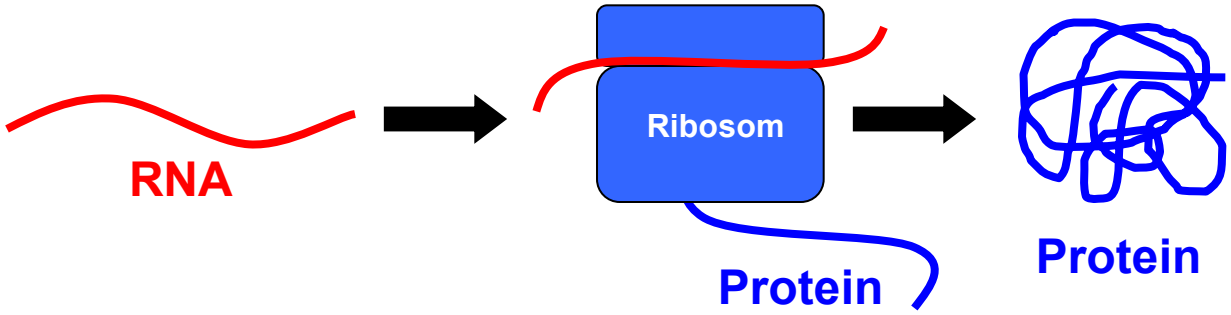
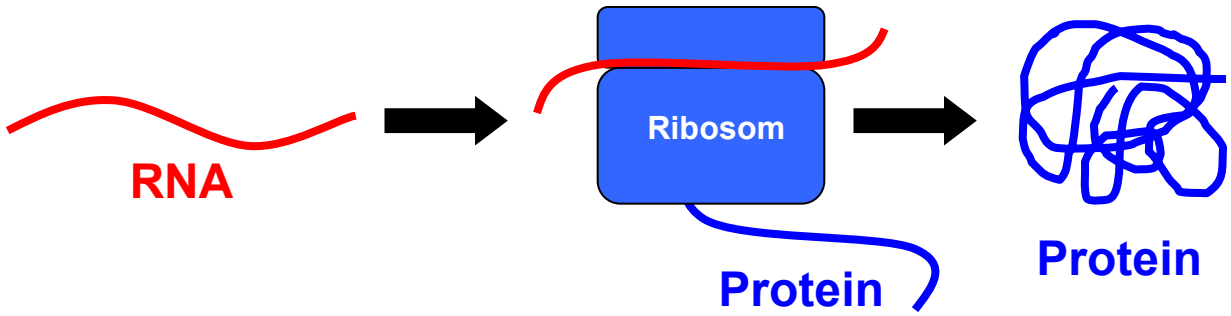


Základní vlastnosti proteinů

Teoretický úvod

Aplikovaná bioinformatika, Jaro 2017

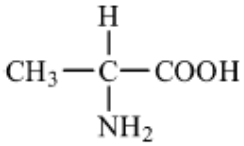
Proteiny

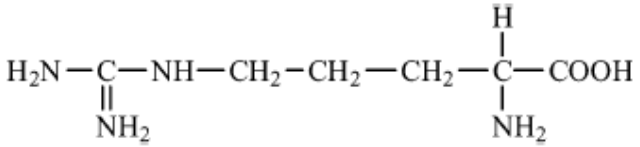


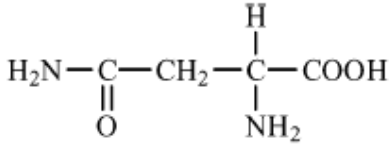
- **Protein, polypeptid, bílkovina.**
- **Lineární polymer aminokyselin spojených peptidovými vazbami.**
- **Funkce: katalytická, regulační, transportní, zprostředkování pohybu, obranná, strukturální, zásobní.**

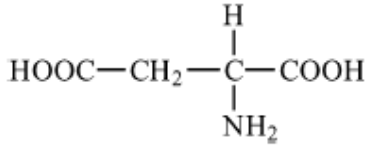
Proteinogenní aminokyseliny

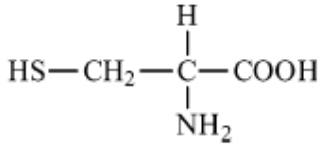
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Cysteine	Cys	C
Glutamic acid	Glu	E

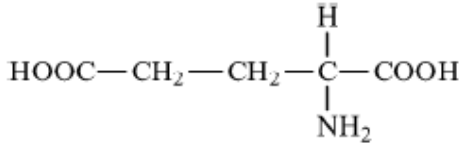






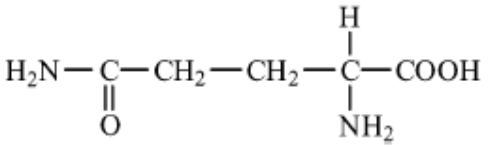


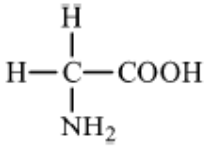


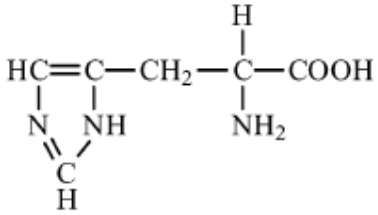


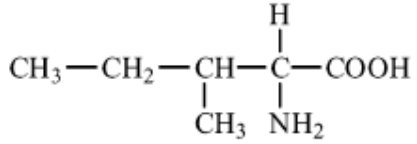
Proteinogenní aminokyseliny

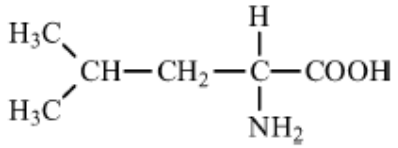
Glutamine	Gln	Q
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L





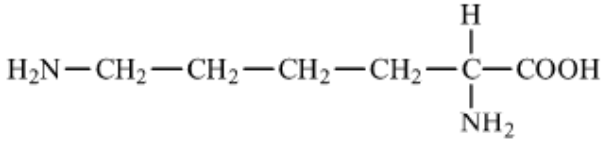


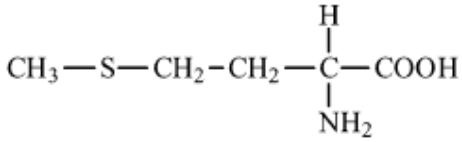


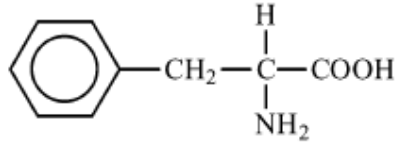


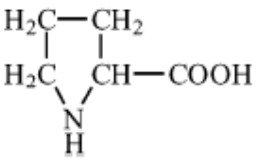
Proteinogenní aminokyseliny

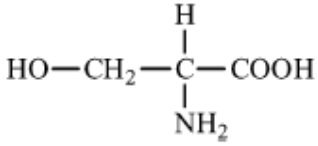
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T

