

Referências Bibliográficas

Introduction to Protein Structure – Branden & Tooze
Capítulos 1 a 5

Fundamentals of Biochemistry
Voet, Voet e Pratt
Capítulos 4 e 6

Biochemistry
Voet & Voet
Capítulo 9

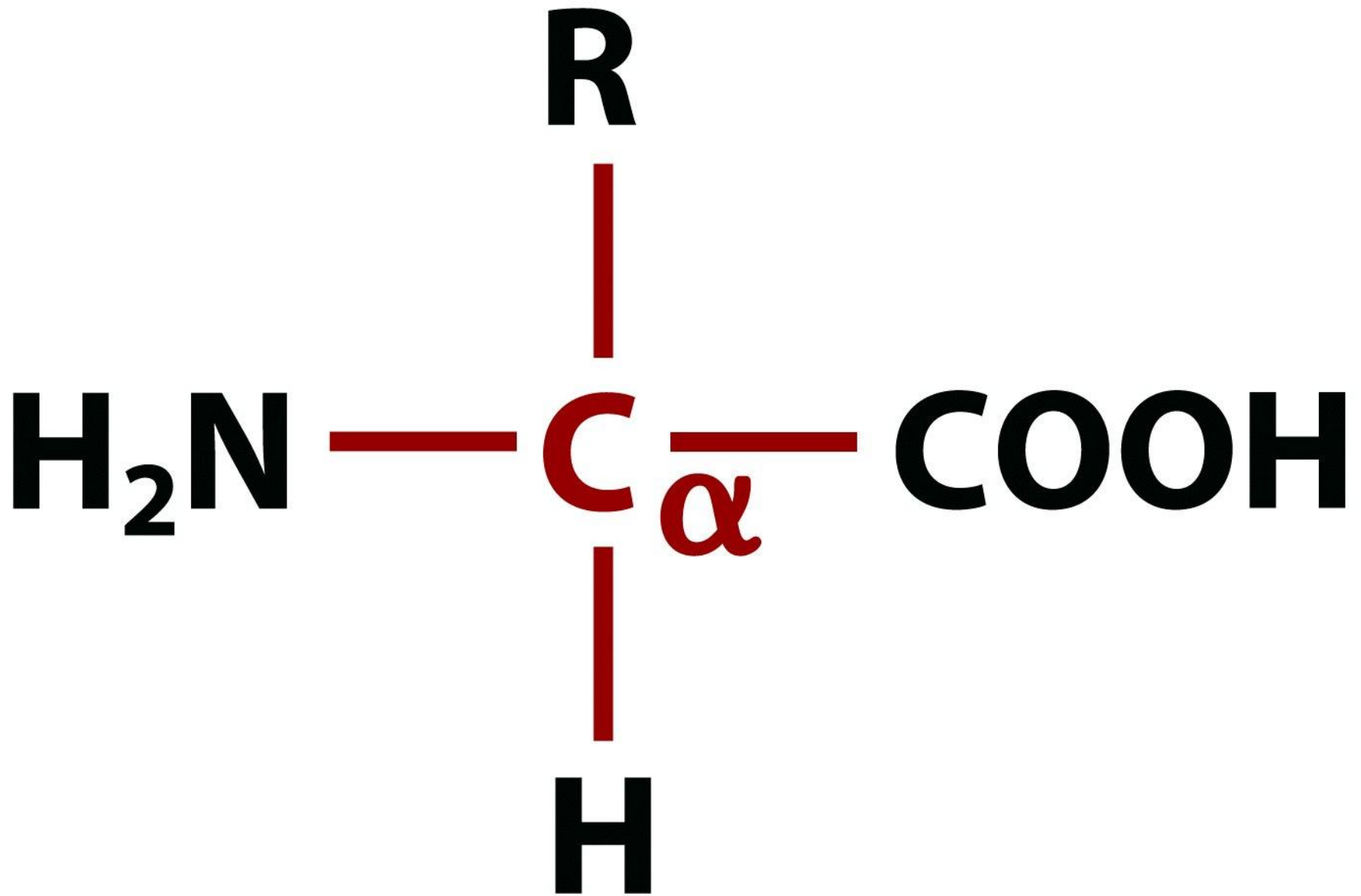
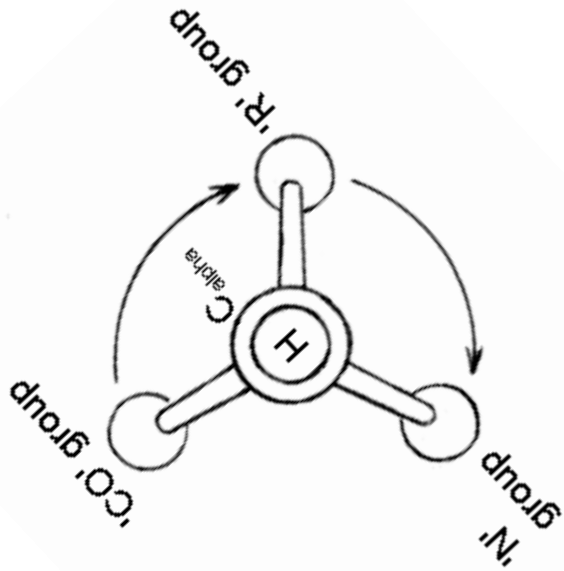


Figure 4-1 Fundamentals of Biochemistry, 2/e
© 2006 John Wiley & Sons

L-aminoácido



<http://www.cryst.bbk.ac.uk/PPS95/index.html>

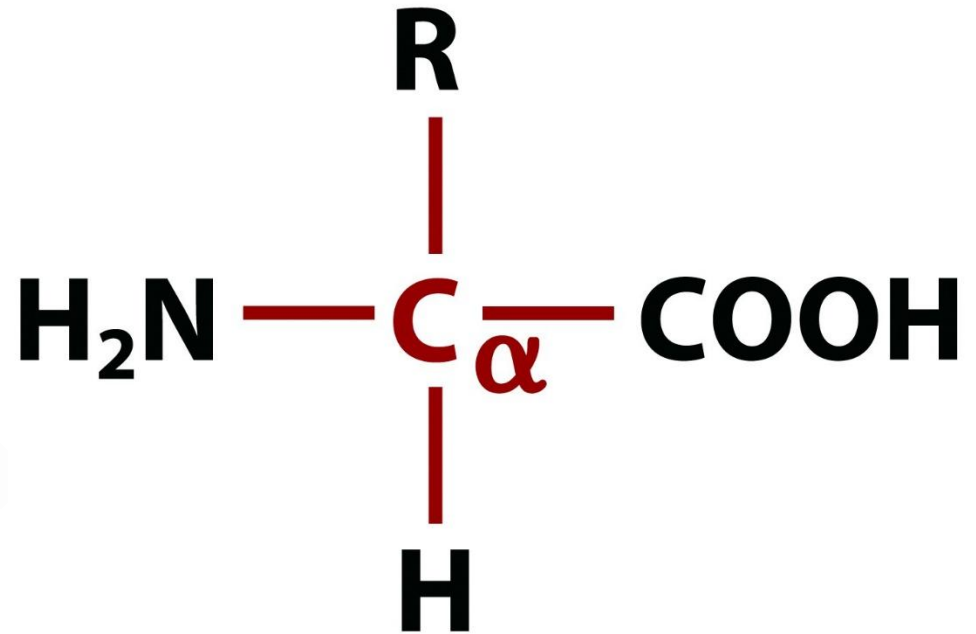


Figure 4-1 Fundamentals of Biochemistry, 2/e
© 2006 John Wiley & Sons

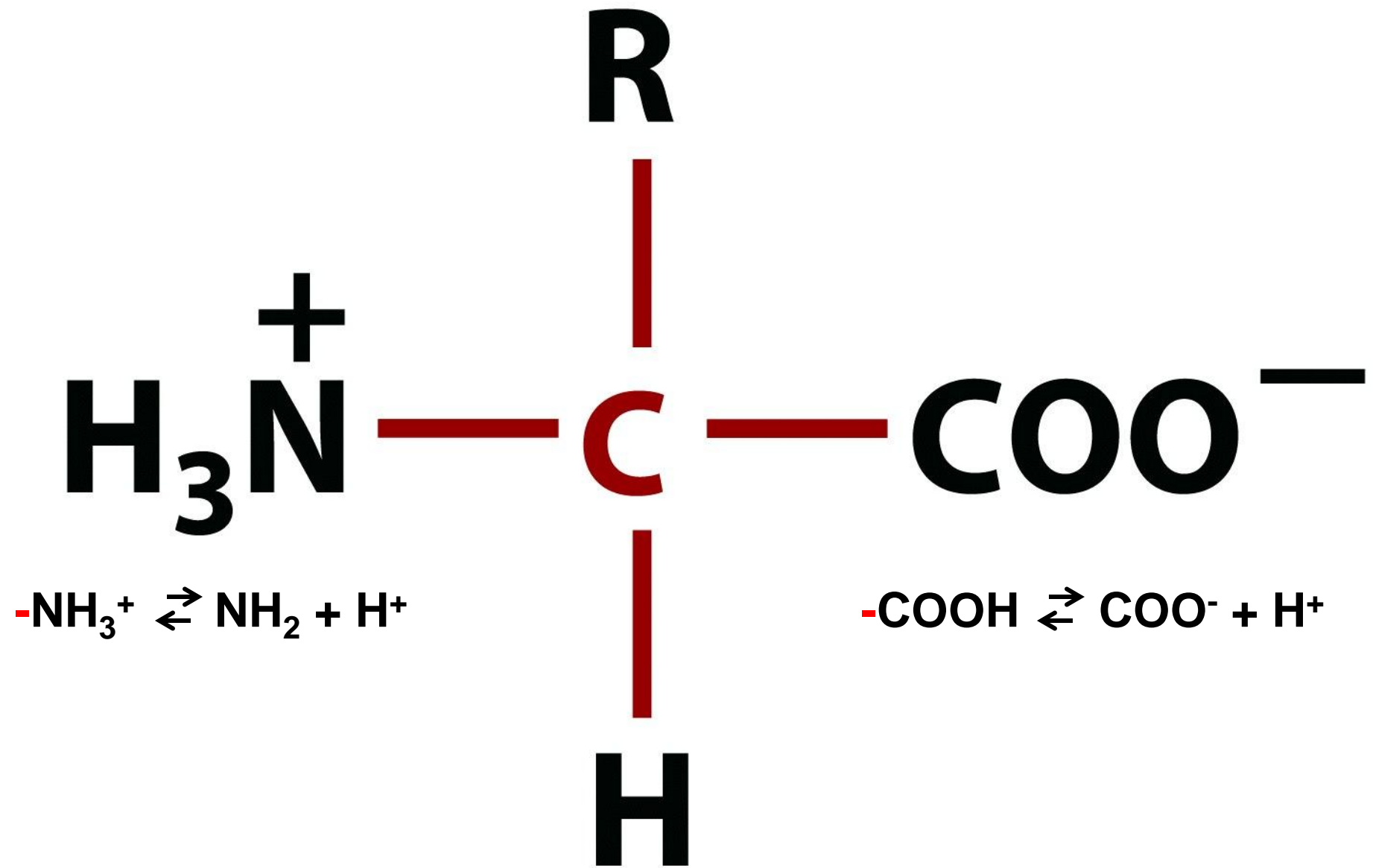


Figure 4-2 Fundamentals of Biochemistry, 2/e
© 2006 John Wiley & Sons

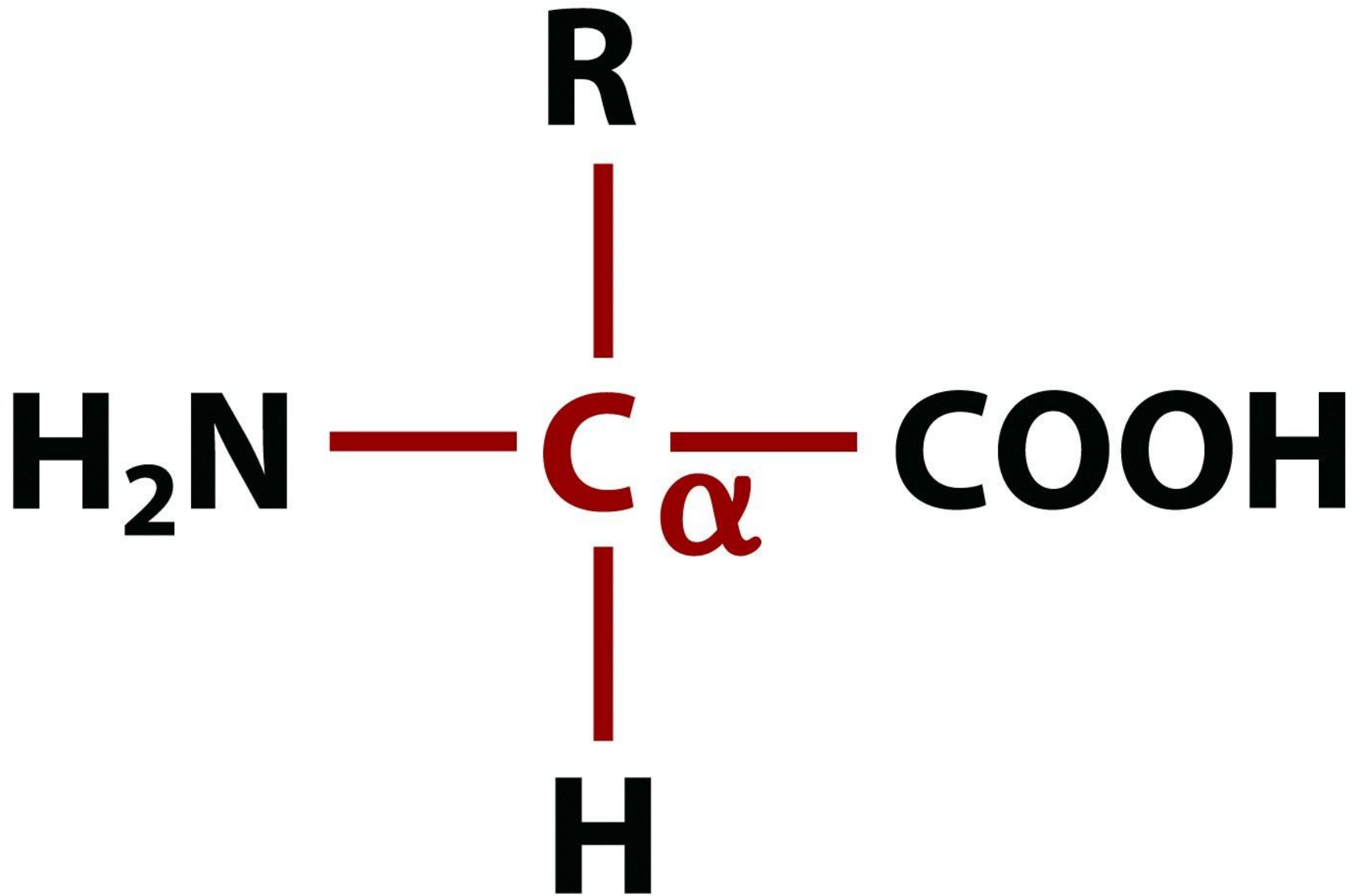


Figure 4-1 Fundamentals of Biochemistry, 2/e
© 2006 John Wiley & Sons

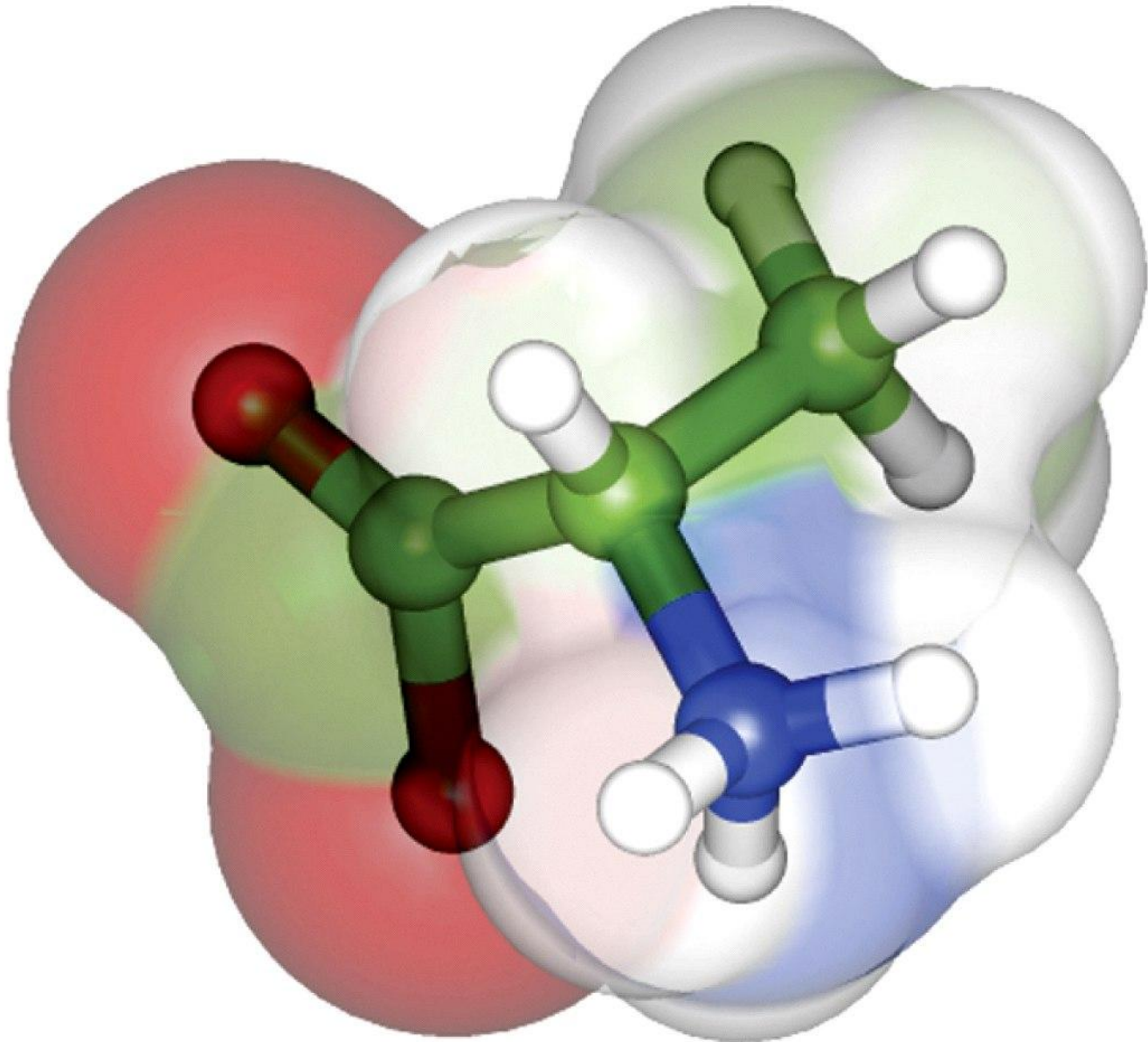


Figure 4-4a Fundamentals of Biochemistry, 2/e
© 2006 John Wiley & Sons

A

Table 4-1 Key to Structure. Covalent Structures and Abbreviations of the “Standard” Amino Acids of Proteins, Their Occurrence, and the pK Values of Their Ionizable Groups

Name, Three-letter Symbol, and One-letter Symbol	Structural Formula ^a	Residue Mass (D) ^b	Average Occurrence in Proteins (%) ^c	pK ₁ α-COOH ^d	pK ₂ α-NH ₃ ⁺ ^d	pK _R Side Chain ^d
<i>Amino acids with nonpolar side chains</i>						
Glycine Gly G	$\begin{array}{c} \text{COO}^- \\ \\ \text{H}-\text{C}-\text{H} \\ \\ \text{NH}_3^+ \end{array}$	57.0	7.2	2.35	9.78	
Alanine Ala A	$\begin{array}{c} \text{COO}^- \\ \\ \text{H}-\text{C}-\text{CH}_3 \\ \\ \text{NH}_3^+ \end{array}$	71.1	7.8	2.35	9.87	
Valine Val V	$\begin{array}{c} \text{COO}^- \\ \\ \text{H}-\text{C}-\text{CH} \\ \quad \diagup \quad \diagdown \\ \text{NH}_3^+ \quad \text{CH}_3 \quad \text{CH}_3 \end{array}$	99.1	6.6	2.29	9.74	

^aThe ionic forms shown are those predominating at pH 7.0 (except for that of histidine^f) although residue mass is given for the neutral compound. The C_α atoms, as well as those atoms marked with an asterisk, are chiral centers with configurations as indicated according to Fischer projection formulas (Section 4-2). The standard organic numbering system is provided for heterocycles.

