

# Advanced Cell Biology. Lecture 15

# Alexey Shipunov

Minot State University

February 18th, 2011





## Outline

## How proteins are controlled



## Outline

## How proteins are controlled

## How proteins are studied



## Outline

## How proteins are controlled

## How proteins are studied

## DNA replication







Previous final question: the answer

Which way of sedimentation employs differences between molecular sizes?

- ▶ Velocity sedimentation: larger proteins sediment faster



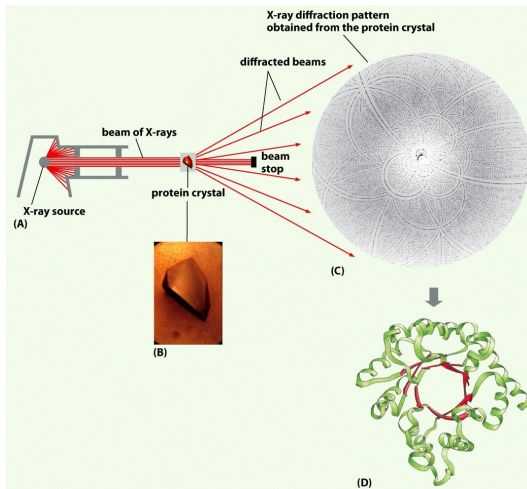
## Safe crackers animation



- ▶ Based on the ability of proteins to form crystals
- ▶ These crystals will scatter X-rays and we will see diffraction pattern, different for different proteins
- ▶ Useful for prediction of protein 3D structure



# X-rays crystallography

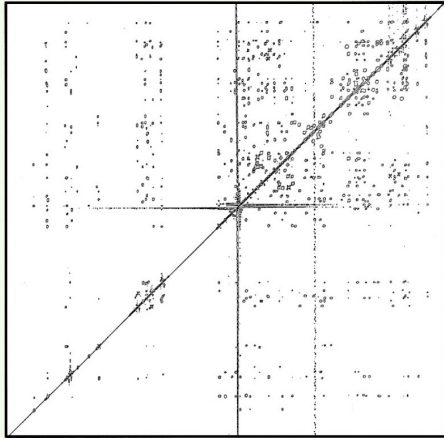




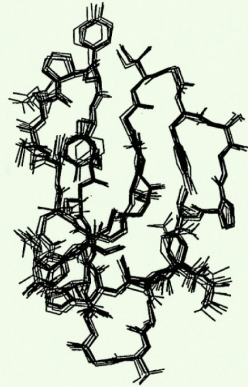
- ▶ Unfortunately, not all proteins form crystals and it is not always possible to obtain enough protein for crystallization
- ▶ Method is based on the response of nuclei to radio waves
- ▶ Again, it is used mostly for understanding conformation of protein



## Nuclear magnetic resonance



(A)



(B)



- ▶ UniProt
- ▶ NCBI protein
- ▶ RCSB



# NCBI Protein database

The screenshot shows the NCBI Protein database homepage in a web browser. The address bar displays <http://www.ncbi.nlm.nih.gov/protein>. The page features a search bar with the text "Protein" and a "Search" button. Below the search bar, there is a section titled "Protein" with the subtitle "Translations of Life". To the right of this section, a dark blue box contains the text: "The Protein database is a collection of sequences from several sources, including translations from annotated coding regions in GenBank, RefSeq and TPA, as well as records from SwissProt, PIR, PRF, and PDB. Protein sequences are the fundamental determinants of biological structure and function." Below this, there are three columns of links: "Using Protein" (Quick Start Guide, FAQ, Help, GenBank FTP, RefSeq FTP), "Protein Tools" (Sequence Revision History, BLAST, LinkOut, E-Utilities, Bink, Batch Entrez), and "Other Resources" (GenBank Home, RefSeq Home, CDD, PeptideMine, Structure). At the bottom, there is a navigation bar with links to "GETTING STARTED", "RESOURCES", "POPULAR", "FEATURED", and "NCBI INFORMATION".

File Edit View History Bookmarks Tools Help

http://www.ncbi.nlm.nih.gov/protein

Protein home

NCBI Resources How To

My NCBI Sign In

Protein  
Translations of Life

Search: Protein Limits Advanced search Help

Search Clear

**Protein**

The Protein database is a collection of sequences from several sources, including translations from annotated coding regions in GenBank, RefSeq and TPA, as well as records from SwissProt, PIR, PRF, and PDB. Protein sequences are the fundamental determinants of biological structure and function.