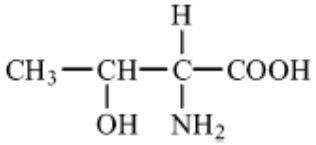






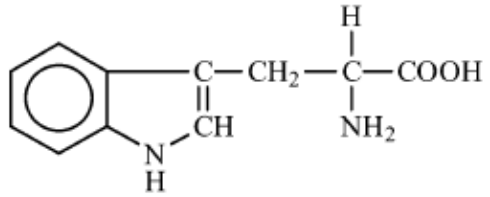


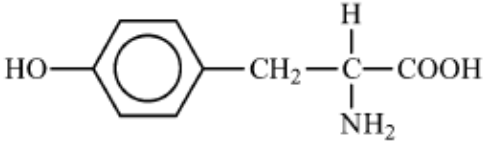


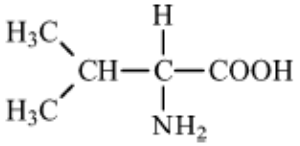


Proteinogenní aminokyseliny

Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V







21. aminokyselina – selenocystein, Sec, U

22. aminokyselina – pyrrolysin, Pyl, O

Part 2, Section B: THE ONE-LETTER SYSTEM (revision and updating of [11])

3AA-20. THE NEED FOR A CONCISE REPRESENTATION OF SEQUENCE

3AA-20.1. General Considerations Regarding the One-Letter System

There are difficulties in using the three-letter system (3AA-14 to 3AA-19) in presenting long protein sequences. A one-letter code is much more concise, and is helpful in summarizing large amounts of data, in aligning and comparing homologous sequences, and in computer techniques for these processes. It may also be used to label residues in three-dimensional pictures of protein molecules.

The possibility of using one-letter symbols was mentioned by Gamow & Ycas [26] in 1958. Sorm *et al.* [27] systematized the idea in 1961 (see, for example, [28]), and Dayhoff and Eck used one-letter symbols derived partly from the code of Sorm *et al.* in their compilations of protein sequences ([29], latest edition [30]). IUB-IUPAC recommendations [11] were approved in 1968 on the basis of proposals of a subcommittee of W. E. Cohn, M. O. Dayhoff, R. V. Eck, and B. Keil, and these recommendations are given here with no substantial change.

```
PLLSASIVSAPVVTSETYVDIPGLYLDVAKAGIRDGKLQVILNVPTPY
ATGNNFPGIYFAIATNQGVVADGCFTYSSKVPESTGRMPFTLVATIDV
GSGVTFVKGQWKSVRGSAMHIDSYASLSAIWGTAAPSSQGSGNQGAET
GGTGAGNIGGGGERDGTFNLPPHIKFGVTALTHAANDQTIDIYIDDDP
KPAATFKGAGAQDQNLGTKVLDSGNGRVRVIVMANGRPSRLGSRQVDI
FKKSYFGIIGSEDGADDDYNDGIVFLNWPLG
```

http://www.chem.qmul.ac.uk/iupac/AminoAcid/index.html

Part 2, Section B: THE ONE-LETTER SYSTEM (revision and updating of [11])

3AA-20. THE NEED FOR A CONCISE REPRESENTATION OF SEQUENCE

3AA-20.1. General Considerations Regarding the One-Letter System

There are difficulties in using the three-letter system (3AA-14 to 3AA-19) in presenting long protein sequences. A one-letter code is much more concise, and is helpful in summarizing large amounts of data, in aligning and comparing homologous sequences, and in computer techniques for these processes. It may also be used to label residues in three-dimensional pictures of protein molecules.

The possibility of using one-letter symbols was mentioned by Gamow & Ycas [26] in 1958. Sorm *et al.* [27] systematized the idea in 1961 (see, for example, [28]), and Dayhoff and Eck used one-letter symbols derived partly from the code of Sorm *et al.* in their compilations of protein sequences ([29], latest edition [30]). IUB-IUPAC recommendations [11] were approved in 1968 on the basis of proposals of a subcommittee of W. E. Cohn, M. O. Dayhoff, R. V. Eck, and B. Keil, and these recommendations are given here with no substantial change.

```
BclB                -------ERDGIFTLPPNIAFGVTALVNSSAPQTIEVFVDDNPKPAATFQGAGTQDANLNTQIVNSGK-G  62
BclC                -------ERDGTFNLPPHIKFGVTALTHAANDQTIDIYIDDDPKPAATFKGAGAQDQNLGTKVLDSGN-G  62
PA-IIL              -------ATQGVFTLPANTRFGVTAFANSSGTQTVNVLVNN--ETAATFSGQSTNNAVIGTQVLNSGSSG  61
BclA                ADSQTSSNRAGEFSIPPNTDFRAIFFANAAEQQHIKLFIGDSQEPAAYHKLTTRDGPREAT--LNSGN-G  67
BclD                -------DRNGNFSLPPNTAFKAIFYANAADRQDLKLFIDDAPEPAATFVGNSEDGVRLFT--LNSKG-G  60
Clustal Consensus             * *.:*.:  * .   .:::  * :.: :.:  :.** .     :.    *  ::*   *

BclB                KVRVVVTANGKPSKIGSRQVDIFK--------KTYFGLVGSEDGGDGDYNDGIAILNWPLG 115
BclC                RVRVIVMANGRPSRLGSRQVDIFK--------KSYFGIIGSEDGADDDYNDGIVFLNWPLG 115
PA-IIL              KVQVQVSVNGRPSDLVSAQVILTN--------ELNFALVGSEDGTDNDYNDAVVVINWPLG 114
BclA                KIRFEVSVNGKPSATDARLAPINGKKSDGSPFTVNFGIVVSEDGHDSDYNDGIVVLQWPIG 128
BclD                KIRIEASANGRQSATDARLAPL----SAGD--TVWLGWLGAEDGADADYNDGIVILQWPIT 115
Clustal Consensus   :::. . .**: *   :  . :              :. : :*** * ****.:..::**:  49
```

One-letter symbol	Three-letter symbol	Amino acid
A	Ala	alanine
B	Asx	aspartic acid or asparagine
C	Cys	cysteine
D	Asp	aspartic acid
E	Glu	glutamic acid
F	Phe	phenylalanine
G	Gly	glycine
H	His	histidine
I	Ile	isoleucine
K	Lys	lysine
L	Leu	leucine
M	Met	methionine
N	Asn	asparagine
P	Pro	proline
Q	Gln	glutamine
R	Arg	arginine
S	Ser	serine
T	Thr	threonine
V	Val	valine
W	Trp	tryptophan
X **	Xaa	unknown or 'other' amino acid
Y	Tyr	tyrosine
Z	Glx	glutamic acid or glutamine (or substances such as 4-carboxyglutamic acid and 5-oxoproline that yield glutamic acid on acid hydrolysis of peptides)

Note on the Choice of Symbols

Initial letters of the names of the amino acids were chosen where there was no ambiguity. There are six such cases: cysteine, histidine. isoleucine, methionine, serine and valine. All the other amino acids share the initial letters A, G, L, P or T, so arbitrary assignments were made. These letters were assigned to the most frequently occurring and structurally most simple of the amino acids with these initials, alanine (A), glycine (G), leucine (L), proline (P) and threonine (T).