^bThe residue masses are given for the neutral residues. For the molecular masses of the parent amino acids, add 18.0 D, the molecular mass of H₂O, to the residue masses. For side chain masses, subtract 56.0 D, the formula mass of a peptide group, from the residue masses.

^cCalculated from a database of nonredundant proteins containing 300,688 residues as compiled by Doolittle, R.F. in Fasman, G.D. (Ed.), *Predictions of Protein Structure and the Principles of Protein Conformation*, Plenum Press (1989).

^dData from Dawson, R.M.C., Elliott, D.C., Elliott, W.H., and Jones, K.M., *Data for Biochemical Research* (3rd ed.), pp. 1–31, Oxford Science Publications (1986).

Table 4-1 Key to Structure. Covalent Structures and Abbreviations of the “Standard” Amino Acids of Proteins, Their Occurrence, and the pK Values of Their Ionizable Groups

Name, Three-letter Symbol, and One-letter Symbol	Structural Formula ^a	Residue Mass (D) ^b	Average Occurrence in Proteins (%) ^c	pK ₁ α-COOH ^d	pK ₂ α-NH ₃ ⁺ ^d	pK _R Side Chain ^d
<i>Amino acids with nonpolar side chains</i>						
Leucine Leu L	$ \begin{array}{c} \text{COO}^- \\ \\ \text{H}-\text{C}-\text{CH}_2-\text{CH} \\ \qquad \quad / \quad \backslash \\ \text{NH}_3^+ \qquad \text{CH}_3 \quad \text{CH}_3 \end{array} $	113.2	9.1	2.33	9.74	
Isoleucine Ile I	$ \begin{array}{c} \text{COO}^- \quad \text{CH}_3 \\ \qquad \quad \\ \text{H}-\text{C}-\text{C}^*-\text{CH}_2-\text{CH}_3 \\ \qquad \quad \\ \text{NH}_3^+ \quad \text{H} \end{array} $	113.2	5.3	2.32	9.76	
Methionine Met M	$ \begin{array}{c} \text{COO}^- \\ \\ \text{H}-\text{C}-\text{CH}_2-\text{CH}_2-\text{S}-\text{CH}_3 \\ \\ \text{NH}_3^+ \end{array} $	131.2	2.2	2.13	9.28	

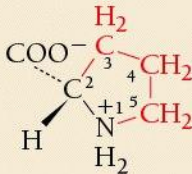
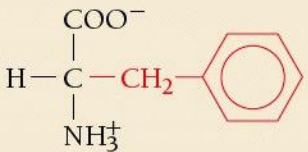
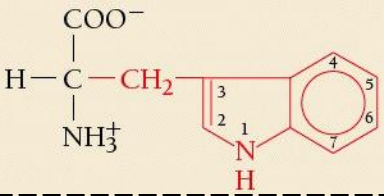
^aThe ionic forms shown are those predominating at pH 7.0 (except for that of histidine^f) although residue mass is given for the neutral compound. The C_α atoms, as well as those atoms marked with an asterisk, are chiral centers with configurations as indicated according to Fischer projection formulas (Section 4-2). The standard organic numbering system is provided for heterocycles.

^bThe residue masses are given for the neutral residues. For the molecular masses of the parent amino acids, add 18.0 D, the molecular mass of H₂O, to the residue masses. For side chain masses, subtract 56.0 D, the formula mass of a peptide group, from the residue masses.

^cCalculated from a database of nonredundant proteins containing 300,688 residues as compiled by Doolittle, R.F. in Fasman, G.D. (Ed.), *Predictions of Protein Structure and the Principles of Protein Conformation*, Plenum Press (1989).

^dData from Dawson, R.M.C., Elliott, D.C., Elliott, W.H., and Jones, K.M., *Data for Biochemical Research* (3rd ed.), pp. 1–31, Oxford Science Publications (1986).

Table 4-1 Key to Structure. Covalent Structures and Abbreviations of the “Standard” Amino Acids of Proteins, Their Occurrence, and the pK Values of Their Ionizable Groups

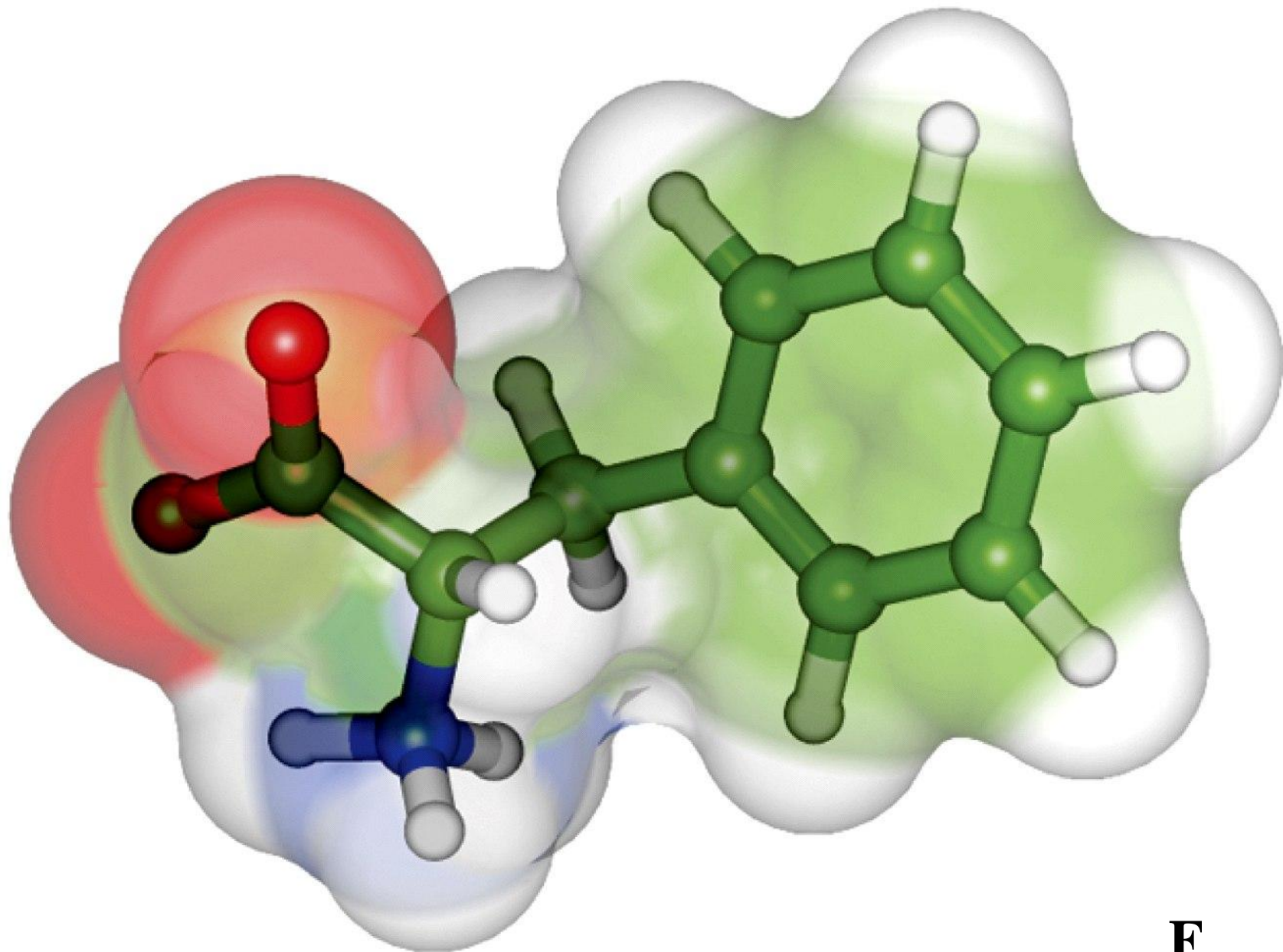
Name, Three-letter Symbol, and One-letter Symbol	Structural Formula ^a	Residue Mass (D) ^b	Average Occurrence in Proteins (%) ^c	pK ₁ α-COOH ^d	pK ₂ α-NH ₃ ⁺ ^d	pK _R Side Chain ^d
<i>Amino acids with nonpolar side chains</i>						
Proline Pro P		97.1	5.2	1.95	10.64	
Phenylalanine Phe F		147.2	3.9	2.20	9.31	
Tryptophan Trp W		186.2	1.4	2.46	9.41	

^aThe ionic forms shown are those predominating at pH 7.0 (except for that of histidine^f) although residue mass is given for the neutral compound. The C_α atoms, as well as those atoms marked with an asterisk, are chiral centers with configurations as indicated according to Fischer projection formulas (Section 4-2). The standard organic numbering system is provided for heterocycles.

^bThe residue masses are given for the neutral residues. For the molecular masses of the parent amino acids, add 18.0 D, the molecular mass of H₂O, to the residue masses. For side chain masses, subtract 56.0 D, the formula mass of a peptide group, from the residue masses.

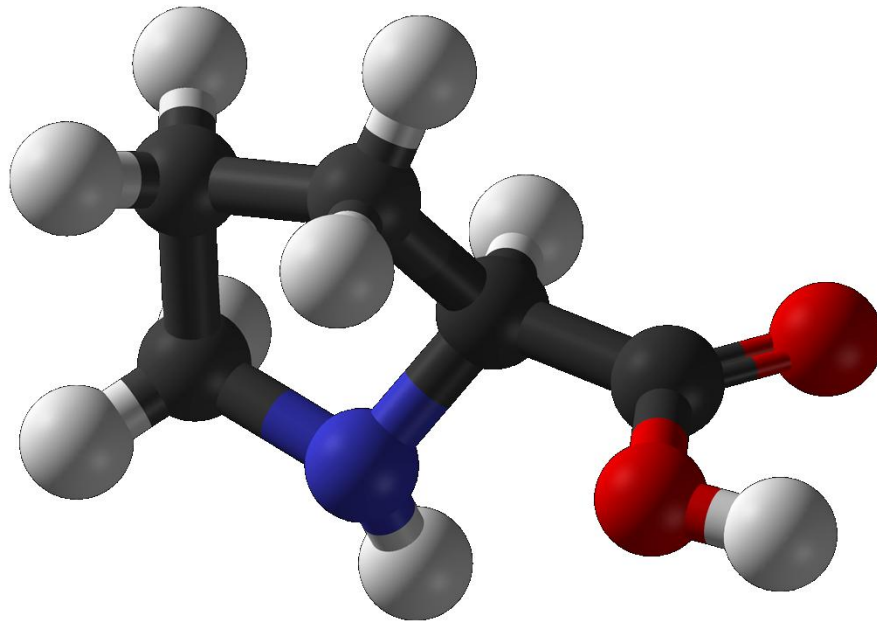
^cCalculated from a database of nonredundant proteins containing 300,688 residues as compiled by Doolittle, R.F. in Fasman, G.D. (Ed.), *Predictions of Protein Structure and the Principles of Protein Conformation*, Plenum Press (1989).

^dData from Dawson, R.M.C., Elliott, D.C., Elliott, W.H., and Jones, K.M., *Data for Biochemical Research* (3rd ed.), pp. 1–31, Oxford Science Publications (1986).



F

Figure 4-4c Fundamentals of Biochemistry, 2/e
© 2006 John Wiley & Sons



P

Table 4-1 (continued)

Name, Three-letter Symbol, and One-letter Symbol	Structural Formula ^a	Residue Mass (D) ^b	Average Occurrence in Proteins (%) ^c	pK ₁ α-COOH ^d	pK ₂ α-NH ₃ ⁺ ^d	pK _R Side Chain ^d
<i>Amino acids with uncharged polar side chains</i>						
Serine Ser S	$ \begin{array}{c} \text{COO}^- \\ \\ \text{H}-\text{C}-\text{CH}_2-\text{OH} \\ \\ \text{NH}_3^+ \end{array} $	87.1	6.8	2.19	9.21	
Threonine Thr T	$ \begin{array}{c} \text{COO}^- \quad \text{H} \\ \quad \quad \\ \text{H}-\text{C}-\text{C}^*-\text{CH}_3 \\ \quad \quad \\ \text{NH}_3^+ \quad \text{OH} \end{array} $	101.1	5.9	2.09	9.10	
Asparagine ^e Asn N	$ \begin{array}{c} \text{COO}^- \quad \quad \text{O} \\ \quad \quad \quad // \\ \text{H}-\text{C}-\text{CH}_2-\text{C} \\ \quad \quad \quad \backslash \\ \text{NH}_3^+ \quad \quad \text{NH}_2 \end{array} $	114.1	4.3	2.14	8.72	

^eThe three- and one-letter symbols for asparagine *or* aspartic acid are Asx and B, whereas for glutamine *or* glutamic acid they are Glx and Z. The one-letter symbol for an undetermined or “nonstandard” amino acid is X.

^fBoth neutral and protonated forms of histidine are present at pH 7.0, since its pK_R is close to 7.0.

Table 4-1 part 4 Fundamentals of Biochemistry, 2/e

© 2006 John Wiley & Sons

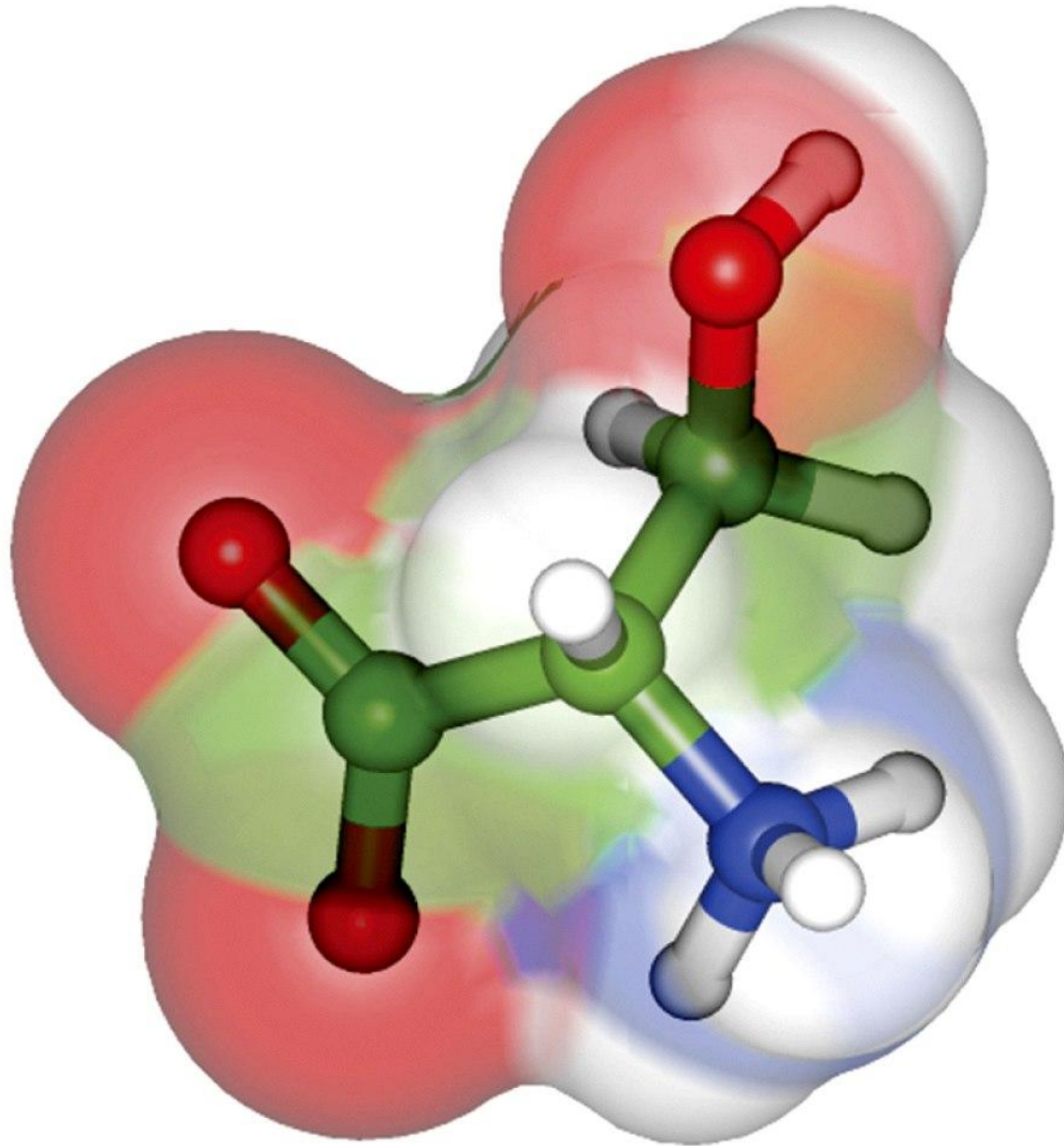


Figure 4-5a Fundamentals of Biochemistry, 2/e
© 2006 John Wiley & Sons

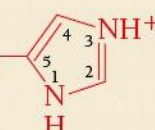
Table 4-1 (continued)

Name, Three-letter Symbol, and One-letter Symbol	Structural Formula ^a	Residue Mass (D) ^b	Average Occurrence in Proteins (%) ^c	pK ₁ α-COOH ^d	pK ₂ α-NH ₃ ⁺ ^d	pK _R Side Chain ^d
<i>Amino acids with uncharged polar side chains</i>						
Glutamine ^e Gln Q		128.1	4.3	2.17	9.13	
Tyrosine Tyr Y		163.2	3.2	2.20	9.21	10.46 (phenol)
Cysteine Cys C		103.1	1.9	1.92	10.70	8.37 (sulfhydryl)

^eThe three- and one-letter symbols for asparagine *or* aspartic acid are Asx and B, whereas for glutamine *or* glutamic acid they are Glx and Z. The one-letter symbol for an undetermined or “nonstandard” amino acid is X.

