6.6.11. 4-Hydroxyproline makes up about 13.5% of the residues within the mammalian collagen family of proteins. Recall that collagen is the main protein of the connective tissue and represents about one-fourth of the total protein content in many animals. Hydroxyproline contributes to the stability of the triple helix and also aids in cross-linking collagen fibers to form larger macromolecular complexes.

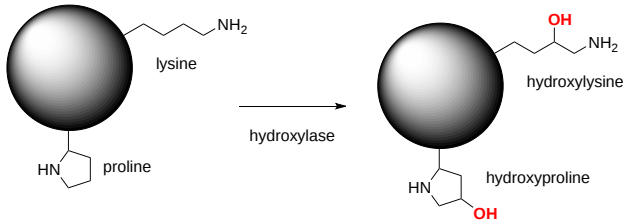


Figure 6.6.11: Hydroxlation of lysine and proline. Xu, Y., et. al. (2014) Int J Mol Sci 15(5):7594-7610.

Protein Methylation

Alkyl substituents are also a common posttranslational modification. Introducing such alkyl groups results in the alteration of the hydrophobicity of the modified protein. The most common type of protein alkylation is protein methylation, which is mediated by methyltransferase enzymes. One-carbon methyl groups are added to nitrogen or oxygen (N- and O-methylation, resp.) on amino acid side chains, increasing protein hydrophobicity or neutralizing a negative charge when bound to carboxylic acids. While N-methylation is typically irreversible, O-methylation is potentially reversible. Methylation occurs so often that its primary methyl donor, S-adenosyl methionine (SAM), is one of the most-used enzymatic substrates after ATP. The methylation of a histone protein is shown in Figure 6.6.12

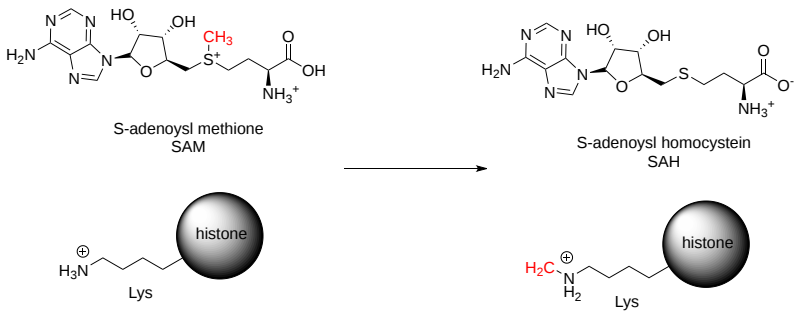


Figure 6.6.12: Methylation of a histone protein by SAM

This modification (not unlike phosphorylation reactions) plays a role in regulating protein-protein interactions. For instance, the arginine methylation of proteins can either inhibit or promote protein-protein interactions depending on the type of methylation. Protein methylation is also a common modification found in ***histone proteins.*** The transfer of methyl groups from S-adenosyl methionine to histones is catalyzed by enzymes known as ***histone methyltransferases***. The N-terminal tails of histones H3 and H4 receive methyl groups on specific lysines. Methylation then determines if gene transcription is activated or repressed, thus leading to different biological outcomes.

Histone methylation was traditionally thought to be irreversible. However, ***histone demethylases*** demonstrate the reversibility of this PTM. The simultaneous removal of one histone methylation and the addition of another can enable transcriptional fine-tuning.

In addition to methylation, histone acetylation, deacetylation, and mono-ubiquitination are essential parts of gene regulation, as shown below in Figure \(\PageIndex{13\). Acetylation removes the positive charge on the histones, thereby decreasing the interaction of the N termini of histones with the negatively charged phosphate groups of DNA. Consequently, condensed chromatin is transformed into a more relaxed structure associated with greater gene transcription levels.

Nonhistone proteins also exhibit methylation as a common PTM, which is used to regulate signal transduction pathways. Furthermore, methylation works in concert with other types of PTMs to exert influence not only on chromatin remodeling but also on gene transcription, protein synthesis, and DNA repair.

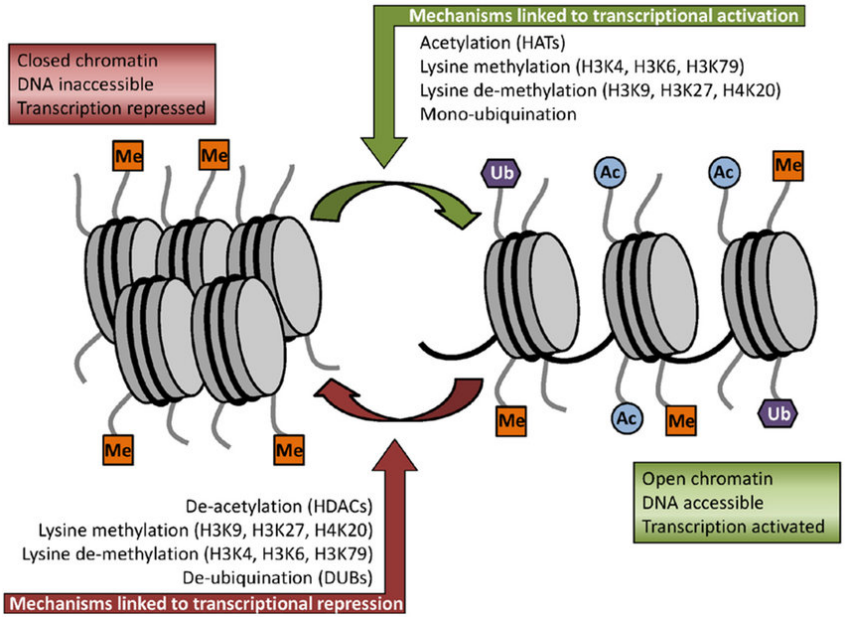


Figure \(\PageIndex{13\): Some of the key histone modifications influencing gene expression (Me: methylation, Ub: ubiquination, Ac: acetylation). Figure from: Ullah, M.F. (2015) Medicines 2:141-156.

6.6.2: Allosteric Regulation

Allosteric regulation fine-tunes most biological processes, including signal transduction, enzyme activity, metabolism, and transport. ***Allostery***, an intrinsic property of a protein, is referred to as the regulation of activity at one site (also known as an ***orthosteric site***) in a protein by a topographically and spatially distant site; the latter is designated as an ***allosteric site. Allosteric***

regulation occurs through binding a modulator (e.g., small molecule or protein) at an allosteric site to engender a conformational change that affects function at the orthosteric site. This effect may cause the re-distribution of the conformational ensemble by either stabilizing an active conformation (***allosteric activation***) or destabilizing an inactive conformation (***allosteric inhibition***) in response to allosteric perturbations (Figure 8.13). Traditionally, the repertoire of allostery was primarily confined to determining the allosteric effects or mechanisms in individual multi-subunit or monomer proteins by conformational transitions. Recently, increasing evidence has indicated that allosteric signals can propagate across several or numerous proteins to sculpt allosteric networks.

Allosteric regulation is particularly important in the cell's ability to adjust enzyme activity based on the surrounding environmental conditions. Feedback control loops, such as feedback inhibition from downstream products or feedforward from upstream substrates, are common allosteric regulatory mechanisms found in nature. Another example of allostery includes oxygen binding to one of the subunits of hemoglobin that prompt cooperative binding to other subunits.

We mentioned before that PTMs can lead to conformational changes in proteins. Other methods to induce such allosteric changes include the binding of small molecules, the binding of proteins, changes in the redox environment of a protein, which affects disulfide bonds, and general changes in protein flexibility and dynamics. Let's look at some examples of allostery brought about by these changes. Many of these examples are presented by Laskowski et al. (https://doi.org/10.1016/j.febslet.2009.03.019).

Allosterism by small molecules: opening and closing active sites

Small molecule binding at an allosteric site can lead to small and large conformation changes in a distal active site. An often observed motion is a partial or full hinge-clamping conformational change that brings critical catalytic groups into a more organized and effective active site, which excludes water as a competing hydrolytic substrate (for example). One example of a small hinge-clamping change is seen in phosphoglycerate dehydrogenase (PGDH, which catalyzes the first step (oxidation by NAD^+) in synthesizing serine. It converts 3-phosphoglycerate into 3-phosphohydropyruvate. The end product of the pathway, serine, binds to an allosteric site on PGDH and inhibits it, a classic example of feedback inhibition of the first reaction of a pathway by the end product. PGDH has a regulatory binding domain (RBD) that binds serine, a substrate binding domain (SBD) and an NAD^+/NADH binding domain (NBD) which is where the allosteric inhibitor binds, plus binding domains for substrate (SBD) and the NAD nucleotide (NBD). Figure 6.6.14 shows the small hinge-bending conformational change on bind conversion of monomeric *apo*-PGDH (1psd) to *holo*-PGDH with serine (not shown) bound in the RBD allosteric domain. The RBD-SBD domains, relative to the NBD domain, undergoes a 15° rotation.

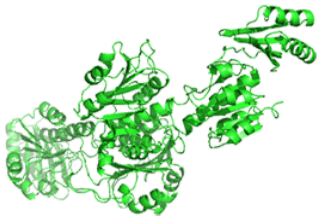


Figure 6.6.14: Small hinge-bending conformational changes on bind conversion of monomeric *apo*-PGDH (1psd) to *holo*-PGDH on binding the allosteric end-product inhibitor serine (not shown)

The functional enzyme is a tetramer (only the monomer is shown in the above figure). The rotation in the full tetramer aligns catalytic residues and closes off the active site.

Conformational changes are ubiquitous on the binding of ligands to proteins. These changes also occur on substrate binding at the active site. It is important to recognize that substrate-induced changes in the active site are **not** allosteric since the conformational changes, no matter how big, are caused by binding at the active site, not at an allosteric site. A classic example of a large conformational change on binding occurs when glucose or glucose analogs bind to the active site of hexokinase, the first enzyme in the glycolytic pathway. Figure 6.6.15 shows the large conformational change on binding a glucose analog to hexokinase (1hkg, no glucose; 2yhx, with glucose analog). The background is shown in black to remind readers that this is **NOT** an example of allosterism, although the changes facilitate catalysis by excluding water, a competing substrate) from the active site after glucose and ATP (not shown) bind. On binding the glucose analog in the active, the active site becomes sequestered ("jaw" clamping down on binding) from the solvent.

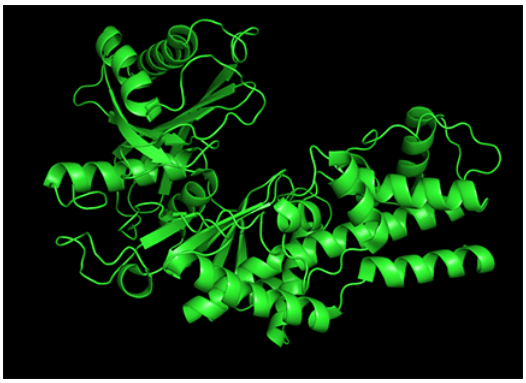


Figure 6.6.15: Large, substrate-induced (non-allosteric) conformational change on binding a glucose analog to (1hkg, no glucose to 2yhx, with a glucose analog.

Allosterism by small molecules: Subtitle electrostatic effects

Often, no large conformation change is evident in the protein. In those cases, subtle rearrangements of key residues in the active site or near it (which might promote access) may result from allosteric effector binding. A simple change in the electrostatic environment might account for the effect. Such appears to be the case with chorismate mutase, an enzyme in the bacterial, fungal, and plant pathways for synthesizing aromatic amino acids tyrosine and phenylalanine. The enzyme catalyzes the conversion of chorismate to prephenate, an intermediate in synthesizing aromatic amino acids. An offshoot pathway takes chorismate to the last aromatic amino acid, tryptophan. The enzyme chorismate mutase is key in the metabolic decision pathway to make tyrosine/phenylalanine or tryptophan. The enzyme is activated by tryptophan and inhibited by tyrosine. The aligned sequences of two crystal structures of the enzyme with the bound allosteric activator tryptophan (Trp 502 in pdb structure 1csm showing a magenta glutamic acid 23) and with bound inhibitor (Tyr 300 in pdb structure 2csmA showing cyan Glu 23) are shown in Figure 6.6.16

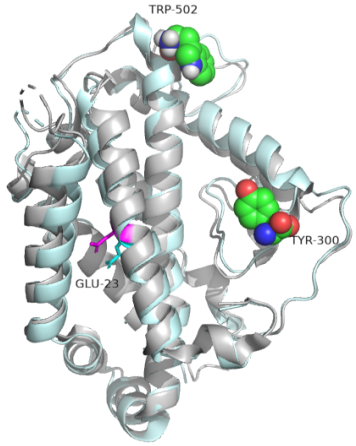


Figure 6.6.16: Aligned structures of chorismate mutase with bound allosteric activator tryptophan (Trp 502 , pdbID 1csm , magenta glutamic acid 23) and with bound inhibitor (Tyr 300 pdbID 2csmA, cyan Glu 23)

No significant changes in the protein's overall structure or active site occur. However, the alignment of glutamic acid 23 is different, which may be the basis for the observed allosteric effects. When active, glutamic acid 23 is buried in the active site pocket, but in the inhibited site, it swings into the binding site. Since the chorismate is a charged dicarboxylic acid, Glu 23 probably repels its binding, inhibiting the enzyme's activity.

Allosterism by phosphorylation

We'll focus on one example of a huge conformational change on phosphorylation of a serine side chain in glycogen phosphorylase (GP). This enzyme catalyzes the phosphorolysis of an acetal link between a terminal glucose on a glucose polymer, glycogen, and the next glucose in the chain. This reaction is not a hydrolysis, in which water acts as a nucleophile to cleave acetal bonds. This is a key reaction in metabolism since it cleaves a major energy storage molecule, glycogen.

The enzyme is a multimer that exists in two major states, a T-state and an R-state glycogen phosphorylase (GP). We explored T and R states and allosterism in their interconversion when we discussed oxygen binding to hemoglobin in Chapter 5.3. Crystal

structures of the T and R-state bound to allosteric activators IMP and AMP are known. High AMP concentrations imply low ATP concentrations and the need to mobilize glycogen reserves. Conversion to the R state is also promoted by the phosphorylation of Serine 14, which activates the enzyme. The allosteric activators AMP and IMP bind to a disordered loop (313-326) and change it to a form promoted by nucleotide binding. which is disordered in the free structure and adopts a conformation dictated mainly by the type of nucleotide that binds at this site. Figure 6.6.17 shows a monomer of the European rabbit glycogen phosphorylase in the R (active) state bound to AMP (3e3n).

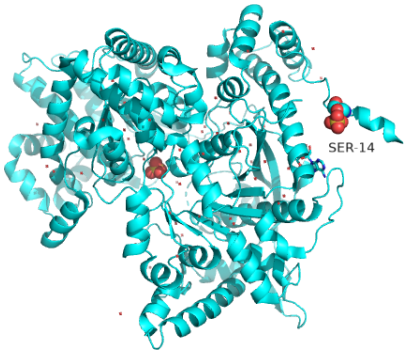


Figure 6.6.17: Monomer of the European rabbit glycogen phosphorylase in the R (active) state bound to AMP (3e3n).

Only one chain of the active homodimer is shown. The helix containing the phospho-Ser 14 swings away from the rest of the protein.

Figure 6.6.18 shows the difference between the T (inactive) state of rabbit muscle glycogen phosphorylase (7P7D) and the R (active) state (3e3n, neither phospho-Ser 14 nor AMP shown). Note the huge conformational state elicited on phosphorylation of Ser 13.

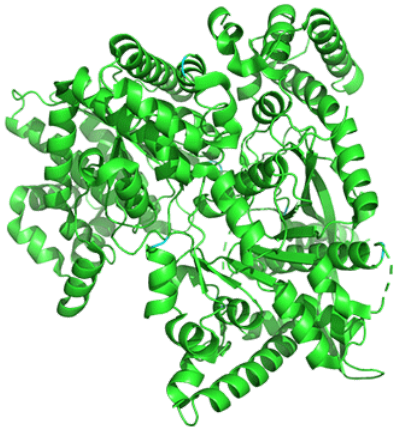


Figure 6.6.18: Conformational changes going from T (inactive) state of rabbit muscle glycogen phosphorylase (7P7D) to the R (active) state (3e3n) on phosphorylation of Ser 13.

