

Harvard University, Fall 2003
Chemistry 170: Chemical Biology
Mondays and Wednesdays 1:00 pm - 2:30 pm, Pfizer Lecture Hall, 12 Oxford Street

Instructors

Professor David R. Liu: liu@chemistry.harvard.edu, Mallinckrodt 303J, 496-1067
Professor Liu's office hours: Wednesdays 2:30-3:30 pm, Mallinckrodt 303J
Head Teaching Fellow: Jeff Doyon, doyon@fas.harvard.edu
Teaching Fellows: Michael Fischbach, fischbac@fas.harvard.edu
Polina Kehayova, kehayova@fas.harvard.edu
Vijay Krishnamurthy, krishn@fas.harvard.edu
Tom Snyder, tsnyder@fas.harvard.edu
Arturo Vegas, vegas@fas.harvard.edu

Prerequisites

A strong background in organic chemistry is required. A basic understanding of biochemistry and molecular biology is assumed although a lack of this understanding can be remedied during the semester by very diligently reviewing the relevant materials (including the suggested textbooks and papers).

Textbooks

No texts are formally required; the references below are recommended and should be available in the Cabot Science Library, the Chemistry library, the Harvard Coop, or from online bookstores. If you are considering pursuing a field related to chemical biology you may want to own some or all of these books.

Strongly recommended:

Creighton, T. E. *Proteins: Structure and Molecular Function*
Blackburn, G. M. and Gait, M. *Nucleic Acids in Chemistry and Biology*
Fersht, A. *Structure and Mechanism in Protein Science*
Silverman, R. B. *The Organic Chemistry of Enzyme-Catalyzed Reactions*

Recommended as excellent references for specific topics covered in this course:

Branden, C. and Tooze, J. *Introduction to Protein Structure*
Sanger, W. *Principles of Nucleic Acid Structure*
Silverman, R. B. *The Organic Chemistry of Drug Design*

Recommended as general references in biological chemistry and biochemistry:

Dugas, H. *Bioorganic Chemistry*
Stryer, L. *Biochemistry*
Voet, D. *Biochemistry*

Lecture notes, journal articles from the primary literature, and lists of supplementary references will be distributed regularly and form the majority of the course material.

Reserve materials and course web site

All lectures will be videotaped and the tapes placed on reserve in the Cabot Science Library in the Science Center. In addition, the Chemistry 170 website will also contain the lecture slides in color PDF format, problem sets and solutions, section papers, and other handouts when possible. To access the website, go to <http://www.courses.fas.harvard.edu/~chem170> and login with username = chem170 and password = fortytwo.

Weekly Help Sessions

Two teaching fellows will be available to answer remaining questions about lecture material each Monday from 7:00-8:00 pm in the Liu Group Room (Mallinckrodt 303L). To gain entrance to Mallinckrodt in the evening hours, enter through Fairchild by showing your Harvard ID to the guard. From Fairchild, take the stairs up to the 4th floor of Fairchild, and cross over to the 3rd floor of Mallinckrodt through the door connecting the two buildings. Unless otherwise noted, there will be no formal presentations by the teaching fellows during these sessions, so please come with your questions. No help session will be held on October 13 (Columbus Day), December 22, or December 29. You are always welcome to contact your teaching fellows outside of these sessions for additional guidance.

Course Requirements

Sections are mandatory and will be held once a week starting Monday, September 22 in Mallinckrodt 217 (in the CCB department center). Sections involve a detailed discussion of two previously distributed journal articles, problem set questions due that week, and recent lectures. Section preference forms will be handed out in class on September 15 and are due by noon on Wednesday, September 17 in Mallinckrodt 303E. If we do not receive your section preference form on time you will be assigned a section.