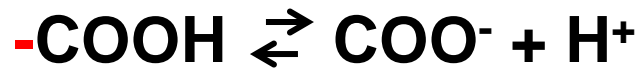
^fBoth neutral and protonated forms of histidine are present at pH 7.0, since its pK_R is close to 7.0.

Table 4-1 (continued)

Name, Three-letter Symbol, and One-letter Symbol	Structural Formula ^a	Residue Mass (D) ^b	Average Occurrence in Proteins (%) ^c	pK ₁ α-COOH ^d	pK ₂ α-NH ₃ ⁺ ^d	pK _R Side Chain ^d
<i>Amino acids with charged polar side chains</i>						
Lysine Lys K	$\begin{array}{c} \text{COO}^- \\ \\ \text{H}-\text{C}-\text{CH}_2-\text{CH}_2-\text{CH}_2-\text{CH}_2-\text{NH}_3^+ \\ \\ \text{NH}_3^+ \end{array}$	128.2	5.9	2.16	9.06	10.54 (ε-NH ₃ ⁺)
Arginine Arg R	$\begin{array}{c} \text{COO}^- \\ \\ \text{H}-\text{C}-\text{CH}_2-\text{CH}_2-\text{CH}_2-\text{NH}-\text{C} \begin{array}{l} \text{NH}_2 \\ \text{NH}_2^+ \end{array} \\ \\ \text{NH}_3^+ \end{array}$	156.2	5.1	1.82	8.99	12.48 (guanidino)
Histidine ^f His H	$\begin{array}{c} \text{COO}^- \\ \\ \text{H}-\text{C}-\text{CH}_2-\text{Imidazole}^+ \\ \\ \text{NH}_3^+ \end{array}$ 	137.1	2.3	1.80	9.33	6.04 (imidazole)
Aspartic acid ^e Asp D	$\begin{array}{c} \text{COO}^- \\ \\ \text{H}-\text{C}-\text{CH}_2-\text{C} \begin{array}{l} \text{O} \\ \text{O}^- \end{array} \\ \\ \text{NH}_3^+ \end{array}$	115.1	5.3	1.99	9.90	3.90 (β-COOH)
Glutamic acid ^e Glu E	$\begin{array}{c} \text{COO}^- \\ \\ \text{H}-\text{C}-\text{CH}_2-\text{CH}_2-\text{C} \begin{array}{l} \text{O} \\ \text{O}^- \end{array} \\ \\ \text{NH}_3^+ \end{array}$	129.1	6.3	2.10	9.47	4.07 (γ-COOH)

^eThe three- and one-letter symbols for asparagine *or* aspartic acid are Asx and B, whereas for glutamine *or* glutamic acid they are Glx and Z. The one-letter symbol for an undetermined or “nonstandard” amino acid is X.

^fBoth neutral and protonated forms of histidine are present at pH 7.0, since its pK_R is close to 7.0.



$$\text{pH} = \text{pK}_a + \log \frac{B}{AH}$$

AH = forma protonada do ácido fraco

B = forma desprotonada do ácido fraco

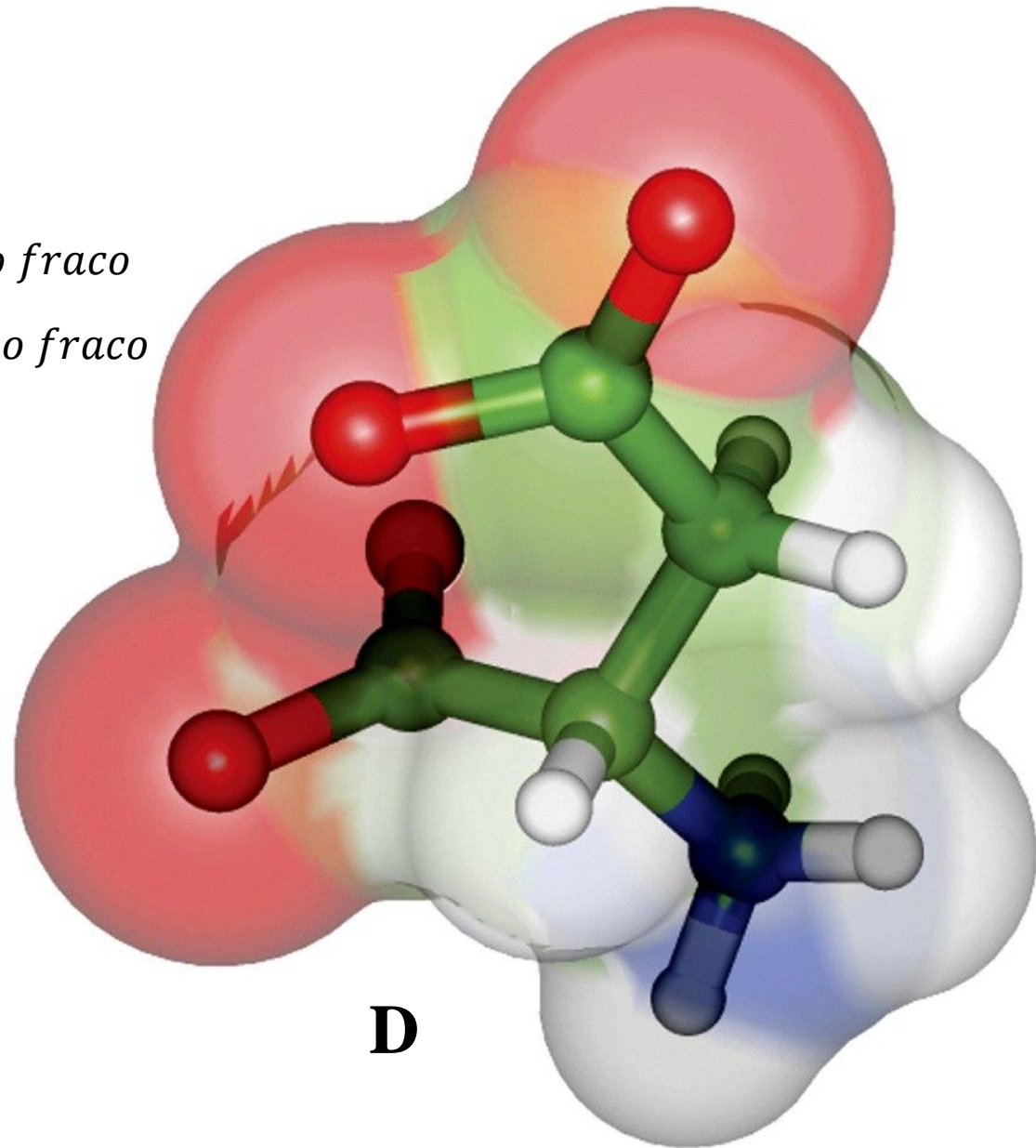


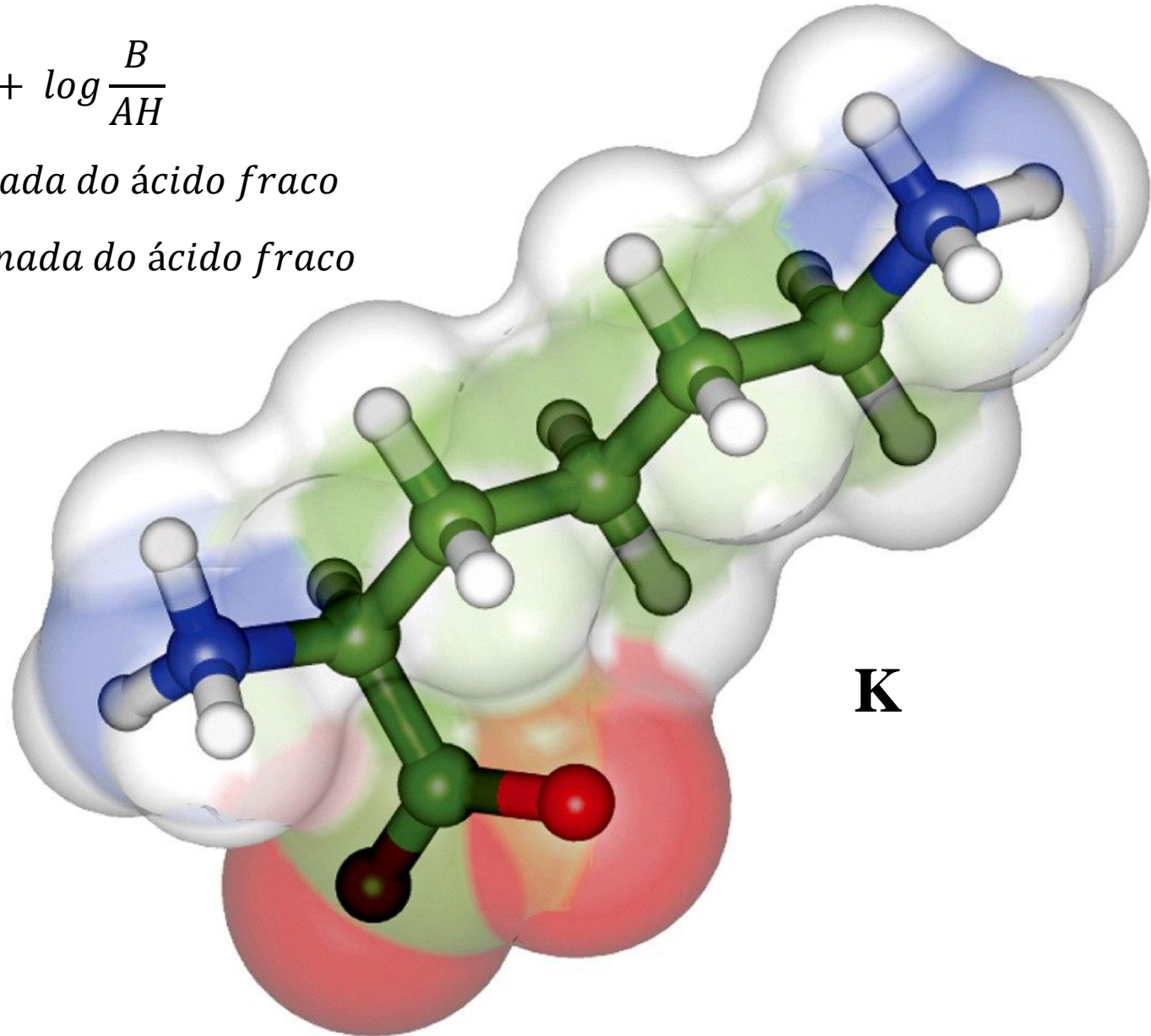
Figure 4-7a Fundamentals of Biochemistry, 2/e
© 2006 John Wiley & Sons



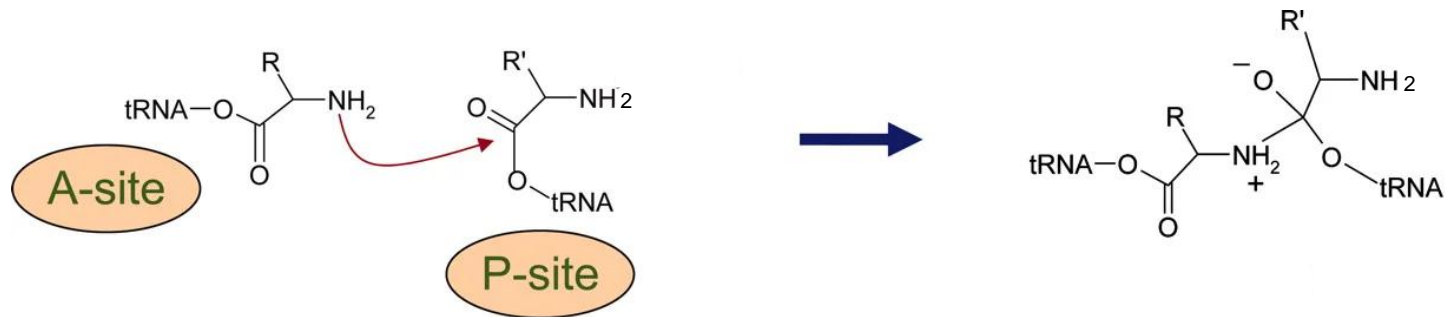
$$\text{pH} = \text{pK}_a + \log \frac{B}{AH}$$

AH = forma protonada do ácido fraco

B = forma desprotonada do ácido fraco



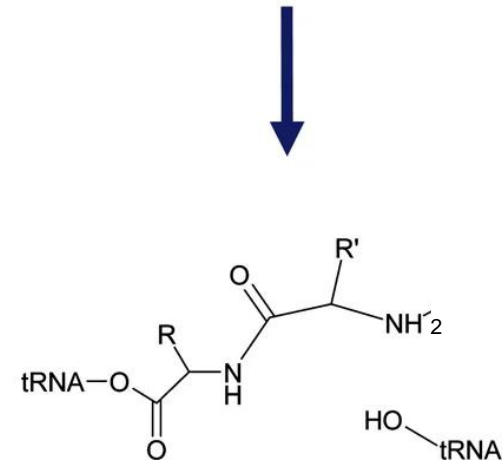
Mecanismo simplificado da formação da ligação peptídica nos ribossomos



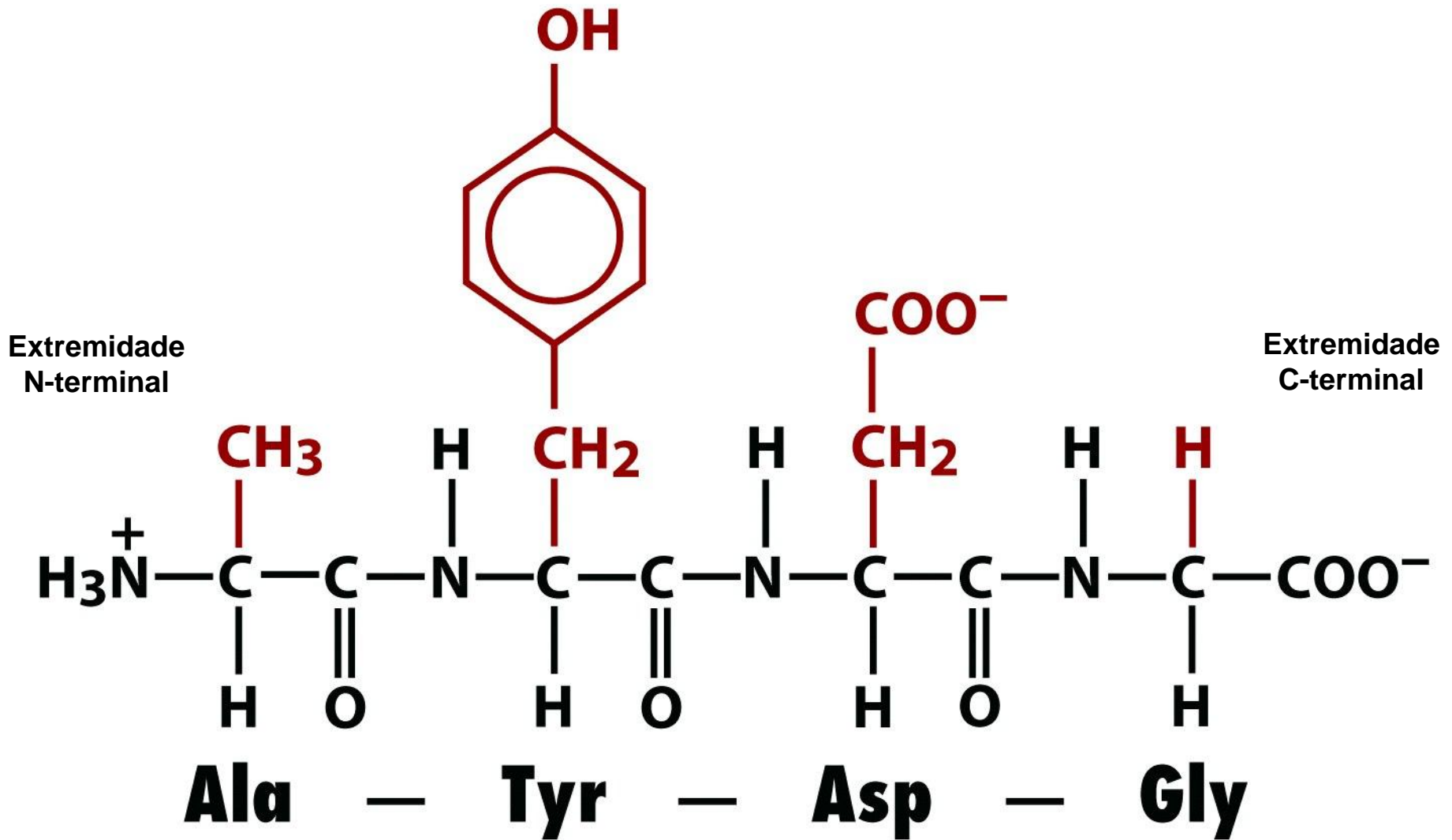
A ligação do aminoácido ao tRNA depende de ATP.

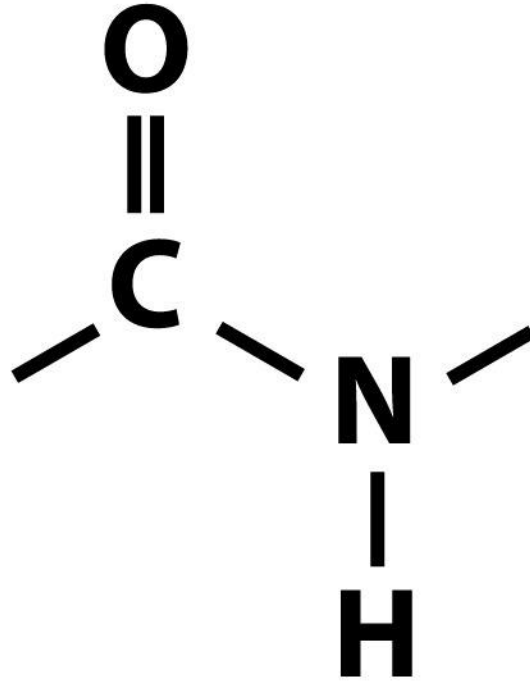
O aminoácido liga-se covalentemente pelo grupo α -carboxila à hidroxila 3 da ribose de um nucleotídeo do tRNA (extremidade 3').