Other assignments were made on the basis of associations that might be helpful in remembering the code, e.g. the phonetic associations of F for phenylalanine and R for arginine. For tryptophan the double ring of the molecule is associated with the bulky letter W. The letters N and Q were assigned to asparagine and glutamine respectively; D and E to aspartic and glutamic acids respectively. K and Y were chosen for the two remaining amino acids, lysine and tyrosine, because, of the few remaining letters, they were close alphabetically to the initial letters of the names. U and O were avoided because U is easily confused with V in handwritten material, and O with G, Q, C and D in imperfect computer print-outs, and also with zero. J was avoided because it is absent from several languages.

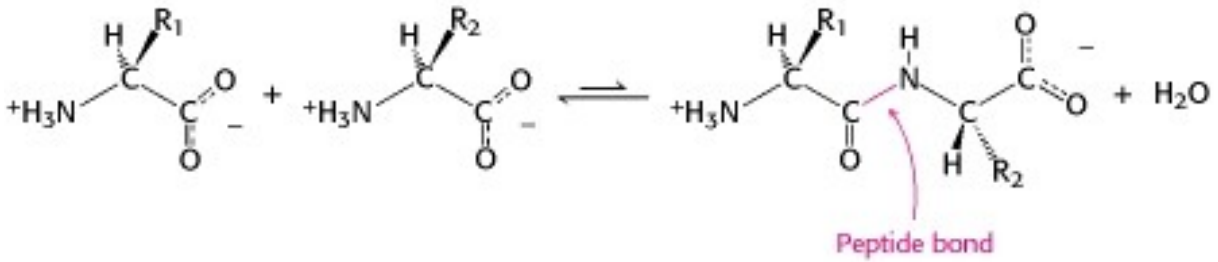
Two other symbols are often necessary in partly determined sequences, so B was assigned to aspartic acid or asparagine when these have not been distinguished; Z was similarly assigned to glutamic acid or glutamine. X means that the identity of an amino acid is undetermined, or that the amino acid is atypical. See the Addendum for an alternative use of X.

21. aminokyselina – selenocystein, Sec, U

22. aminokyselina – pyrrolysin, Pyl, O

http://www.chem.qmul.ac.uk/iupac/AminoAcid/index.html

Peptidová vazba



NH_2—COOH + NH_2—COOH

N-konec NH_2—COOH C-konec

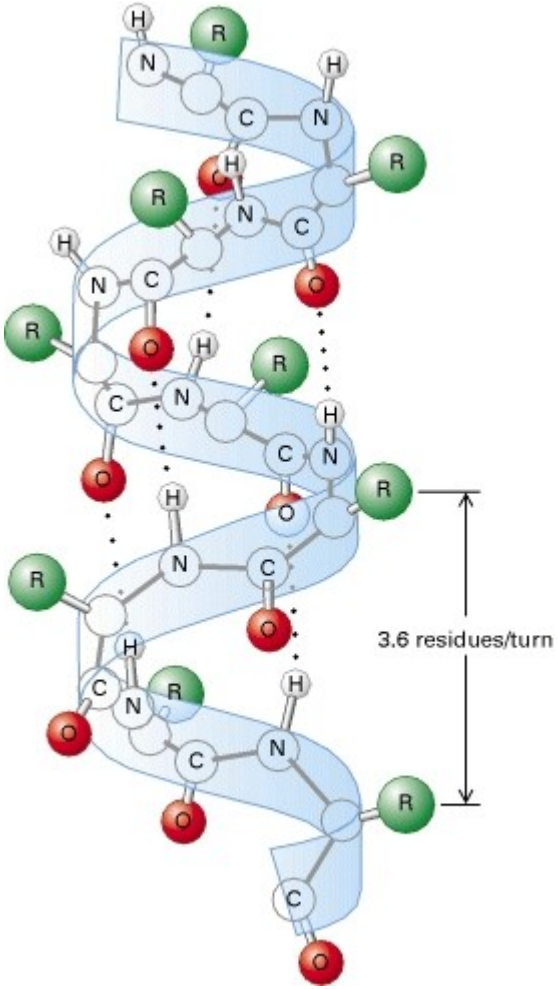
N MTWCKTMIDQGRSWPHCYYGMAA
DTYYKKLTPGHTQVGITILMGAC
GCCCGTGCRNMSDETGCWWCGTA
HSPGCTDEQLRCGLVCGT C

N → C

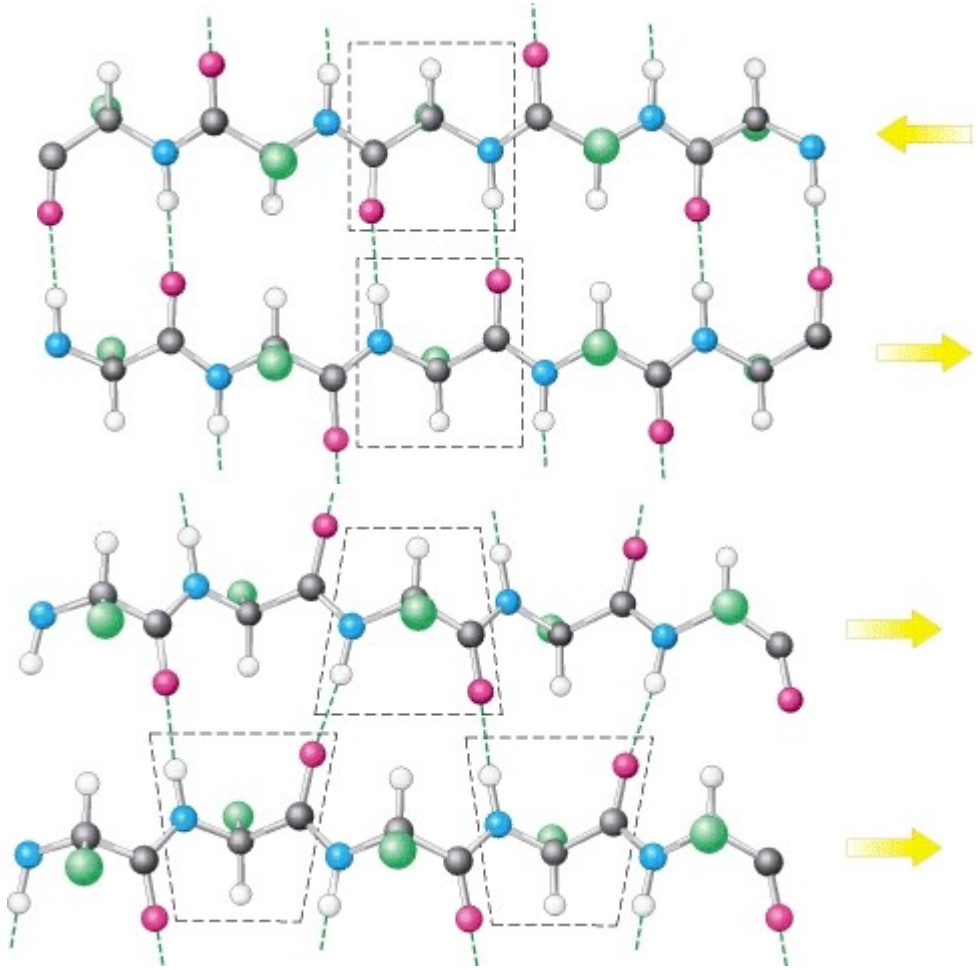
Proteinogenní aminokyseliny

- Stavební jednotky proteinů: α-L-aminokyseliny.
- 20 standardních proteinogenních aminokyselin.
- Alifatické (Gly, Ala, Val, Leu, Ile).
- Sirné (Cys, Met).
- S OH skupinou (Ser, Thr).
- Kyselé a z nich odvozené (Glu, Gln, Asp, Asn).
- Bazické (Lys, Arg).

