

# IV107 Bioinformatika I

## Přednáška 2

Katedra informačních technologií  
Masarykova Univerzita Brno

Podzim 2018



## Bioinformatika

- ▶ zpracování hromadných molekulárně–biologických dat
- ▶ posledních cca. 30 let
- ▶ data: genomika a proteomika
  - ▶ sekvence
  - ▶ struktury
  - ▶ interakce a jiné funkce
  - ▶ expresní data
- ▶ anotace cca 1000 prokaryotických genomů  
(<http://www.cbs.dtu.dk/services/GenomeAtlas/>)
- ▶ Galerie sekvenovaných genomů  
(<http://www.genomenewsnetwork.org/>)

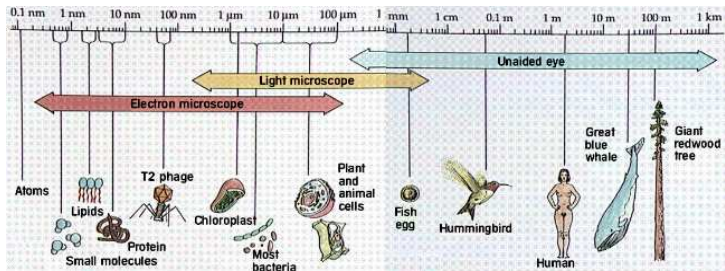


# DNA websites

- ▶ <http://www.dnafb.org/>
- ▶ <http://www.dnalc.org/>
- ▶ <http://www.dnai.org/>



# Hierarchie biologických struktur



# Outline

## Struktura DNA

Z historie

DNA je nosičem genetické informace

DNA má strukturu antiparalelní spirály

## Centrální dogma a genová exprese

### Replikace DNA

### Transkripce

### Translace

Genetický kód

## Struktura proteinů

Primární struktura (Sekvence)

Sekundární struktura

Terciární struktura (domény)

Kvartérní struktura (komplexy)

Funkce proteinů

Nástroje mol.biologie, genomiky a proteomiky

# Kořeny genetiky

Gregor J. Mendel

V pokusech s rostlinami si všiml, že potomství dvou rodičů nezávisí na jejich vzhledu /**fenotyp**/, nýbrž na jakýchsi symbolicky popsaných faktorech, které do značné míry odpovídají dnešnímu označení gen /**genotyp**/

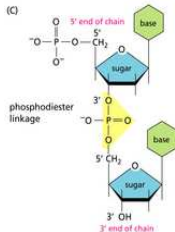
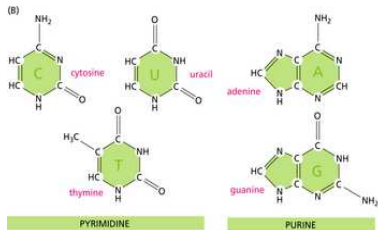
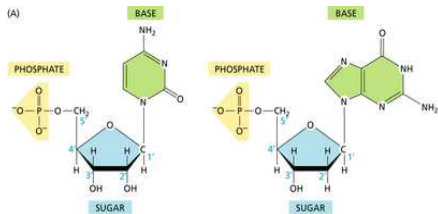


# Molekula DNA

- ▶ Objevena 1869 , považována za příliš jednoduchou
- ▶ Polymerická molekula, monomerem je:
  - ▶ deoxyribonukleotid /DNA/
  - ▶ ribonukleotid /RNA/
- ▶ Symbolické značení A,C,G,T,U
- ▶ Spirálovitá prostorová struktura