Allosterism by small molecules: control of quaternary structure

The binding of small molecules at allosteric sites may promote or inhibit the formation of a protein's correct functional multimeric structure. One example is ATP phosphoribosyltransferase (APRT), which catalyzes the first step in synthesizing histidine. As we saw with phosphoglycerate dehydrogenase (PGDH), APRT has three domains. I and II comprise the active site that is located between them, and a regulatory domain, III (the functional protein is a dimer). As with PGDH, APRT is allosterically inhibited by the end product of the pathway, histidine. On histidine binding to the inactive form (1nh7), a large conformational change results (1nh8), as shown in Figure 6.6.19.

Figure 6.6.19: Conformational changes on pathway end product histidine binding to the inactive form (1nh7) form of ATP phosphoribosyltransferase (APRT) inactivating the enzyme (1nh8)

This change drives the formation of a hexamer - $(dimer)_3$, which closes off the active site and inactivates the enzyme. The enzyme ribonuclease reductase behaves similarly.

Allosteric by protein binding

Another protein can bind to an enzyme and activate it by promoting an allosteric change. An example is the binding of a regulatory protein, CyclinA, to cyclin-dependent kinase 2 (CDK2). A large conformational change occurs on binding cyclin to one side of CDK2's catalytic cleft, activating the enzyme by altering active site geometry and making the active site more accessible. CDK enzymes are involved in cell-cycle progression. For more control of the cell cycle, the binding of cyclin and the phosphorylation of CDK2 are required for activity. On binding cyclin, the active site is more available for substrate (ATP) binding. Next, the amino acid that needs to be phosphorylated for activation, Thr 16, is made accessible. Figure 6.6.20 shows the conformational change in apo-CDK2 (1hc1) on cyclin binding, which allows ATP binding (1fin).

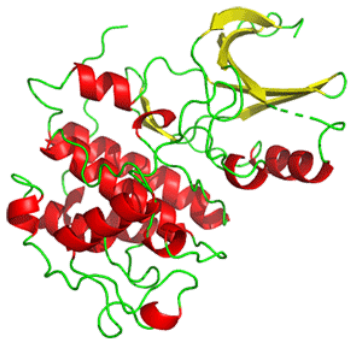


Figure 6.6.20 show the conformational change in apo-CDK2 (1hc1) on cyclin binding which allows ATP binding (1fin).

Cyclin (the regulatory protein) is shown in gray. ATP is shown in spacefill. CDK2 is shown as a colored cartoon.

Allosterism by disulfide bonds

It should make sense that cleaving an intrachain or interchain disulfide, whose presence constrains a specific protein conformation, should significantly alter protein structure and function. One example is the protein botulinum neurotoxin type A, a neurotoxin and Zn protease. It cleaves proteins like synaptobrevins, syntaxin, and SNAP-25 in the neurosynapse required for neurotransmitter release.

It has a catalytic (Zn peptidase/protease) and a translocation domain in a single protein chain that contains two disulfide bonds when synthesized protein (PDB code 3bta). On cleavage at a select peptide bond, it is split into A and B chains, which remain connected by a single disulfide bond, with the A chain containing the catalytic domain. The B chain contains the translocation domain, which effectively still blocks the active site of the peptidase/protease domain. The disulfides are cleaved when the protein enters an endosome, which contains a more reducing environment. The separate A chain now separates for the inhibiting B domain and expresses protease activity. The active protein then enters the cytoplasm.

Figure 6.6.21 shows the domain structure and disulfide bonds of the unprocessed form. The N-terminal Zn proteinase domain (Peptidase_M27, amino acids 1-409) is followed by an intrachain disulfide between C429 and C 453.

Figure 6.6.21: The domain structure and disulfide bonds of the unprocessed form of botulinum neurotoxin type A

The mature form is proteolyzed between amino acids 447 and 448 (within the sequences connected by the disulfide (429-453). The protein remains inactive until the disulfide bond is cleaved.

Figure 6.6.22 shows an interactive iCn3D model of botulinum neurotoxin type A (3bta).

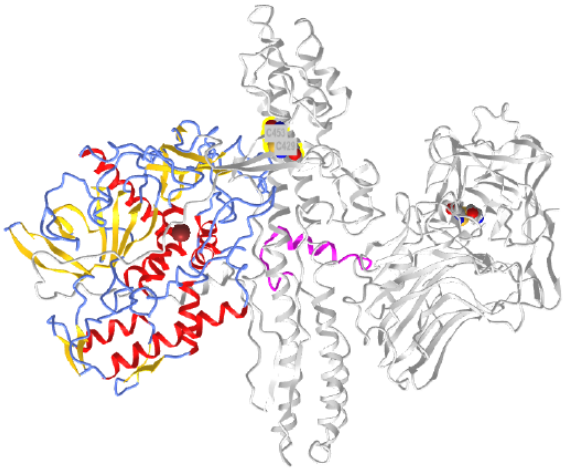


Figure 6.6.22: Botulinum neurotoxin type A (3bta) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...CfSsBH94jcPyb8

The N-terminal Zn peptidase/protease domain is shown in color, and the other regions of the protein are shown in gray. The two disulfide bonds are shown in spacefill, with CPK color. On cleavage of the disulfide bond between C429 and C453, the Zn protease domain, which is inhibited in the presence of the constraining disulfides, is released and activated.

Allosterism by changes in protein flexibility and population shifts

Sometimes, crystal structure shows little difference between the apo- and holo-forms of the protein containing a separate allosteric binding site. One example is the enzyme dihydrodipicolinate synthase (DHDPS) which catalyzes the first step in the bacterial and plant pathways for lysine synthesis. DHDPS catalyzes the synthesis of 4-hydroxy-2,3,4,5-tetrahydro-2-dipicolinic acid from aspartate-β-semialdehyde and pyruvate. One explanation for allosterism in the protein is that it occurs on time scales too short to be seen in crystal or NMR structures. Rather some transient conformation might cause allosterism. Such changes are detectable in molecular dynamics simulation of the protein in the presence and absence of the allosteric effector molecule.

This possibility suggests that the allosteric effector might change the distribution of possible transient conformational states within a population of such states. If an allosteric effector bound one of the transient states in the population, that state would be "removed" from the ensemble of states, shifting the dynamic equilibrium among the populated states to produce more of the specific conformation that binds the allosteric effector. This sounds like the T to R state conversion in the MWC model oxygen binding to hemoglobin and the allosterism discussed above with the enzyme glycogen phosphorylase, only a much shorter time scale. This model is consistent with a conformational selection model in which the allosteric molecule binds preferentially to a different preexisting conformational state that exists transiently.

6.6.3: Zymogen Activation

A ***zymogen,*** also called a ***proenzyme***, is an inactive precursor of an enzyme. A zymogen requires a biochemical change (such as a hydrolysis reaction revealing the active site or changing the configuration to reveal the active site) to become an active enzyme.

Protease enzymes secreted by the pancreas are initially synthesized as zymogens. The pancreas secretes zymogens to help prevent the enzymes from inappropriately digesting proteins in the pancreatic cells in which they are synthesized. Enzymes like Trypsin are

synthesized as proenzymes. For trypsin, trypsinogen is an inactive precursor translated by the rough endoplasmic reticulum and transported to the Golgi apparatus for sorting. Trypsinogen is always co-synthesized and packed with a pancreatic secretory trypsin inhibitor (PSTI) that inhibits premature activation. Thus, there are two mechanisms in place to maintain the inactivity of the protease within the pancreas: (1) synthesis of the zymogen or proenzyme form, and (2) co-expression of a trypsin inhibitor protein that will bind and inhibit any prematurely cleaved trypsin until it has reached the small intestine. An animation showing structural differences between bovine trypsinogen (magenta) with just amino acids 10-15 of the pre-sequence (spacefill orange) and mature trypsin (cyan) is shown in Figure 6.6.23

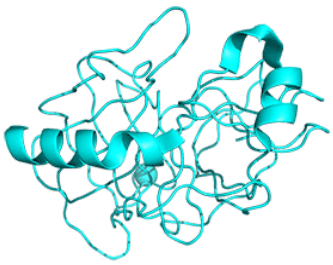


Figure 6.6.23: Animation showing structural differences between bovine trypsinogen (magenta) with just amino acids 10-15 of the pre-sequence (spacefill orange) and mature trypsin (cyan)

During packaging within the Golgi system, the trypsinogen and other digestive enzymes condense into core particles and are packed in zymogen granules. The condensed enzymes are stable, and minimal activation occurs within the zymogen granules. Once the pancreatic cells receive secretory stimuli, these zymogen granules are released into the lumen of the pancreatic duct, which carries the digestive enzymes into the duodenum. Once in the duodenum, enteropeptidase activates trypsinogen by removing the 7-10 amino acids trypsinogen activation peptide (TAP) from the N-terminal region (see above). Removal of TAP induces a conformational change, resulting in active trypsin. TAP is immunologically distinct from the same sequence within trypsinogen, thereby allowing the detection of trypsinogen activation *in situ*.

Once activated, trypsin will cleave and activate other zymogen proteases and lipases in the duodenum. These include the activation of elastase, chymotrypsin, carboxypeptidase, and lipase, as shown in Figure 6.6.24 Zymogens are also found in other cellular processes as well. For example, caspases, intracellular proteases, are activated similarly during cellular apoptosis or programmed cell death. The process of blood clotting also involves the activation of zymogens.

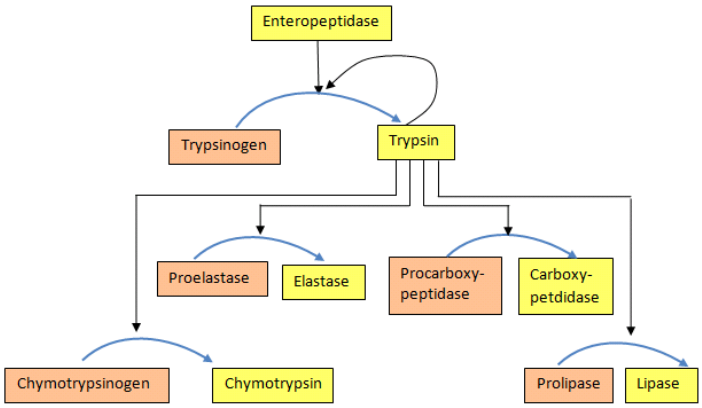


Figure 6.6.24: Activation of proenzymes

6.6.4: Isozymes

Isozymes (also known as **isoenzymes**) are enzymes that differ in amino acid sequence but catalyze the same chemical reaction. These enzymes usually display different kinetic parameters (e.g. different K_M or K_{cat} values), or different regulatory properties. Isozymes permit the fine-tuning of metabolism to meet the particular needs of a given tissue or developmental stage. Often, isozymes are coded for by homologous genes duplicated within the genome and then diverged over time. *Isozymes* should not be confused with **allozymes,** ***allelic variants of the same gene locus*** found within a population. Allozymes represent enzymes from different alleles of the same gene. Isozymes represent enzymes from different genes that process or catalyze the same reaction or

from the same gene that produces, through differential splicing of primary RNA transcripts, sequences of similar but different sequences. We will focus on isozymes within this section. Part of the regulation derived from the production of different isozymes can arise from the differential expression of isozymes.

Isozymes are usually the result of gene duplication. Over evolutionary time, if the function of the new variant remains *identical* to the original, then it is likely that one or the other will be lost as mutations accumulate, resulting in a pseudogene. However, if the mutations do not immediately prevent the enzyme from functioning but instead modify either its function or its pattern of expression. The two variants may both be favored by natural selection and become specialized for different functions. For example, they may be expressed at different stages of development or in different tissues. Some isozymes may also arise from convergent evolution and may not share a high degree of sequence homology or common ancestry.

6.6.4.1: The Cyclooxygenase I and II (Cox-1 and Cox-2) Isozymes

As an example of isozymes, we will discuss **cyclooxygenases COX-1** and **COX-2**, which are also called **Prostaglandin Synthases.** They regulate a key step in prostaglandin and thromboxane synthesis and are the targets of **nonsteroidal anti-inflammatory drugs (NSAIDs)**, such as aspirin, ibuprofen, naproxen, and Celebrex (Figure 6.6.25). Prostaglandins (PG) are a group of physiologically active lipid compounds called eicosanoids with diverse hormone-like effects in animals. Prostaglandins have been found in almost every tissue in humans and other animals. They are derived enzymatically from the fatty acid arachidonic acid. Every prostaglandin contains 20 carbon atoms, including a 5-carbon ring. They are a subclass of eicosanoids and the prostanoid class of fatty acid derivatives.

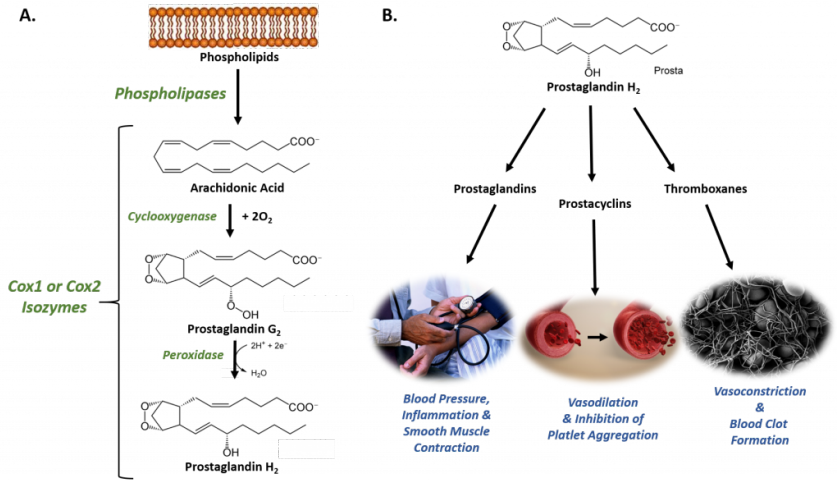


Figure 6.6.25: The Role of COX-1 and COX-2 in Prostaglandin Biosynthesis. (A) COX-1 and COX-2 are bifunctional enzymes that mediate the cyclooxygenase and peroxidase reactions that convert Arachidonic Acid to Prostaglandin H_2. (B) Some of the physiological effects that prostaglandins, prostacyclins, and thromboxanes have on biological processes.Images from: Администрация Волгоградской области , scientific animations, and Fuzis

Cycloxygenases are a class of enzymes called dioxygenases, which incorporate both atoms of O_2 into a substrate. We will explore the mechanism of cyclooxygenase in more detail in Chapter 13.5: Biological Oxidation and Reduction Reactions.

The structural differences between prostaglandins account for their different biological activities. A given prostaglandin may have different and even opposite effects in different tissues. The ability of the same prostaglandin to stimulate a reaction in one tissue and inhibit the same reaction in another tissue is determined by the type of receptor to which the prostaglandin binds. They act as autocrine or paracrine factors, with their target cells present near the site of their secretion. Prostaglandins differ from endocrine hormones in that they are not produced at a specific site but in many places throughout the human body and tend to act locally once secreted. Prostaglandins are implicated in various physiological processes such as gastrointestinal cytoprotection, hemostasis, thrombosis, and renal hemodynamics.

Through their role in vasodilation, prostaglandins are also involved in inflammation and can trigger the onset of a fever or the sensation of pain. They are synthesized in the walls of blood vessels. They prevent needless clot formation and regulate the contraction of smooth muscle tissue. The prostacyclins, a special class of prostaglandins, are powerful, locally-acting vasodilators and inhibit the aggregation of blood platelets. Conversely, thromboxanes (produced by platelet cells) are vasoconstrictors and facilitate platelet aggregation. Their name comes from their role in clot formation or thrombosis.

The cyclooxygenases COX-1 and COX-2 regulate the first two steps in prostaglandin and are bifunctional enzymes containing two active sites. The first active site performs the *bis*-oxygenation and cyclization of arachidonic acid, whereas the second active site mediates a peroxidase (reduction) reaction to form PGH_2. The enzyme is an example of a dioxygenase, which uses O_2 as a substrate. The enzyme contains a heme cofactor. The functional enzymes of both COX-1 and COX-2 are homodimers of 70 kDa subunits, each having an N-terminal epidermal growth factor domain, a membrane binding domain, and a C-terminal catalytic domain. The cyclooxygenase active site is on the opposite side of the peroxidase active site in the catalytic domain.

Figure 6.6.26 shows an interactive iCn3D model of arachidonic acid bound to V349I murine COX-2 (6OFY). Just one chain of the dimer is shown.

