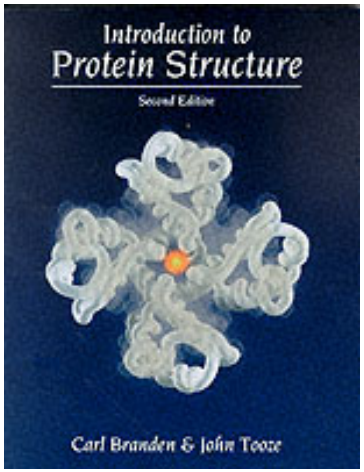
CE405/505: Protein Engineering

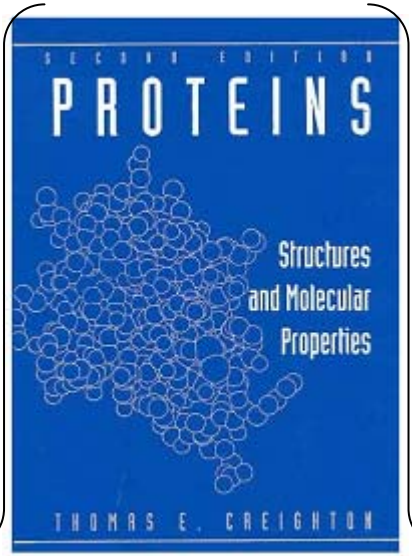
MWF 9 – 9:50 am, 214 Norton Hall

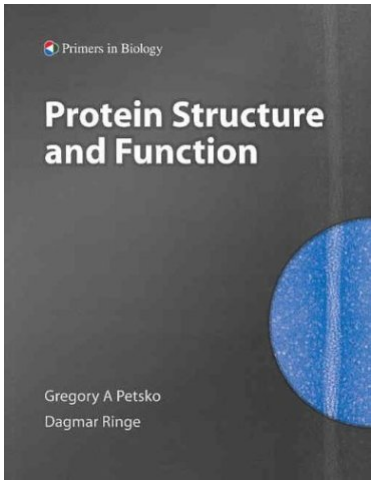
Instructor: Sheldon Park (905 Furnas Hall)

Office Hours: Tu, Th 4 – 5:00 pm

(If not in, try 815 Furnas Hall)







Grade

- Homework (7 minus 1) : 5% each
- Mid terms (2) : 15% each
- Pop quizzes (*n*) : 5%
- Final exam : 35%

Pre-requisites—"Recommended"

- Chemistry, Biochemistry, Molecular Biology

Textbook

- Introduction to Protein Structure, 2nd Ed. Branden & Tooze
- Other references:
  - Proteins: Creighton
  - Protein Structure and Function, Petsko & Ringe

What is the class about?

Properties intrinsic to protein molecules

- Protein molecules share common physical and chemical attributes
- Structural organization
  - hierarchy, classification, databases
- Sequence-structure relationship
  - stability, core packing, amino acid alphabet, folding

Properties important to macromolecular interactions

- molecular interaction, receptor ligand interaction, interface

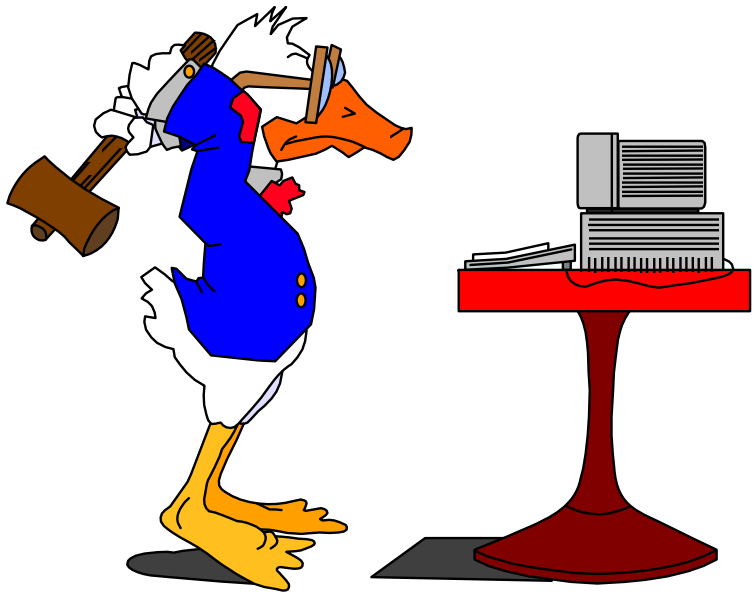
How do we identify and leverage the information that exists in known structures to accelerate the discovery of new molecules?