# Nukleotidy





# Zastoupení nukleotidů v různých organismech

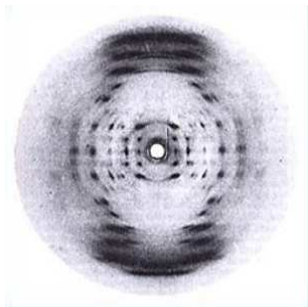
Organizmus	%A	%G	%C	%T	$\frac{A+G}{T+C}$	$\frac{A+T}{G+C}$
E.coli	24.7	26.0	25.7	23.6	1.03	0.93
S.lutea	13.4	37.1	37.1	12.4	1.04	0.35
S.cerevisae	31.3	18.7	17.1	32.9	1.09	1.79
H.sapiens	30.9	19.9	19.8	29.4	1.00	1.52

Dedukce: Nukleotidy se vyskytují v párech

From A.L.Lehninger (1970), Biochemistry, Worth Publishers, New York



# Difrakce na molekule DNA



Difrakce krystalické DNA z roku 1952 od Rosalind Franklin  
Svědčí o periodicitě 0.3nm a 3.4nm



# Struktura DNA



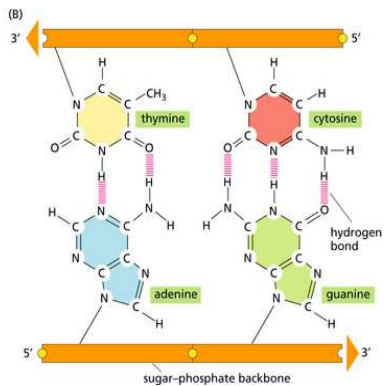
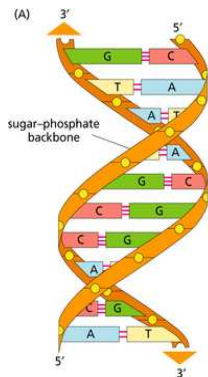
Tento obrázek a jedna stránka textu pomohli autorům k Nobelově ceně. Přínos Rosalind Franklin ve formě pečlivého zkoumání krystalů v té chvíli oceněn nebyl.



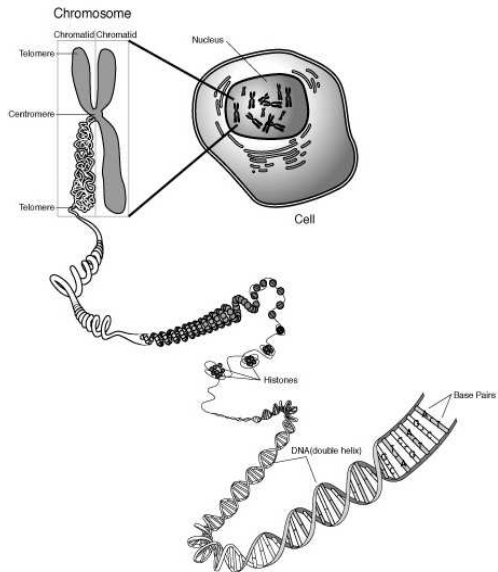
Poštovní známka k 50. výročí objevení struktury DNA



# Struktura DNA



# Struktura DNA



# Uspořádání DNA v jádře

Uspořádání	Počet nukleotidů	Relativní délka k jádru
Lineární	10000	600000
Plošné	25 mil	240
Prostorové	62.5 mld	0.1

Prostorové uspořádání DNA se 12000000 nukleotidy v jádře o rozměrech 0.0034 mm (rozměry jednoho páru nukleotidů dle modelu Watsona a Cricka jsou  $1.36 \times 0.34nm$ ).