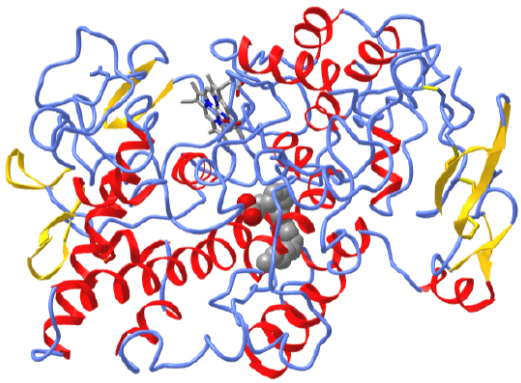


Figure 6.6.26: Arachidonic acid and heme bound to V349I murine COX-2 (6OFY) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?gUtX3uAtwoozBvJUA

The heme is shown in stick while the arachidonic acid is shown in spacefill. Note that arachidonic acid is very kinked and not extended due to its four cis double bonds. Three additional amino acids that play key roles in NSAID inhibition of Cox are highlighted. These include Ser 530 (in model 531), which is covalently acetylated by aspirin, Arg 120 (in model 121), and Tyr 355 (in model 356). These play a key role in the binding of NSAIDs and access to the arachidonic acid binding site.

The COX-1 enzyme is widely distributed in many tissues where it is constitutively expressed. Expression of the COX-2 isoform (shown in Figure 6.6.2), on the other hand, is normally undetectable in most tissues (except for the central nervous system, kidneys, and seminal vesicles). Various inflammatory and mitogenic stimuli induce COX-2. A third isoform named COX-3 was recently identified as a COX-1 splicing variant. This new variant may play a role in processes such as fever and pain. Additionally, a high level of COX-2 expression is found in several forms of cancer. For example, COX-2 overexpression is related to poor prognosis in certain breast cancers and endometrial adenocarcinomas.

COX-2, unlike COX-1, is induced in inflammatory cells and activated by various inflammatory and mitogenic stimuli. Under these conditions, COX-2 activity leads to the production of prostanoid mediators that trigger important inflammatory processes. Although inflammation is initially a necessary process to fight infection, if it is maintained or remains uncontrolled, it can provoke chronic pathologies and tissue damage. This is why COX proteins are key targets for anti-inflammatory and pain management. Their actions are illustrated in Figure 6.6.27.

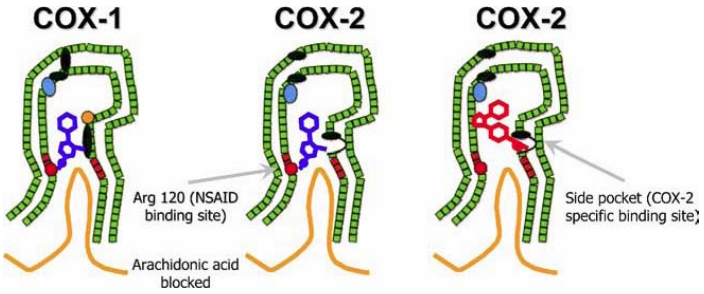


Figure 6.6.27: **NSAID Inhibition of COX-1 and COX-2 Enzymes.** The schematic representation of the COX-1 (large green figure) active site inhibited by a nonselective NSAID (central blue figure). The NSAID blocks the entrance channel to COX-1. The binding and transformation of arachidonic acid (bottom yellow figure) within COX-1 is prevented. *Middle Lower Panel* shows the

inhibition of COX-2 by a nonselective NSAID (central blue figure). The nonselective binding uses an amino acid residue, Arg120, that is conserved in both enzymes. The *right panel* shows the inhibition of COX-2 by COX-2 selective NSAIDs (central red figure). The COX-2 side pocket allows specific binding of the COX-2 selective NSAIDs. The entrance channel to COX-2 is blocked. The bulkier COX-2-selective NSAID will not fit into the narrower COX-1 entrance channel, allowing COX-1 to remain active. Upper Figure by Saiz, M., Gonzalez, R., & Garcia, E., (2019) Protopedia and Lower Figure by Meek, I.L., et al. (2010) Pharmaceuticals 3(7):2146-2162.

6.6.4.2: Nonsteroidal Antiinflammatory Drugs (NSAIDs)

Clinically available NSAIDs can be separated into three different classes based on their mechanism of action:

- **ASPIRIN:** - Acts to irreversibly inhibit COX 1 & COX-2 by covalent acetylation of Ser 530 in the active site. Most notably, low doses of aspirin can suppress platelet COX-1 activity by 95% or more, an effect that is permanent for the lifetime of the platelet since platelets lack DNA and cannot synthesize new enzymes. Due to aspirin's antithrombotic properties at low doses, this treatment has been found to have cardioprotective effects and is often prescribed for patients at high risk of myocardial infarction. All other NSAIDs interact with COX isoforms reversibly and produce variable COX inhibition (ranging from 50% to 95%) in a time-dependent fashion based on how quickly they are metabolized in the body.
- **NON-SELECTIVE COX INHIBITORS:** Different non-selective NSAIDs have varying inhibitory effects against COX-1 & COX-2 (Figure 8.3). The two most commonly used over-the-counter drugs in this group (ibuprofen & naproxen) produce reversible platelet inhibition ranging from 50 to 95% in a reversible, time-dependent manner (see below). These NSAIDs may be insufficient to provide cardio-protection throughout a commonly used dosing interval and are not commonly used for this purpose. Ketorolac (Toradol ®), an NSAID most commonly used in a hospital setting to treat moderately severe pain, is classified as a non-selective NSAID. However, it is, arguably, a very selective COX-1 inhibitor (Figure 8.3). Inhibition of COX-1 can result in unwanted side effects, such as gastrointestinal discomfort and, in severe cases, ulceration.
- **COXIBS:** Selective COX-2 inhibitors were designed and marketed to treat pain and inflammation while avoiding the GI side effects. However, soon after they were introduced into the market, their use led to the first reported incidence of increased cardiovascular events (myocardial infarction and stroke) in 2004. Rofecoxib (Vioxx ®), one of the most selective COX-2 inhibitors, was removed from the market because of mounting evidence for significant cardiovascular toxicity (Drazen, 2005). Celecoxib (Celebrex ®) is currently the only FDA-approved coxib available in the US, and it has been given a black box warning indicating the potential risk of cardiovascular toxicity. It has a 10-20-fold selectivity for COX-2 over COX-1. Etoricoxib (Arcoxia ®) is a second coxib with ~10^6-fold selectivity for COX-2 over COX-1 that is available outside the United States.

Figure 6.6.28shows the Selectivity and Treatment Efficacy of Nonsteroidal Anti-inflammatory Drugs (NSAIDs).

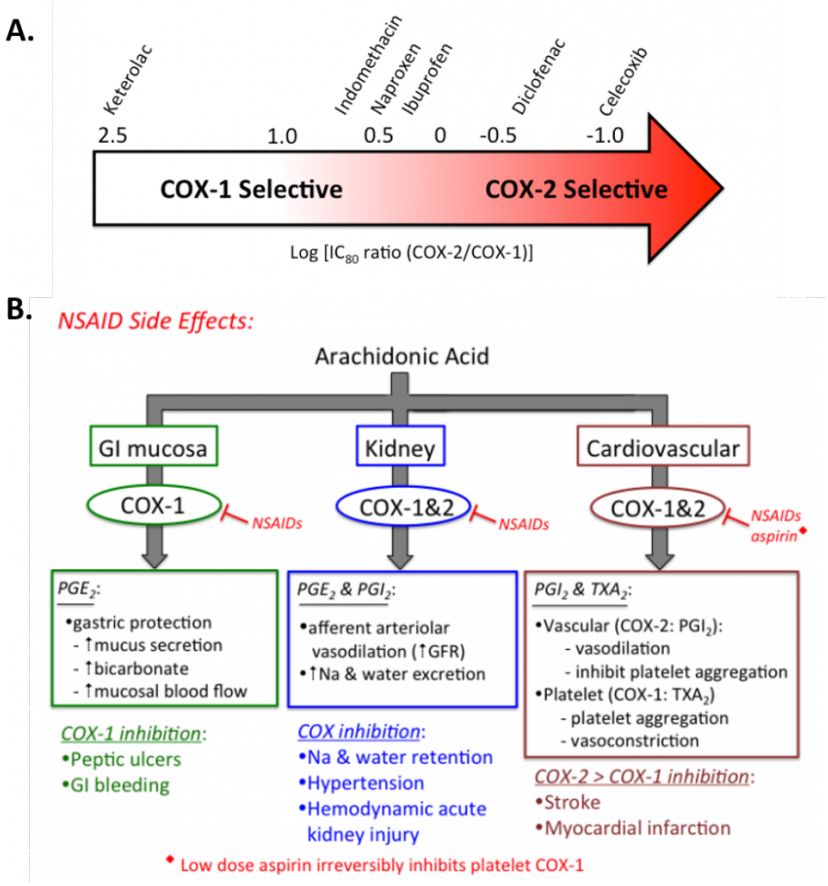


Figure 6.6.28: Selectivity and Treatment Efficacy of Nonsteroidal Anti-inflammatory Drugs (NSAIDs). (A) Relative COX-1 & COX-2 selectivity for commonly used non-aspirin NSAIDs. Celecoxib (Celebrex ®) is the only COX-2 selective NSAID on the market in the US. Adapted from Danelich et al (2015). (B) Major physiological roles for COX-1 & COX-2, and mechanisms underlying drug-induced side effects. PGI_2: prostacyclin, TXA_2: thromboxane.*Figures from: Clarkson, C.W. (2018) TUSOM Pharmwiki*

The structures of ibuprofen (Advil), naproxen, aspirin, Celebrex, and acetaminophen (Tylenol, which reduces fever and pain but not inflammation, so it's not an NSAID) are shown in Figure 6.6.29

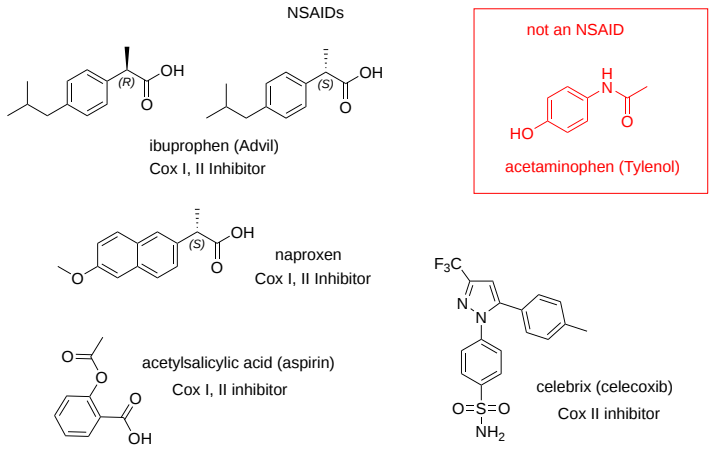


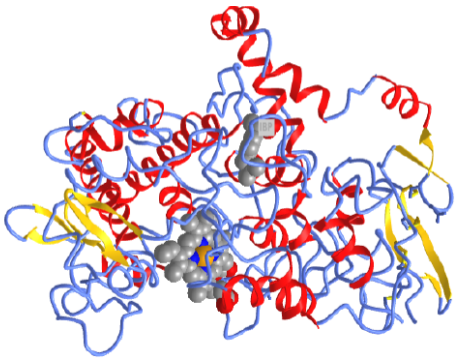
Figure 6.6.29: Structures of ibuprofen, naproxen, celebrix (NSAIDs) and acetaminophen (not an NSAID)

Acetominophen - Tylenol

Although this extremely popular drug relieves pain and reduces fever, it does **not** reduce peripheral inflammation and is hence **not classified as an NSAID**. It does have effects on the central nervous system. It does **not** appear to bind to the active site of COX-1 or COX-2 (no PDB structures available), but it can reduce their activities probably either upstream or downstream of

the pathways these enzymes are part of. It does decrease prostaglandin production in the CNS, which reduces CNS pain and fever. It might act as a reducing substrate in the peroxidase site. Many proteins bind this drug. which is understandable given its relatively simple structure shown in Figure 6.6.5. A metabolite of it, p-aminophenol, can be esterified to arachidonic acid to produce the fatty acid amide, which can work through cannabinoid receptors to reduce pain.

Figure 6.6.30 shows an interactive iCn3D model of the NSAID ibuprofen bound to cyclooxygenase-2 (4PH9). Just one chain of the functional dimer is shown.



NCBI iCn3D Figure 6.6.30: Ibuprofen bound to cyclooxygenase-2 (4PH9). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...syYYjSoeZRGNU8

Ibuprofen is sold as the racemic mixture of R and S enantiomers. Still, it was thought that only the S isomer's inhibition of Cox II led to its antiinflammatory effects, while the other isomer had no effects and no side effects. The crystal structures of IBP bound to both Cox I and II are known. At least in the mouse Cox2, the S-isomer binds more tightly than the R-isomer. Arg-120 and Tyr-355 at the entrance of the cyclooxygenase channel are important in the action of IBP in Cox2 (which are labeled in Figure 6.6.6.

As mentioned above, aspirin inhibits COXs through the covalent acetylation of a key serine side chain. The other NSAIDs inhibit prostaglandin H(2) synthase through blockage of the channel for arachidonic acid binding. This should make you think they act as competitive inhibitors, which is true for ibuprofen and a methyl flurbiprofen derivative. However, the effect of two others NSAIDs, aclofenac and flurbiprofen, seems to have an added time component not found in simple competitive inhibition model. They are called slow tight-binding inhibitors, and a simple model describing it is shown in Figure 6.6.31.

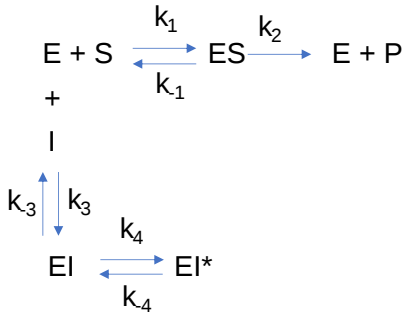


Figure 6.6.31: Model for slow binding inhibition

A time step suggests a conformational change after initial binding. Yet crystal studies of the four inhibitors mentioned above show that the NSAID-occupied active sites are essentially the same, suggesting no major global conformational change. It may be that the rate involves slow hydrogen bonding around two key residues, Arg 120 and Tyr 355.

Slow binding and slow tight-binding inhibition

All the enzyme inhibition models we discussed in Chapter 6.4 are based on reversible rapid equilibrium binding of an inhibitor to E (competitive), ES (uncompetitive), and E and ES (mixed/noncompetitive). Equations can be derived for **slow binding** and **slow tight-binding** inhibition. In these models, the degree of inhibition at a fixed inhibitor concentration changes with time to establish an "equilibrium" among all species. The model above works for both. Both types of inhibitors (slow and slow-tight) inhibit enzymes in a time-dependent fashion. In slow binding, it takes longer to establish equilibrium among the three

species, E, EI, and EI*. In slow binding, the steps determining the concentration of EI (k_3I and k_{-3}) are assumed to be fast compared to those involved in determining the concentration of EI* (k_4 and k_{-4}). When [I] >> [E], the following dissociation constants for pure competitive ($K_{IS} = K_C$) and for the slow binding inhibition (K_I), which accounts for mass balance, hold.

$$K_{IS} = K_C = \frac{[E][I]}{[EI]} \tag{6.6.1}$$

$$K_I^* = \frac{[E][I]}{[EI] + [EI^*]} \tag{6.6.2}$$

In initial rate Michaelis-Menten kinetics, v_0 is measured as a function of [S]. The initial rate v_0 is determined by measuring [P] vs. t at the beginning of the reaction when the substrate is not depleted. Valid v_0 values require that [P] vs t curves are linear. The slope of that line is the initial velocity, v_0. This is true for uninhibited and reversibly inhibited (competitive, uncompetitive, and mixed) reactions. However, with slow binding inhibition, the [P] vs t curves bend with time as the slower formation of EI* occurs. It's a bit like the pre-steady state assumption in enzyme kinetics. At very short times (msec), the P vs. t curves bend as a steady state emerges. It's the same with slow binding inhibition. This is illustrated in Figure 6.6.32

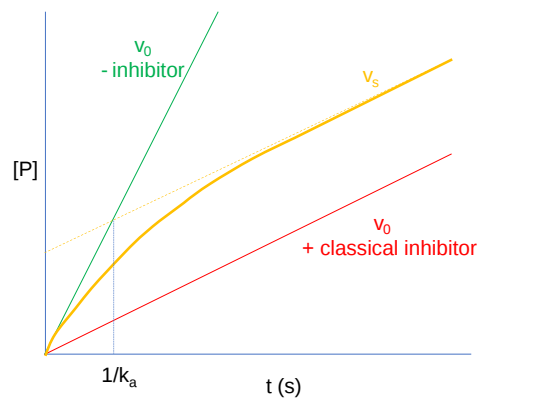


Figure 6.6.32: P vs. t curves for uninhibited (green), classical inhibition (red), and slow binding inhibition (gold)

A progress curve equation showing [P] vs. t for slow binding in inhibition is shown below without inhibition.

$$[P] = v_s t + \frac{(v_o - v_s)\left(1 - e^{-k_a t}\right)}{k_a} + C \tag{6.6.3}$$

In this equation, k_a is the apparent first-order rate constant to the development of the steady state at a given substrate and inhibitor concentration, v_0 is the very first initial rate, and v_s is the steady-state rate.