**Using Protein**

[Quick Start Guide](#)  
[FAQ](#)  
[Help](#)  
[GenBank FTP](#)  
[RefSeq FTP](#)

**Protein Tools**

[Sequence Revision History](#)  
[BLAST](#)  
[LinkOut](#)  
[E-Utilities](#)  
[Bink](#)  
[Batch Entrez](#)

**Other Resources**

[GenBank Home](#)  
[RefSeq Home](#)  
[CDD](#)  
[PeptideMine](#)  
[Structure](#)

You are here: NCBI > Proteins > Protein

Write to the Help Desk

**GETTING STARTED**  
NCBI Education

**RESOURCES**  
Chemicals & Biossays

**POPULAR**  
PubMed

**FEATURED**  
GenBank

**NCBI INFORMATION**  
About NCBI

http://www.ncbi.nlm.nih.gov/protein#



# RCSB database

File Edit View History Bookmarks Tools Help

http://www.rcsb.org/pdb/home/home.do

RCSB Protein Data Bank

**PDB**  
PROTEIN DATA BANK

A MEMBER OF THE **PDZ**  
An Information Portal to Biological Macromolecular Structures  
As of Tuesday Feb 15, 2011 at 4 PM PST there are 71264 Structures | PDB Statistics

Contact Us | Print

PDB ID or Text PDB ID lookup or Text search of the complete structure file Search Advanced Search

**MyPDB** Hide  
Login to your Account  
Register a New Account

**Home** Hide  
News & Publications  
Usage/Reference Policies  
Deposition Policies  
Website FAQ  
Deposition FAQ  
Contact Us  
About Us  
Careers  
External Links  
Sitemap  
New Website Features

**Deposition** Hide  
All Deposit Services  
Electron Microscopy  
X-ray | NMR  
Validation Server  
Biocytic Beamline  
Related Tools

**Search** Hide  
Advanced Search  
Latest Release  
New Structure Papers  
Sequence Search  
Chemical Components  
Unreleased Entries  
Browse Database  
Histograms

**Tools** Hide  
Download: Entries | Ligands  
Compare Structures  
FTP Services

**A Resource for Studying Biological Macromolecules**

The PDB archive contains information about experimentally-determined structures of proteins, nucleic acids, and complex assemblies. As a member of the wwPDB, the RCSB PDB curates and annotates PDB data according to agreed upon standards.

The RCSB PDB also provides a variety of tools and resources. Users can perform simple and advanced searches based on annotations relating to sequence, structure and function. These molecules are visualized, downloaded, and analyzed by users who range from students to specialized scientists.

**Hide Welcome Message**

**Featured Molecules** Hide  
List View of Archive By: Title | Date | Category

**Structural View of Biology**

**Infrastructure & Communication**

**Molecule of the Month: Integrin**

Our bodies are composed of approximately ten trillion cells, which pose challenging problems for structure and communication. All of these cells must be connected strongly together, to allow us to stand and walk. The infrastructure holding us together, however, must also be malleable enough to allow repairs, to allow us to heal from wounds. These many cells must also communicate with each other, ensuring that each plays its own proper part. Many different molecules in our bodies are involved in this complex infrastructure of support and communication, and Integrins play a central role.

**Customize This Page**

**New Features** Hide  
Redesigned Help System  
Latest features released:  
Website Release Archive: ...

**RCSB PDB News** Hide  
Weekly | Quarterly | Yearly

2011-02-15  
**Upcoming Meeting: AAAS Meeting and Family Days**  
RCSB PDB and SBAB will promote molecular explorations of biology at the AAAS Annual Meeting (Feb 17-21, Washington DC). [more...](#)

• Create High Resolution Images  
• Store Personal Annotations With MyPDB  
• NO Science Glycoprotein Reducing Results