# Outline

## Struktura DNA

Z historie

DNA je nosičem genetické informace

DNA má strukturu antiparalelní spirály

## Centrální dogma a genová exprese

### Replikace DNA

### Transkripce

### Translace

Genetický kód

## Struktura proteinů

Primární struktura (Sekvence)

Sekundární struktura

Terciární struktura (domény)

Kvartérní struktura (komplexy)

Funkce proteinů

Nástroje mol.biologie, genomiky a proteomiky

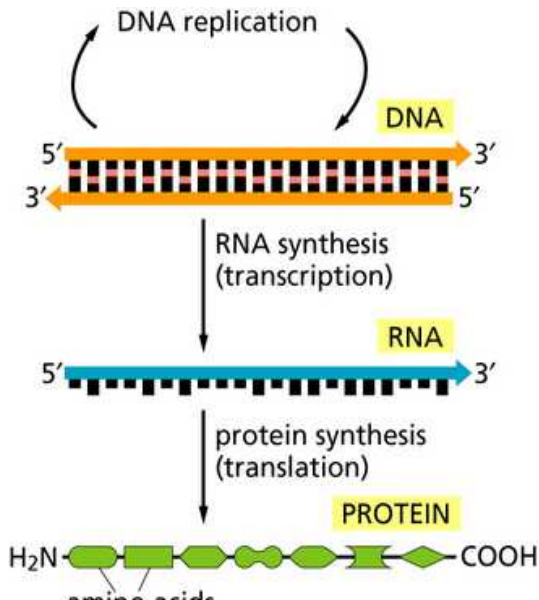
# Základní pravidla

- ▶ DNA → RNA → protein
  - ▶ retroviry: RNA → DNA
- ▶ jeden gen → jeden protein
  - ▶ RNA geny kódují RNA
  - ▶ alternativní sestřih umožňuje tvorbu několika proteinů
  - ▶ post-translační modifikace umožňují vznik různých forem jednoho typu proteinu





# Expresse genů



# Outline

## Struktura DNA

Z historie

DNA je nosičem genetické informace

DNA má strukturu antiparalelní spirály

## Centrální dogma a genová exprese

## Replikace DNA

## Transkripce

## Translace

Genetický kód

## Struktura proteinů

Primární struktura (Sekvence)

Sekundární struktura

Terciární struktura (domény)

Kvartérní struktura (komplexy)

Funkce proteinů

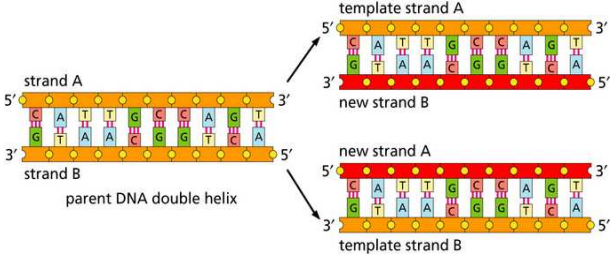
Nástroje mol.biologie, genomiky a proteomiky

# Replikace DNA

Hlavní enzymy	helikáza, DNA polymeráza
Templát	DNA
Substrát	deoxyribonukleotidy
Produkt	DNA



# Replikace DNA



# Outline

## Struktura DNA

Z historie

DNA je nosičem genetické informace

DNA má strukturu antiparalelní spirály

## Centrální dogma a genová exprese

## Replikace DNA

## Transkripce

## Translace

Genetický kód

## Struktura proteinů

Primární struktura (Sekvence)

Sekundární struktura

Terciární struktura (domény)

Kvartérní struktura (komplexy)

Funkce proteinů

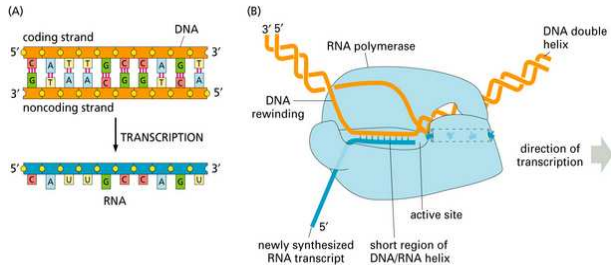
Nástroje mol.biologie, genomiky a proteomiky