Sekundární struktura

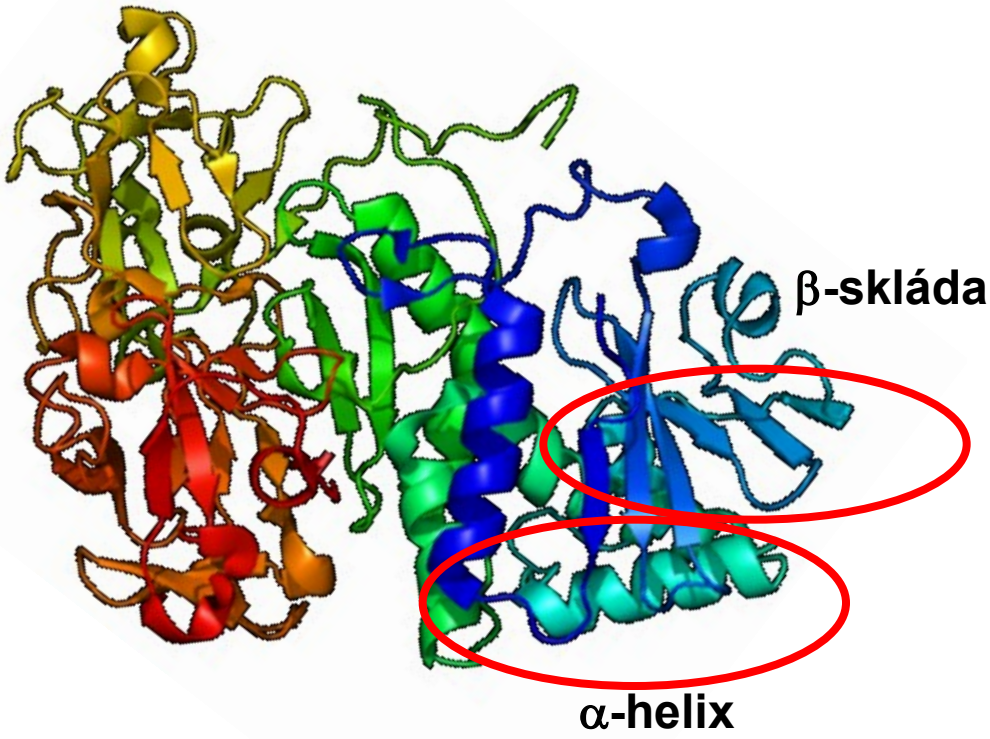


α-helix



β-skládaný list

Terciární struktura



Jak se chovají proteiny? (v laboratoři)

- **Limitované množství proteinu (cena, dostupnost).**
- ***In vitro* mohou rychle ztrácet aktivitu (nutná správná sekundární, terciární a někdy i kvarterní struktura).**
- **Mohou být nestabilní (některé velmi nestabilní) mimo své optimální prostředí v buňce (organismu).**
- **K výrazné destabilizaci a denaturaci může docházet již za laboratorní teploty (25 C).**

…zlobí.

Práce s proteiny

- **Většina savčích proteinů začíná denaturovat již při teplotách nad 40 C. Při teplotě 95 C dochází k úplné denaturaci téměř všech proteinů během několika minut. K výrazné destabilizaci a denaturaci může docházet již za laboratorní teploty (25 C).**

S proteiny pracujeme „na ledu“.

Práce s proteiny

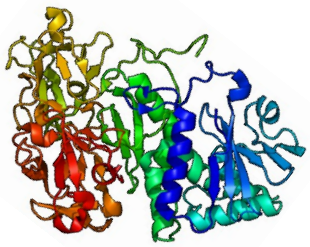
- **Proteiny jsou štěpeny proteasami a peptidasami. Optimum těchto enzymů je 37 C, za nižší teploty se jejich aktivita snižuje (ale jsou aktivní i při 4 C). Proteasy se do vzorku dostanou neopatrnou manipulací a jsou také produkovány mikroorganismy.**

Minimalizace kontaminace vzorku (ochranné pomůcky) a použití inhibitorů proteas.

Práce s proteiny

- **Všechny vzorky jsou dříve či později kontaminovány bakteriemi. Mikroorganismy si na proteinech pochutnají i při 4 C a produkují proteasy.**

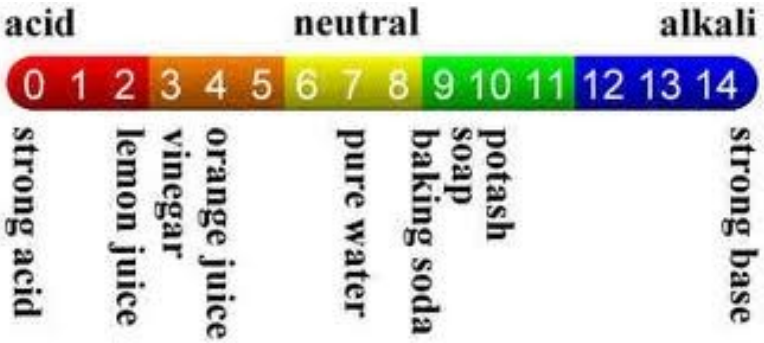
Přídavek antibakteriálních látek (0,02-0,05 % azid sodný)



Práce s proteiny

- **Proteiny jsou aktivní (a stabilní) v určitém rozmezí pH. A to může být pro některé proteiny velmi úzké… Fyziologické pH pro většinu proteinů je cca 7,2-7,4. Silně kyselé nebo zásadité prostředí proteiny denaturuje.**

Nutné kontrolované prostředí – pufry o vhodném pH.



Práce s proteiny

- **Proteiny jsou aktivní (a stabilní) v určitém rozmezí pH. A to může být pro některé proteiny velmi úzké… Fyziologické pH pro většinu proteinů je cca 7,2-7,4. Silně kyselé nebo zásadité prostředí proteiny denaturuje.**

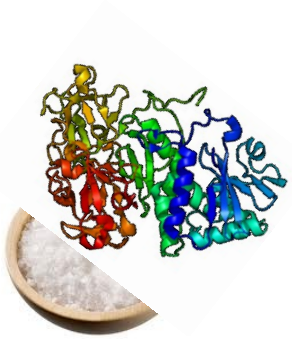
Nutné kontrolované prostředí – pufry o vhodném pH.