Developing intuition into protein structure by studying examples

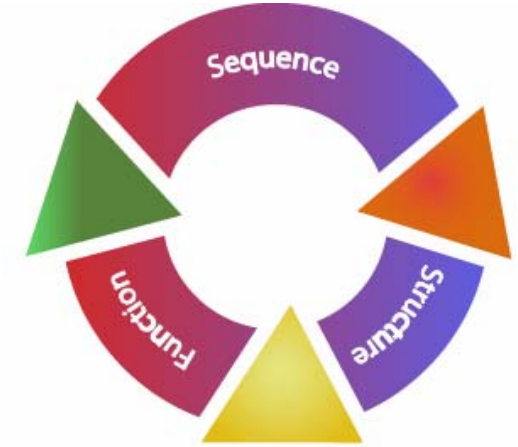
- Mutagenesis
- High resolution structure
- Knowledge-based and computational protein design (low throughput)
- Engineering by directed evolution (high throughput)

Using computer to study and visualize protein molecules

- Bioinformatics
- Molecular graphics
- Homology modeling
- MD simulations



Protein Engineering



The art of creating proteins with novel structural and functional properties

To study sequence-structure-function relationship in protein

- systematic and deliberate engineering of mutant proteins is essential to infer fundamental principles
- applies detailed physicochemical analysis to protein structures to identify relevant structural parameters and develop new metrics for functional correlation

Biotechnology and industrial applications

- molecular biology applications (polymerase, ligase, protease, etc)
- detergent (subtilisin)
- starch-processing industry (alpha amylase, glucose isomerase)
- waste treatment, paper manufacturing, organic synthesis, …

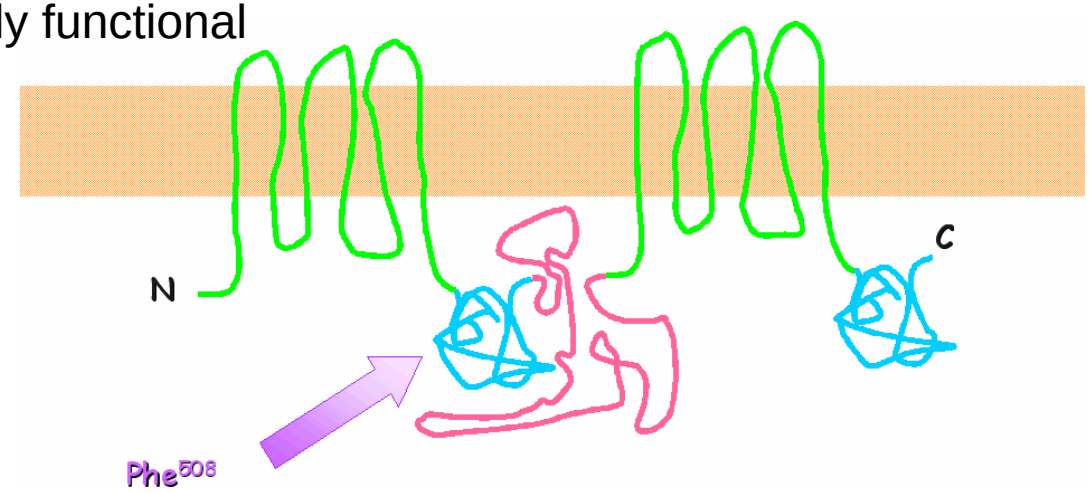
Cystic fibrosis

Cystic fibrosis is a hereditary disease that affects growth, breathing, digestion and reproduction, and cause early death

The disease is caused by a mutation in cystic fibrosis transmembrane conductance regulator (CFTR), which is a transporter of chloride ions across the epithelial cell membrane