# Transkripce

Hlavní enzymy	RNA polymeráza
Templát	DNA
Substrát	ribonukleotidy
Produkt	RNA



# Transkripce přenáší část genetické informace z DNA na mobilní mRNA



# Outline

## Struktura DNA

Z historie

DNA je nosičem genetické informace

DNA má strukturu antiparalelní spirály

## Centrální dogma a genová exprese

## Replikace DNA

## Transkripce

## Translace

Genetický kód

## Struktura proteinů

Primární struktura (Sekvence)

Sekundární struktura

Terciární struktura (domény)

Kvartérní struktura (komplexy)

Funkce proteinů

Nástroje mol.biologie, genomiky a proteomiky

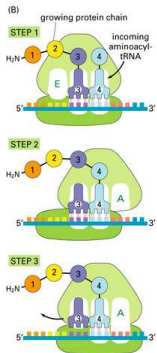
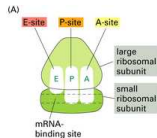


# Translace

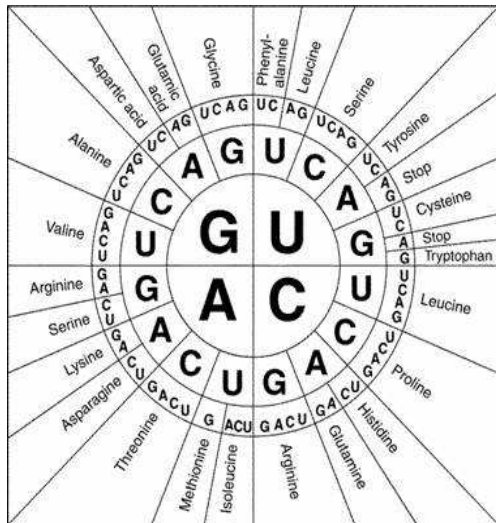
Hlavní enzymy	ribozom
Templát	RNA
Substrát	aminokyseliny
Produkt	protein (bílkovina, peptid)



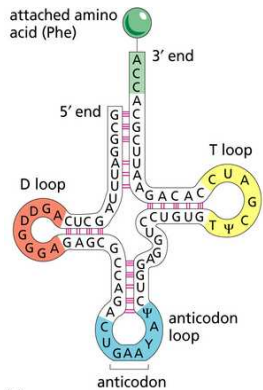
# Translace probíhá na ribozomech



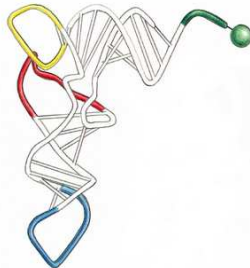
# Triplety nukleotidů kódují jednotlivé aminokyseliny



# Kód sprostředkovává tRNA



(A)



(B)

# Kontrolní otázky

1. DNA určitého organismu obsahuje 17% tymínu (T). Jaký je obsah cytozínu (C)?
2. Jaká je komplementární sekvence k sekvenci 5'-ACGT-3'?  
Co je na těchto sekvencích zajímavé? Jaké vlastnosti by mohl mít protein, který se váže na DNA s takovou sekvencí?



# Outline

## Struktura DNA

Z historie

DNA je nosičem genetické informace

DNA má strukturu antiparalelní spirály

## Centrální dogma a genová exprese

### Replikace DNA

### Transkripce

### Translace

Genetický kód

## Struktura proteinů

Primární struktura (Sekvence)

Sekundární struktura

Terciární struktura (domény)

Kvartérní struktura (komplexy)

Funkce proteinů

Nástroje mol.biologie, genomiky a proteomiky

# Struktura proteinů

Primární	sekvence aminokyselin
Sekundární	hlavně $\alpha$ -helix a $\beta$ -struktura
Terciární	3-D uspořádání
Domény	jedna nebo víc na protein
Kvartérní	komplexy skládající se z více podjednotek