Esta reação de formação da ligação peptídica ocorre no ribossomo.



Estrutura Primária





Unnumbered figure pg 131 Fundamentals of Biochemistry, 2/e
© 2006 John Wiley & Sons

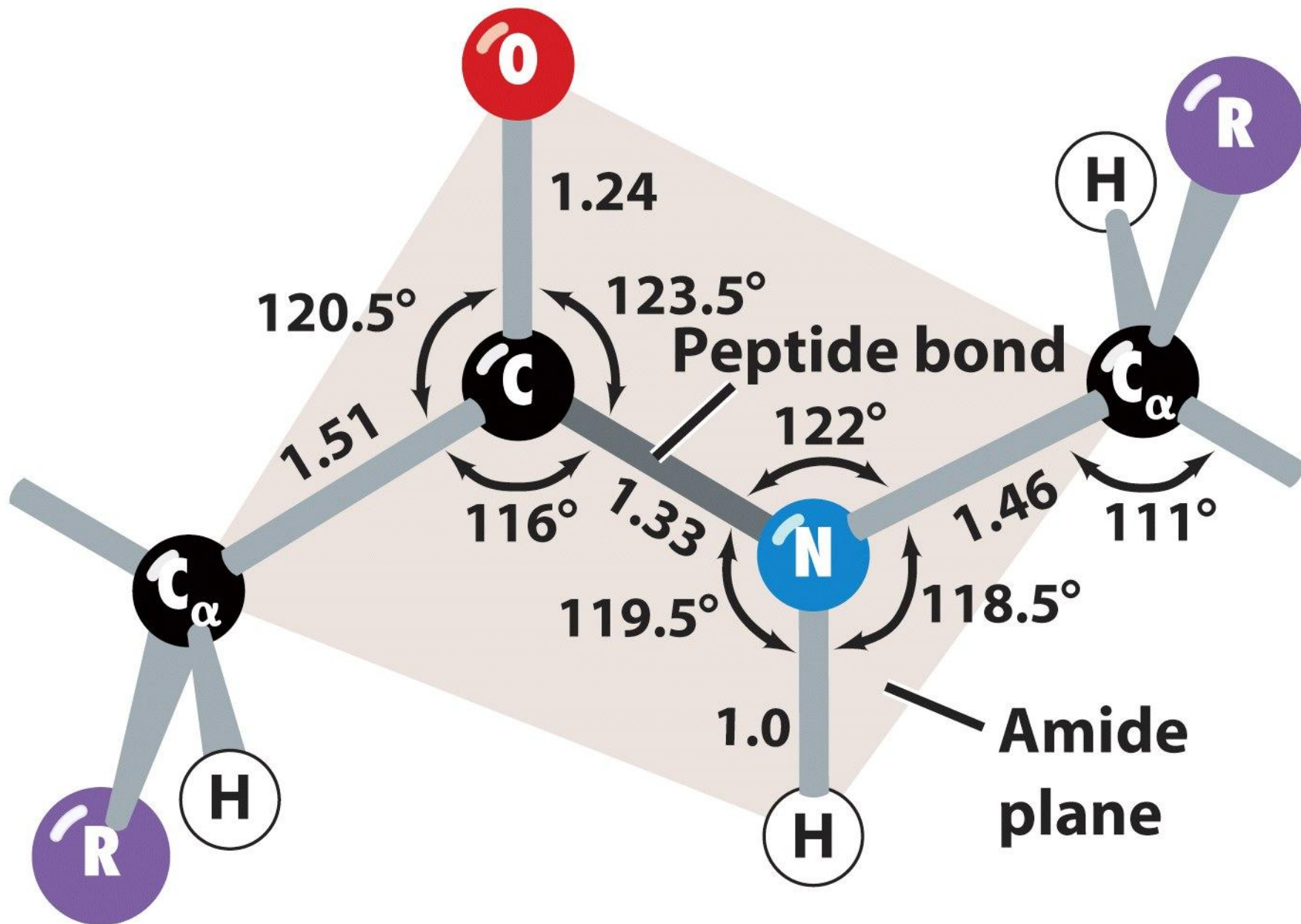
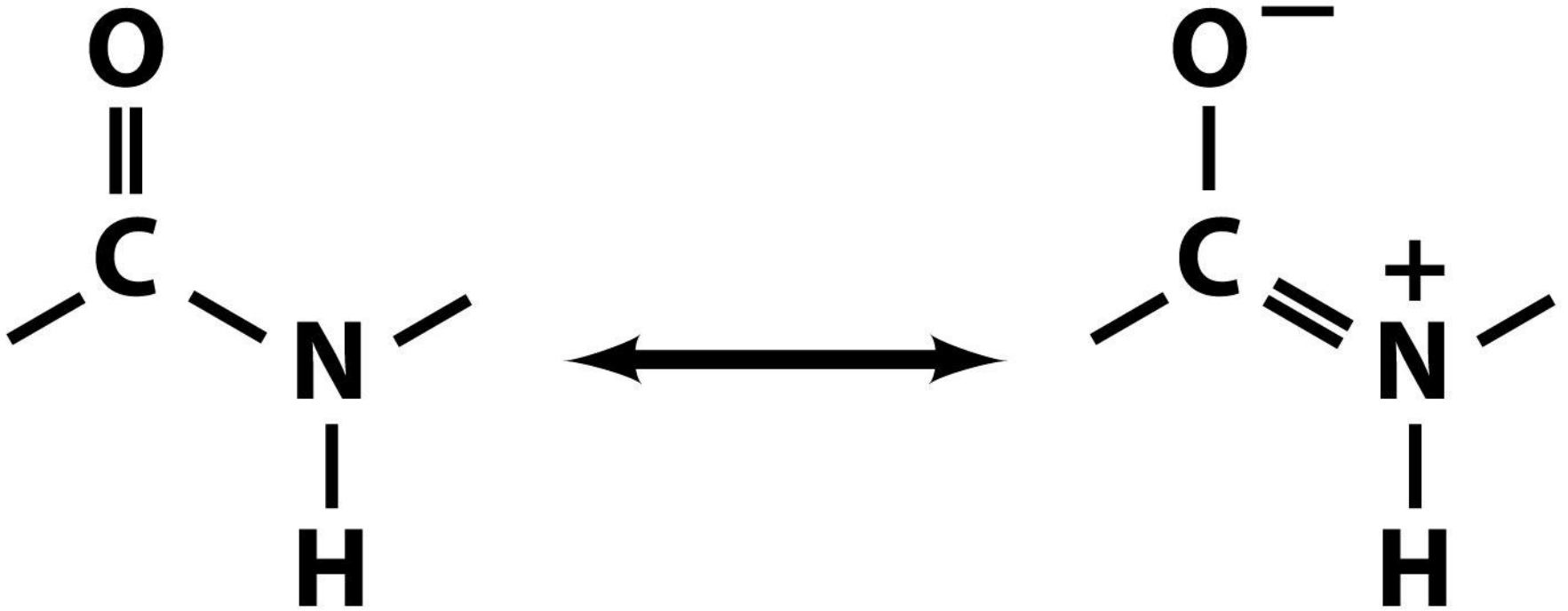


Figure 6-2 Fundamentals of Biochemistry, 2/e
 © 2006 John Wiley & Sons



Unnumbered figure pg 131 Fundamentals of Biochemistry, 2/e
© 2006 John Wiley & Sons

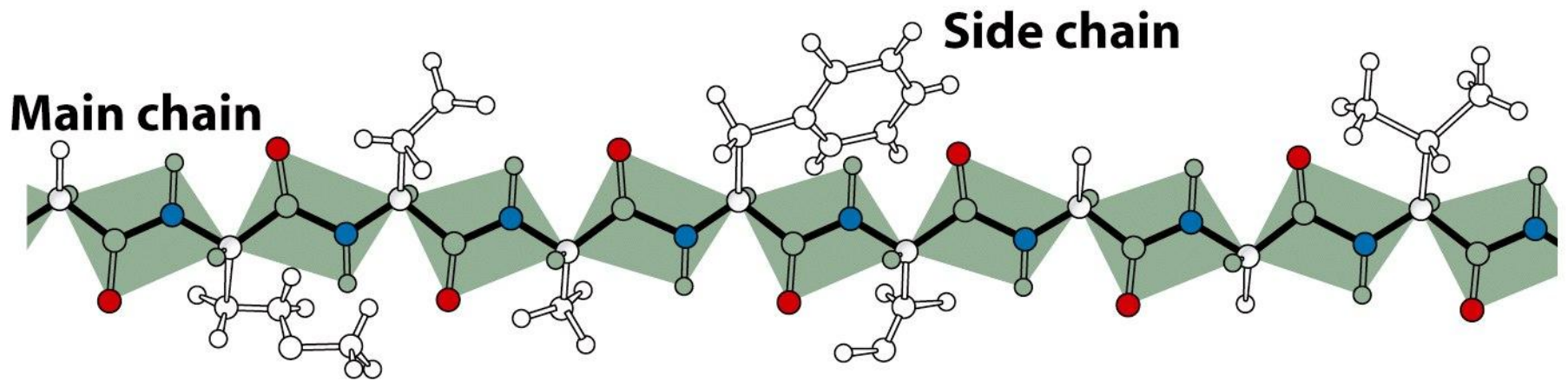


Figure 6-3 Fundamentals of Biochemistry, 2/e
© 2006 John Wiley & Sons

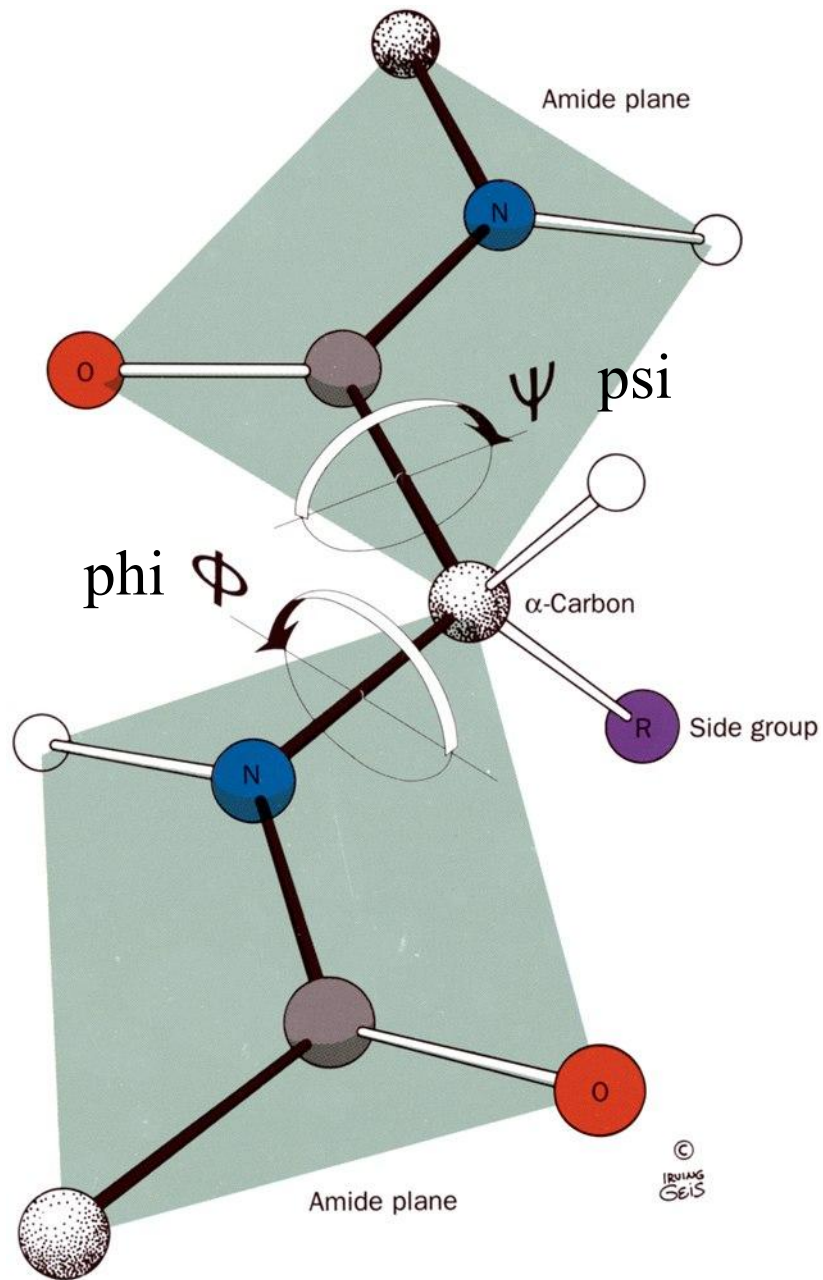


Figure 6-4 Fundamentals of Biochemistry, 2/e

Plote de Ramachandram

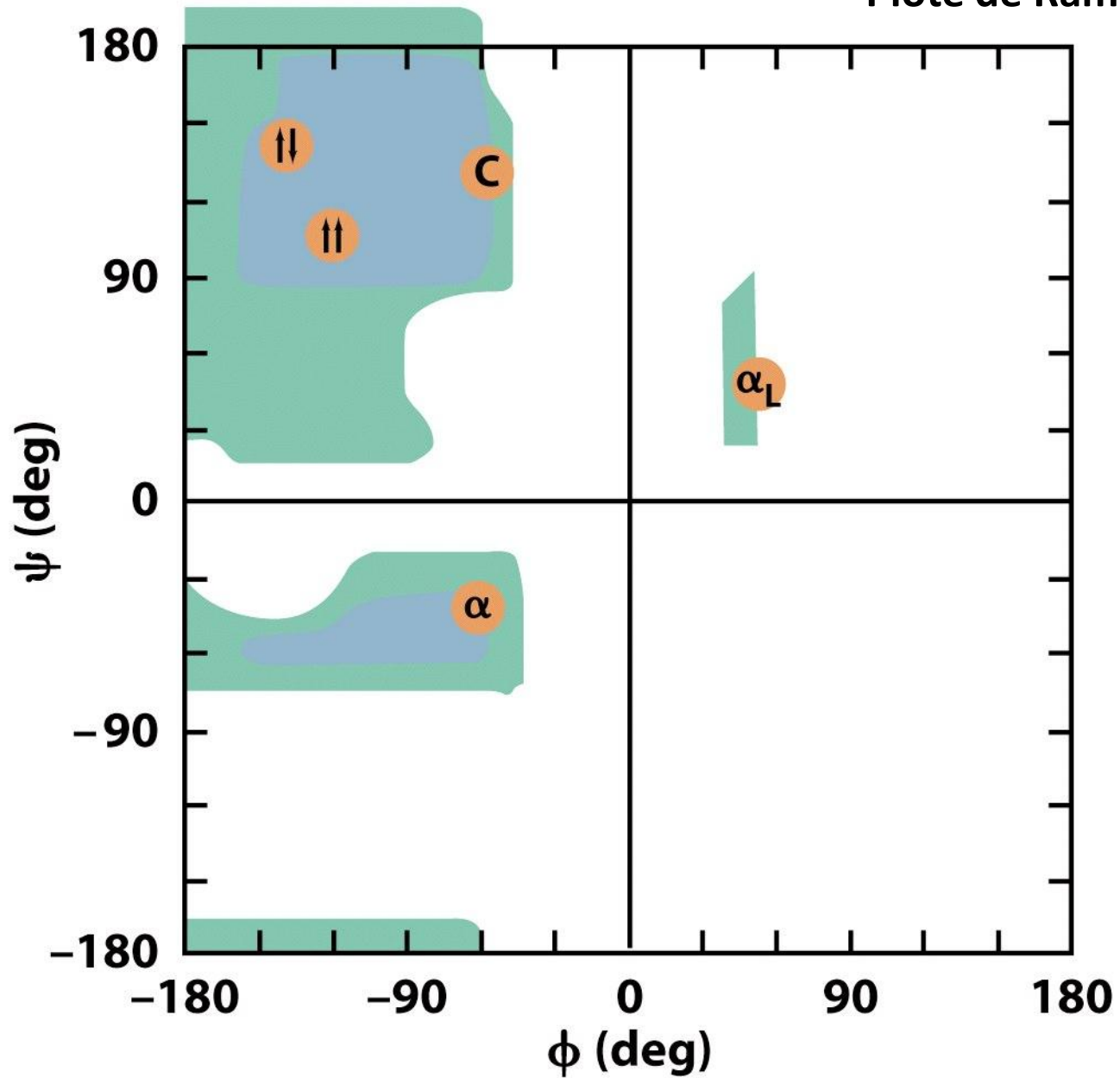


Figure 6-6 Fundamentals of Biochemistry, 2/e
© 2006 John Wiley & Sons

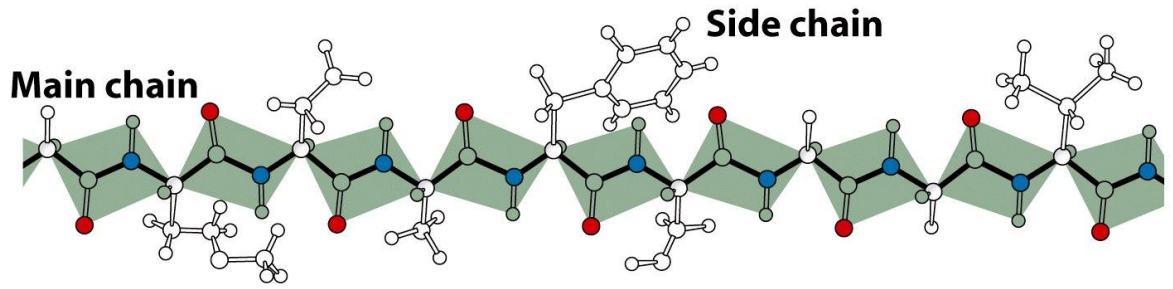
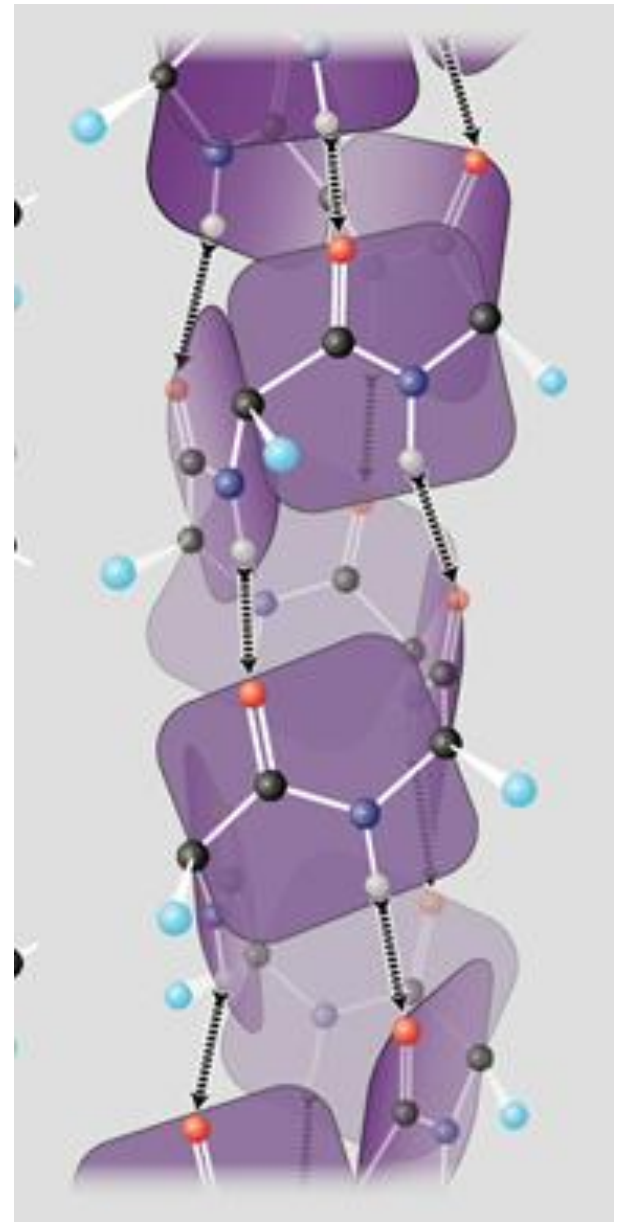