**wwPDB News** Hide  
Statement on Retraction of PDB Entries

2011-02-14  
**Special Symposium Celebrating the 40th Anniversary of the PDB**

• Time-stamped Copies of PDB Archive Available via FTP  
• Full wwPDB News



# UniProt database

The screenshot shows the UniProt website in a web browser. The browser's address bar displays 'http://www.uniprot.org/'. The website has a navigation bar with links for 'Downloads', 'Contact', and 'Documentation/Help'. Below this is a search bar with tabs for 'Search', 'Blast', 'Align', 'Retrieve', and 'ID Mapping'. The 'Search' tab is active, showing a 'Search in' dropdown set to 'Protein Knowledgebase (UniProtKB)' and a 'Query' input field. Below the search bar, the 'WELCOME' section states the mission of UniProt: 'The mission of UniProt is to provide the scientific community with a comprehensive, high-quality and freely accessible resource of protein sequence and functional information.' To the right, the 'NEWS' section features a 'UniProt release 2011\_02 - Feb 8, 2011' announcement, detailing automatic annotation of UniProtKB/TrEMBL using PDB-derived data and cross-references to nextProt. Below the news, the 'SITE TOUR' section includes a thumbnail image of the website and text encouraging users to learn how to make best use of the tools and data. The 'PROTEIN SPOTLIGHT' section is also visible at the bottom. On the left, a table titled 'What we provide' lists the components of the UniProt database: UniProtKB (Protein knowledgebase, consisting of two sections: Swiss-Prot, which is manually annotated and reviewed, and TrEMBL, which is automatically annotated and is not reviewed), UniRef (Sequence clusters, used to speed up sequence similarity searches), UniParc (Sequence archive, used to keep track of sequences and their identifiers), and Supporting data (Literature citations, taxonomy, keywords and more).

**WELCOME**

The mission of UniProt is to provide the scientific community with a comprehensive, high-quality and freely accessible resource of protein sequence and functional information.

**What we provide**

UniProtKB	Protein knowledgebase, consists of two sections: ★ Swiss-Prot, which is manually annotated and reviewed. ★ TrEMBL, which is automatically annotated and is <b>not</b> reviewed. Includes <a href="#">Complete Proteome Sets</a> .
UniRef	Sequence clusters, used to speed up sequence similarity searches.
UniParc	Sequence archive, used to keep track of sequences and their identifiers.
Supporting data	<a href="#">Literature citations, taxonomy, keywords and more.</a>

**Getting started**


- [Text search](#)
- [Sequence similarity searches \(BLAST\)](#)

**NEWS**

**UniProt release 2011\_02 - Feb 8, 2011**  
 Automatic annotation of UniProtKB/TrEMBL using PDB-derived data ? Cross-references to nextProt

- › [Statistics for UniProtKB:](#)
- › [Swiss-Prot - TrEMBL](#)
- › [Forthcoming changes](#)
- › [News archives](#)

**SITE TOUR**



Learn how to make best use of the tools and data on this site.

**PROTEIN SPOTLIGHT**



- ▶ Both DNA strands may act as a template for the synthesis of other strand
- ▶ Template hypothesis was first expressed by Nikolaj Koltsov in 1927
- ▶ Replication is semiconservative



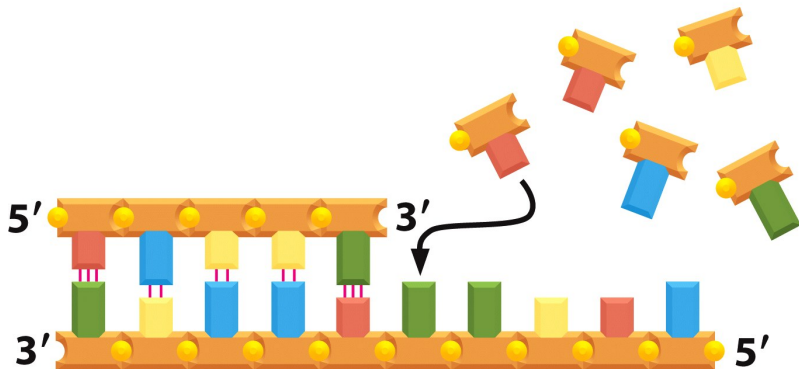


Figure 6-2 Essential Cell Biology 3/e (© Garland Science 2010)



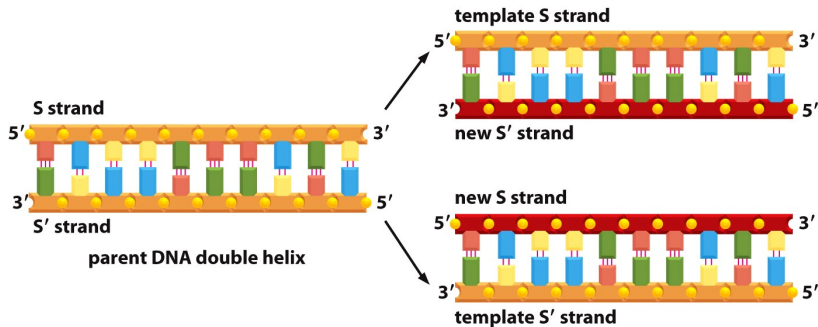


Figure 6-3 Essential Cell Biology 3/e (© Garland Science 2010)



- ▶ Theoretically, three variants of replications were possible: semiconservative, dispersive and conservative
- ▶ Experiment was based on two bacterial cultures which grew on different media: with normal nitrogen  $^{14}\text{N}$ , and with heavy nitrogen,  $^{15}\text{N}$
- ▶ After growing for 20' on heavy medium, bacteria produce DNA molecules with intermediate weight *only*
- ▶ That ruled out conservative hypothesis.