Problem sets will consist of questions posed in the previous week's lecture notes together with more in-depth questions handed out in advance of the relevant lectures. These questions are designed to foster deeper thinking about topics introduced in lecture and to help students prepare for taking examinations (below). Attending lectures and completing suggested readings will provide sufficient background to complete problem sets. *Written answers to the previous week's problem set questions (covering two lectures) are due each Monday at 1:00 (before lecture). Problem sets will be graded for both accuracy and effort.*

Two take-home examinations will be distributed during the semester on Wednesday, October 29, 1:00 pm (midterm), and on Monday, January 5, 1:00 pm (final). Each examination will be cumulative in material covered and will be due in Professor Liu's office 48 hours (midterm) or 72 hours (final) after distribution. Students are free to use their own class notes, handouts, and the library during this time, but are not allowed to collaborate with other students.

An original research proposal in chemical biology is due at the end of the course. Research proposals will be presented during public poster sessions held Thursday, January 15 and Friday, January 16. A proposal topic must be submitted to your teaching fellow in the form of a brief outline or description no later than Monday, November 24 for approval and guidance. If this proposal is approved, a revised version of your proposal including suggestions by your TF is due on Monday, December 8. If your first proposal is rejected, a proposal on a new topic is due December 8. A brief written version of the proposal is also due with your poster presentation. See the attached handout for details regarding the format of the outline and proposal.

The course grade will be determined as:

- Section participation and problem sets (20%)
- Midterm examination (20%), due October 31, 2003 at 1:00 pm
- Final examination (30%), due January 8, 2004 at 1:00 pm
- Research proposal (30%), due in outline November 24, revised outline due December 8, 2003, written proposal and poster presentation due January 15-16, 2004

Chemistry 170, Fall 2003 Course Outline and Take-Home Lessons

Lecture dates are subject to change.

Physical properties of nucleic acids (Lectures 1 & 2: September 15 & 17)

*What is the chemical basis for nucleic acid structure and for the specificity behind nucleic acid hybridization?
What are the geometric, chemical, and biological properties of the common forms of DNA and RNA structure?*

- Nomenclature
- Hydrogen bonding
- Ionization states
- Anti* and *syn* conformers
- Tautomerization
- Ribose conformers
- Interactions holding DNA together
- Forces determining nucleic acid structure
- Modern solid-phase DNA synthesis
- Peptide nucleic acids (PNAs)
- DNA and RNA structural diversity
 - DNA polymorphism
 - RNA polymorphism
 - Roles of DNA and RNA forms in biology
 - Higher order DNA and RNA structures
 - Chemical etiology of nucleic acids
 - Unnatural base pairs
 - Insights into DNA replication

Physical properties of proteins (Lecture 3: September 22)

*What chemical principles determine protein structure at the molecular level?
How can our understanding of these chemical principles explain the macromolecular properties of proteins?*

- Post-translational modification
- Post-translational rearrangement
 - Histidine decarboxylase
 - Phenylalanine ammonia lyase
- Amino acid properties
- Determinants of protein structure
- Hydrophobic packing in phage λ repressor
- Protein conformational space
- Secondary and higher order structures

Novel biosyntheses of proteins (Lecture 4: September 24)

*How can proteins with nonproteinogenic amino acids be biosynthesized in vitro and in vivo?
At what level can we interpret a given set of mutagenesis data?
How are proteins biosynthesized nonribosomally?*

- Chemical approaches to studying proteins
- Incorporation of unnatural amino acids into proteins in vitro
 - Chorismate mutase
 - Ligand-gated ion channel receptors
 - Staphylococcal* nuclease
 - Site-specific chemical proteolysis
- Nonribosomal peptide synthesis
 - Mechanism
 - Structural basis for amino acid specificity
 - Increasing structural diversity
 - Peptide cyclization by NRPSs

Mechanistic enzymology (Lectures 5 & 6: September 29, October 1)

What chemical factors contribute to rate acceleration by enzymes?