# Skládání proteinů a druhý kód

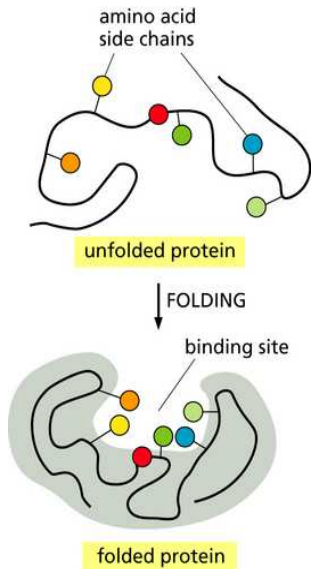
Mnoho čerstvě syntetizovaných proteinů se poskládá do své finální podoby bez pomoci dalších faktorů.

- ▶ demonstrováno na konkrétním proteinu (Anfinsen, 1966, ribonukleáza A)
- ▶ terciární struktura je zakódována v primární struktuře (kód dodnes neobjeven!)

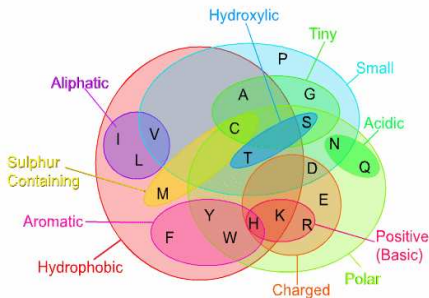




# Protein folding



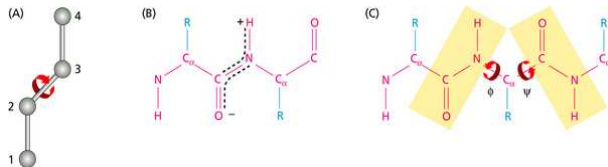
# Fyzikálně–chemické vlastnosti aminokyselin



## Amino Acids

- A** alanine (ala)
- R** arginine (arg)
- N** asparagine (asn)
- D** aspartic acid (asp)
- C** cysteine (cys)
- Q** glutamine (gln)
- E** glutamic acid (glu)
- G** glycine (gly)
- H** histidine (his)
- I** isoleucine (ile)
- L** leucine (leu)
- K** lysine (lys)
- M** methionine (met)
- F** phenylalanine (phe)
- P** proline (pro)
- S** serine (ser)
- T** threonine (thr)
- W** tryptophan (trp)
- Y** tyrosine (tyr)

# Aminokyseliny mohou být vůči sobě různě orientovány

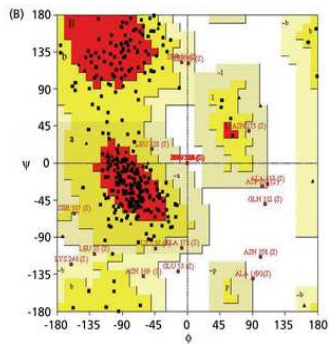
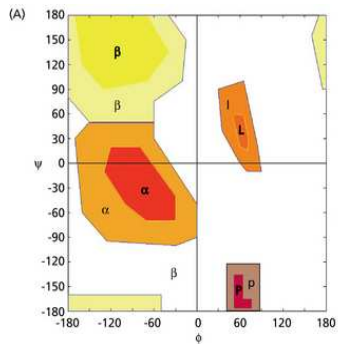


Torzní úhly:  $\phi$ ,  $\psi$ ,  $\omega$

cis ( $\omega = 0^\circ$ )

trans ( $\omega = 180^\circ$ )