The most common mutation is deletion of an amino acid phenylalanine 508 in the nucleotide binding domain of CFTR

The mutation causes the protein to be degraded before reaching the destination, although the folded protein is mostly functional

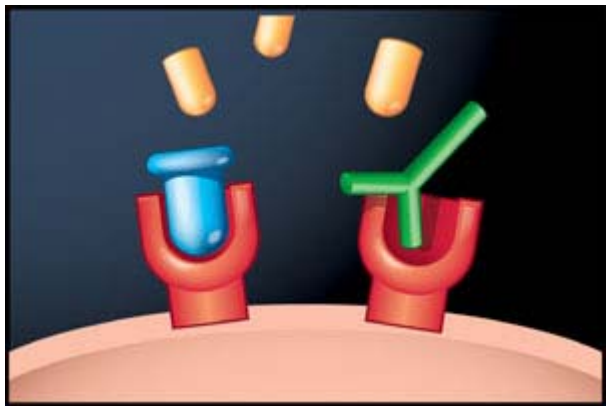


Therapeutic Proteins

One of three major classes of therapeutic drugs
(the other two are small molecules and RNA)

The therapeutic protein market was valued at $37b in 2003 and expected to top $90b by 2010

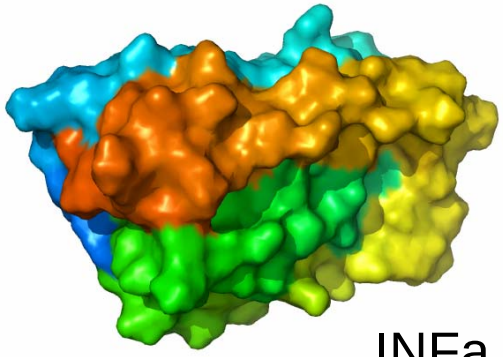
Therapeutic proteins are used to treat patients with cancer, heart attacks, strokes, cystic fibrosis, diabetes, anaemia, haemophilia



Erythropoietins

Procrit treats chemotherapy related anemia, chronic kidney disease not on dialysis, anemia related to AZT treatment (HIV)

Epogen (Amgen) treats anemia in patients with kidney disease on dialysis—exact copy of the hormone secreted by the kidney—tells bone marrow to make red blood cells

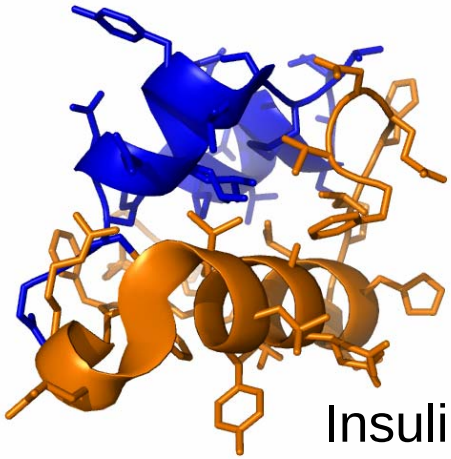


INFa

Interferons

Intron A (Schering) mimics the activity of an interferon that is secreted in response to viral infection. Used to treat hepatitis B and C

Roferon A is used to treat hepatitis C but not B



Insulin

Insulins

Extracted from beef or pork pancreas glands

Produced from recombinant DNA

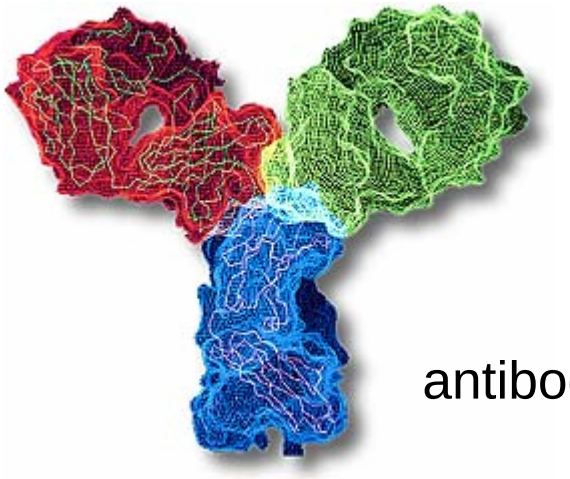
Monoclonal Anibodies

Remicade: chimeric monoclonal antibody composed of human constant and murine variable region, produced by a recombinant cell line, blocks TNFa , is used to treat rheumatoid arthritis, psoriatic arthritis