*What are the mechanisms by which common cofactors catalyze reactions?
How does chorismate mutase accelerate the Claisen rearrangement of chorismate to prephenate?
What are current theories explaining how OMP decarboxylase achieves its remarkable rate enhancement?
How do squalene cyclase and oxidosqualene cyclase catalyze the formation of triterpenes?*

- Components of enzymic catalysis
 - Transition state stabilization
 - Proximity effects and prepaying entropy
 - General acid/general base catalysis
 - Nucleophilic and electrophilic catalysis
 - Electrostatic catalysis and modulating the dielectric
- Cofactor catalysis
 - NADH
 - FADH₂
 - Pyridoxal phosphate
 - Thiamine pyrophosphate
- Chorismate mutase
- Orotidine 5'-monophosphate decarboxylase
 - Kinetic isotope effects
- Enzyme-catalyzed mechanisms of triterpene formation
 - Chemical studies on squalene and oxidosqualene cyclases
 - Substrate folding
 - Cyclization events
 - Rearrangement
 - Biological studies on squalene and oxidosqualene cyclases
 - Structural studies
 - Enzyme labeling studies
 - Site-directed mutagenesis

De novo protein design (Lecture 7: October 6)

*What approaches and basic principles have researchers used to design proteins?
What are the types of protein structures we can make by de novo design approaches?
What are the current limitations of de novo protein design?
How does a molten globule differ from a typical native protein?*

- The protein design problem
- Designability of protein folds
- Minimalist approaches to protein design
- Empirical approaches to protein design
- Molten globules
- Context effects: Janus and Chameleon
- Computational design of proteins
 - Structural motifs
 - Receptors
 - Catalysts
- Synthetic enzymes and receptors

Rational protein engineering (Lecture 8: October 8)

*What are the approaches that researchers have taken towards successfully engineering proteins rationally?
What factors contribute to the complexity of proteins and limit our ability to rationally engineer them?*

- Fiddling with proteins ("protein terrorism"): the uphill climb
- Dehydrogenases that use different cofactors
- Trypsin proteases with new specificities
- Engineering a protein ligase
- Nuclear hormone receptor engineering
- Engineering new activities into existing active sites
- Engineering homing endonucleases

October 13: Columbus Day

Molecular evolution (Lectures 9 & 10: October 15, 20)

What are the requirements to evolve a molecule?

What are the advantages and disadvantages of the various ways to screen or select for desired properties?

How have researchers evolved proteins with enhanced activities or new substrate specificities?

How have researchers evolved proteins and nucleic acids with entirely new activities?

Concepts in molecular evolution

- Rational and combinatorial approaches to problem solving

- Implementation of evolutionary approaches

 - Sources of diversity

 - Methods to amplify molecules

 - Tagging and compartmentalization

 - Picking desired molecules

- In vitro* versus *in vivo* methods; screens versus selections

Evolving enhanced natural function

- β-lactamase and cephalosporinase

- Oxidosqualene cyclase

- Subtilisin

- Green fluorescent protein (GFP)

- Aminoacyl-tRNA synthetases

- Human interferon-β

- Tetracycline repressor

Evolving new function

- Remodeling proteins

 - Protein-protein interface of hGH

 - Minimizing protein A

- Erythropoietin peptide mimetics

- Protease-resistant peptide ligands by mirror image phage display

Nucleic acid catalysis (Lecture 11: October 22)

What essential biological processes are known to be catalyzed by RNA?

What are common mechanisms of RNA (or DNA) catalysis?

What reactions beyond those found to be ribozyme-catalyzed in nature have been catalyzed by laboratory-evolved nucleic acids?

Natural catalytic RNA

- Tetrahymena* self-splicing intron

- Hepatitis delta virus

- The ribosome

Laboratory-evolved catalytic RNA

- Aminoacyl-tRNA synthetase

- RNA ligase

- RNA polymerase

- Continuous evolution of an RNA ligase

- Diels-Alderase

Catalytic DNA

- Evolution with modified nucleotides

Emerging Roles of RNA (Lecture 12: October 27)

How have researchers and nature used ligand-binding RNA aptamers to regulate RNA function?

What is RNAi and how is it thought to work?