- ▶ Theoretically, three variants of replications were possible: semiconservative, dispersive and conservative
- ▶ Experiment was based on two bacterial cultures which grew on different media: with normal nitrogen  $^{14}\text{N}$ , and with heavy nitrogen,  $^{15}\text{N}$
- ▶ After growing for 20' on heavy medium, bacteria produce DNA molecules with intermediate weight *only*
- ▶ That ruled out conservative hypothesis. How to rule out dispersive hypothesis?—Explanation has been given on lecture.



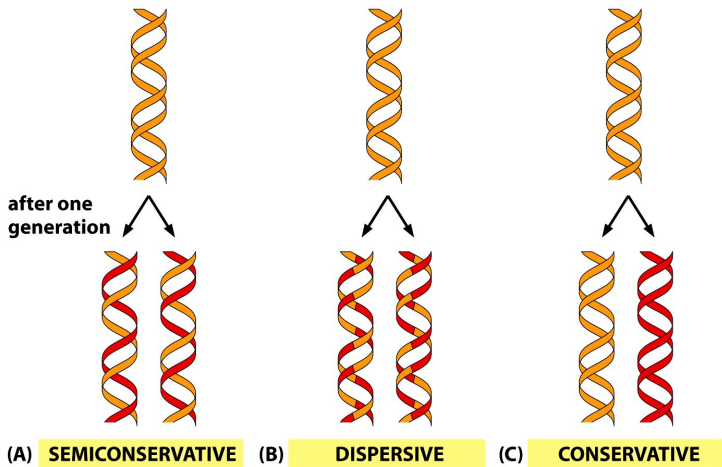


Figure 6-6 Essential Cell Biology 3/e (© Garland Science 2010)



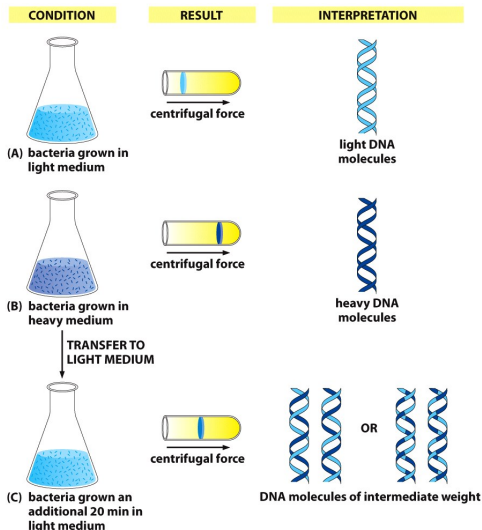


Figure 6-8 Essential Cell Biology 3/e (© Garland Science 2010)



- ▶ DNA double helix should open for replication: this is a replication origin place
- ▶ When replication starts, these openings will grow and form replication forks which are visible under microscope



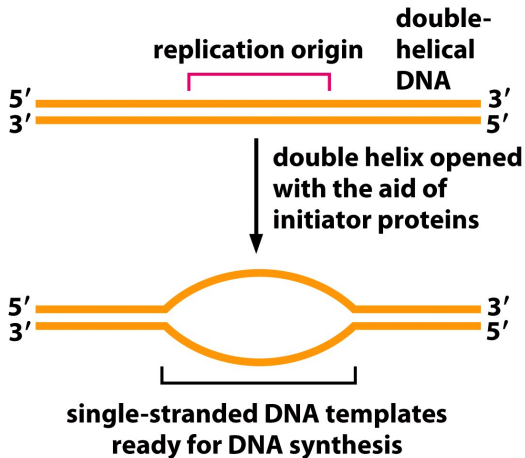


Figure 6-5 Essential Cell Biology 3/e (© Garland Science 2010)



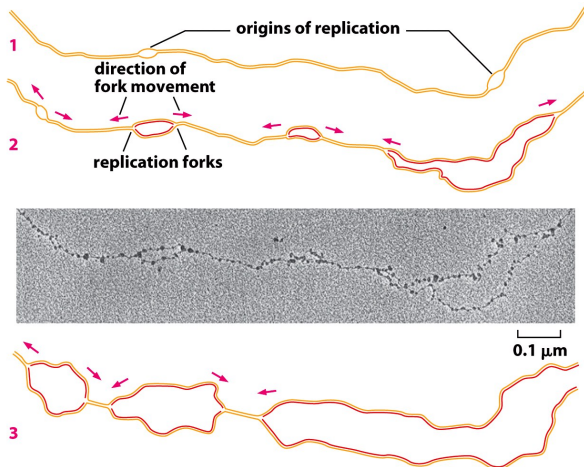


Figure 6-9 Essential Cell Biology 3/e (© Garland Science 2010)