Figure 6-3 Fundamentals of Biochemistry, 2/e
© 2006 John Wiley & Sons



α Hélice

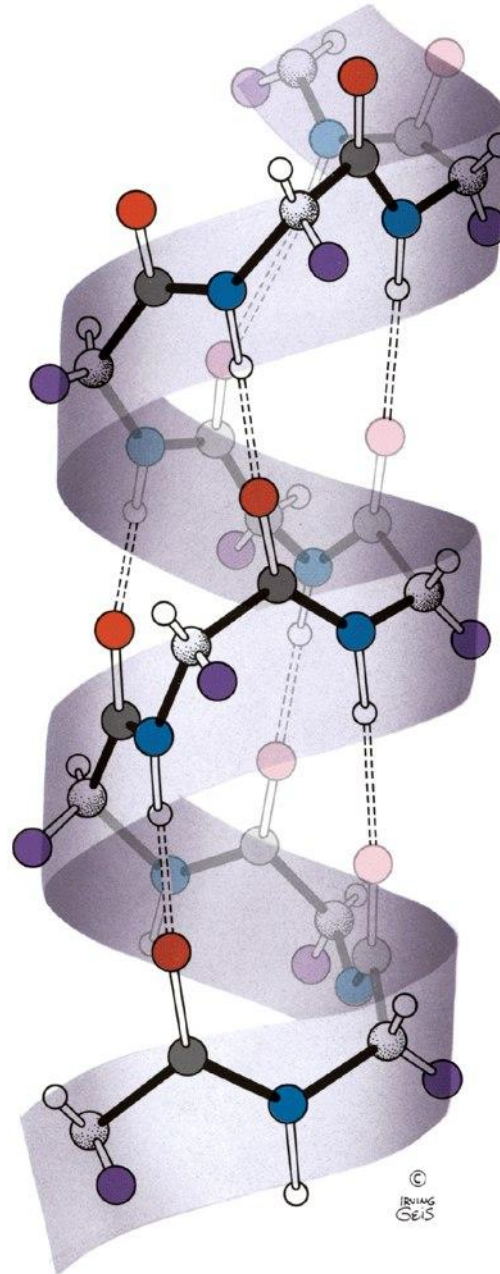


Figure 6-7 Fundamentals of Biochemistry, 2/e

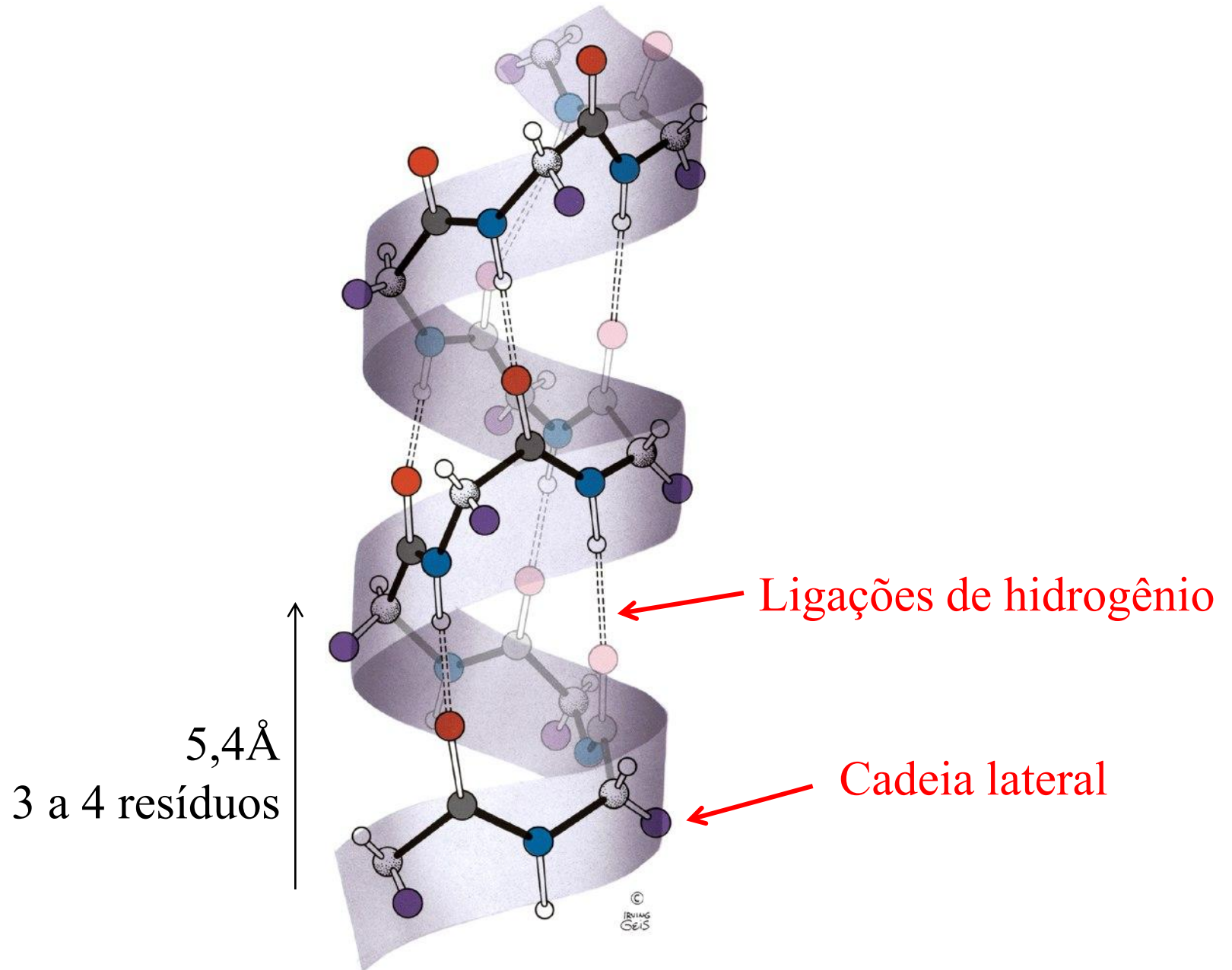


Figure 6-7 Fundamentals of Biochemistry, 2/e

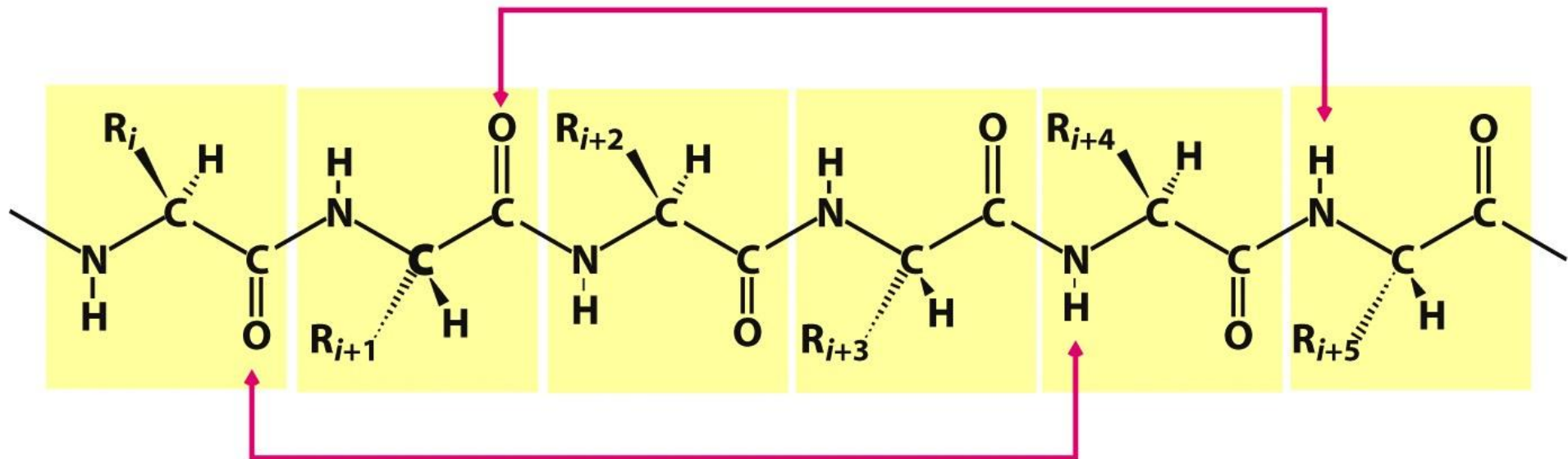
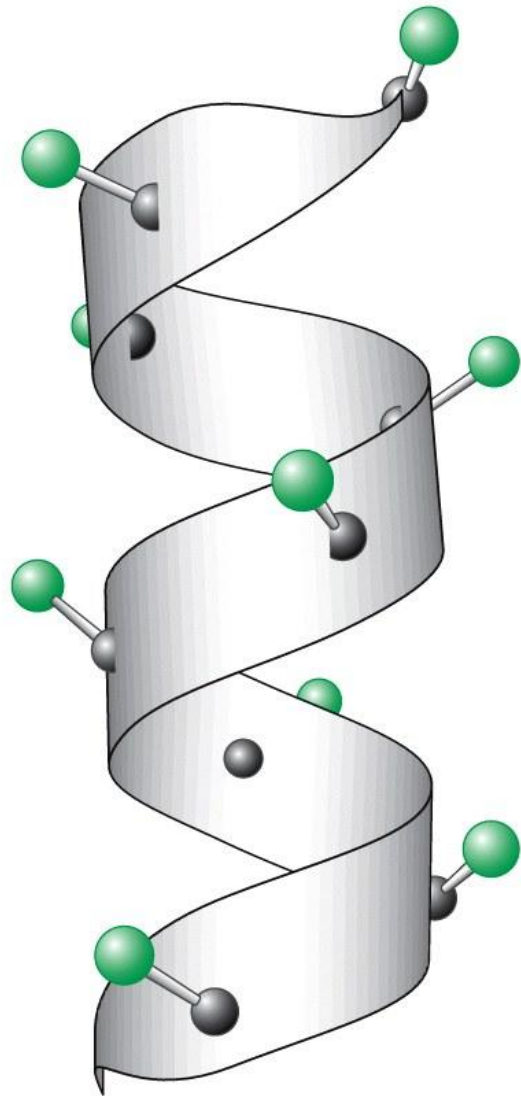
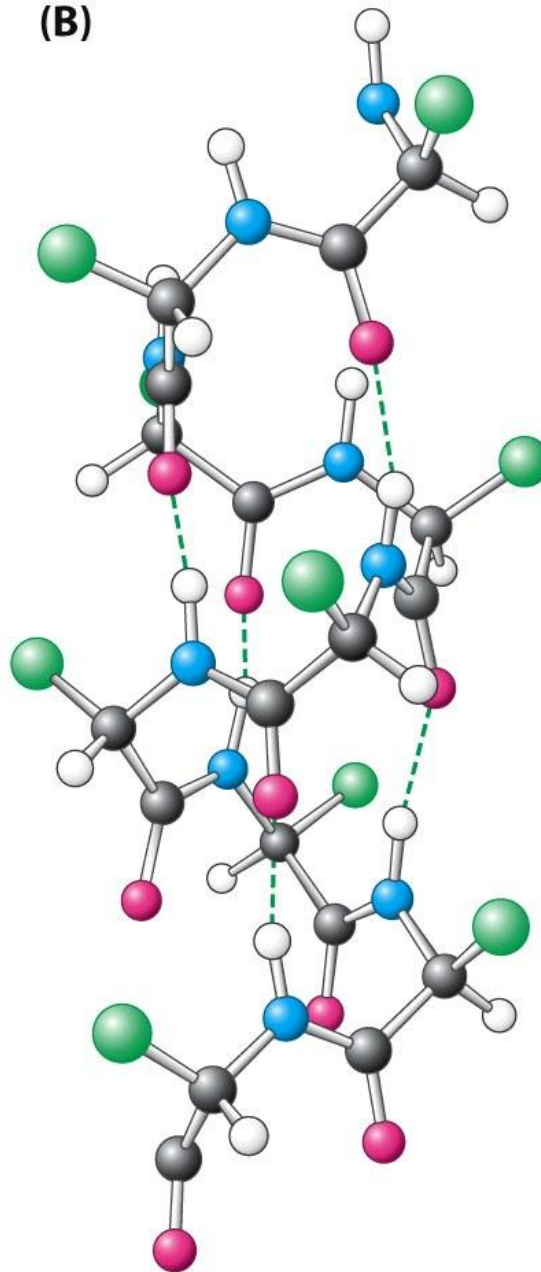


Figure 2.25
Biochemistry, Seventh Edition
© 2012 W. H. Freeman and Company

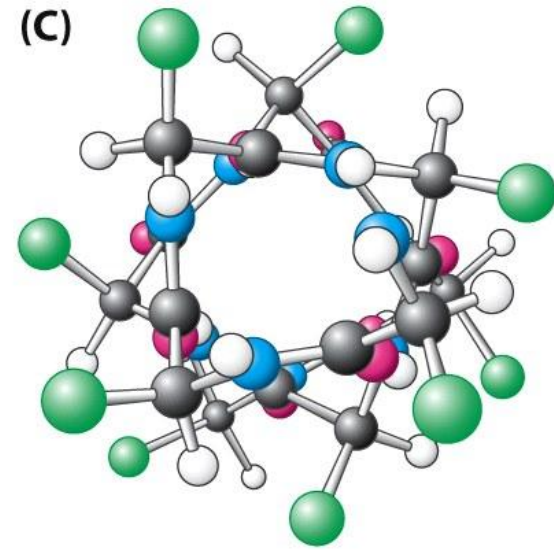
(A)



(B)



(C)



(D)

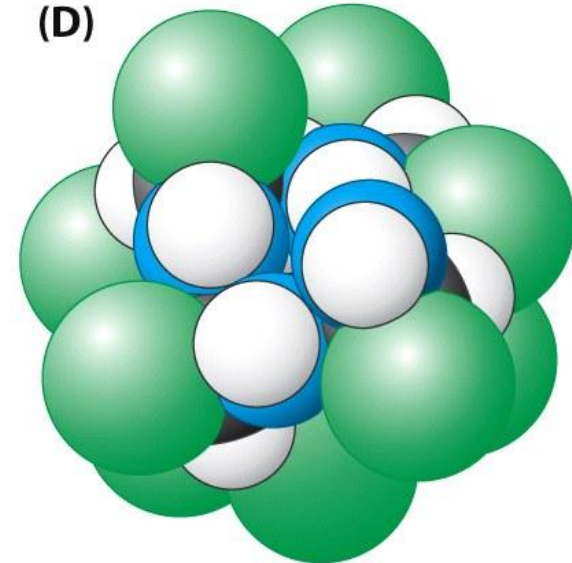
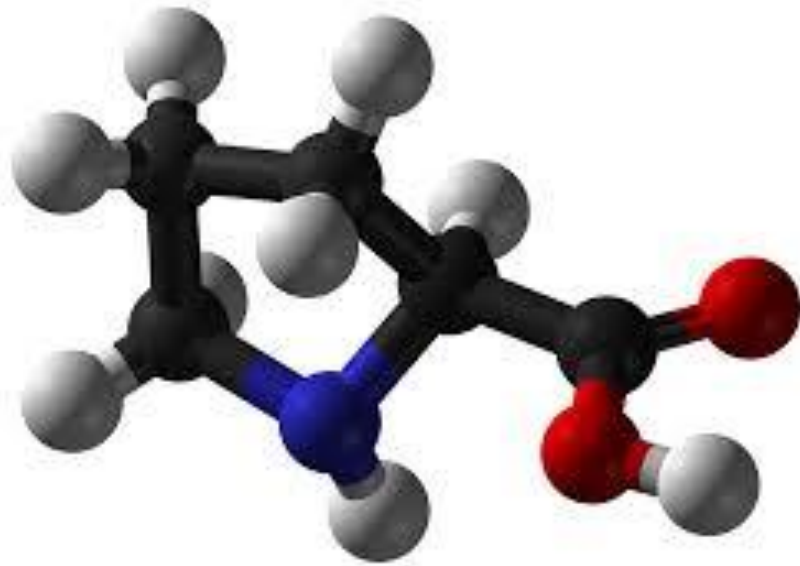