Slow tight-binding inhibition occurs when the rate constants for net production of EI* are fast compared to the step for forming EI. In this case, even small amounts of I will produce EI* as the reaction is pulled to EI*, and more complicated equations must be used to determine product vs time equations.

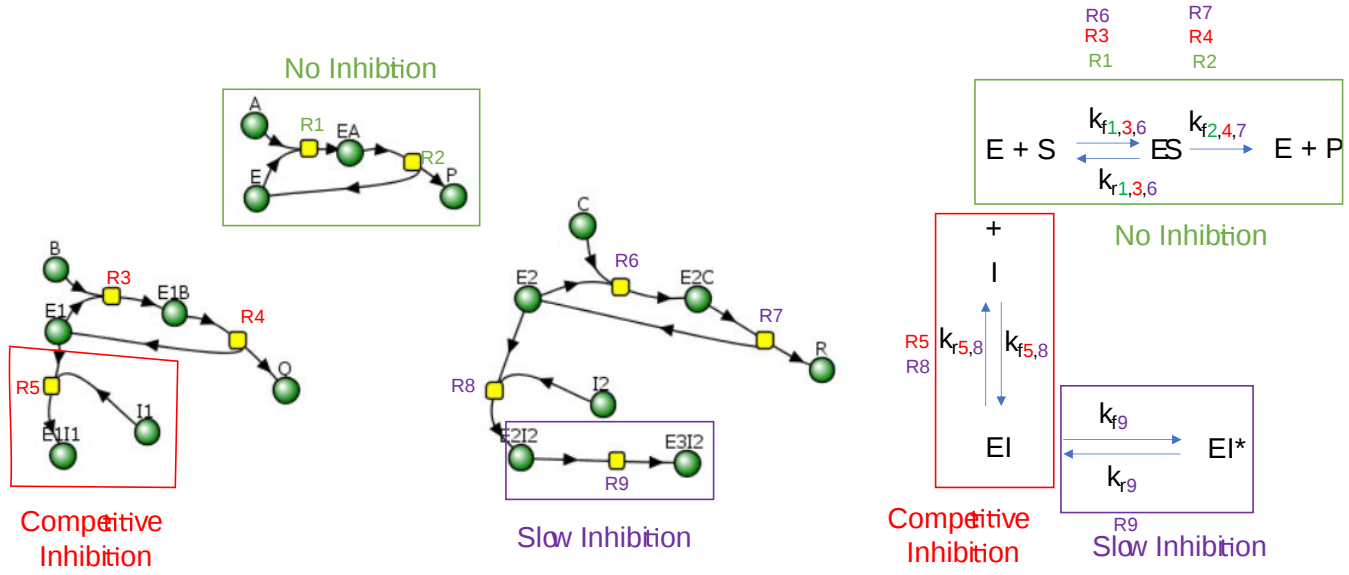
Other chemical models could also account for slow binding inhibition. These include a slow binding rate constant k_4, a slow isomerization of EI after fast binding of I, or a slow binding to a specific, low population conformation of enzyme in a process called conformational selection. Slow-binding inhibitors of enzymes (such as neural acetylcholine esterase) are known.

Here is a Vcell simulation showing progress curves for an uninhibited reaction, one in the presence of a competitive inhibitor and one in the presence of a slow (competitive) inhibitor.

VCell MODEL

Comparison No Inhibition, Competitive Inhibition, and Slow (Competitive) Inhibition

Vcell reaction diagram with all equations based on mass action (not Michael-Menten kinetics)



Concentrations

- A, B and C (substrates), t_0 = 1 uM
- I_1 and I_2 (inhibitors) = 1 uM
- E, E_1 and E_2 (enzymes) at t_0 = 0.1 uM
- P, Q and R (products) at t_0 = 0 uM

Rate Constants

All forward (k_f) and reverse (k_r) rate constants are set initially to 1 except for

- k_{r2} for all reactions = 0 and k_{f4} = 1.

Select Load [model name] below

Load SlowInhibition

Select **Start** to begin the simulation.

Only plots of P, Q and R (products) are initially shown. Select **Plot** to change Y axis min/max, then **Reset** and **Play** | Select **Slider** to change which constants are displayed | Select **About** for software information.

Move the sliders to change the constants and see changes in the displayed graph in real-time.

Time course model made using Virtual Cell (Vcell), **The Center for Cell Analysis & Modeling**, at **UConn Health**. Funded by NIH/NIGMS (R24 GM137787); Web simulation software (miniSidewinder) from Bartholomew Jardine and Herbert M. Sauro, University of Washington. Funded by NIH/NIGMS (RO1-GM123032-04)

6.6.4.3: Coxibs and the Thromboxane/Prostacyclin Imbalance Hypothesis

Previous research indicates that in the cardiovascular system, more significant inhibition of COX-2 vs. COX-1 (as produced by COX-2 selective “coxibs”) can tip the normal balance between the effects produced by prostacyclin & thromboxane, resulting in an increased likelihood for platelet aggregation and vasoconstriction. These effects can help to explain the higher incidence of myocardial infarction and stroke observed when these drugs have been used clinically. The mechanisms involved are illustrated in Figure 6.6.33.

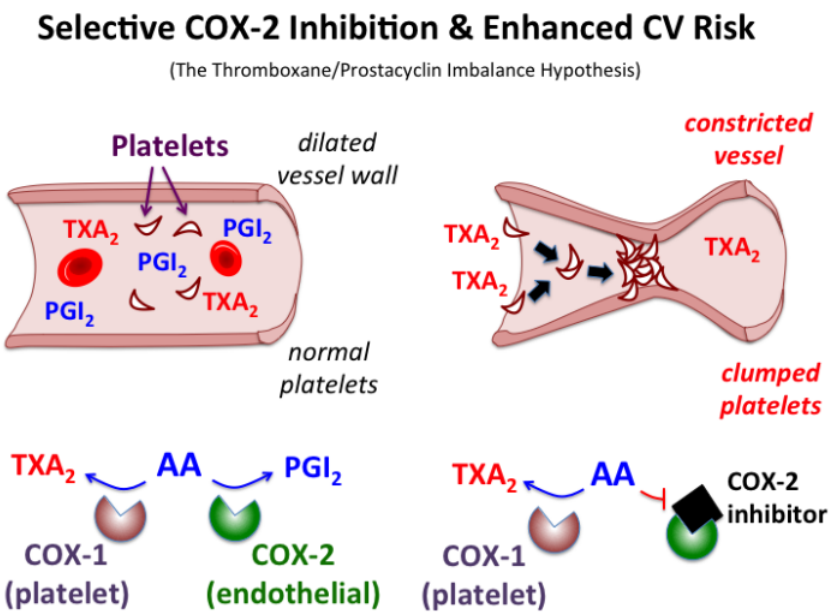


Figure 6.6.33: **COX-2 Inhibitors & Cardiovascular Risk.** *Figure from: Clarkson, C.W. (2018) TUSOM Pharmwiki*

The left graphic illustrates the normal balanced effect between prostacyclin (PGI_2) and Thromboxane (TXA_2). PGI_2 is produced primarily by COX-2 activity in the endothelial cell wall of blood vessels. PGI_2 produces vasodilation and inhibits platelet activation. In contrast, TXA_2 is produced primarily by COX-1 activity inside platelets, producing vasoconstriction and enhanced platelet aggregation. Normal vascular homeostasis is maintained when there is a balanced effect of both PGI2 & TXA2. However, when the balance is tipped in favor of TXA_2 formation after selective inhibition of COX-2 (right graphic), vasoconstriction and platelet clumping are more likely to occur, potentially causing an increased risk for cardiovascular events such as myocardial infarction and stroke. Overall, the COX-1/COX-2 isozyme example sheds light on the complexity of biological systems and the ability for slight adjustments in gene expression to create varied and tissue-specific responses.

6.6.5: Summary

This chapter explores the multiple layers of regulation that contribute to the dynamic and complex proteome, far exceeding the number of protein-coding genes. It highlights the critical roles of post-translational modifications (PTMs), allosteric regulation, zymogen activation, and the expression of isozymes in fine-tuning protein function in response to cellular and environmental cues.

Proteome Complexity and Post-Translational Modifications (PTMs):
The chapter begins by illustrating how approximately 20,000 human genes generate a vast array of proteoforms through alternative splicing and an extensive repertoire of PTMs. PTMs, which include phosphorylation, acetylation, glycosylation, ubiquitination, sumoylation, oxidation, methylation, and more, dramatically diversify protein structure, stability, activity, and interactions. Both enzymatic and nonenzymatic PTMs contribute to the formation of up to a million different protein variants, enabling precise control over cellular functions. Detailed tables and figures outline common PTMs, the amino acid residues they target, and the enzymes involved.

Detection and Analysis of PTMs:
Advances in analytical techniques, particularly mass spectrometry and antibody-based methods, now allow for large-scale and site-specific identification of PTMs. The chapter discusses how these tools, along with methods such as Western blotting and protein sequencing, have become essential for mapping the dynamic landscape of PTMs and understanding their biological roles.

Allosteric Regulation:
Beyond PTMs, the chapter examines allosteric regulation—how binding events at sites distant from an enzyme's active site can induce conformational changes that modulate activity. Examples include small molecule binding that triggers hinge-clamping motions, subtle electrostatic rearrangements, and changes in protein quaternary structure. The concept of conformational selection is also discussed, emphasizing how shifts in the ensemble of protein conformations can regulate activity and signal transduction.

Zymogen Activation:
To prevent unwanted enzymatic activity, many enzymes are synthesized as inactive precursors known as zymogens. The chapter

describes how zymogen activation, through specific proteolytic cleavage or conformational changes, ensures that enzymes such as digestive proteases or caspases become active only under appropriate conditions. This mechanism is crucial for maintaining cellular integrity and preventing self-digestion.

Isozymes and Differential Regulation:
Isozymes, or isoenzymes, are different forms of an enzyme that catalyze the same reaction but possess distinct kinetic properties, regulatory mechanisms, and tissue-specific expressions. These variations, often arising from gene duplication or alternative splicing, allow organisms to finely tune metabolic processes. The chapter concludes with a detailed discussion of cyclooxygenases (COX-1 and COX-2), key enzymes in prostaglandin synthesis that serve as primary targets for nonsteroidal anti-inflammatory drugs (NSAIDs). The distinct roles and regulatory mechanisms of COX isozymes illustrate how subtle differences in enzyme structure and expression can have significant physiological consequences, including effects on cardiovascular health and inflammation.

Overall, this chapter provides a comprehensive overview of the sophisticated strategies employed by cells to regulate protein function post-synthesis. By integrating discussions on PTMs, allosteric modulation, zymogen activation, and isozymes, the chapter underscores the complexity and precision of cellular regulation, setting the stage for further exploration of metabolic control and signal transduction in biochemistry.

6.6.6: Some References

1. Shen, Q., Wang, G., Li, S., Liu, X., Lu, S., Chen, Z., Song, K., Yan, J., Geng, L, Huang, Z., Huang, W., Chen, G., and Zhang, J. (2016) ASDv3.0: Unraveling allosteric regulation with structural mechanisms and biological networks. Nucleic Acids Research 44(D1):D527-D535. Available at: https://academic.oup.com/nar/article/44/D1/D527/2503129
2. Wikipedia contributors. (2020, January 19). Isozyme. In *Wikipedia, The Free Encyclopedia*. Retrieved 03:42, May 19, 2020, from https://en.Wikipedia.org/w/index.php?title=Isozyme&oldid=936548836
3. Wikipedia contributors. (2020, April 30). COX-2 inhibitor. In *Wikipedia, The Free Encyclopedia*. Retrieved 07:00, May 22, 2020, from en.Wikipedia.org/w/index.php?title=COX-2_inhibitor&oldid=954080651
4. Clarkson, C.W. (2018) Major Side Effects of NSAIDs and COX-2 Selective Inhibitors. *TUSOM Pharmwiki*. Available at: http://tmedweb.tulane.edu/pharmwiki/doku.php/nsaid_side_effects?do=
5. Santos, A.L. and Lindner, A.B. (2017) Protein Porsttranslational Modifications: Roles in Aging and Age-Related Disease. *Oxidative Medicine and Cellular Longevity*, Article ID: 5716409. Available at: https://www.hindawi.com/journals/omcl/2017/5716409/#copyright
6. Wikipedia contributors. (2020, May 12). Protein phosphorylation. In *Wikipedia, The Free Encyclopedia*. Retrieved 17:38, May 25, 2020, from en.Wikipedia.org/w/index.php?title=Protein_phosphorylation&oldid=956228409
7. Szylveszter, K.P., Németh, T., and Mócsai, A. (2019) Tyrosine Kinases in Autoimmune and Inflammatory Skin Diseases. Front. Immunol. 10.3389(2019.01862). Available at: https://www.frontiersin.org/articles/10.3389/fimmu.2019.01862/full
8. Wikipedia contributors. (2020, May 7). Histone. In *Wikipedia, The Free Encyclopedia*. Retrieved 21:12, May 25, 2020, from en.Wikipedia.org/w/index.php?title=Histone&oldid=955458038
9. Wikipedia contributors. (2020, May 13). Mucin. In *Wikipedia, The Free Encyclopedia*. Retrieved 01:10, June 7, 2020, from en.Wikipedia.org/w/index.php?title=Mucin&oldid=956387296
10. Wikipedia contributors. (2020, March 15). Allosteric regulation. In *Wikipedia, The Free Encyclopedia*. Retrieved 04:07, June 7, 2020, from en.Wikipedia.org/w/index.php?title=Allosteric_regulation&oldid=945637073
11. Dixit, Ajay. Dawra, Rajinder K. Dudeja, Vikas. Saluja, Ashok K. (2016). Role of trypsinogen activation in genesis of pancreatitis.
Pancreapedia: Exocrine Pancreas Knowledge Base, DOI: 10.3998/panc.2016.25
12. Coll-Martinez, B., and Crosas, B. (2019) How the 26S Proteasome Degrades Ubiquitinated Proteins in teh Cell. *Biomolecules* 9(9):395. Retrieved from: https://www.mdpi.com/2218-273X/9/9/395
13. Liu, W., Tang, X., Qi, X., Ghimire, S., Ma, R., Li, S., Zhang, N., and Si H. (2020) The Ubiquitin Conjugating Enzyme: An Important Ubiquitin Transfer Platform in Ubiquitin-Proteosome System. *Int J Mol Sci* 21(8):2894. Retrieved from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7215765/

6.7: Ribozymes - RNA Enzymes

Learning Goals (ChatGPT o1, 1/30/25)

- **Foundations of Catalysis by RNA (Ribozymes):**
  - Explain the concept that catalysis is not exclusive to proteins and that RNA molecules can fold into defined secondary and tertiary structures to form active catalytic centers.
  - Describe the basic principles of electron pushing in chemical reactions, including the roles of electron donors (sources) and electron acceptors (sinks), and how these principles apply to ribozyme catalysis.
- **Classification and Structural Diversity of Ribozymes:**
  - List and classify the different types of ribozymes, including small self-cleaving ribozymes (hammerhead, viroid, hairpin, and riboswitches) as well as larger self-splicing introns (Group I and Group II) and RNP complexes (e.g., RNase P, spliceosome, ribosome).
  - Compare the sizes, structures, and functional roles of small versus large ribozymes, including their dependence on metal ions and other cofactors for catalytic activity.
- **Mechanisms of Ribozyme Catalysis:**
  - Illustrate the general mechanism of RNA-catalyzed phosphodiester bond cleavage, including the roles of general acid and base catalysis, SN2 transesterification, and transition state stabilization.
  - Analyze specific reaction mechanisms for well-characterized ribozymes (e.g., hammerhead, hairpin) by mapping out electron flow (arrow-pushing) and identifying critical nucleotides (e.g., G8, G12) that facilitate catalysis.
- **Functional Examples and Applications:**
  - Examine the catalytic strategies of ribozymes such as self-cleaving reactions in viroids and hairpin ribozymes, noting how they process their own RNA or cleave external substrates (cis versus trans catalysis).
  - Discuss the roles of larger ribozymes (Group I and II introns, spliceosomal introns, RNase P, and the ribosome) in RNA processing and gene expression, and compare their mechanisms to those of small self-cleaving ribozymes.
- **Evolutionary Implications:**
  - Evaluate the hypothesis that ribozymes may have preceded protein enzymes in early evolution, considering their dual role as both genetic material and catalysts.
  - Discuss the significance of ribozyme catalysis in the RNA world hypothesis and how modern ribozymes (and engineered artificial ribozymes) inform our understanding of early biocatalysis.
- **Advanced Functional Ribozymes:**
  - Explore specialized ribozymes such as the methyltransferase ribozyme (MTR1), detailing how small molecule cofactors (e.g., O6-methylguanine) assist in methyl transfer reactions and expand the catalytic repertoire of RNA.
  - Discuss how ribozymes integrate cofactor binding, structural dynamics, and catalytic mechanisms to perform complex chemical transformations.
- **Analytical and Structural Techniques:**
  - Identify and discuss the experimental methods (such as X-ray crystallography, cryo-EM, and interactive structural modeling via tools like iCn3D) used to elucidate ribozyme structures and catalytic mechanisms.
  - Interpret interactive models and figures to understand the active site organization and the role of key residues in ribozyme function.