How have small RNAs been implicated in gene-specific translational inhibition?

- Allosteric ribozymes

- Natural riboswitches

- RNA interference (RNAi)

 - Discovery

 - Mechanism

- Translational inhibition

microRNAs
Listeria thermosensor
tmRNA

(End of midterm material)

October 29: Midterm handed out at 1:00 pm in the Pfizer Lecture Hall; no lecture

DNA damage (Lecture 13: November 3)

*What are the most reactive sites on DNA and RNA to nucleophilic, electrophilic, or photochemical reaction?
What are the chemical mechanisms behind mutagenesis?*

Spontaneous
 Deamination
 Depurination
Oxidative damage mechanisms
Ionizing irradiation damage mechanisms
UV irradiation damage mechanisms
Chemical abuse of DNA
 Alkylation and damage by electrophiles
 Nucleophilic attack of DNA
 Crosslinking DNA
 Metabolically activated DNA damage agents

Mechanisms of DNA repair (Lecture 14: November 5)

*What are the mechanisms by which enzymes repair lesions to DNA?
How does the cell distinguish between correct and incorrect bases when repairing mismatched (but otherwise normal) DNA?*

Four systems for DNA repair
Direct reversal of alkylation
 Ada
 AlkB
Base excision repair
 Uracil glycosylase
 MUG
 Oxoguanine glycosylase
Methyl-directed mismatch repair
 MutS finds the mismatches
 MutL is the switch
 MutH is the endonuclease

Novel natural proteins (Lectures 15 & 16: November 10, 12)

*What properties make each of these four classes of proteins unique?
What is the chemical basis for each of their unique properties?
How have researchers exploited these properties toward novel uses of these proteins?*

Green fluorescent protein (GFP)
 Chromophore formation
 Protein structure
 Generating new GFP variants
 Red fluorescent protein
Inteins: self-splicing proteins
 Structural basis for protein splicing
 Chemical mechanisms of splicing
 Applications of inteins
 Protein purification and tagging
 Expressed protein ligation
Catalytic antibodies
 Using immunological diversity

- Recruiting catalytic groups
- Reactive immunization
- An antibody aldolase
- Evolution of antibodies *in vitro*

Prions

- The protein only hypothesis
- Prion transmission and propagation
- Structure and proposed mechanisms of fibril formation

Metabolic Engineering (Lecture 17: November 17)

What methods have been used to introduce orthogonal chemical handles on the outside of cells?

What are the drawbacks and advantages of these methods?

How is erythromycin biosynthesized in mechanistic detail?

How can the gene structure of a polyketide synthase be used to partially predict the polyketide's structure?

Engineering cell surface glycoproteins

- Decorating cell surfaces with ketones
- Targeting cells with toxins
- Creating artificial virus receptors

Engineering polyketide synthases

- Erythromycin biosynthesis
- Engineering new polyketide synthases
- Combining synthetic substrates
- Libraries of polyketides from libraries of polyketide synthases

Challenges to metabolic engineering of polyketides

Molecular mechanisms of drug action and drug resistance (Lecture 18: November 19)

How do β -lactams kill bacteria?

What is the molecular basis for resistance to β -lactams?

How does vancomycin inhibit bacterial growth?

What are the components and molecular mechanisms behind vancomycin resistance?

What are the chemical bases for Viagra's pharmacological properties?

Penicillin and β -lactams

Mechanisms of β -lactam resistance

Vancomycin

Mechanisms of vancomycin resistance

- VanA, VanH, VanX, VanS, VanR

- New mechanisms of glycopeptide action

Chemical approaches to overcoming vancomycin resistance

How ViagraTM works

- Development of ViagraTM

- Mechanism of sildenafil action

Rational design of macromolecular ligands (Lectures 19 & 20: November 24, December 1)

How were polyamides designed to sequence specifically recognize DNA?

How do the latest generation of DNA-binding polyamides differ in binding properties from that of distamycin, and what is the molecular basis of these differences?