- **Pufr, tlumivý roztok, ústojný roztok, ústoj:**
  látka (směs látek) schopná udržovat stabilní pH po přídavku silné kyseliny nebo zásady do systému.
  Příklad: slabá kyselina/její sůl, HA/A^-.
  „Přirodní“ x syntetické pufry. Pufry nesmí interagovat s proteiny nebo interferovat s jejich funkcí!

Práce s proteiny

- **Proteiny vyžadují pro svou aktivitu (a stabilitu) určitou koncentraci solí (iontů). Vysoká i nízká koncentrace solí může způsobovat agregaci a precipitaci. Proteiny většinou nejsou stabilní v čisté vodě.**

Nutná optimalizovaná koncentrace soli v roztoku.



Práce s proteiny

- **Při práci s nízkými koncentracemi proteinů (< 1 mg/ml) se může výrazně projevit ztráta způsobená vazbou na stěny použité nádobky (zkumavky).**

Pokud je to možné, lze použít inertní proteiny (BSA, cca 2 mg/ml), které vazbě zabrání.

Práce s proteiny

Protein LoBind Tubes

Při použití těchto zkumavek dochází k minimálním ztrátám proteinů. V těchto zkumavkách nedochází k redukci aktivity enzymů ani k denaturačním efektům způsobeným vazbou na povrch běžných zkumavek. Příště již žádné trápení se saturací povrchu BSA nebo silikonizací běžných plastových zkumavek pro práci s malými množstvími proteinů. Zkumavky jsou vyrobeny speciální technologií z nejčistšího polypropylenu.

Specifikace:

- minimální ztráty proteinů (méně než 3%, BSA 1 µg/ml)
- vnitřní povrch zkumavek nemá žádný potah (např. silikonem)
- bez DNA, DNas, RNas a PCR inhybitorů (PCR clean)
- vyráběné jako Safe-Lock zkumavky ve velikostech 0,5 ml, 1,5 ml a 2 ml
- vhodné i pro centrifugaci při vysokých otáčkách

Práce s proteiny

- **Proteiny mohou být rovněž poškozeny mechanicky při příliš energickém míchaní nebo třepání!**

Nutná opatrná a jemná manipulace s proteiny.

Vortex

Míchačka

Skladování proteinů

	Storage Condition			
Characteristic	Solution at 4°C	Solution in 25-50% glycerol or ethylene glycol at -20°C	Frozen at -20° to -80°C or in liquid nitrogen	Lyophilized (usually also frozen)
Typical shelf life	1 month	1 year	Years	Years
Requires sterile conditions or addition of antibacterial agent	Yes	Usually	No	No
Number of times a sample may be removed for use	Many	Many	Once; repeated freeze-thaw cycles generally degrade proteins	Once; it is impractical to lyophilize a sample multiple times

Skladování proteinů

- **Lyofilizace – mrazová sublimace.**
- **Odstranění vody ze zmraženého vzorku za sníženého tlaku.**
- **Nedostatek vody zabraňuje růstu mikroorganismů a inhibuje enzymy (proteasy).**
- **Nepoškozuje vzorek v takovém rozsahu jako jiné způsoby dehydratace (vysoká teplota, vysoušedla).**
- **Jednoduchá rehydratace.**

Lyophilized (usually also frozen)
Years
No
Once; it is impractical to lyophilize a sample multiple times

Skladování proteinů

Characteristic	Storage Condition			
	Solution at 4°C	Solution in 25-50% glycerol or ethylene glycol at -20°C	Frozen at -20° to -80°C or in liquid nitrogen	Lyophilized (usually also frozen)
Typical shelf life	1 month	1 year	Years	Years
Requires sterile conditions or addition of antibacterial agent	Yes	Usually	No	No
Number of times a sample may be removed for use	Many	Many	Once; repeated freeze-thaw cycles generally degrade proteins	Once; it is impractical to lyophilize a sample multiple times

Sterilní zkumavky, sterilizace filtrací.
Inhibitory proteas.

www.piercenet.com

Skladování proteinů

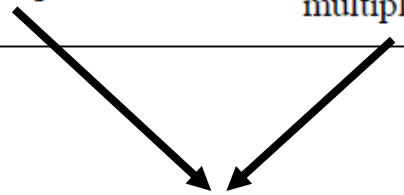
Kryoprotektanty zabraňují tvorbě krystalků ledu a poškození proteinu.

	Storage Condition			
Characteristic	Solution at 4°C	Solution in 25-50% glycerol or ethylene glycol at -20°C	Frozen at -20° to -80°C or in liquid nitrogen	Lyophilized (usually also frozen)
Typical shelf life	1 month	1 year	Years	Years
Requires sterile conditions or addition of antibacterial agent	Yes	Usually	No	No
Number of times a sample may be removed for use	Many	Many	Once; repeated freeze-thaw cycles generally degrade proteins	Once; it is impractical to lyophilize a sample multiple times

www.piercenet.com

Skladování proteinů

	Storage Condition			
Characteristic	Solution at 4°C	Solution in 25-50% glycerol or ethylene glycol at -20°C	Frozen at -20° to -80°C or in liquid nitrogen	Lyophilized (usually also frozen)
Typical shelf life	1 month	1 year	Years	Years
Requires sterile conditions or addition of antibacterial agent	Yes	Usually	No	No
Number of times a sample may be removed for use	Many	Many	Once; repeated freeze-thaw cycles generally degrade proteins	Once; it is impractical to lyophilize a sample multiple times



Nutné připravit několik alikvotů (částí zásobního roztoku) proteinu.

Skladování proteinů

	Storage Condition			
Characteristic	Solution at 4°C	Solution in 25-50% glycerol or ethylene glycol at -20°C	Frozen at -20° to -80°C or in liquid nitrogen	Lyophilized (usually also frozen)
Typical shelf life	1 month	1 year	Years	Years
Requires sterile conditions or addition of antibacterial agent	Yes	Usually	No	No
Number of times a sample may be removed for use	Many	Many	Once; repeated freeze-thaw cycles generally degrade proteins	Once; it is impractical to lyophilize a sample multiple times

Proteiny mohou být lyofilizací nebo zamražením nevratně poškozeny!