By achieving these learning goals, students will be well-equipped to understand the chemical and biological principles underpinning ribozyme catalysis, appreciate the structural and mechanistic diversity of RNA enzymes, and recognize their evolutionary significance and practical applications in modern biochemistry.

6.7.1: Ribozymes

Any molecule that displays any of the catalytic motifs seen in the earlier chapters (general acid/base catalysis, electrostatic catalysis, nucleophilic catalysis, intramolecular catalysis, and transition state stabilization) can be a catalyst. So far, we have only examined protein catalysts. These can fold to form unique 3D structures, with active sites with appropriate functional groups or nonprotein "cofactors" (metal ions, vitamin derivatives) participating in catalysis. There is nothing special about the ability of proteins to do this. RNA can also form secondary and tertiary structures, as seen in Chapter 8. RNA molecules that act as enzymes are called **ribozymes**.

We are presenting the section before Chapter 8 for a few reasons. Most readers have encountered the structures of RNA and DNA before. You most likely know about three different types of RNA: ribosomal RNA (rRNA), transfer RNA (tRNA), and messenger RNA (mRNA). Likewise, you know from introductory biology classes the essential dogma of biology, DNA, the holder of the genetic code) is transcribed into RNA, which is translated into a protein sequence. Finally, most have studied (even at the high school level) that DNA of many species has exons and introns (intervening sequences), the latter of which are spliced out of RNA transcripts to form mature RNA. In this section, we will discuss the catalytic properties of ribozymes, so the introductory background we just mentioned, although important, takes a "second" seat to the chemistry of catalysis, which is the main topic of Chapter 6.

There are 12 classes of ribozymes

- small self-cleaving RNAs (9 classes)
- Group I introns
- Group II introns
- Ribonuclease P

The large ribonucleoprotein nanoparticles, the spliceosome and ribosome, are also functionally ribozymes.

The term ribozyme is used for RNA that can act as an enzyme. Ribozymes are mainly found in selected viruses, bacteria, plant organelles, and lower eukaryotes. Ribozymes were first discovered in 1982 when Tom Cech’s laboratory observed Group I introns acting as enzymes. This was shortly followed by the discovery of another ribozyme, Ribonuclease P, by Sid Altman’s laboratory. Both Cech and Altman received the Nobel Prize in chemistry in 1989 for their work on ribozymes.

Ribozymes can be categorized based on size. Small ones, which usually don't require metal ions for activity, vary from 30-150 nucleotides, while large ones can be a few thousand nucleotides in length. This translates into approximate molecule weights (using this formula for single-stranded RNA: (# nucleotides x 320.5) + 159.0) into 9800 for a 30-mer and 640,000 for a 2000 mer, typical of small and very large proteins/protein complexes, respectively. into two groups depending upon their size – small and large. Large ribozymes, which require metal ions for activity, can vary in size from a few hundred to several thousand nucleotides. Examples of small ones include hammerhead, viroid, hairpin, and riboswitch ribozymes. Large ones include type I and II self-spicing introns, bacterial ribonuclease P, and the RNA in spliceosomes and ribosomes. Many are not true enzymes since they catalyze their cleavage, although some can cleave presented RNA substrates. Large ones act as true catalysts.

Since RNAs can carry genetic information and act as enzymes, they probably evolved before proteins, which require nucleic acids for their synthesis. In addition, DNA required a special enzyme (ribonucleotide reductase) encoded by DNA to reduce the 2'OH to a 2'H. Artificial ribozymes have been made to catalyze many reactions that require protein enzymes. We will explore some ribozymes in several classes.

6.7.2: Small self-cleaving RNAs

We'll consider four small self-cleaving RNAs- hammerhead, viroid, hairpin ribozymes, and the glucosamine-6-phosphate riboswitch (glmS). All catalyze the cleavage of an internal phosphodiester bond (cis catalysis) or in a presented substrate (trans catalysis) by a transesterification reaction. Internal cis catalysis reactions cleave the ribozyme into two fragments, activating their catalytic activity. They don't act as true catalysts since they engage in only one cleavage cycle. Trans catalysis, in which a substrate RNA binds to and is cleaved by the ribosome, would be considered true catalysis.

The cleavage reaction in an internal cis cleavage is a S_N2 trans-esterification reaction, as shown in Figure 6.7.1.

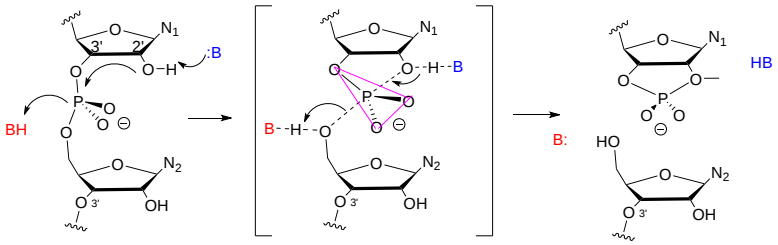


Figure 6.7.1: SN2 trans-esterification cleavage of an internal phosphodiester bond in small RNA ribozymes

In the reaction, an adjacent general base (:B) abstracts a proton of the C'2 OH of nucleotide N_1. The resulting 2' O^- acts as a nucleophile in a S_N2 reaction and attacks the δ^+ phosphorous in the phosphodiesterase bond, forming a pentavalent, trigonal pyramidal sp^3d hybridized intermediate/transition state, which collapses, breaking the phosphodiesterase bond between nucleotide N_1 and N_2. This is an inline mechanism: the incoming nucleophilic O in C2' and the exiting one on the O of 5' CH_2OH of nucleotide 2 are axial, separated by 180^0. A general acid, BH, facilitates the departure of the exiting nucleophile by its protonation. Bound metal ions may facilitate the reaction and can be considered cofactors but might be more involved in maintaining a catalytically active structure. Self-cleaving small RNAs are also found in humans and may be part of long noncoding RNAs.

a. Hammerhead RNA

The hammerhead ribozyme is a small RNA ribozyme with a conserved core with three helical stems. It has a structure similar to the head of a hammerhead shark. One predicted secondary structure of a hammerhead ribozyme is shown below in Figure 6.7.2

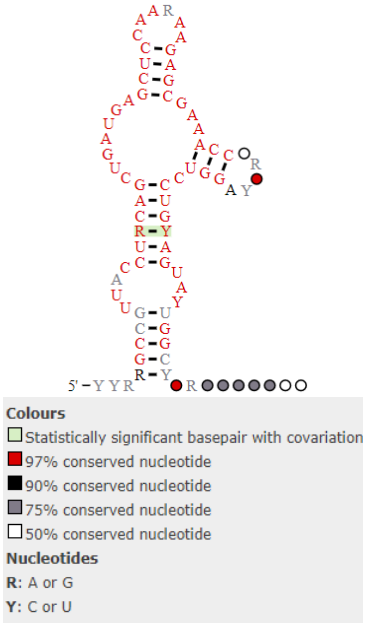


Figure 6.7.2: Predicted secondary structure and sequence conservation of the HH9 ribozyme found conserved from lizard to human genomes.

A possible trigonal pyramidal intermediate/transition state in the Hammerhead ribosome is shown in Figure 6.7.3.

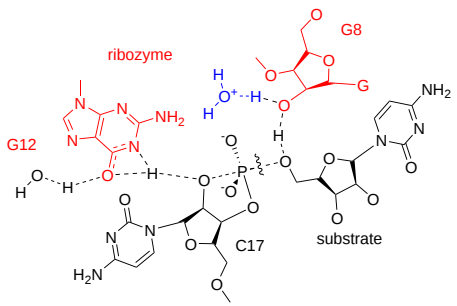
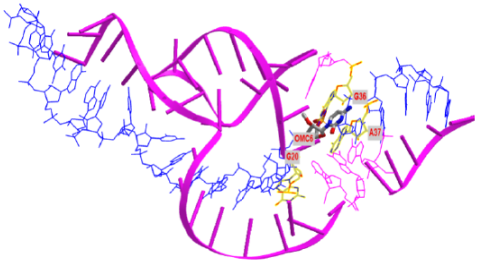


Figure 6.7.3: Role of G8 and G12 in active site of full length Schistosoma mansoni hammerhead ribozyme. After Martick and Scott, Cell. 126, 2006 DOI:https://doi.org/10.1016/j.cell.2006.06.036

A deprotonated G-12 in the ribozyme probably acts as a general base that activates the 2'-OH to form the incoming nucleophile that attacks the trans-substrate RNA. The 2′-OH of G-8 in the ribozyme appears to hydrogen bond to the 5′-O of the departing nucleophile in the substrate where bond scission occurs. The ribozyme increases the rate by 1000-fold.

Figure 6.7.4 shows an interactive iCn3D model of the full-length Schistosoma manson catalytically active hammerhead ribozyme (3dz5). This is an example of a ribozyme that acts in trans as the cleaved phosphodiester bond is a bound RNA single-stranded substrate.



NCBI iCn3D Figure 6.7.4: Full-length Schistosoma manson catalytically active hammerhead ribozyme (3dz5). (Copyright; author via source).

Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/icn3d/share.html?qNBw2MY1RZ7xKogH7

The numbering system is a bit different in the iCn3D model above. PDB requires sequential numbering, whereas ribozyme sequences are numbered in discontinuous ways. Core residues that are conserved are given common numbers. However, the different projecting double-stranded RNA regions vary among ribozymes and are numbered differently. The G8 and G12 in Figure 6.7.3 are numbered G20 and G36, respectively.

b. ***Viroids***

Viroids are small, single-stranded circular RNAs that infect plant cells. They are not packaged with viral capsid protein. Some enter cells along with viruses and are called **virusoids** or **viroidlike satellite RNAs**. Intrastrand pairing occurs, and they are synthesized as tandem repeats containing multiple adjacent copies of the viroid. These repeats are cut and ligated to form the individual mature viroid by internal ribozyme sequences. One example is the hepatitis delta virus (HDV), a satellite of the hepatitis B virus. A possible mechanism of catalysis of the hepatitis delta virus ribozyme involving general acid/base catalysis is shown in Figure 6.7.5.

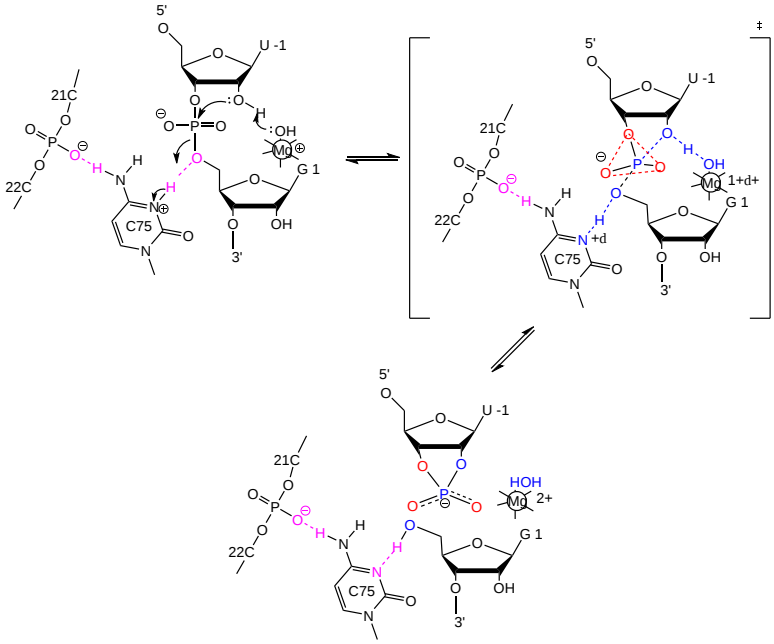


Figure 6.7.5: General acid/base catalysis by Hepatitis delta virus (HDV) ribozyme

c. ***Hairpin ribozyme***

The satellite RNA of plant viruses encodes hairpin ribozymes. They are about 50 nucleotides long and can cleave themselves internally or, in a truncated form, can cleave other RNA strands in a transesterification reaction. The structure consists of two domains, stem A required for binding (self or other RNA molecules) and stem B, required for catalysis. Self-cleavage in the hairpin ribozyme occurs in stem A between an A and G bases (which are splayed apart) when the 2' OH on the A attacks the phosphorous in the phosphodiester bond connecting A and G to form a pentavalent intermediate.

Rupert et al. solved the crystal structure of a hairpin ribozyme with a non-cleavable substrate analog containing a 2'-O-CH_3 group on the ribose. This acts as a nucleophile in the transesterification cleavage of RNA.

A38 in Stem B appears to be able to interact with the products (the cleaved A now in the form of a cyclic phosphodiester with itself) and the departing G. With a transition state pentavalent analog of the sessile A-G bond in which the phosphodiester linking A and G in the substrate is replaced with a pentavalent vanadate bridge between A and G. This is Illustrated in Figure 6.7.6.

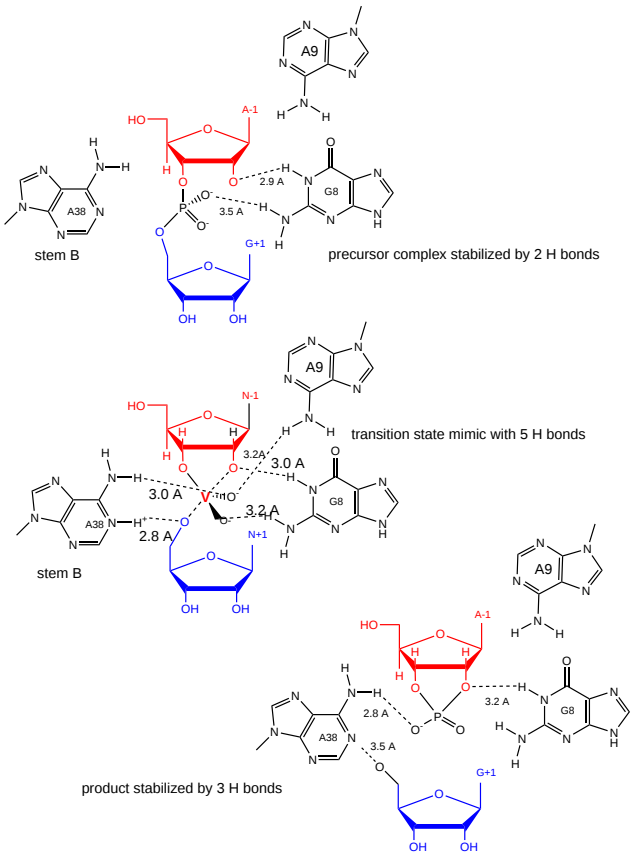


Figure 6.7.6: Transition state binding stabilization of hairpin ribozyme (adapted from Rupert et al, Science, 298 pg 1423 2002)

However, A 38 does not appear to react with the sessile A -G groups in the normal substrate, indicating that the main mechanism used by this ribozyme is transition state binding. Since RNA molecules have fewer groups available for acid/base and electrostatic catalysis (compared to protein enzymes), ribozymes, presumably the earliest type of biological catalyst, probably use transition state binding as their predominant mode of catalytic activity.

Figure 6.7.7 shows an interactive iCn3D model of the hairpin ribozyme in the catalytically-active conformation (1M5K).