# Ramachandran diagram

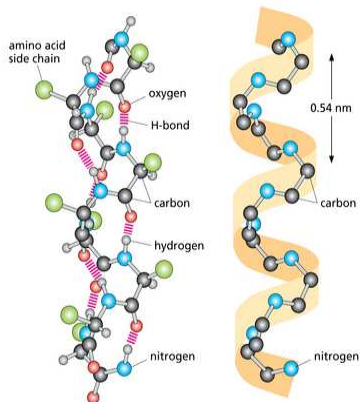


# $\alpha$ -helix

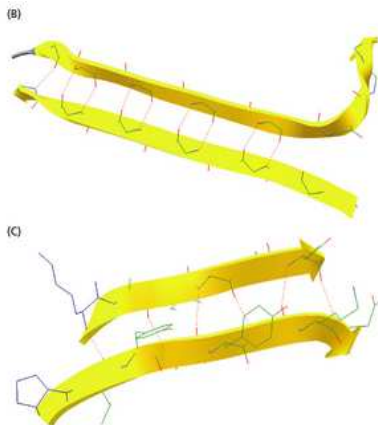
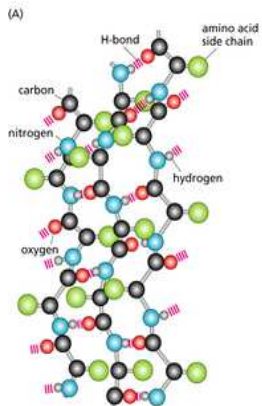
(A)



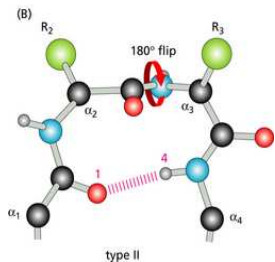
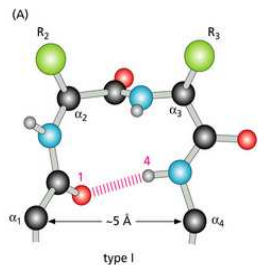
(B)