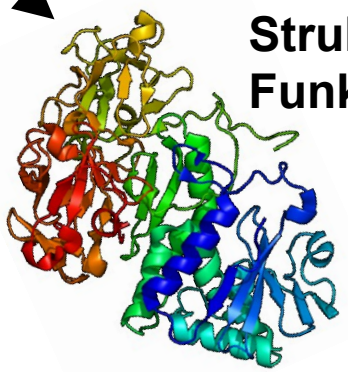
www.piercenet.com

Práce s proteiny

Protein je správně sbalený, aktivní, v dostatečném množství a koncentraci

Optimální podmínky

Gen → Protein



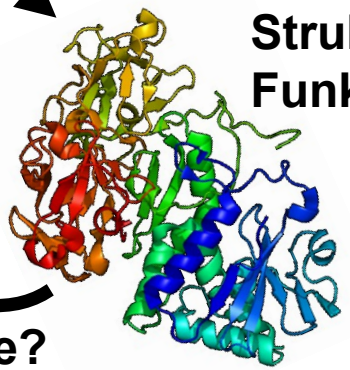
Struktura
Funkce

Práce s proteiny

Protein je správně sbalený, aktivní, v dostatečném množství a koncentraci

Optimální podmínky

Gen → Protein

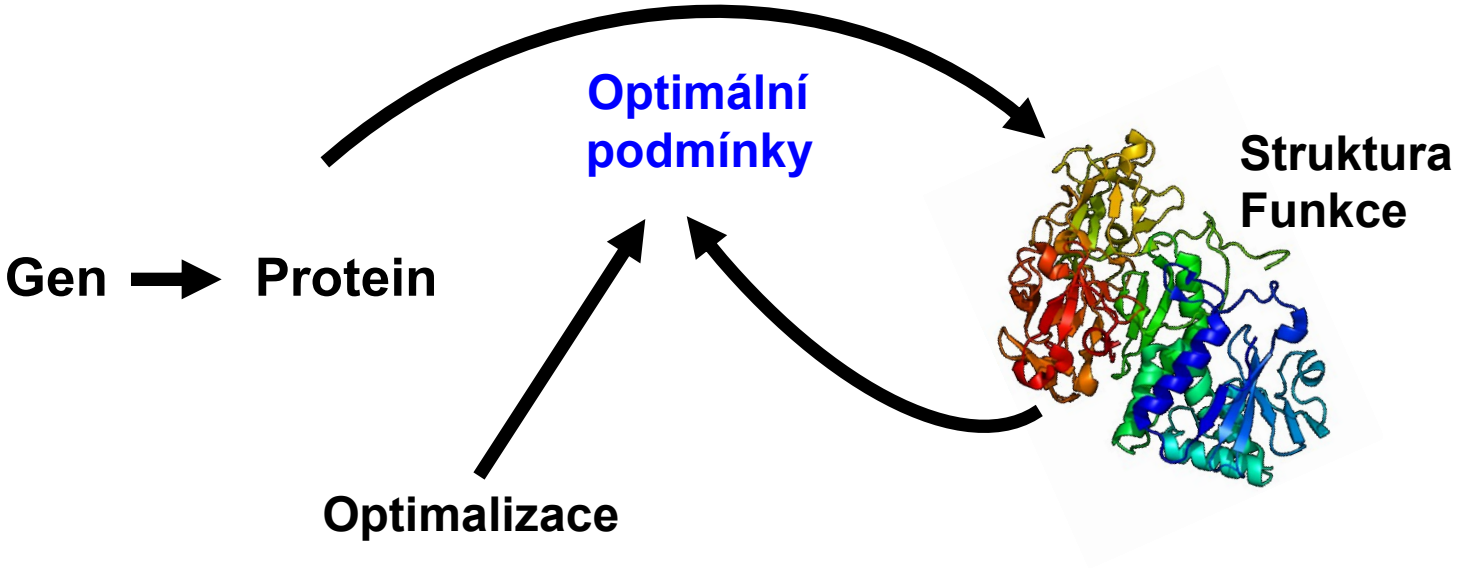


Struktura
Funkce

Hydrofobní interakce?
Sirné (disulfidové) můstky?
Oligomerizace?
Nutné kofaktory?
Přirozené prostředí (kompartment) v buňce?