What are the principles that drive rational drug design?

How were saquinavir and ritanovir developed?

Sequence-specific DNA binding polyamides

- Minor groove binding

- Distamycin: A:T and T:A binding

- Designing C:G and G:C binding

- Improving binding by entropy prepayment

- Cracking the A:T/T:A degeneracy

- Targeting genes *in vivo* with polyamides

 - Blocking DNA binding by proteins

 - Inducing DNA cleavage

 - Inhibiting nucleosome translocation

Rational drug design of protein binding ligands

- Structure-based drug design
- HIV and HIV protease
 - Development of saquinavir
 - Development of ritanovir
 - Cyclic ureas: a newer approach
 - Resistance to protease inhibitors
- Qualitative principles of rational drug design

November 26: no lecture (pre-Thanksgiving break)

Combinatorial approaches to small molecule discovery (Lecture 21: December 3)

*How are combinatorial approaches being applied to the synthesis and discovery of small molecules?
What are the chemical requirements for using combinatorial chemistry in small molecule discovery and what are the most common solutions that satisfy these requirements?*

- Solid phase synthesis
 - Resins
 - Linkers
- Sources of diversity
 - Building block
 - Stereochemical
 - Skeletal rearrangements
- Tagging strategies
 - Spatially separated synthesis
 - Direct methods
 - Deconvolution
 - Encoding with tags
 - Halogenated aromatics
 - DNA tags
 - Radiotags
- Picking desired molecules
- Target-oriented synthesis of a library of benzodiazepines
- Diversity-oriented synthesis of a library of polycyclic compounds
- Incorporating evolution-based elements into combinatorial small molecule discovery
 - Dynamic combinatorial chemistry
 - Target-directed combinatorial chemistry
 - DNA-templated synthesis

Chemical genetics (Lecture 22: December 8)

What similarities and differences emerge when comparing small molecules and genetic mutations as entries into studying biological processes?

How was FK506BP discovered and what is the molecular basis for its immunosuppressive properties?

What principles were used to apply chemical genetics to the study of protein tyrosine kinases?

- Forward and reverse traditional genetics and chemical genetics
- Chemical genetics in immunosuppression
 - FK506 and cyclosporin A
 - Elucidation of the target of FK506
 - The function of FK506 binding protein
- Forward chemical genetics of mitosis: monastrol
- Reverse chemical genetics in protein kinases
 - Developing orthogonal kinase-ATP pairs
 - Probing Src and Fyn kinase targets
 - Engineering kinase-specific inhibitors

Genomics (Lecture 23: December 10)

What are the most common methods for studying genes and gene expression on a genome-wide scale?

How do these methods work operationally?

What insights into biological systems have emerged from the use of genomics?

What challenges lie ahead?

- Introduction to genomics
- Synthetic oligonucleotide arrays
 - Synthesizing the array
 - Case studies using synthetic DNA arrays
 - Expression monitoring in yeast
 - Kinase inhibition
- cDNA arrays
 - Printing the array
 - Case studies using cDNA arrays
 - Heat shock and phorbol ester response
 - Genome-wide expression in the diauxic shift
 - Genomics of tuberculosis treatment
- Probing protein-DNA interactions with DNA arrays
- Transfected cell microarrays

Proteomics (Lecture 24: December 15)

What is the need for proteomics?

How are proteins separated and identified on an organism-wide scale?

How can the function of proteins and their interaction partners be determined on this scale?

What new technologies drive proteomics?

- Introduction to proteomics
- Protein separation and identification
 - Two-dimensional electrophoresis
 - Mass spectrometry of proteins
 - Case study: phosphorylated proteins
- Determining protein function
 - Activity-based proteome profiling
 - General approach to identifying protein function
 - Proteome profiling
- Revealing protein interactions
 - Yeast two hybrid experiments
 - Yeast *n*-hybrid
 - Proteome-wide protein-protein interaction mapping
- Proteomics-oriented technologies