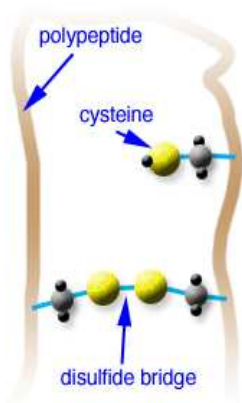
# $\beta$ -skládany list



# Otočky nebo smyčky (angl. turns)



# Disulfidický (cysteinový) mostík

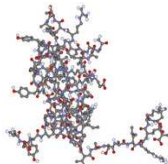




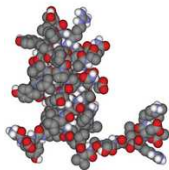
# Zobrazení 3-D struktury proteinu



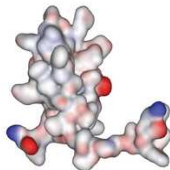
wire-frame



ball and stick



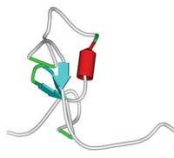
space-filling



surface

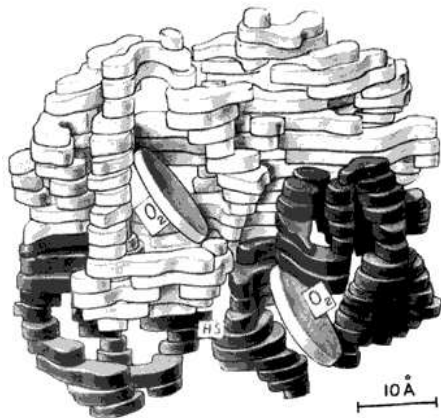


C $\alpha$  representation



$\alpha/\beta$  schematic

# Kvarterní struktura hemoglobinu

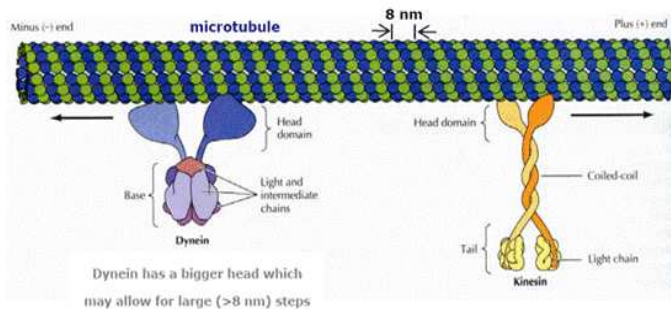


# Dílčí funkce proteinů

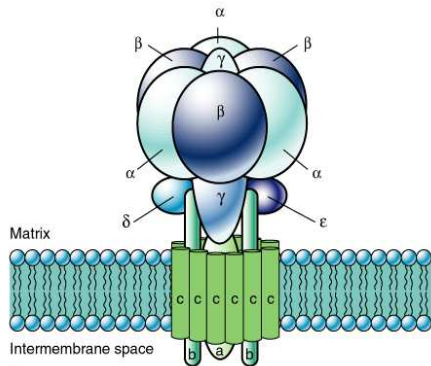
- ▶ Enzymy
  - ▶ katalyzátory snižují energetický práh reakcí
  - ▶ substrát se proměňuje v produkt
  - ▶ aktivní místo
- ▶ Interakce protein-protein
- ▶ Interakce protein-DNA
- ▶ Interakce protein-ligand
- ▶ Transdukce signálu, regulace
- ▶ Strukturální proteiny (vlákna, glykoproteiny)
- ▶ Transportní
- ▶ Póry a kanály
- ▶ Pumpy
- ▶ Motory



# Motory



# ATPáza

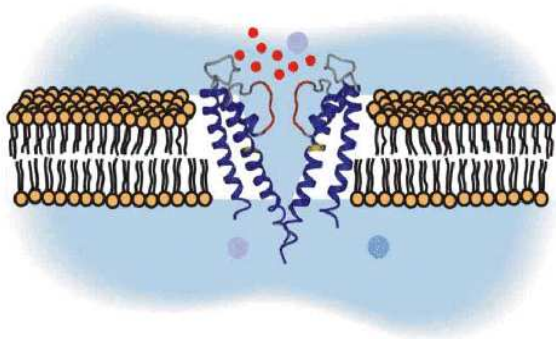


(b)

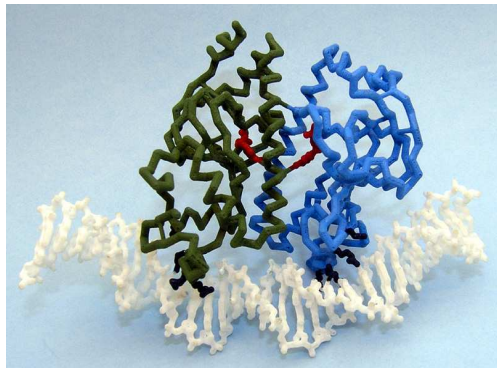
Copyright 1999 John Wiley and Sons, Inc. All rights reserved.



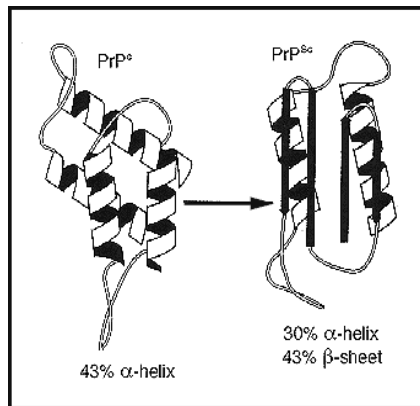
# Kanál umožňuje iónům proniknout membránou



# Receptor cyklického AMP (dimer a DNA)



# Prion ve formě $PrP^c$ a $PrP^{Sc}$





Nástroje mol.biologie, genomiky a proteomiky



# Outline

Příloha



# For Further Reading

## Projekt SIMAP

<http://www.czechnationalteam.cz/view.php?nazevclanku=simap&cisloclanku=2007020002>

[http://www.rozhlas.cz/leonardo/priroda/\\_zprava/321214](http://www.rozhlas.cz/leonardo/priroda/_zprava/321214)