Práce s proteiny

Protein je správně sbalený, aktivní, v dostatečném množství a koncentraci

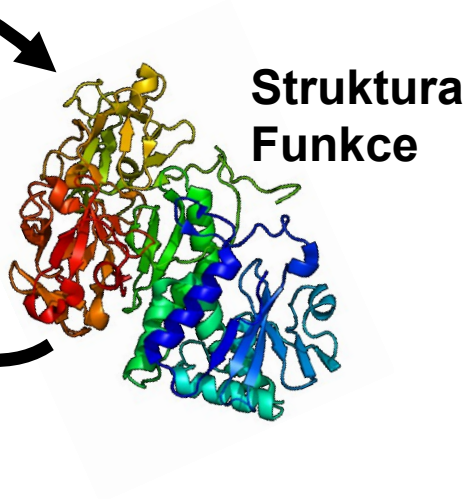


Práce s proteiny

Protein je správně sbalený, aktivní, v dostatečném množství a koncentraci

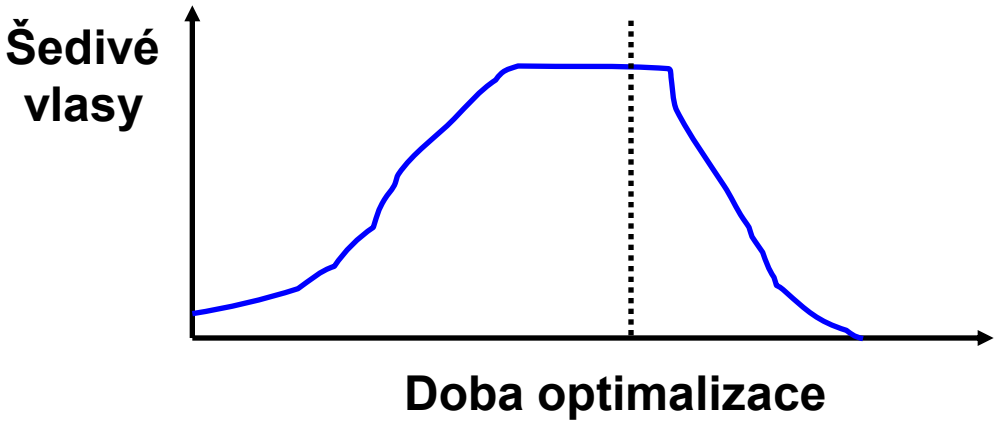
Optimální podmínky

Gen → Protein



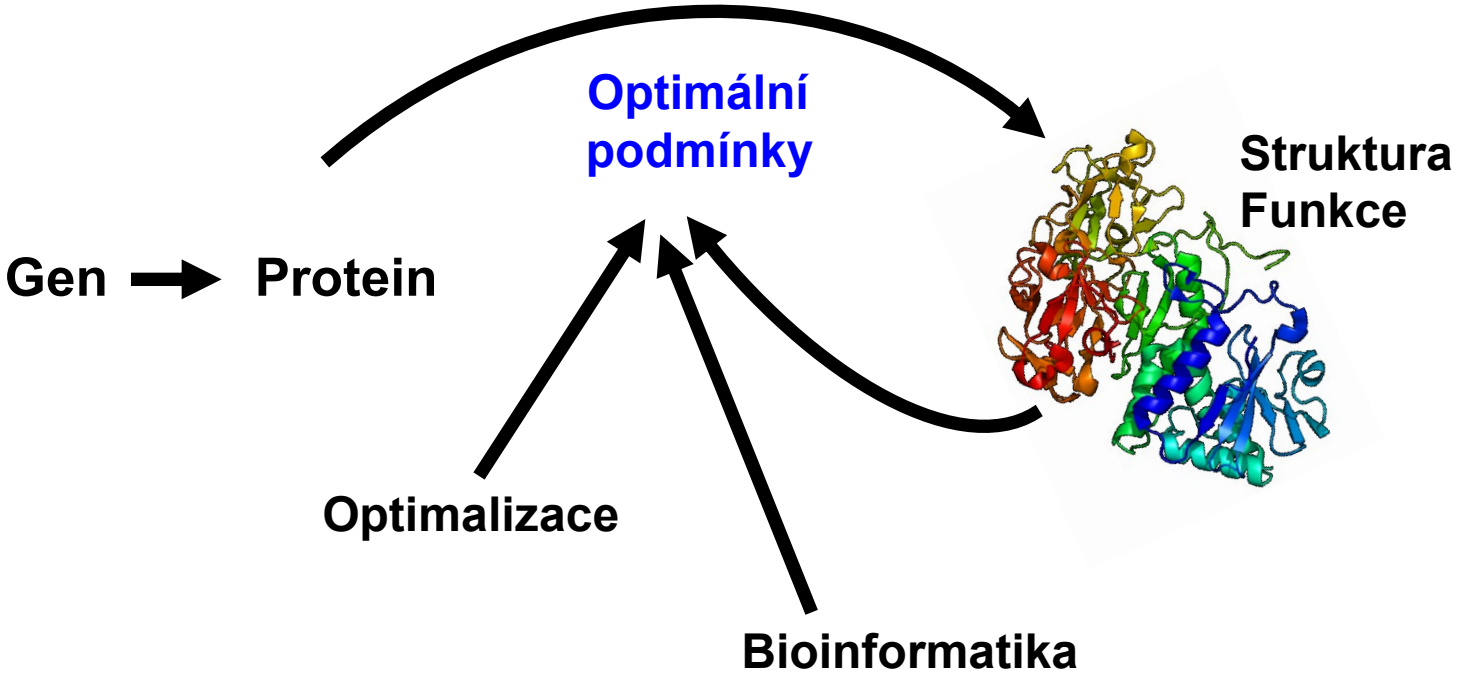
Struktura
Funkce

Optimalizace



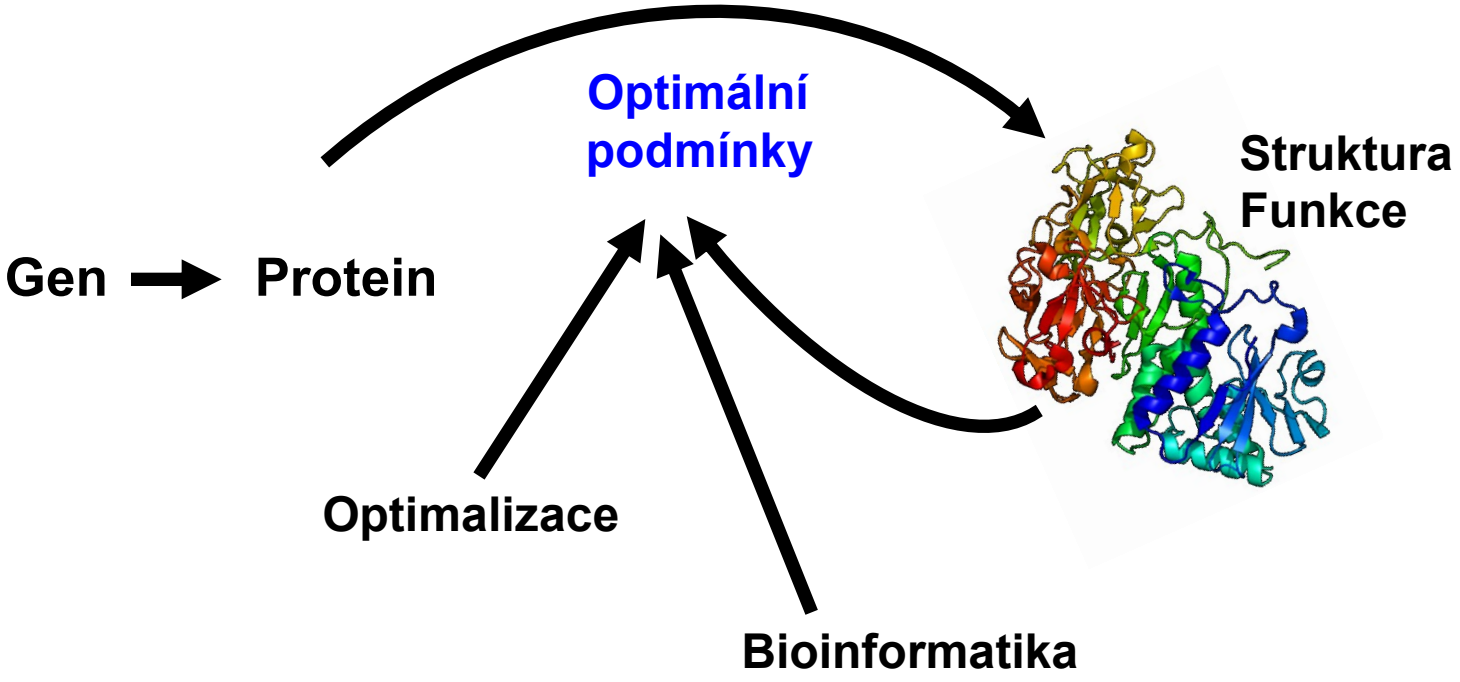
Práce s proteiny

Protein je správně sbalený, aktivní, v dostatečném množství a koncentraci



Práce s proteiny

Protein je správně sbalený, aktivní, v dostatečném množství a koncentraci




```
PLLSASIVSAPVVTSETYVD
IPGLYLDVAKAGIRDGKLQV
ILNVPTPY
```

Predikce vlastností

Predikce vlastností proteinů

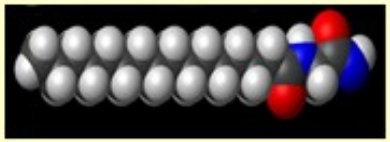
ExPASy Bioinformatics Resource Portal

wolf PSORT

psort

ProtParam

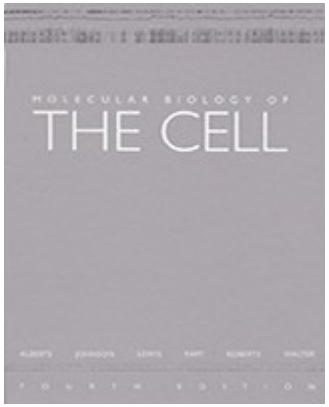
BLAST



TermiNator

ACI IMPBIO

Použitá a doporučená literatura



Molecular Biology of the Cell, 4th edition

Bruce Alberts, Alexander Johnson, Julian Lewis, Martin Raff, Keith Roberts, and Peter Walter

New York: Garland Science; 2002.

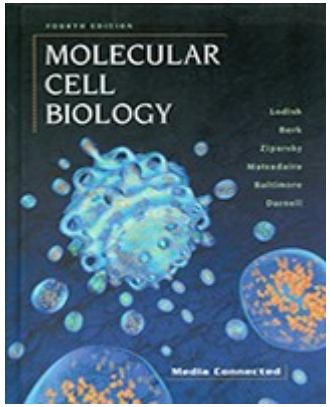
ISBN-10: 0-8153-3218-1ISBN-10: 0-8153-4072-9

Molecular Cell Biology, 4th edition

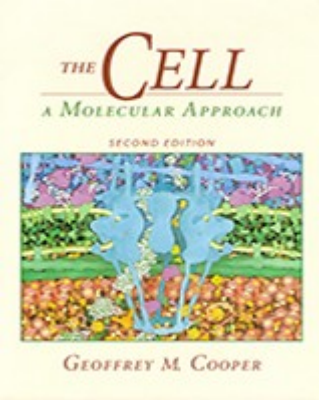
Harvey Lodish, Arnold Berk, S Lawrence Zipursky, Paul Matsudaira, David Baltimore, and James Darnell

New York: W. H. Freeman; 2000.