Figure 2.24
Biochemistry, Seventh Edition
© 2012 W. H. Freeman and Company



Prolina

Plote de Ramachandram

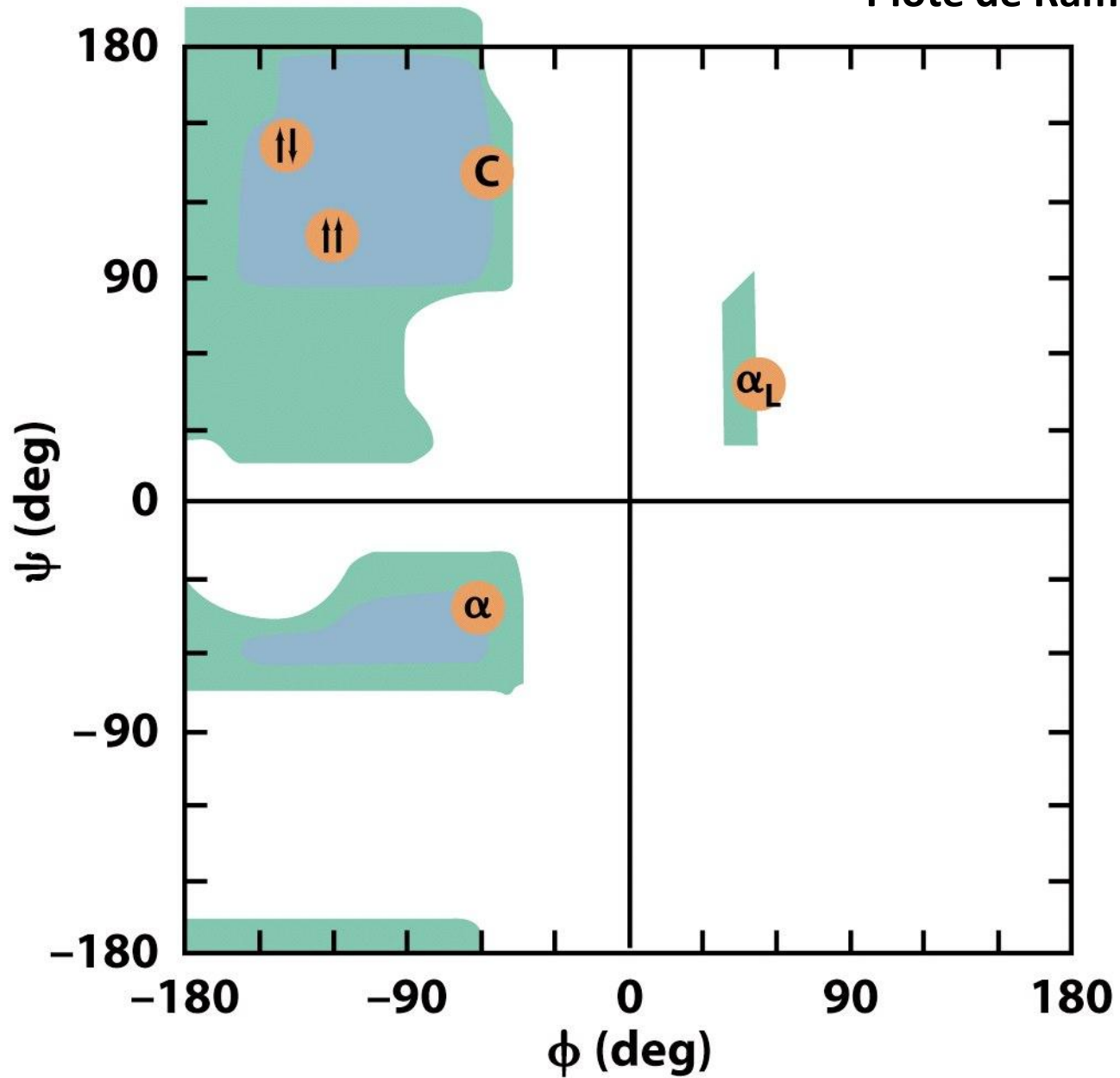


Figure 6-6 Fundamentals of Biochemistry, 2/e
© 2006 John Wiley & Sons

Fita β

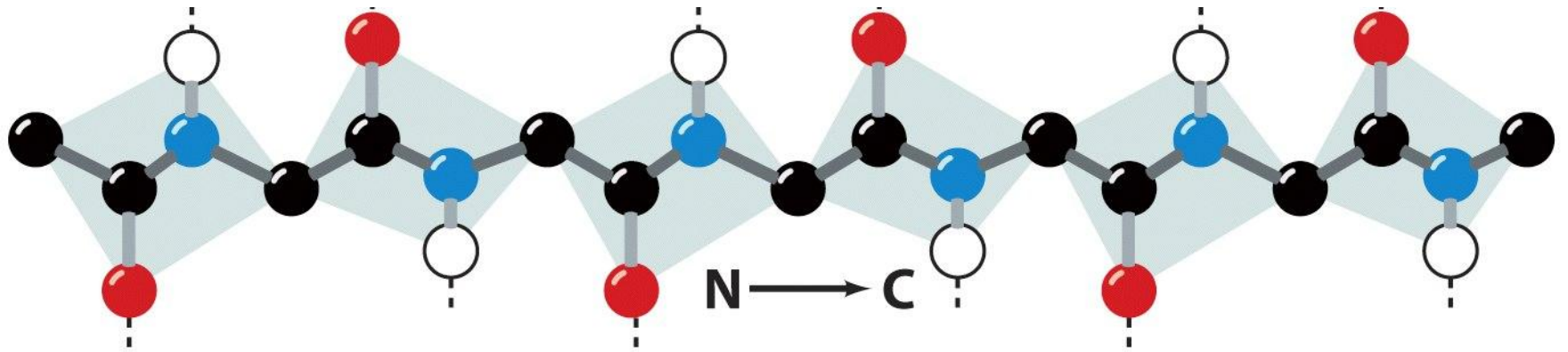


Figure 6-9a Fundamentals of Biochemistry, 2/e

Folha β β sheet

Antiparallel

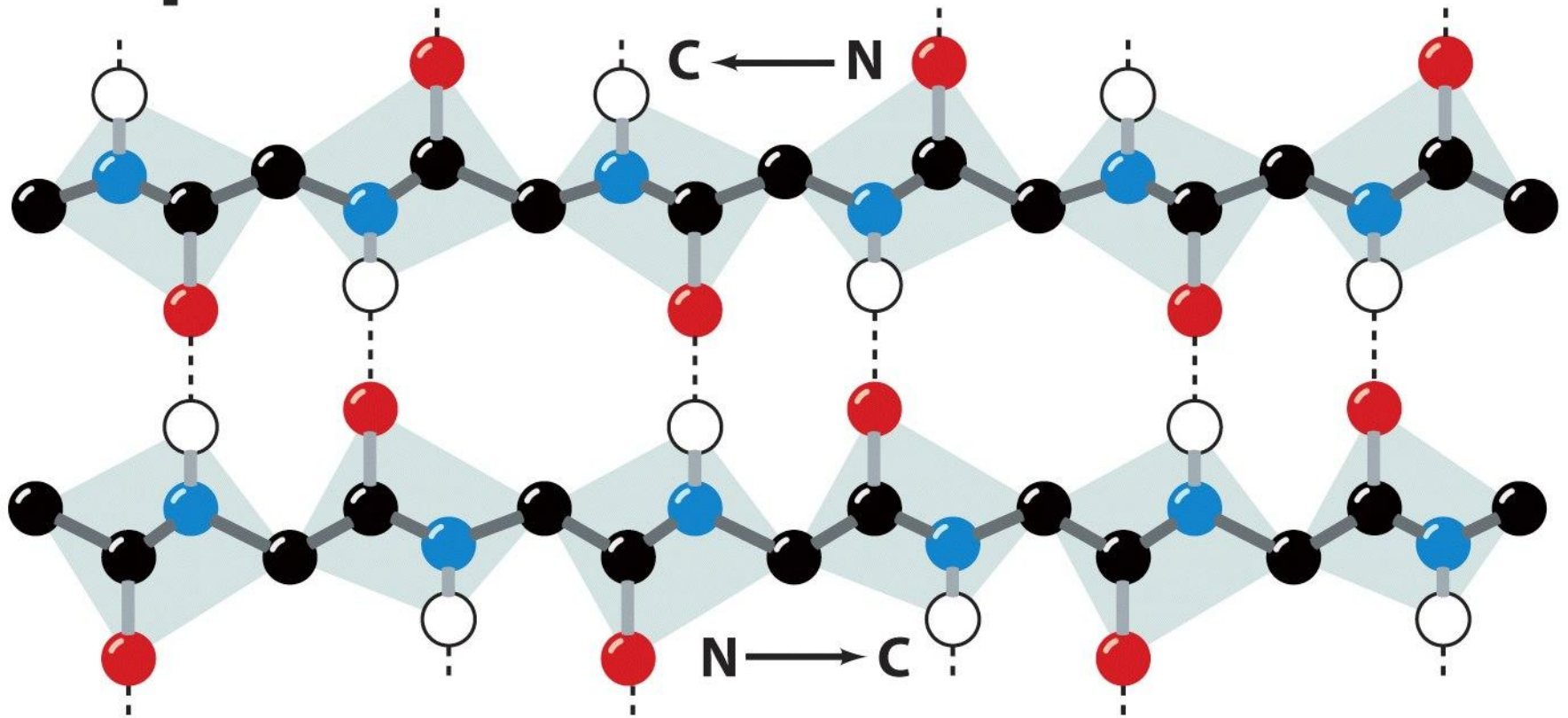


Figure 6-9a Fundamentals of Biochemistry, 2/e

Folha β β sheet

Parallel

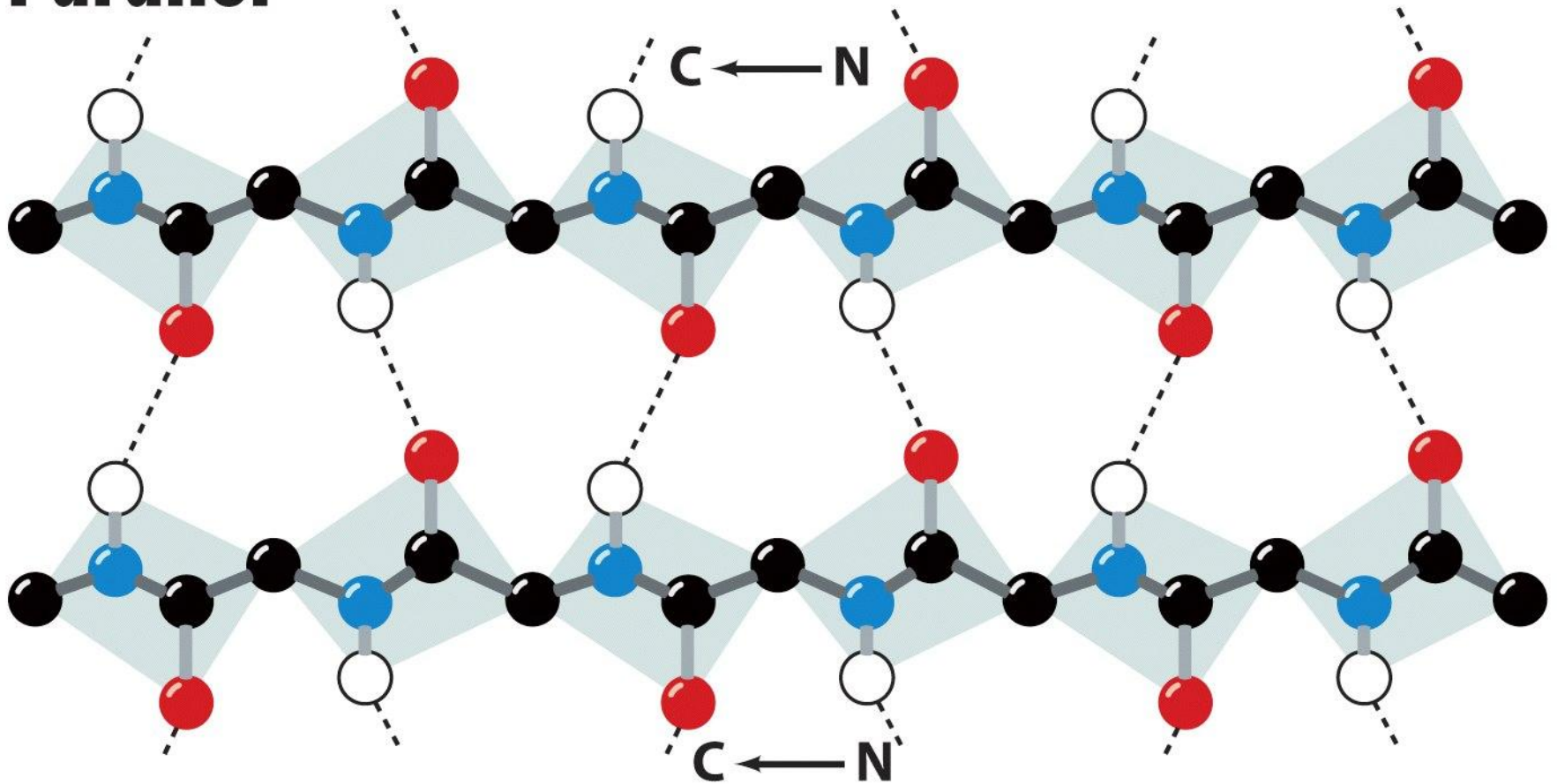
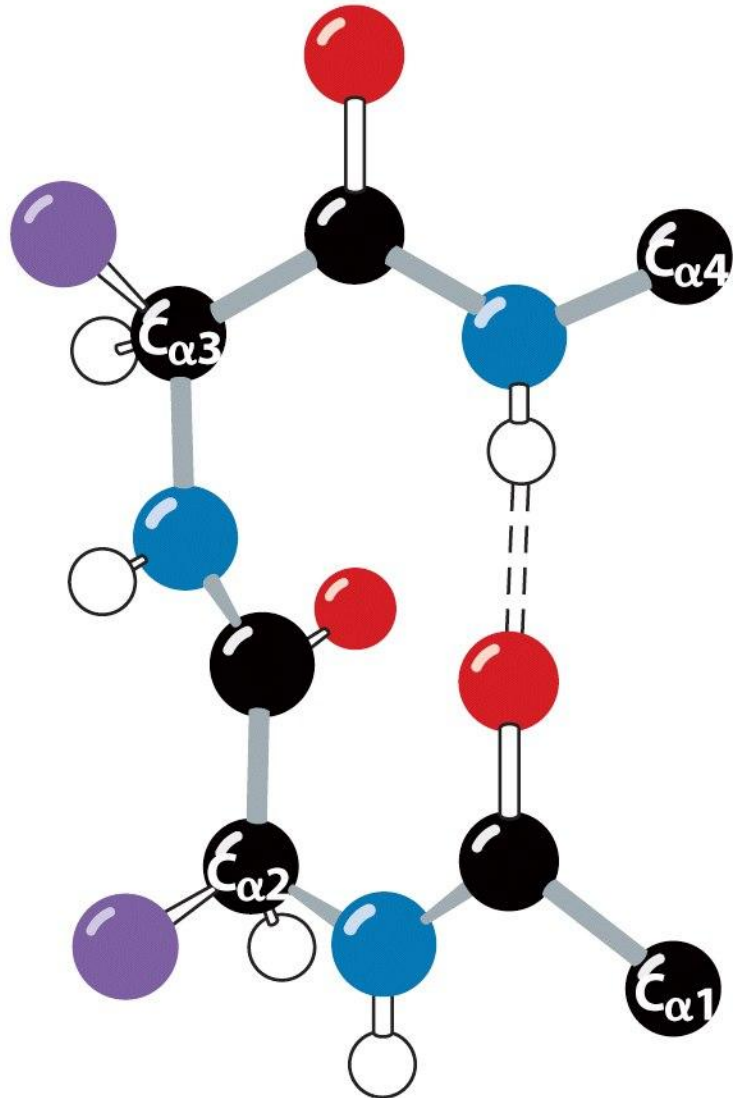


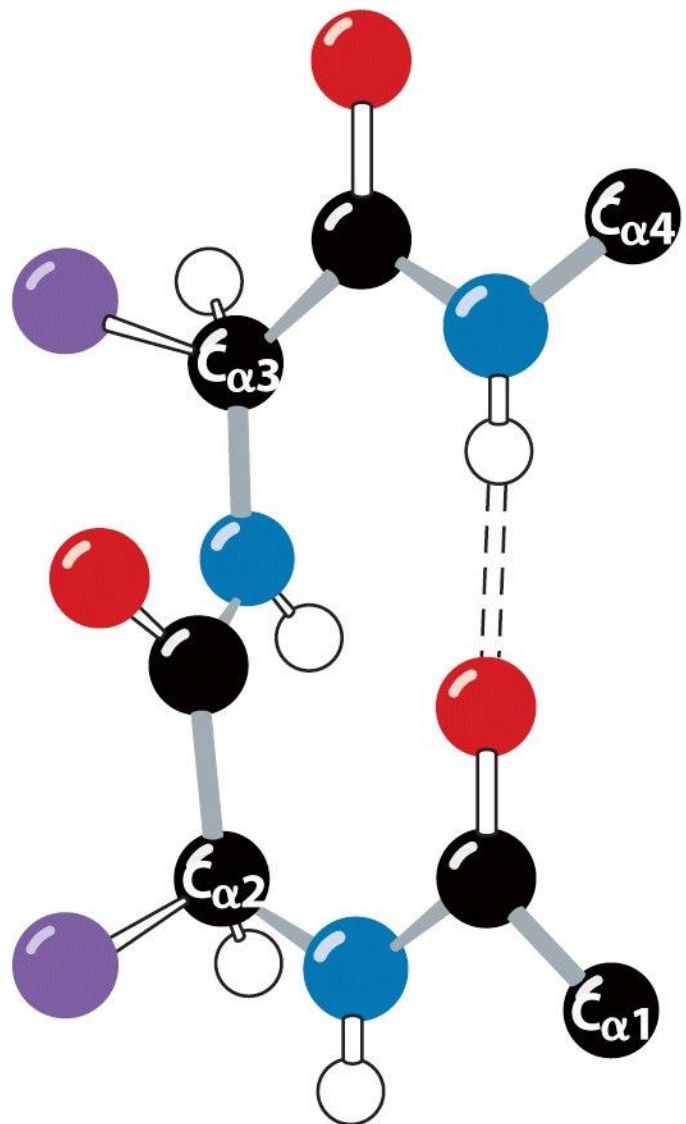
Figure 6-9b Fundamentals of Biochemistry, 2/e