Herceptin is the approved antibody for the treatment of metastatic breast cancer for patients whose tumors overexpress the HER2 protein

Humira (Abbott): recombinant human IgG1 monoclonal antibody, reduces the signs and symptoms of and progression of join damage in adult patients with moderate to severe rheumatoid arthritis by blocking the action of TNFa

Erbitux (ImClone): monoclonal antibody to treat advanced colorectal cancer



antibody

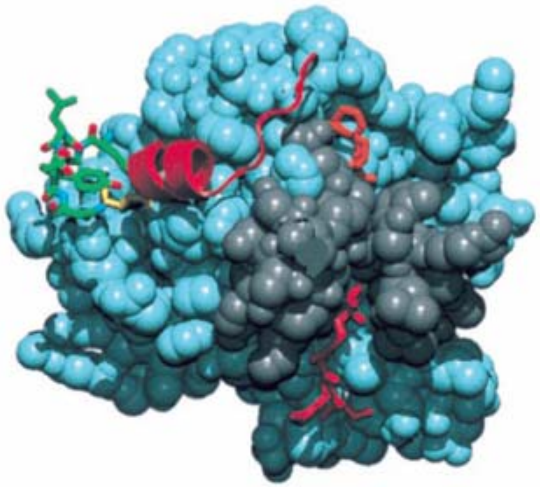
Blood factors

NovoSeven : recombinant Factor VIIa (FVIIa) (serine protease) used in trauma and haemorrhage

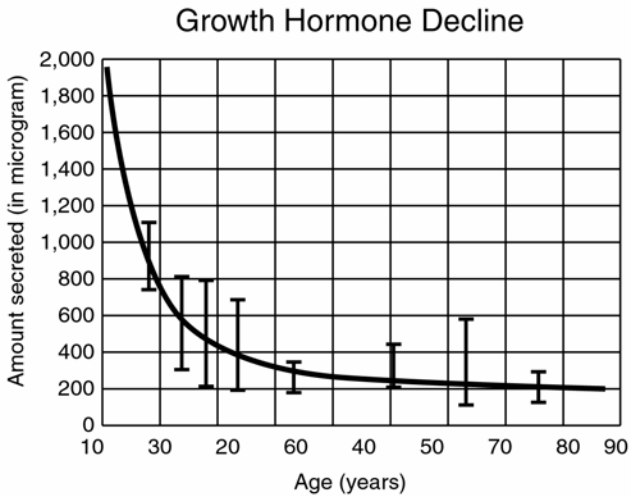
Colony stimulating factors

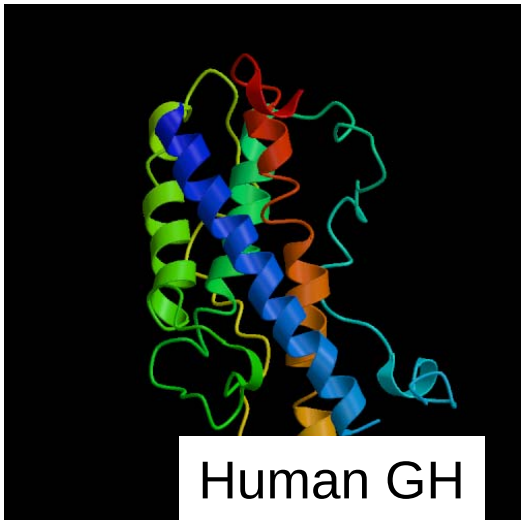
Neupogen: used to increase white blood cells, to decrease the risk of infection, in conditions such as cancer; bone marrow transplant; pre-chemotherapy blood cell collection; and severe chronic neutropenia including congenital neutropenia, cyclic neutropenia, and idiopathic neutropenia