ISBN-10: 0-7167-3136-3



Použitá a doporučená literatura



The Cell, 2nd edition, A Molecular Approach

Geoffrey M Cooper

Boston University

Sunderland (MA): Sinauer Associates; 2000.

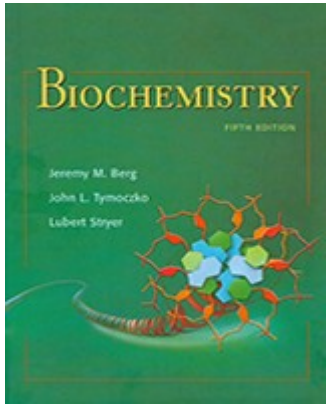
ISBN-10: 0-87893-106-6

Biochemistry, 5th edition

Jeremy M Berg, John L Tymoczko, and Lubert Stryer

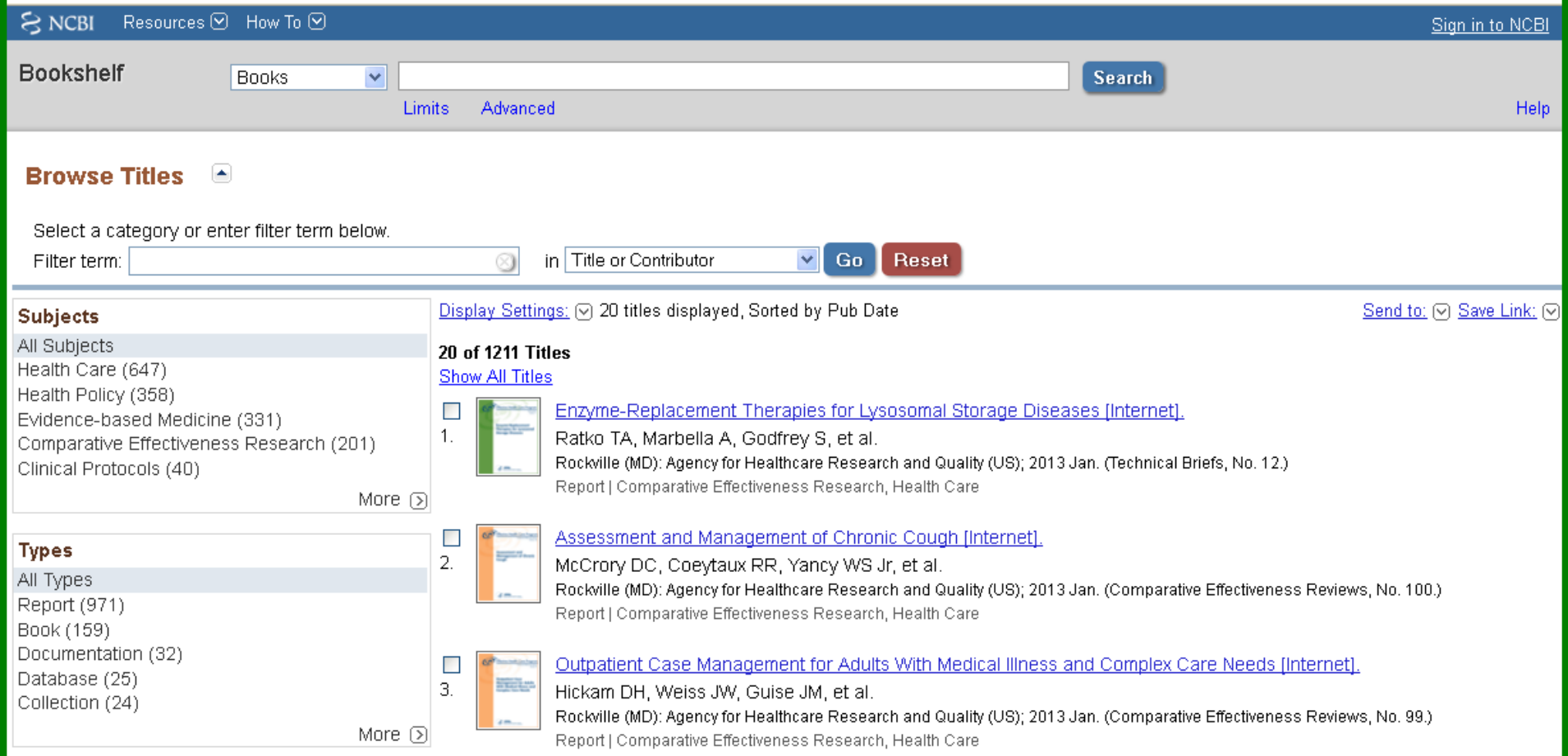
New York: W H Freeman; 2002.

ISBN-10: 0-7167-3051-0



Použitá a doporučená literatura

http://www.ncbi.nlm.nih.gov/books/browse/



Použitá a doporučená literatura

NCBI Resources How To

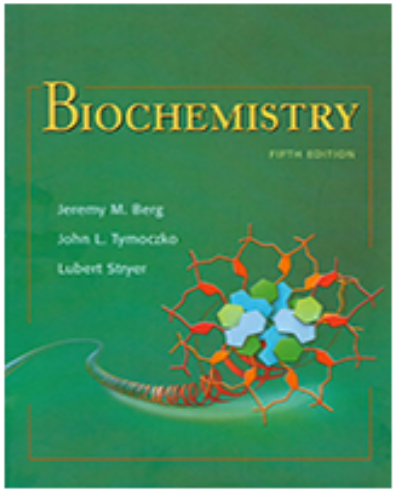
Bookshelf

This Book Proteins

Limits Advanced

By agreement with the publisher, this book is accessible by the search feature, but cannot be browsed.

Bookshelf ID: NBK21154



Biochemistry, 5th edition

Jeremy M Berg, John L Tymoczko, and Lubert Stryer.

▸ Author Information

New York: W H Freeman; 2002.

ISBN-10: 0-7167-3051-0

Copyright and Permissions Cite this Page

Použitá a doporučená literatura

Brian W. Matthews. Hydrophobic Interactions in Proteins, in Encyclopedia of Life Sciences (ELS), John Wiley & Sons, Ltd: Chichester, 2001.

Nomenclature and Symbolism for Amino Acids and Peptides: http://www.chem.qmul.ac.uk/iupac/AminoAcid/index.html

Irwin H. Segel, Leigh D. Segel. pH and Buffers, in Encyclopedia of Life Sciences (ELS), John Wiley & Sons, Ltd: Chichester, 2002.

Tech Tip #43 Protein stability and storage: www.piercenet.com