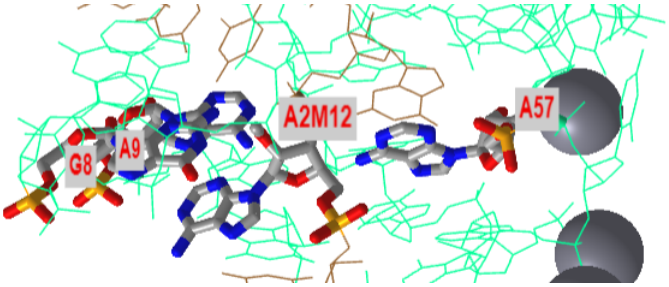


Figure 6.7.7: Hairpin ribozyme (human) in the catalytically-active conformation (1M5K). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...goBWNHUbLU7my8

The hairpin ribozyme is shown in cyan sticks, and the inhibitor substrate is shown in brown sticks. The inhibitor contains a 2'-O-methyl adenosine (A2M12), so it can not be cleaved and instead acts as an inhibitor. The A38 shown in the catalytic mechanism is labeled A57 in the iCn3D.

d. ***Glucosamine-6-phosphate riboswitch (glmS)***:

A novel use of ribozymes was recently reported by Winkler et al. They discovered that the 5' end of the mRNA of the gene glmS (from Gram-positive bacteria) is a ribozyme. The GlmS gene encodes glucosamine-6-phosphate synthetase (GlmS), which catalyzes the reaction of fructose-6-phosphate and glutamine to glucosamine-6-phosphate (GlcN6P) and glutamate. This is the first committed step in bacterial cell wall synthesis. Glucosamine-6-phosphate binds to the ribozyme (3' end of the mRNA) and acts as a cofactor, leading to self-cleavage of the ribozyme. What an amazing mechanism for pathway inhibition. At high GlcN6P concentrations, it binds to the ribozyme, inhibiting its own synthesis. G40 in the active site appears to act as a general base. Figure 6.7.8 shows an interactive iCn3D model of GlmS Ribozyme Bound to Its Catalytic Cofactor, glucosamine 6 phosphate (GlcN6P) (**2NZ4).**

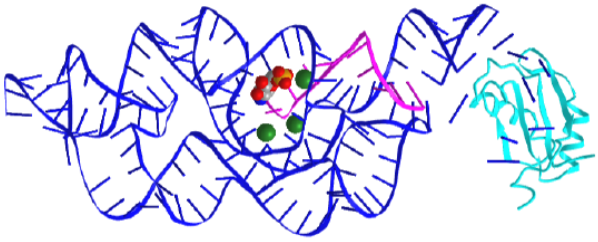


Figure 6.7.8: GlmS ribozyme bound to its catalytic cofactor glucosamine 6 phosphate (GlcN6P) (2NZ4) (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...J19zdfwEujH4c8

Riboswitches are discussed in greater detail in Chapter 28.1: Regulation of Gene Expression in Bacteria.

6.7.3: Group I and Group II Introns

Introns present in RNA molecules must be removed to form mature RNA. About 80% of introns in humans are less than 200 nucleotides long, but some can be 10,000 or longer in length. Before we discuss introns, we'll provide a quick background on RNA splicing. There are two major types of self-splicing introns, Groups I and II. Other introns are removed by a ribonucleoprotein called the spliceosome. Some call these Group III introns. A simple two-step mechanism for the self-splicing Group I and II introns is shown in Figure 6.7.9.

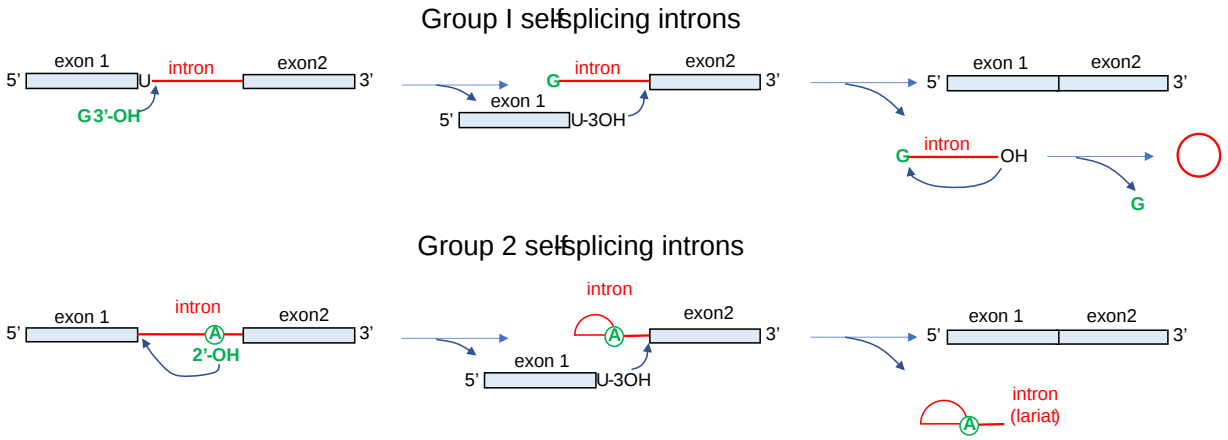


Figure 6.7.9: Self-splicing of Group I and II introns. (after N.K. Tanner /FEMS Microbiology Reviews 23 (1999)

Both required first a scission of the RNA strand followed by ligation of the two exons to form the mature RNA. Note that Group I introns require an external guanosine nucleophile, and the removed intron forms a circular RNA when removed. In contrast, in Group II introns, an internal A residue acts as the first nucleophile in the scission reaction, and the intron, on removal, forms a branched lariat structure.

A simple two-step mechanism for Group II introns which are spliced out from pre-mRNA in the nucleus by a ribonucleoprotein complex called the **splicesome** is shown in Figure 6.7.10.

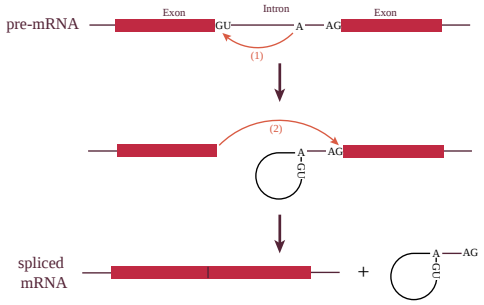


Figure 6.7.10: Two steps of canonical RNA processing, from pre-mRNA to spliced RNA and the branched lariat intron. https://commons.wikimedia.org/wiki/F...g_reaction.svg. Creative Commons Attribution-Share Alike 3.0 Unported

Note that the mechanism is extremely similar to the auto-removal of Group II introns, suggesting an evolutionary relationship between the two.

Figure 6.7.11 below shows a more detailed view of the catalytic cycle of the spliceosome. Five small ribonucleoproteins (U1, U2, U4/U6, and U5 snRNPs) assemble on the nuclear pre-mRNA and facilitate the intron removal, but the main mechanism involves ribozyme activity.

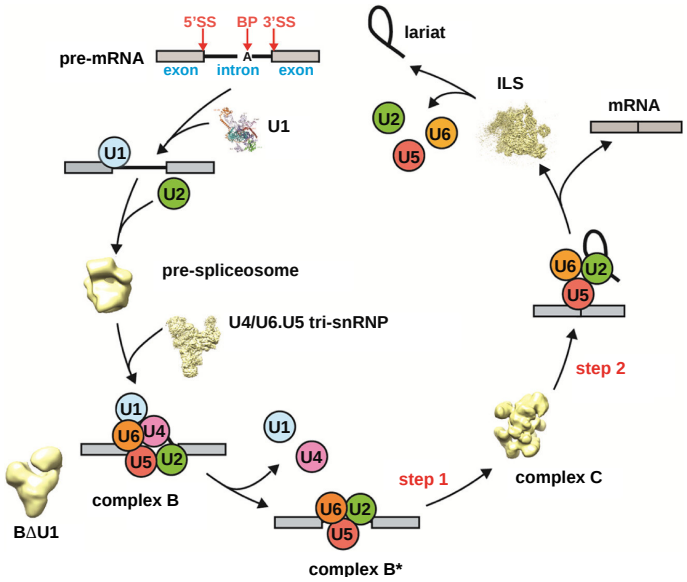


Figure 6.7.11: Step-wise spliceosome assembly from its U-snRNP components. Nguyen et al. Current Opinion in Structural Biology 2016, 36:48–57. http://dx.doi.org/10.1016/j.sbi.2015.12.005. CC BY license (http://creativecommons.org/licenses/by/4.0/).

6.7.3.1: Group I introns

These are found in bacteria, lower eukaryotes (including mitochondrial and chloroplast RNA), and higher plants and are in ribosomal RNA (rRNA), mRNA, and tRNA. They are also found in Gram-positive bacteriophages (viruses that attack bacteria). As shown in Figure 6.7.9, they require guanosine as a cofactor and have a single active site for both scission and ligation to produce the mature mRNA, tRNA, or rRNA. Mg^2 is not needed for catalysis per se but to maintain the correct tertiary structure of the ribozyme with the correct secondary structure. In Group I introns, a guanosine cofactor initiates the **splicing reaction**. They have one active site that catalyzes the initial cleavage of the phosphodiester bond and the final religation after cleavage.

The Group I catalytic core from Tetrahymena thermophila has two domains. A cleft is formed between them when they pack which can bind the short helix containing the 5' splice site. and the guanosine cofactor. This "active" site is pre-formed without substrates, similar to the active sites of protein enzymes. Figure 6.7.12 shows the secondary structure and the reaction of the group I intron ribozyme from *Tetrahymena*.

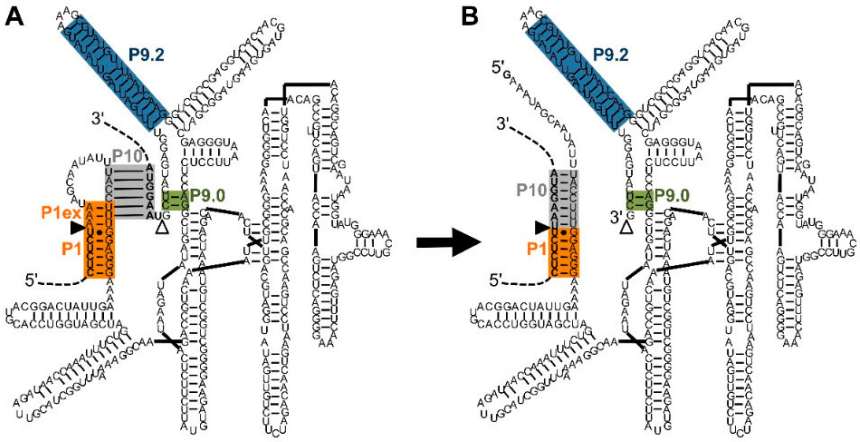


Figure 6.7.12: Secondary structure and reaction of the natural, *cis*-splicing group I intron ribozyme from *Tetrahymena*. Müller. *Molecules* **2017**, *22*(1), 75; https://doi.org/10.3390/molecules22010075. Creative Commons Attribution (CC-BY) license (http://creativecommons.org/licenses/by/4.0/).

The intron runs between the two triangles, which show the 5' (filled triangle) and 3' (open triangle) splice sites. The orange P1 helices contain the 5′-splice site (G:U). The P10 helix (grey), P9.0 helix (green), and the P9.2 helix (blue) are structured to present the appropriate 3′-splice site. During splicing (from (**A**) to (**B**)), the P1 helix extension (orange) is opened to expose the 3′-hydroxyl group of the terminal uridine at the 5′-splice site. The P10 helix then facilitates a conformational change, in which the 3′-exon

(upper dashed line) is positioned adjacent to the 5′-exon (lower dashed line), allowing the nucleophilic attack of the 3′-uridine the 3′-splice site, joining 5′-exon and 3′-exon.

Figure 6.7.13 shows an interactive iCn3D model from cryoEM of the full-length holo L-16 ScaI Tetrahymena ribozyme (7EZ2).

Figure 6.7.13: Holo L-16 ScaI Tetrahymena ribozyme (7EZ2) (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...PhK3BSbFLDnDt8

Three sets of coplanar bases are found in the active site. These include C262, A263, and G312 (top layer, brown), G264, C311, and the ωG - labeled G3 in the iCn3D model- (cyan layer) and A261, A265, and U310 (bottom layer red). The ωG - labeled G3 - is the nucleophile.

The structure is nearly identical to the apo-form of the ribozyme, with just an internal guide RNA sequence undergoing a large change and the guanosine binding site undergoing a small shift on binding RNA substrates.

6.7.3.2: Group 2 Introns

Group II introns are found in mRNA of bacteria and some Archaea, in rRNA, tRNA, and mRNA of chloroplasts and mitochondria, and fungi, plants, and protists. No Class 2 introns appear to be found in eukaryotic genomes. Some of these introns are in gene-encoding proteins, but most are in bacterial noncoding sequences. In Group II introns, an adenosine cofactor initiates the splicing reaction, as shown in Figure 6.7.9.

What's especially interesting about group II introns is that they can reinsert into DNA. Hence, they can be considered mobile genetic elements. A **maturase/reverse transcriptase** enzyme (562 amino acids) is associated with the intron, which helps stabilize the active site of the intron for the **reversible** session and ligation of the intron. Reintegration of the excised branched lariat intron into DNA is called retrotransposition (copy/paste). Mitochondrial and chloroplast Group 2 introns have lost mobility and act as classic introns.

The structures of a group II intron from Thermosynechococcus vestitus (a cyanobacteria) before and after integration have been determined. A branch-site domain VI helix swings 90°, enabling DNA integration. The maturase/revere transcriptase protein assists excision of the intron through the interaction of domain VI of that intron that positions the key adenosine for branched lariat formation during forward splicing, as shown in Figure 6.7.9. The changes in the structure of the group II intron retroelement before (6ME0) and after DNA integration (6MEC) are shown in Figure 6.7.14.

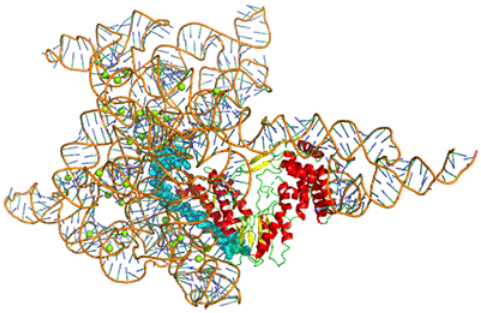
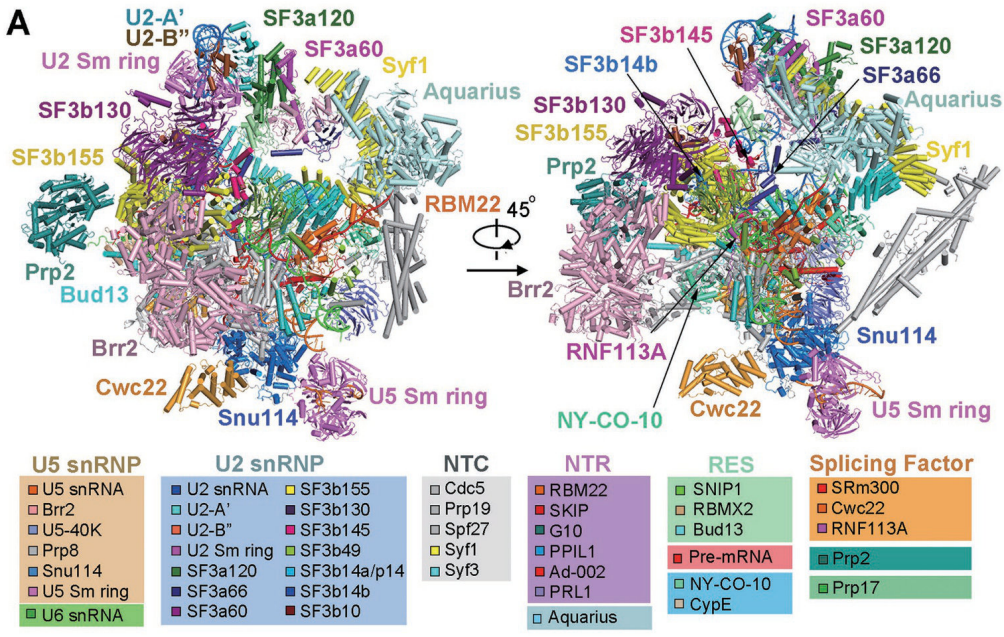


Figure 6.7.14: Changes in the structure of the group II intron retroelement before (6ME0) and after DNA integration (6MEC)

The target DNA before integration is shown in cyan spacefill and orange spacefill after integration. The protein is shown as a colored cartoon. The Group II intron is 867 nucleotides, and the sense target DNA is 47 nucleotides in length.