- ▶ **DNA is synthesized only in 5'-to-3' direction.** Structure of DNA and complexity of replication do not allow the other direction.
- ▶ Therefore, fork is asymmetrical: one strand is replicating smoothly whereas other strand (lagging) replicated by "leaps", still 5'-to-3'
- ▶ Every "leap" produce one Okazaki fragment which are later joined



└ DNA replication

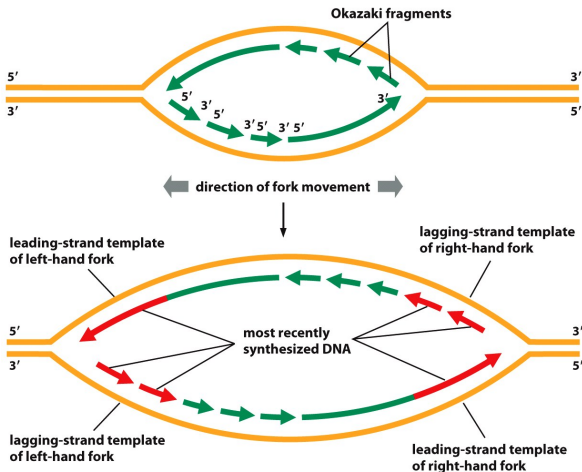


Figure 6-12 Essential Cell Biology 3/e (© Garland Science 2010)



- ▶ DNA polymerase is the main replication enzyme with error rate of  $\approx 0.00000001$
- ▶ Every time DNA polymerase adds new nucleotide, it checks if previous was correctly placed (if not, it removes the wrong nucleotide)
- ▶ Two different binding sites in DNA polymerase work for synthesis and proofreading
- ▶ Proofreading goes in opposite, 3'-to-5' direction



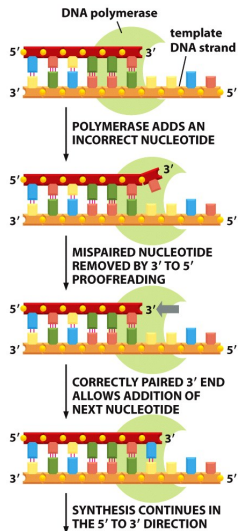


Figure 6-13 Essential Cell Biology 3/e (© Garland Science 2010)



## DNA polymerase movie



## RNA primers

- ▶ DNA polymerase cannot start nucleotide chain itself
- ▶ Instead, primase enzyme synthesizes small RNA primer ( $\approx 10$  nucleotides) which is used as a starting point
- ▶ Primase error rate is  $\approx 0.0001$  (!)
- ▶ In lagging strand, RNA sites are interleaved with DNA fragments
- ▶ Then RNA sites are removed, and DNA ligase joins fragments together



## RNA primers

- ▶ DNA polymerase cannot start nucleotide chain itself
- ▶ Instead, primase enzyme synthesizes small RNA primer ( $\approx 10$  nucleotides) which is used as a starting point
- ▶ Primase error rate is  $\approx 0.0001$  (!)
- ▶ In lagging strand, RNA sites are interleaved with DNA fragments
- ▶ Then RNA sites are removed, and DNA ligase joins fragments together
- ▶ Why primer is RNA?



## DNA replication

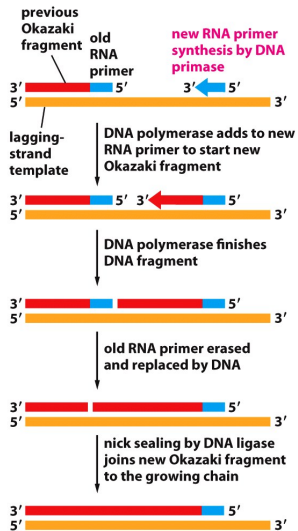


Figure 6-16 Essential Cell Biology 3/e (© Garland Science 2010)







## Final question (2 points)

Why cells use RNA as DNA replication primers?



- ▶ DNA replication is a semiconservative process
- ▶ DNA replication could go only in one direction
- ▶ Proofreading and RNA priming are helping in replication



## For Further Reading



A. Shipunov.

*Advanced Cell Biology* [Electronic resource].

2011—onwards.

Mode of access: [http:](http://)

[//ashipunov.info/shipunov/school/biol\\_250](http://ashipunov.info/shipunov/school/biol_250)



B. Alberts et al.

*Essential Cell Biology*. 3rd edition.

Garland Science, 2009.

*Chapter 6*, pages 198–208.