Neulasta: prescribed to reduce the risk of infection in patients with some tumors receiving chemotherapy that may decrease the number of infection-fighting white blood cells



NovoSeven





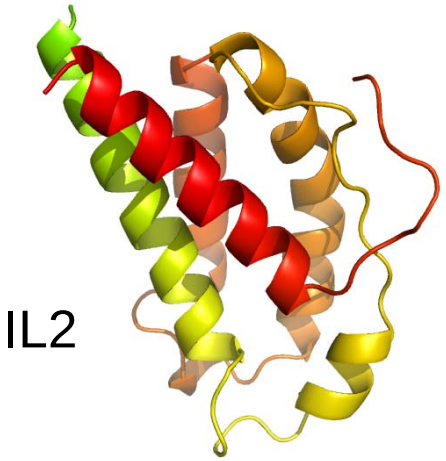
Human GH

Growth hormones

Genotropin, Nutropin, Norditropin: recombinant human growth hormones

Interleuekins

Proleukin IL-2 therapy is used to treat metastatic kidney cancer and metastatic melanoma (skin cancer)



IL2

Growth factors

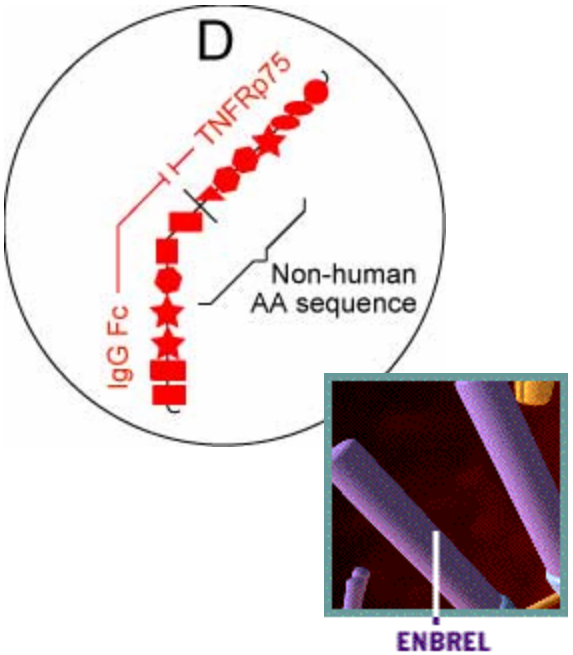
Fiblast (Scios): basic fibroblast growth factor used to treat stroke

Therapeutic vaccines

OncoVAX: vaccine therapy targeted to treat people suffering from Stage II colon cancer. Administered after surgery, OncoVAX can help reduce the recurrence of colon cancer by training a patient's own immune system to attack cancer cells.

Enbrel

recombinant human soluble TNF alpha receptor and is used to treat rheumatoid arthritis



Protein Engineering Strategy

There are two general strategies for protein engineering

Rational design

A scientist applies lessons learned from studying the structure and function of existing proteins to rationally engineer changes

knowledge-based approach

computational protein design

In a **knowledge-based** approach, an expert scientist applies his/her “intuition” to decide what to engineer and how to do it

The information may be chemical, statistical, biochemical, and the success of the approach depends critically on the judgment of the scientist

Computational protein design

A computer algorithm is used to identify amino acid sequences that satisfy various energy criteria

Pros

»Powerful algorithms have been developed to rapidly screen astronomically large number of sequences on a realistic time scale

»A computer code objectively evaluates amino acid sequences using a metric that is extensible to other systems, and thus places protein design on a scientific footing

Cons

»The algorithms require a carefully crafted energy force field which is used to judge the quality of various sequences

»Computational protein design requires a significant setup in terms of code development, force field optimization (and people are very protective of their code)

Directed Evolution

This method mimics natural evolution, in which a generation of diversity and a selection based the survival of the fittest leads to engineered protein molecules with desired characteristics

The diversity is generated either through random mutation or through DNA recombination

The method is inherently high throughput and requires an efficient method of generating a genetically diverse pool of individuals as well as a method to screen or select the pool

On the other hand, many of the discoveries involve mutations that one would not have expected and therefore could not have designed rationally.