6.7.4: Spliceosomal Introns

A comparison of Figure 6.7.9 through Figure 6.7.11 shows the similarities between spliceosomal introns and group II self-splicing introns. Spliceosomal and group II self-splicing introns are structurally and mechanistically homologous, right down to the stereochemistry of the splicing reaction. In eukaryotes, introns in pre-mRNA are removed by splicing and subsequent exon ligation, releasing the intron as a branched lariat molecule. This reaction is performed by the spliceosome, a large nuclear ribonucleoprotein (RNP) complex. Spliceosomes remove introns and splice the exons of most **nuclear genes**. They are composed of 5 kinds of small nuclear RNA (**snRNA**) molecules and over 100 different protein molecules. The RNA — not the protein — catalyzes the splicing reactions. The molecular details of the reactions are similar to those of Group II introns, which has led to speculation that this splicing machinery evolved from them. Figure 6.7.15 shows two views of the cryo-EM structure of the human-activated spliceosome (the Bact complex)



C

Figure 6.7.15: Cryo-EM structure of the human-activated spliceosome (the Bact complex). Zhang et al. https://www.nature.com/articles/cr201814.pdf. Creative Commons Attribution 4.0 Unported License. http://creativecommons.org/licenses/by/4.0/

There are 52 proteins, three small nuclear RNAs (snRNA), and one pre-mRNA. The total molecular mass is 1.8M. U2, U5, and U6 snRNAs are colored marine, orange, and green, respectively. Pre-mRNA is colored red. Figure 6.7.16 shows just the structural changes in RNA and protein components that occur between the early Bact complex (left panel) and the mature Bact complex (right panel).

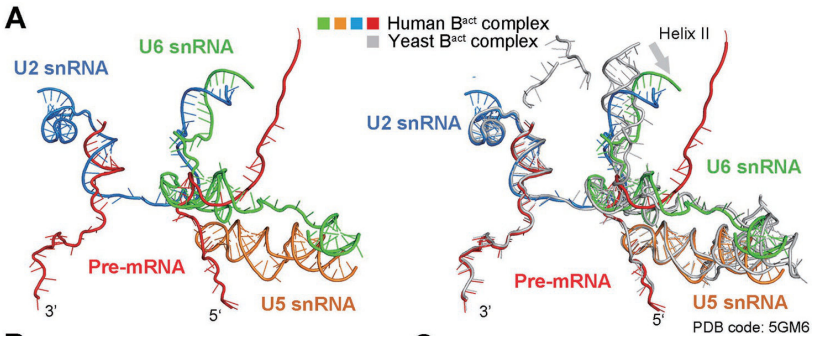


Figure 6.7.16: Structural changes in RNA and protein components that occur between the early Bact complex (left panel) and the mature Bact complex of the human spliceosome.

6.7.5: Ribonuclease P

This enzyme is a ribonucleoprotein that cleaves RNA through the catalytic action of one essential RNA subunit that displays ribozyme activity. It's found in most organisms. As with many ribozymes, the activity is increased 2-3 fold with bound proteins that stabilize the folded ribozyme and help bind the preferred substrate, pre-tRNA. Figure 6.7.17 shows bacterial RNase P ribozyme in complex with tRNA (3q1r).

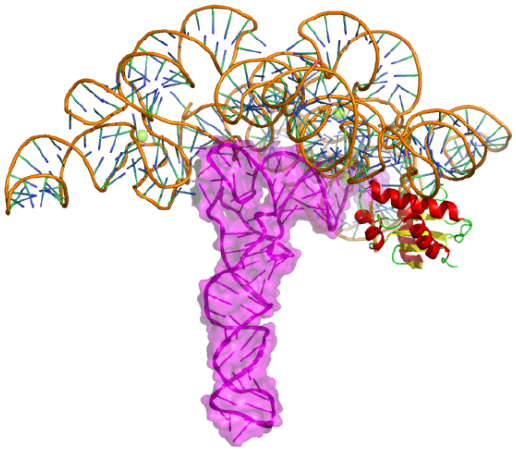


Figure 6.7.17: Bacterial RNase P ribozyme in complex with tRNA (3q1r).

The tRNA is shown as a magenta cartoon/surface, and the protein is in cartoon form. The enzyme cleaves the 5' head end of the precursors of transfer RNA (tRNA) molecules. The enzyme is a heterodimer in bacteria with one RNA and protein subunit.

Figure 6.7.18 shows a possible transition state with key residues involved in binding two Mg^{2+} ions in the active site. These ions are essential for catalysis as they stabilize the pentavalent intermediate/transition state.

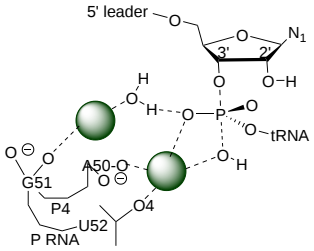


Figure 6.7.18: Transition state of RNase P with key RNA active site

Figure 6.7.19 shows an interactive iCn3D model of the active site with key residues labeled of bacterial RNase P ribozyme in complex with tRNA (3q1r).

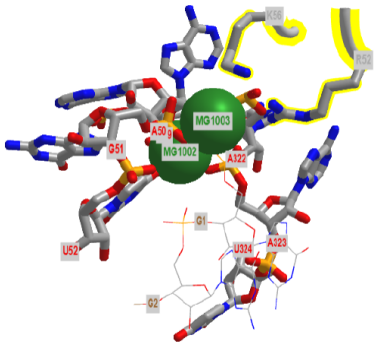


Figure 6.7.19: Bacterial RNase P holoenzyme in complex with tRNA (3q1r) (Copyright; author via source). Click the image for a popup or use this external link:https://structure.ncbi.nlm.nih.gov/i...6SjVMa8g6K47x5

6.7.6: RNA Polymerase Ribozyme

If primordial RNA acted as a metabolic enzyme catalyst and the holder of the genetic information, it would need RNA polymerase activity. Artificial ribozymes with class I RNA ligase activity have been made. Figure 6.7.20 shows an interactive iCn3D model of the active site region of the Class I ligase ribozyme-substrate preligation complex, C47U mutant, Mg^{2+} bound (3R1L).

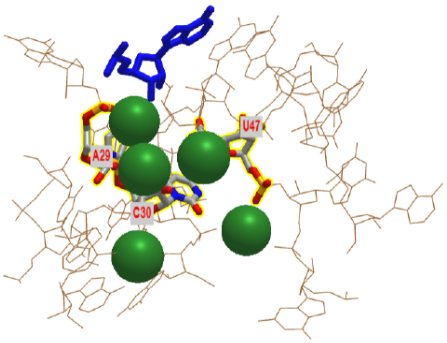


Figure 6.7.20: Active site of the Class I ligase ribozyme-substrate preligation complex, C47U mutant, Mg^{2+} bound (3R1L). (Copyright; author via source). Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...EEQV3Mzc8Jzdv6

The blue stick represents just the 3' terminal adenosine end of the target substrate (5'UCCAGUA3') to which a new nucleoside would be added. The brown represents the active site region of the ribozyme. Three catalytic residues A29, C30, and C47 have been identified in the actual ribozyme. In iCn3D model is of a mutant, C47U, which has no catalytic activity. The green sphere represents Mg^{2+} ions.

Figure 6.7.21 shows the active site residues and how they might facilitate stabilization of the pentavalent intermediate/transition state and the similarity of the active site to a protein RNA polymerase.

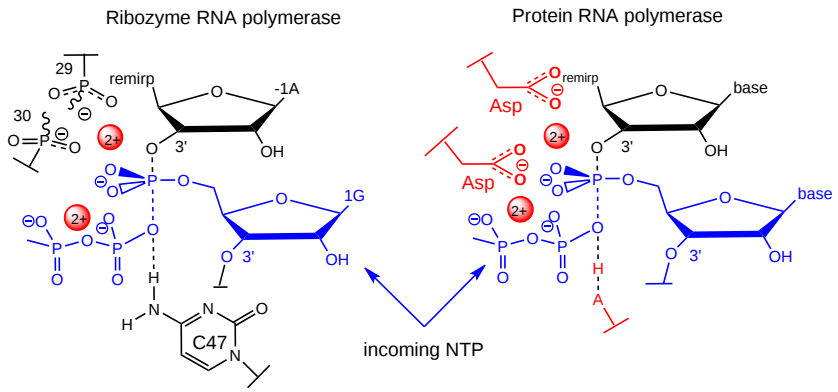


Figure 6.7.21: Comparison of Transition State Models of Ribozyme RNA ligase and a protein RNA Polymerases (after Schechner et al)

A divalent Mg^{2+} in the active site of the ribozyme enhances the nucleophilicity of the 3-OH on the primer, which attached the terminal phosphate of the G(1)TP substrate to form a pentavalent intermediate. The Mg cation is stabilized by oxygens on P 29 and 30 of the ribozyme. The Mg ion also stabilizes the developing charge in the transition state and in the charge in the intermediate. Stabilization of analogous divalent cations in the protein polymerase occurs through Asp side changes in the protein.

6.7.7: Ribosome

Protein synthesis from mRNA templates occurs on a ribosome, a nanomachine composed of proteins and ribosomal RNAs (rRNA). The ribosome is composed of two very large structural units. The smaller unit (termed 30S and 40S in bacteria and eukaryotes, respectively) coordinates the correct base pairing of the triplet codon on the mRNA with another small adapter RNA, transfer or tRNA, that brings a covalently connected amino acid to the site. Peptide bond formation occurs when another tRNA-amino acid molecule binds to an adjacent codon on mRNA. The tRNA has a cloverleaf tertiary structure with some intrastrand H-bonded secondary structure. The last three nucleotides at the 3' end of the tRNA are CpCpA. The amino acid is esterified to the terminal 3'OH of the terminal A by a protein enzyme, aminoacyl-tRNA synthetase.

Covalent amide bond formation between the second amino acid and the first, forming a dipeptide, occurs at the peptidyl transferase center on the larger ribosomal subunit (50S and 60S in bacteria and eukaryotes, respectively). The ribosome ratchets down the mRNA, so the dipeptide-tRNA is now at the **P** or **Peptide site**. It awaits a new tRNA-amino acid at the **A** or **Amino site**. The figure below shows a schematic of the ribosome with bound mRNA on the 30S subunit and tRNAs covalently attached to amino acid (or the growing peptide) at the A and P sites, respectively. Figure 6.7.22 shows a cartoon model of the prokaryotic ribosome with bound mRNA, tRNAs, and the P and A sites.

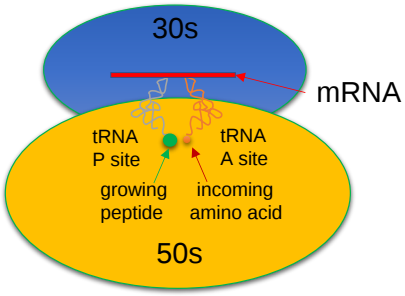


Figure 6.7.22: Cartoon model of the prokaryotic ribosome with bound mRNA, tRNAs, and the P and A sites.

We present another more detailed model of the ribosome complex illustrating protein synthesis in Figure 6.7.23

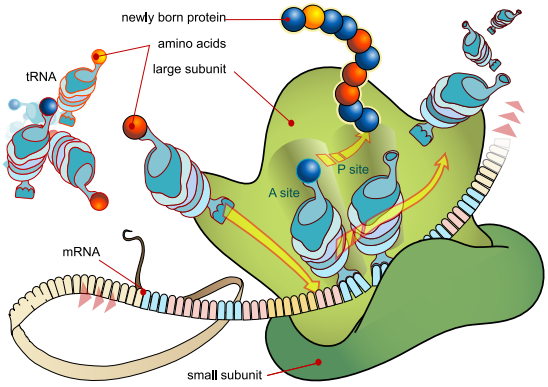


Figure 6.7.23: Ribosome mRNA translation. https://en.Wikipedia.org/wiki/File:R...slation_en.svg. public domain by its author, LadyofHats

A likely mechanism (derived from crystal structures with bound substrates and transition state analogs) for forming the amide bond between a growing peptide on the P-site tRNA and the amino acid on the A-site tRNA is shown below. Catalysis does not involve any ribosomal proteins (not shown) since none is close enough to the peptidyl transferase center to provide amino acids that could participate in general acid/base catalysis, for example. Hence, the rRNA must act as the enzyme (i.e., it is a ribozyme). Initially, it was thought that a proximal adenosine with a perturbed pKa could, at physiological pH, be protonated/deprotonated and hence act as a general acid/base in the reaction. However, none was found. The most likely mechanism to stabilize the oxyanion transition state at the electrophilic carbon attack site is precisely located water, which is positioned at the oxyanion hole by H-bonds to uracil

2584 on the rRNA. The cleavage mechanism involves the concerted proton shuffle shown below. In this mechanism, the substrate (Peptide-tRNA) assists its own cleavage in that the 2'OH is in position to initiate the protein shuttle mechanism. (A similar mechanism might occur to facilitate hydrolysis of the fully elongated protein from the P-site tRNA.) Of course, all of this requires perfect positioning of the substrates. Isn't that what enzymes do best? The main mechanisms for catalysis of peptide bond formation by the ribosome (as a ribozyme) are intramolecular catalysis and transition state stabilization by the appropriately positioned water molecule. These processes are illustrated in Figure 6.7.24

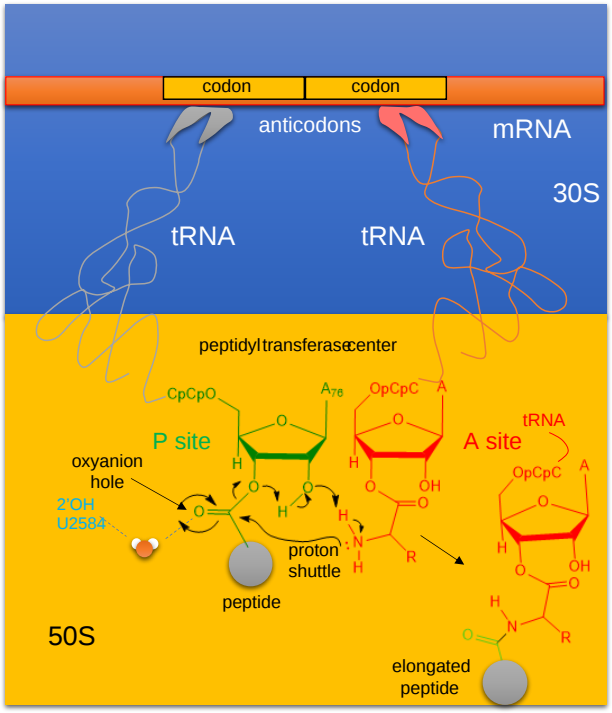


Figure 6.7.24: Mechanism Peptide Bond Formation by the Ribosome

The crystal structure of the eukaryotic ribosome has recently been published (Ben-Shem et al). It is significantly larger (40%) with a mass of around $3x10^6$ Daltons. The 40S subunit has one rRNA chain (18) and 33 associated proteins, while the larger 60S subunit has three rRNA chains (25S, 5.8S and 5S) and 46 associated proteins. The larger size of the eukaryotic ribosome facilitates more interactions with cellular proteins and greater regulation of cellular events. Figure 6.7.25 shows the two copies of the 80S yeast ribosome (4v88) presented to humble readers and authors alike.Figure 6.7.25: The 80S yeast ribosome (4v88). Each subunit is given a different color.

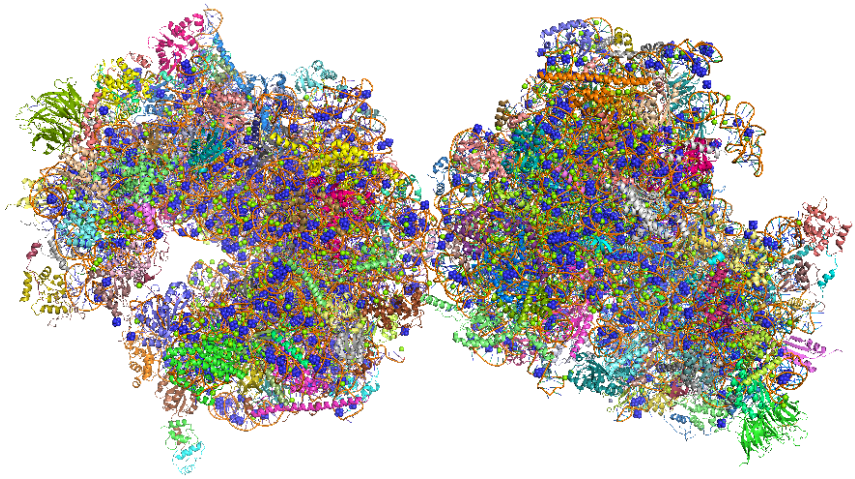


Figure 6.7.25: Structure of the yeast 80S ribosome.