Reverse Turn, Loop, Alça

(a) Type I



(b) Type II



Estrutura Secundária

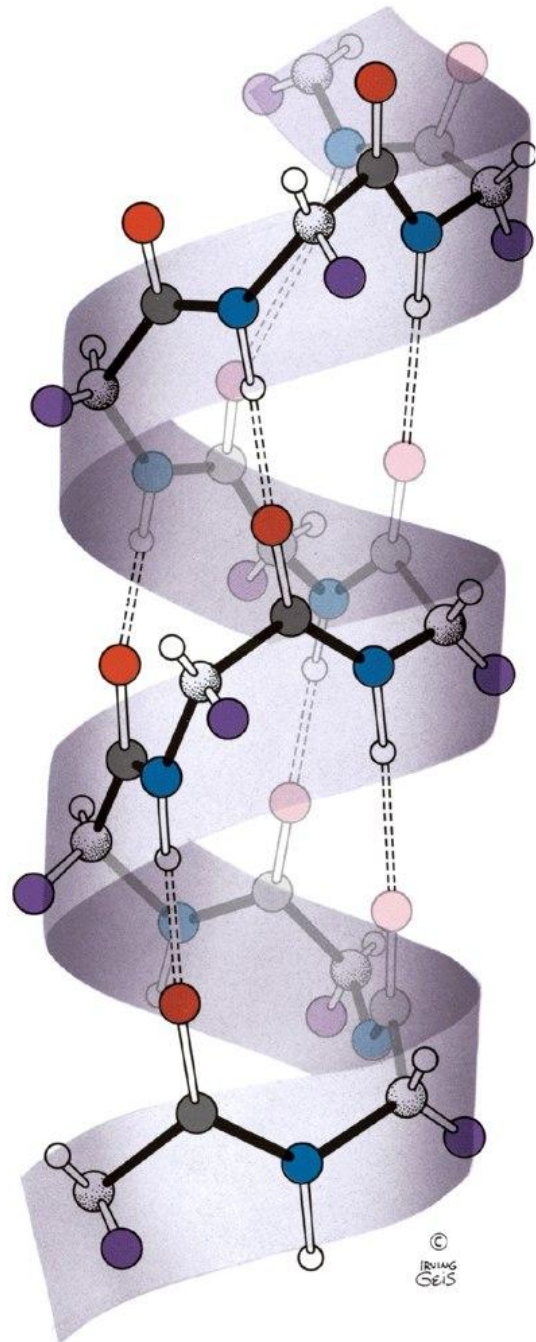


Figure 6-7 Fundamentals of Biochemistry, 2/e

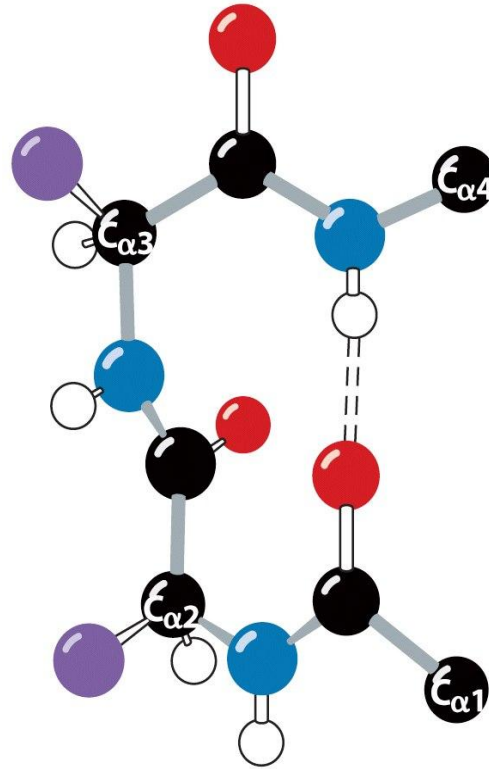
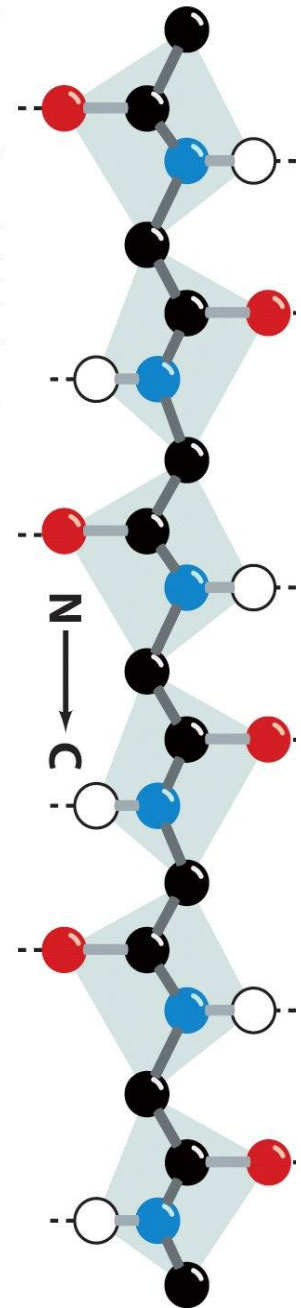


Figure 6-19 Fundamentals of Biochemistry, 2/e

Figure 6-9a Fundamentals of Biochemistry, 2/e



α -hélice



alça



fita β



Motivos (motifs)

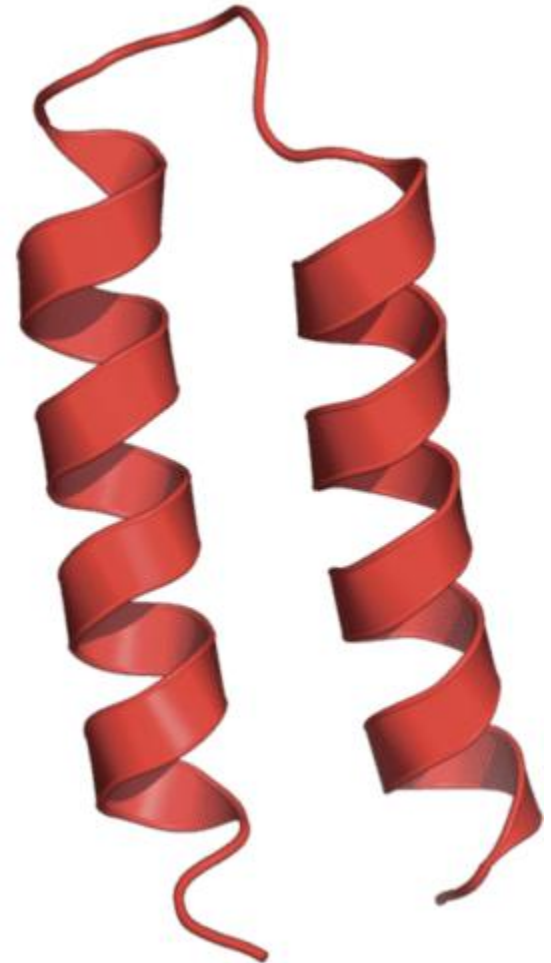
grampo β
 β - hairpin



β - α - β



Grampo α
 α - α



Motivos (motifs)

Greek Key Estilo Grego

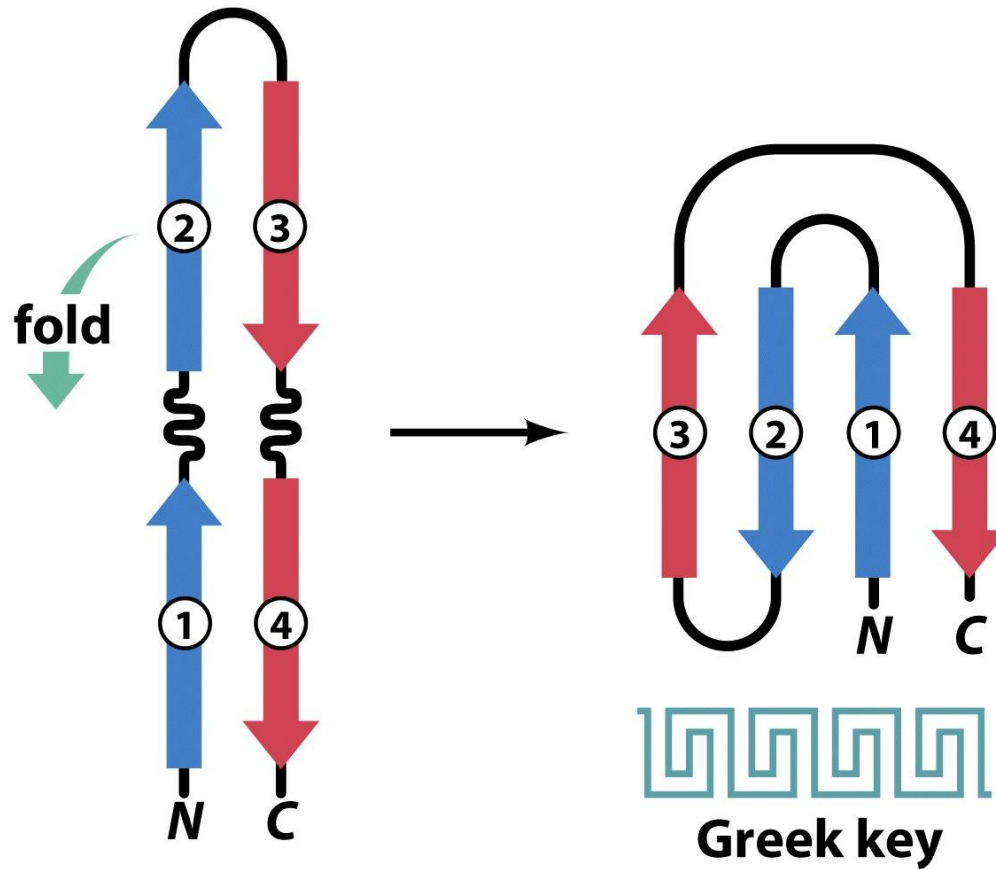
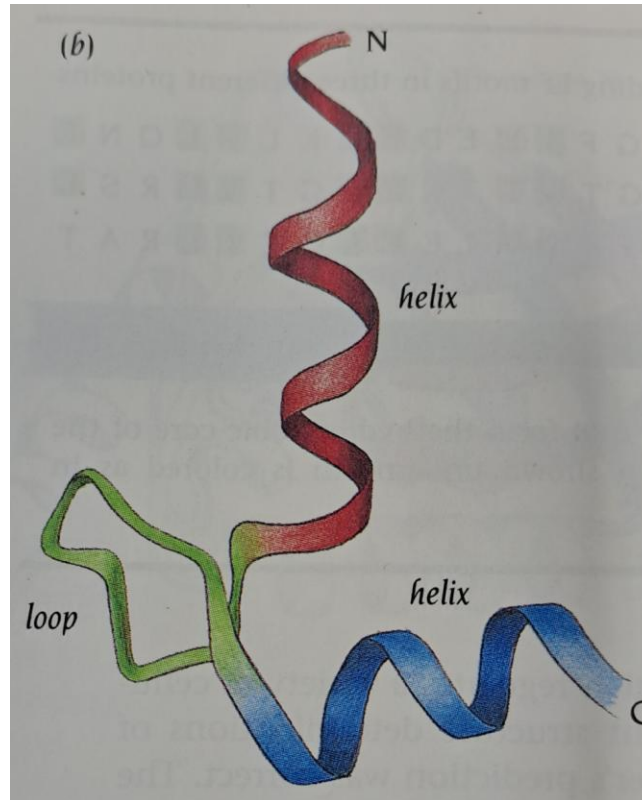


Figure 6-29d Fundamentals of Biochemistry, 2/e
© 2006 John Wiley & Sons

Motivos (motifs)

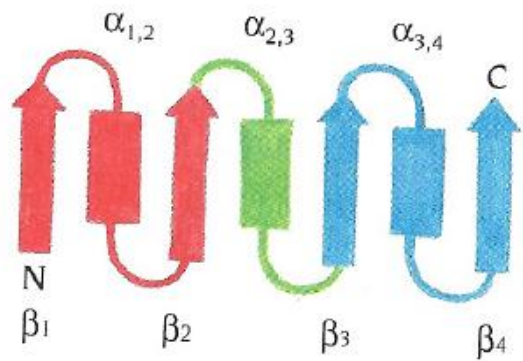
Helix-Loop-Helix Hélice-alça-hélice EF hand



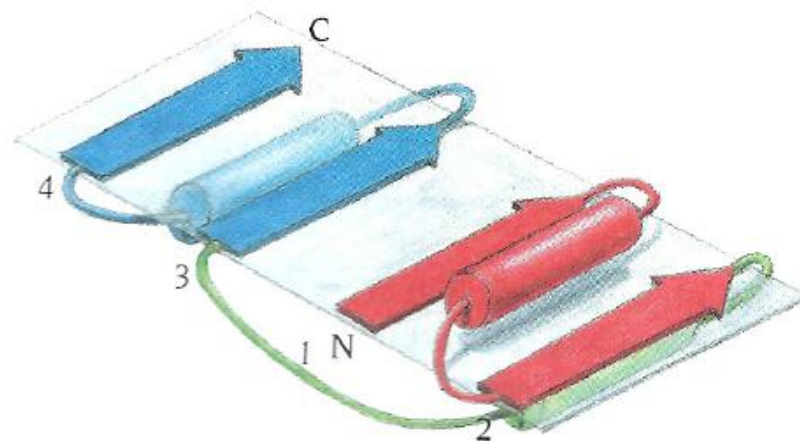
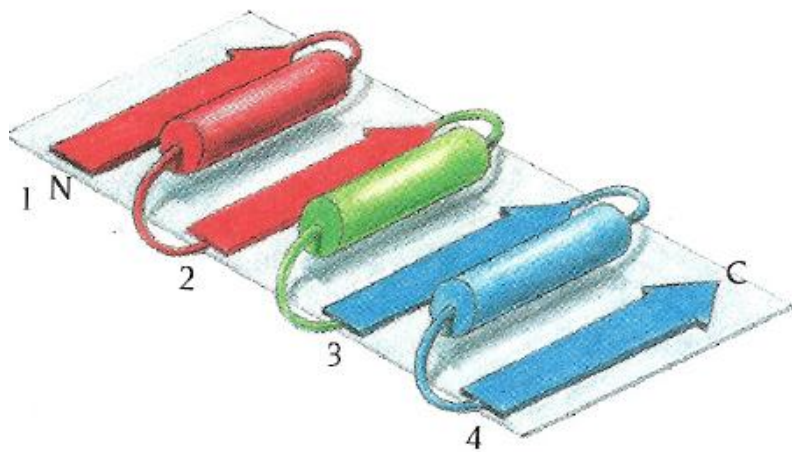
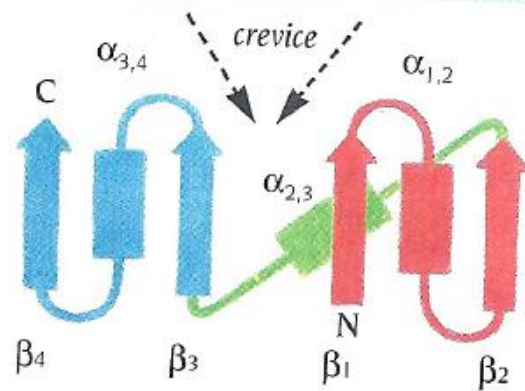
β - α - β



(a) motif 1 motif 2



(b) motif 2 motif 1



Closed barrel

Barrel α/β

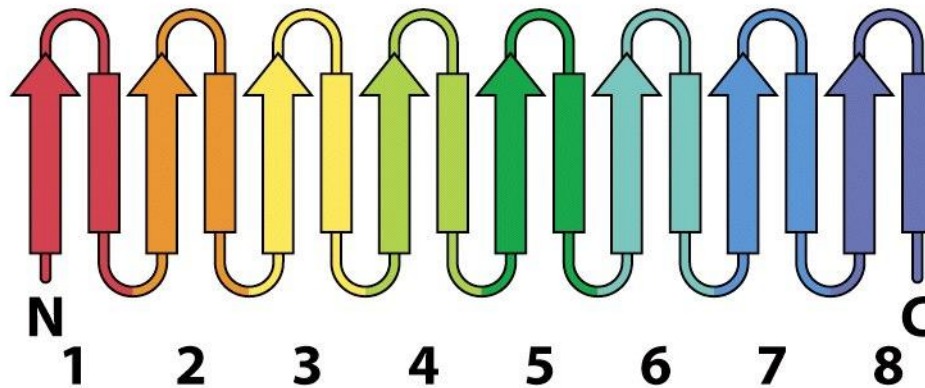
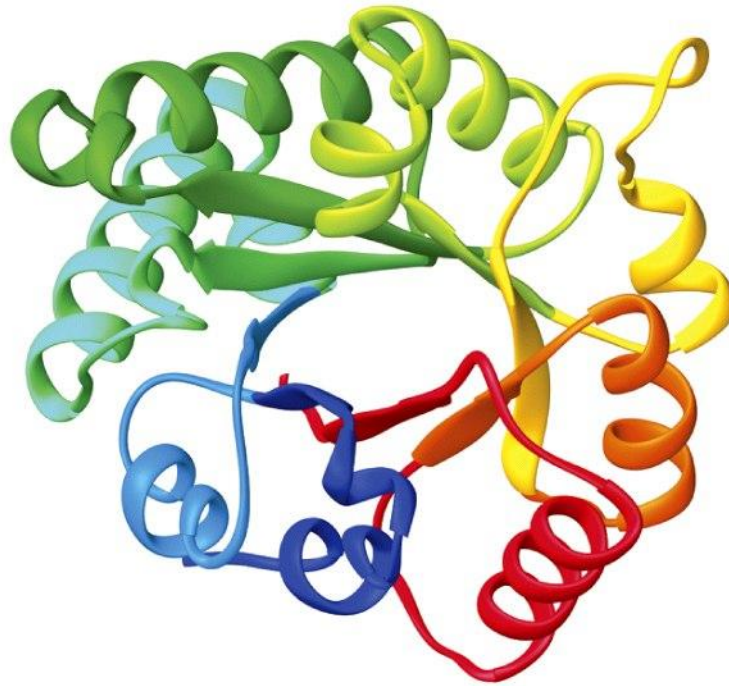
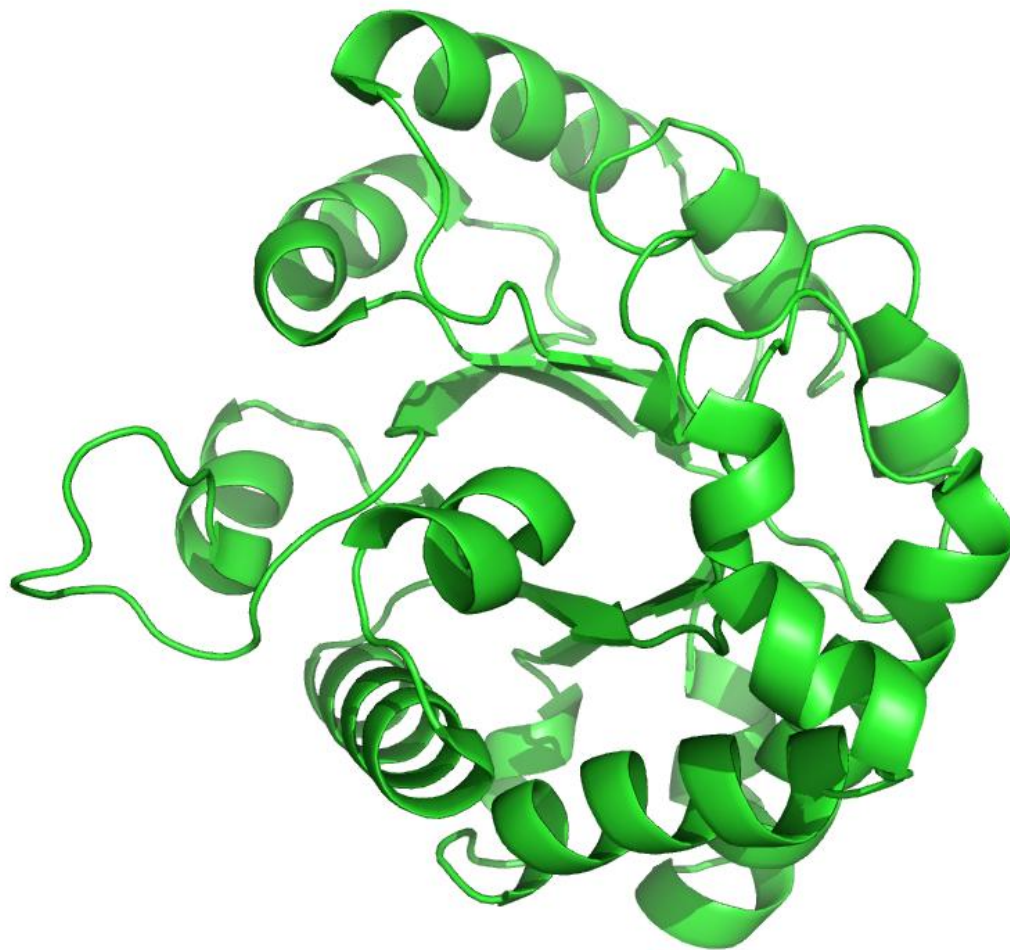