6.7.8: Ribozyme methyltransferase

The ribozymes described above and generally found in nature catalyze phosphoryl transfer reactions and, for the ribosome, peptide bond formation. In vitro evolution can be used to drive new enzymatic functionalities, which would have been required in a RNA-only world that preceded the use of proteins as catalysts. RNA ribozymes have only four bases that can be employed in binding and catalytic steps, compared to 20 amino acids that can serve the same function in proteins. However, as in the case of protein, small molecule cofactors that bind to a potential ribozyme might facilitate greater catalytic efficiency and an expanded repertoire of reaction types. Indeed, we have seen above how small molecules can bind to riboswitches. Figure 6.7.26 shows the reaction and structure of a methyltransferase one ribozyme (MTR1) that acts as a methyltransferase. The small ligand, O^6-methylguanine, binds to the ribozymes and acts as a cofactor in the methylation of adenine 63 in the RNA.

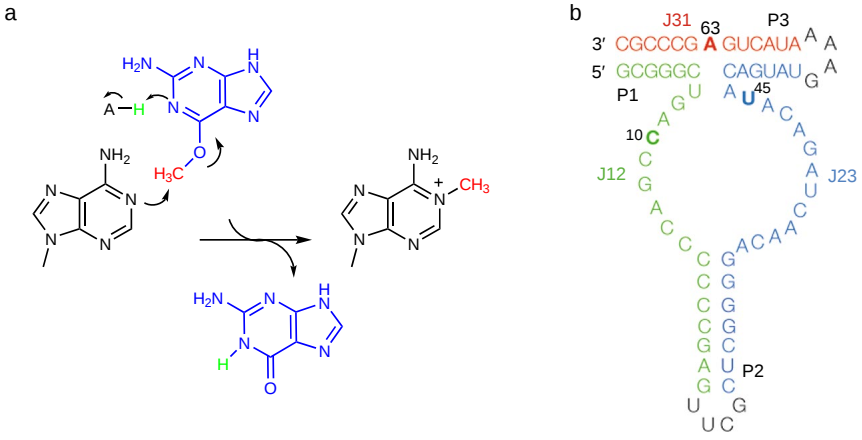


Figure 6.7.26: Deng, J., Wilson, T.J., Wang, J. *et al.* Structure and mechanism of a methyltransferase ribozyme. *Nat Chem Biol* **18**, 556–564 (2022). https://doi.org/10.1038/s41589-022-00982-z. Creative Commons Attribution 4.0 International License. Creative Commons Attribution 4.0 International License, http://creativecommons.org/licenses/by/4.0/

Panel a shows the chemical reaction in which the methyl group of the small ligand O^6-methylguanine is transferred to N1 of Adenine 63 in the RNA.

Panel b shows the sequence of the MTR1 ribozyme as crystallized for the experiments. The RNA is a three-way junction composed of three arms P1, P2 and P3. A GNRA tetraloop has been added to the end of the P3 helix so that the entire ribozyme comprises a single RNA strand. Subsections of the strands are named J12 (colored green), J23 (colored blue), and J31 (colored red).

X-ray crystal structures of the ribozyme in the presence of the cofactor O^6-methylguanine were determined. The final structure contained guanine and an A63 methylated adenosine (1MA), implying the methyl group of the O^6-methylguanine had transferred to A63, leaving guanine bound in the active site. The structure of the guanine bound to the ribozyme is shown in Figure 6.7.27.

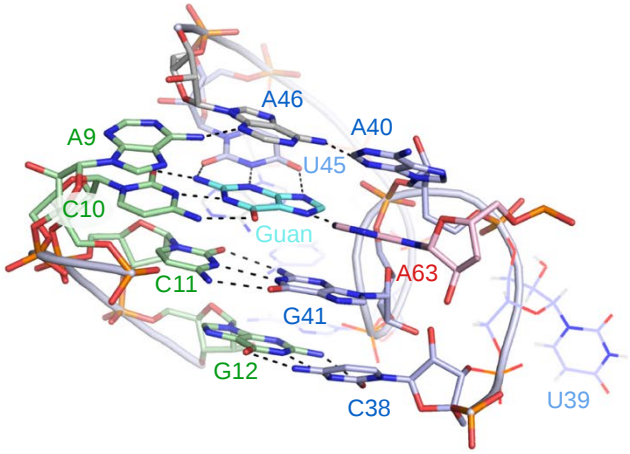


Figure 6.7.27: The four planes of nucleobase interactions in the core of the ribozyme. The four planes are composed of (from bottom to top) G12•C38, C11•G41, exogenous guanine hydrogen bonded to C10, U45, and A63, and the triple interaction A9•A46•A40. Deng, J et al., ibid.

Figure 6.7.28 shows a hypothetical reaction mechanism.

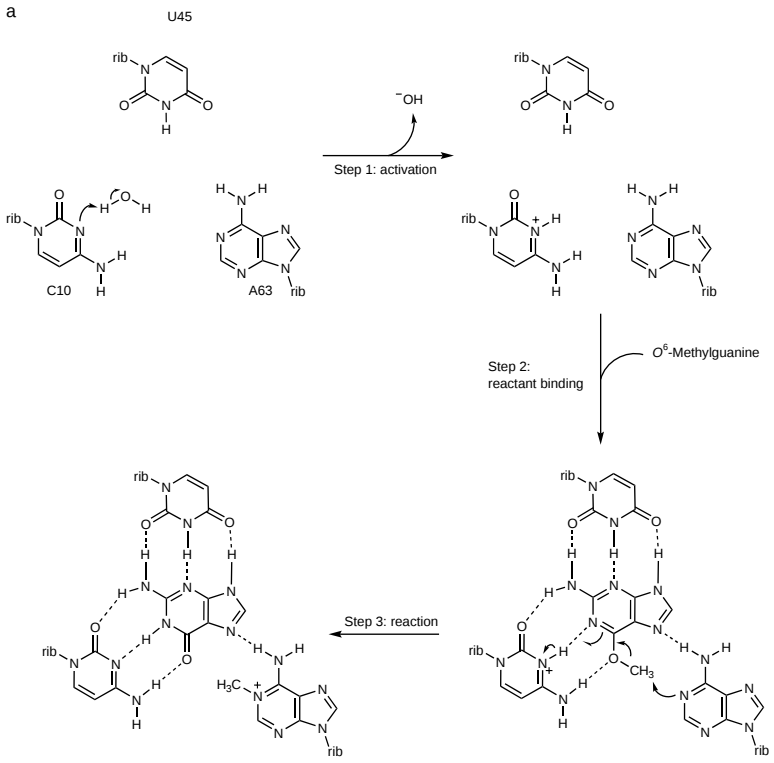


Figure 6.7.28: A proposed catalytic mechanism for the ribozyme methyltransferase. Deng, J et al., ibid.

In step 1, the nucleobase of C10 becomes protonated, and in step 2, the O^6-methylguanine becomes bound. However, these two steps are likely coordinated, as the binding will raise the pK_a of the cytosine.

The methyl transfer reaction occurs in step 3 by the nucleophilic attack of A63 N1 on the methyl group of O^6-methylguanine and the coordinated breakage of the guanine O6–C bond. This involves a train of electron transfers, the movement of the proton from C10 N3 to guanine N1, and a concomitant shift of the positive charge from C10 to the N^1-methyladenine at position 63. In principle, the guanine can now be released as a product, although there is no evidence that this occurs with the present form of the ribozyme. Regeneration of active ribozyme would also require an exchange of the substrate strand to place unmethylated adenine at position 63.

The proposed mechanism is fully consistent with the structure of the MTR1 riboswitch and the effect of the substitutions at C10 and U45 on activity. The complete loss of methylation activity of the C10U variant is fully consistent with the proposed role as a general acid in addition to ligand binding.

Figure 6.7.29 shows an interactive iCn3D model of a methyltransferase ribozyme (7V9E).

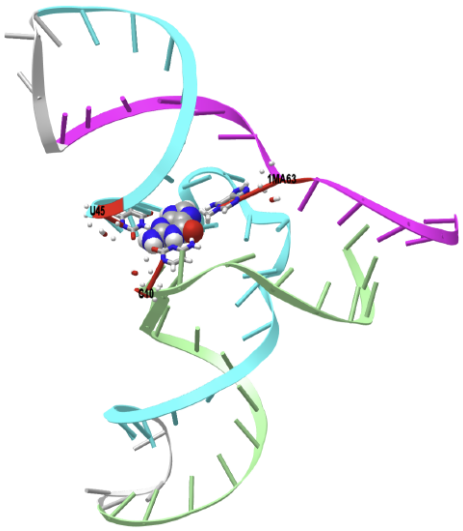


Figure 6.7.19: A methyltransferase ribozyme (7V9E). (Copyright; author via source).
Click the image for a popup or use this external link: https://structure.ncbi.nlm.nih.gov/i...LtyVP4DBnkEgH6

The parts of the ribozyme are named J12 (light green), J23 (cyan), and J31 (magenta). The active site residues, C10, U45 and 1MA63 (1-methyladenosine) are shown in CPK-colored sticks and labeled (disregard small separated spheres). The Ba^{2+} ion in the crystal structure is not displayed. The guanine is shown in spacefill.

6.7.9: Summary

This chapter examines the fascinating world of ribozymes—RNA molecules that not only store genetic information but also function as catalysts. It emphasizes that the capacity for catalysis is not unique to proteins; RNA can fold into complex secondary and tertiary structures that support catalytic activity. The chapter is organized around several key themes:

1. **Introduction to Ribozymes and Catalytic Principles:**
   The chapter opens by discussing the fundamental concept that any molecule capable of forming defined three-dimensional structures and containing functional groups—such as those needed for general acid/base, electrostatic, nucleophilic, or transition state stabilization catalysis—can act as a catalyst. Ribozymes illustrate this principle by employing RNA's limited chemical repertoire to drive electron flow from donors (sources) to acceptors (sinks) during bond making and breaking.

2. **Classification and Diversity:**
   Ribozymes are classified based on size and function. The chapter highlights that there are twelve classes of ribozymes, including small self-cleaving RNAs (e.g., hammerhead, viroid, and hairpin ribozymes, as well as riboswitches like glmS) and large catalytic RNAs such as Group I and Group II introns, ribonuclease P, and the catalytic RNAs of the spliceosome and ribosome. These examples demonstrate both cis (self-cleavage) and trans catalytic activities, illustrating the diversity of mechanisms by which RNA enzymes can operate.

3. **Mechanisms of Catalysis:**
   Detailed mechanistic insights are provided into how ribozymes achieve catalysis. The chapter explains the SN2 transesterification mechanism typical of self-cleaving ribozymes, where a general base activates a 2'-OH nucleophile that attacks an adjacent phosphodiester bond, leading to bond cleavage and formation of a pentavalent transition state. Examples such as the hammerhead ribozyme illustrate how conserved nucleotides (e.g., G8 and G12) and even subtle conformational arrangements contribute to efficient catalysis.

4. **Functional and Structural Examples:**
   Using interactive models and structural diagrams, the chapter explores specific ribozyme examples:

   - **Hammerhead Ribozymes:** Discussed in terms of their conserved secondary structure and active site architecture, which enable efficient self-cleavage and even trans-cleavage of RNA substrates.
   - **Viroids and Hairpin Ribozymes:** These small RNAs demonstrate the ability of ribozymes to function in viral and satellite RNA contexts.

- **Glucosamine-6-phosphate Ribozyme (glmS):** A unique example where the ribozyme is embedded in an mRNA and uses its own catalytic activity, modulated by binding of a small molecule cofactor, to regulate gene expression.
- **Group I and II Introns and Spliceosomal RNPs:** These larger ribozymes exemplify how RNA catalysis is integrated into essential processes such as RNA splicing and ribosome function, hinting at an evolutionary legacy from an RNA world.

5. **Evolutionary and Functional Significance:**
   The chapter considers the evolutionary implications of ribozymes, noting that RNA-based catalysis may have preceded protein enzymes. This dual functionality of RNA—as both genetic material and catalyst—supports the RNA world hypothesis and underscores the importance of ribozymes in early biochemistry.

6. **Integration with Modern Techniques:**
   Finally, the chapter highlights the methods used to study ribozymes, such as X-ray crystallography, cryo-electron microscopy, and interactive modeling tools like iCn3D. These techniques have revealed detailed insights into ribozyme structures, active sites, and catalytic mechanisms, bridging the gap between theoretical models and experimental observations.

Overall, this chapter provides a comprehensive overview of ribozyme catalysis, demonstrating how RNA can serve as an enzyme through diverse structural strategies and catalytic mechanisms. It sets the stage for deeper exploration of RNA's roles in modern biology and the evolutionary transition from RNA to protein catalysts.

6.8: Cofactors and Catalysis - A Little Help From My Friends

Learning Goals (DeepSeek, 1/30/25)

- **Understanding Cofactor Categories and Functions:**
  - Define and differentiate between metal cofactors and coenzymes, including the concept of prosthetic groups versus loosely bound coenzymes.
  - Explain how vitamin-derived coenzymes (e.g., TPP, FAD, NAD^+, PLP) and metal ions (e.g., iron in heme) contribute to enzyme catalysis by acting as electron sources or sinks.
- **Mechanistic Role in Electron Pushing:**
  - Describe the principles of electron flow in chemical reactions, identifying what constitutes an electron "source" and an electron "sink" in reaction mechanisms.
  - Illustrate, using arrow-pushing techniques, how cofactors facilitate electron transfer during bond-making and bond-breaking.
- **Specific Cofactor Mechanisms:**
  - Analyze the role of thiamine pyrophosphate (TPP) in the decarboxylation of α-keto acids and discuss its chemical mechanism in enzymes like pyruvate dehydrogenase.
  - Explain how FAD functions in redox reactions, particularly in the hydride transfer during the oxidation of succinate, and how NAD^+/$NADP^+$ participates as a dissociable electron acceptor in oxidation reactions.
  - Describe the versatile roles of pyridoxal phosphate (PLP) in catalyzing reactions such as decarboxylation, β-elimination, racemization, and transamination, emphasizing the importance of Schiff base formation and electron sink properties.
- **Integration of Cofactor Function with Enzyme Structure:**
  - Discuss how cofactors bind to specific residues (e.g., coordination of heme by histidines) and how this interaction is critical for proper enzyme function.
  - Evaluate how cofactor binding influences enzyme conformation, stability, and catalytic activity.
- **Clinical and Biological Implications:**
  - Relate vitamin deficiencies (e.g., lack of thiamine or niacin) to impaired coenzyme function and the resulting metabolic or disease states.
  - Discuss how alterations in cofactor binding or function can serve as regulatory mechanisms in cellular metabolism.
- **Analytical and Structural Approaches:**
  - Utilize interactive structural models (e.g., iCn3D) to examine cofactor binding sites within enzymes, and interpret how structural data supports mechanistic insights.
  - Compare reaction mechanisms that do and do not require cofactors (e.g., spontaneous decarboxylation vs. TPP-facilitated decarboxylation) to appreciate the catalytic advantages conferred by cofactors.

Achieving these goals will provide students with a comprehensive understanding of how cofactors serve as essential components in enzymatic catalysis by mediating electron flow, thereby linking fundamental chemical principles with the regulation of biological processes.

6.8.1: Cofactors and Electron Pushing: Sources and Sinks

To make and break bonds, electrons have to be moved. In drawing reaction mechanisms, we showed how electrons move from "sources" to "sinks." In many enzyme-catalyzed reactions, vitamin derivatives are used as substrates or "cofactors" or "coenzymes" to facilitate the flow of electrons in bond-making and breaking. The section focuses on cofactors, which facilitate the flow of electrons from the substrate to the product. We will see these enzymes in more detail in specific chapter sections.

Cofactors are molecules that bind to enzymes and are required for catalytic activity. They can be divided into two major categories: ***metals*** and ***coenzymes. Metal cofactors*** commonly found in human enzymes include iron, magnesium, manganese, cobalt, copper,