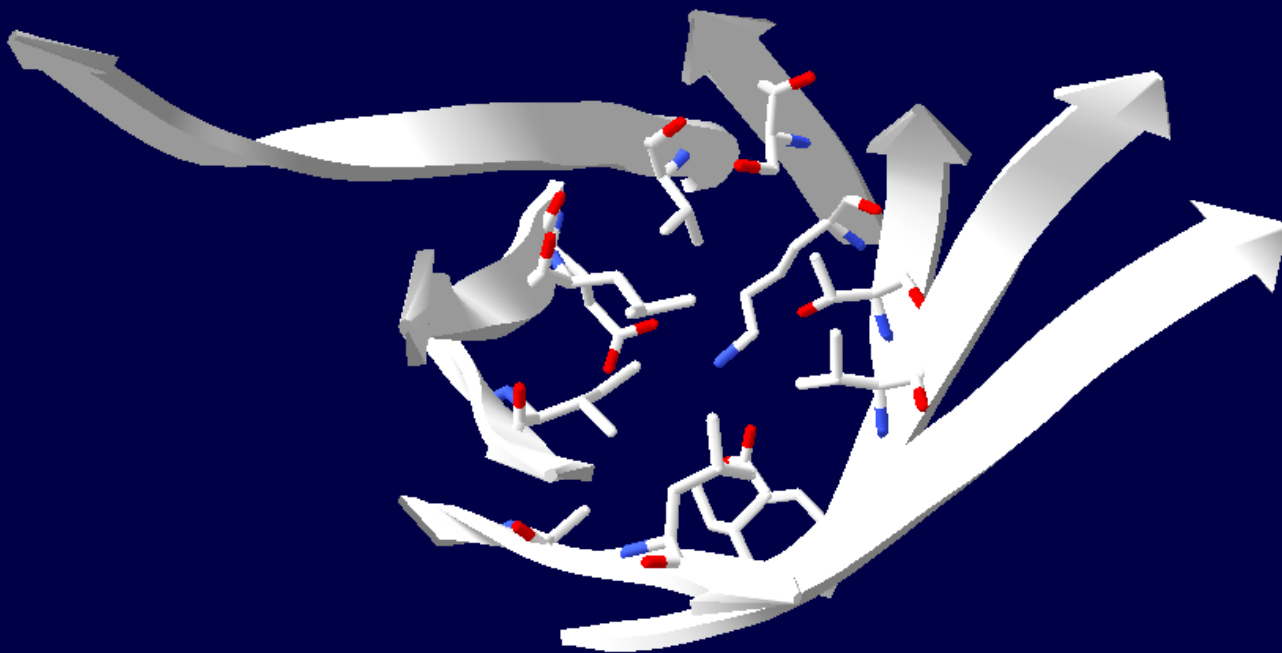
Figure 6-30c Fundamentals of Biochemistry, 2/e

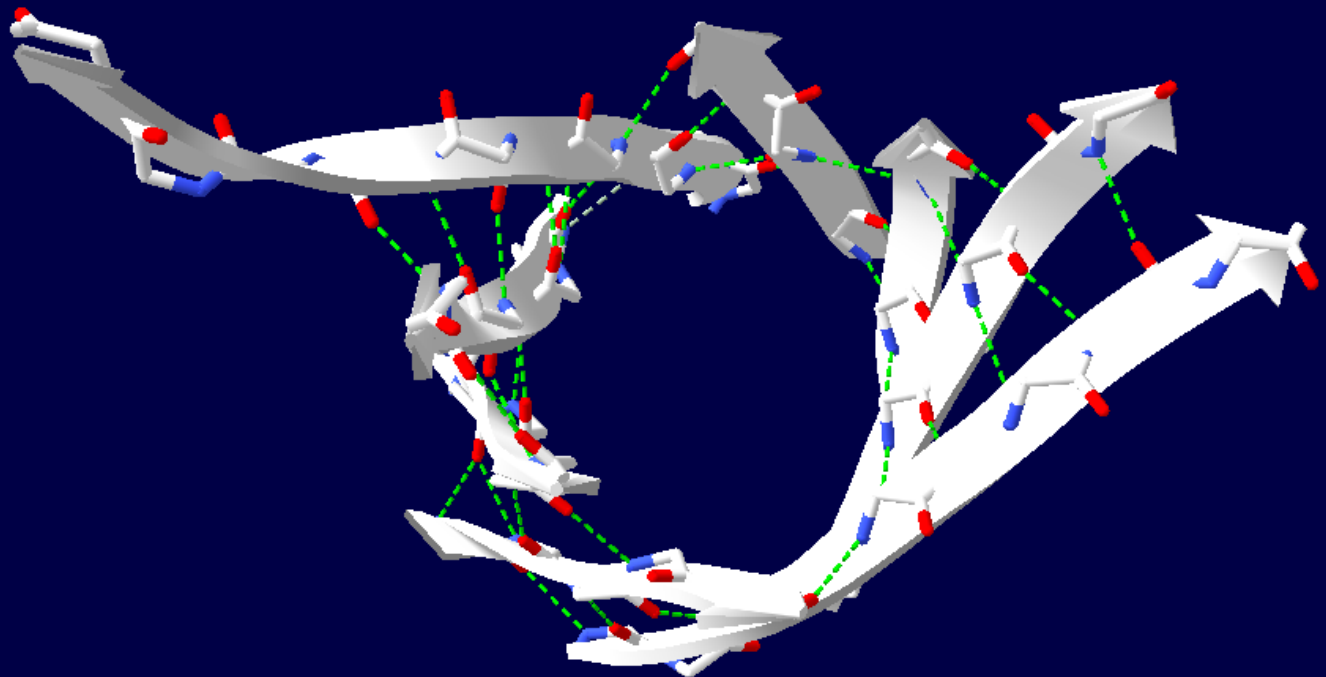
TIM barrel

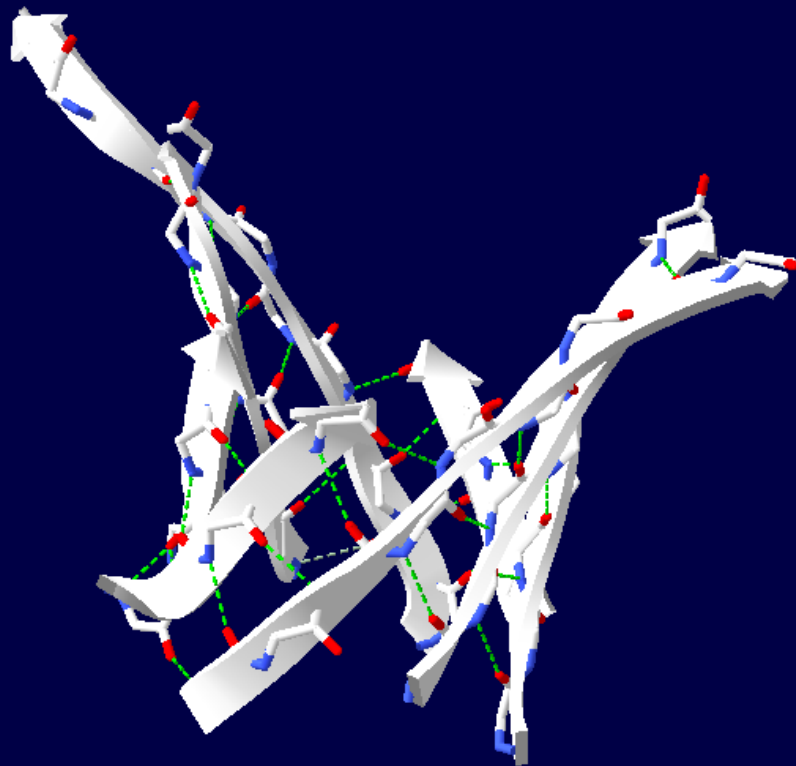
Barril (β/α)₈

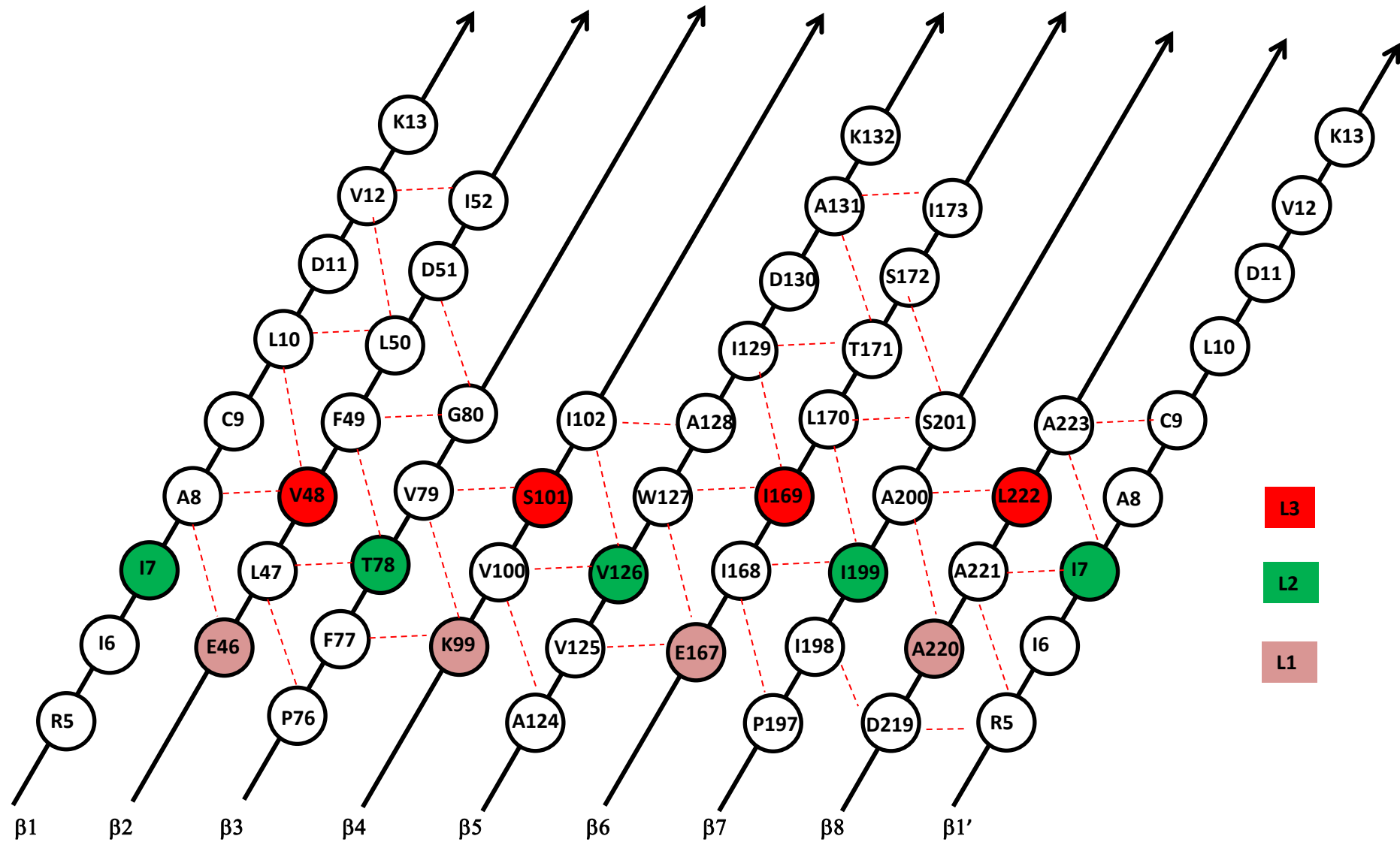


PDB 1VXN









Open twisted sheet

Folha aberta torcida

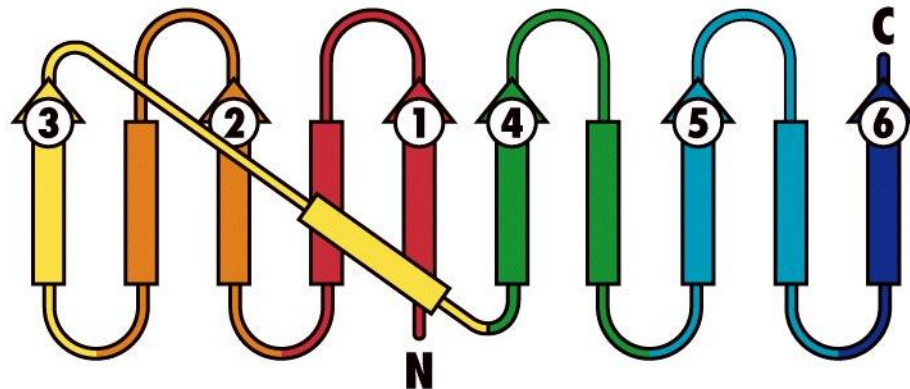
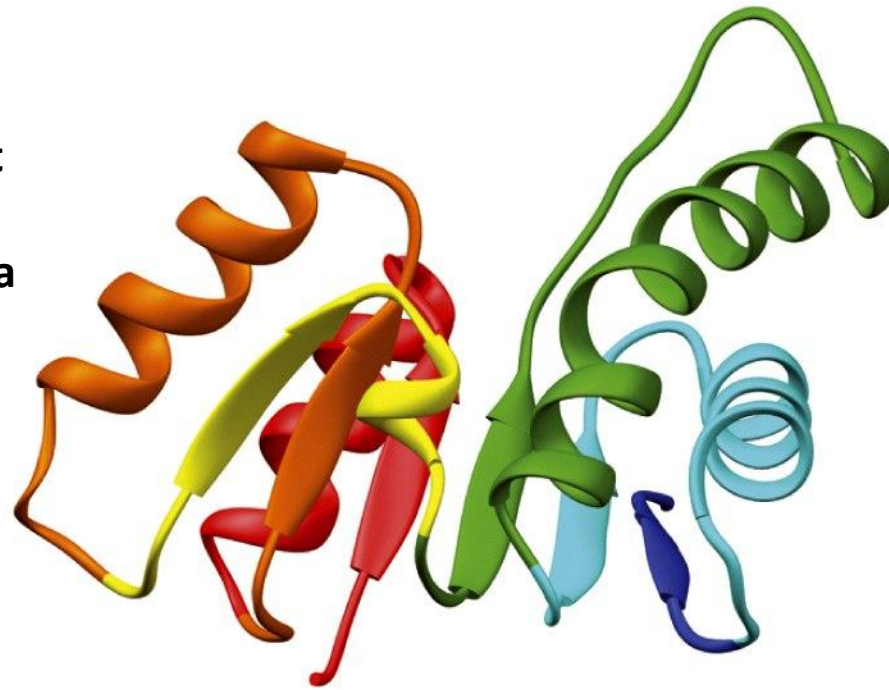
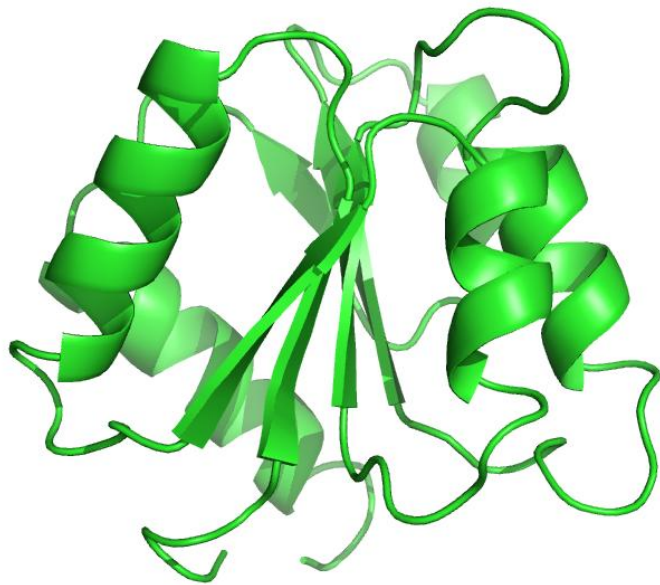


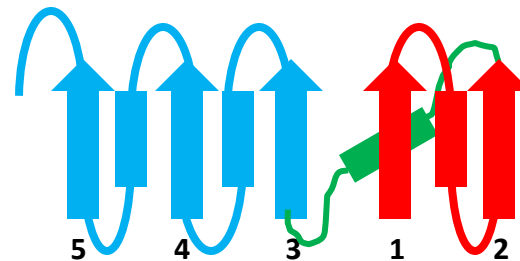
Figure 6-28c Fundamentals of Biochemistry, 2/e
© 2006 John Wiley & Sons

Open twisted sheet

Folha aberta torcida



Three layer sandwich
Sanduiche de 3 camadas (aba)
Rossmann fold



PDB 1CHY

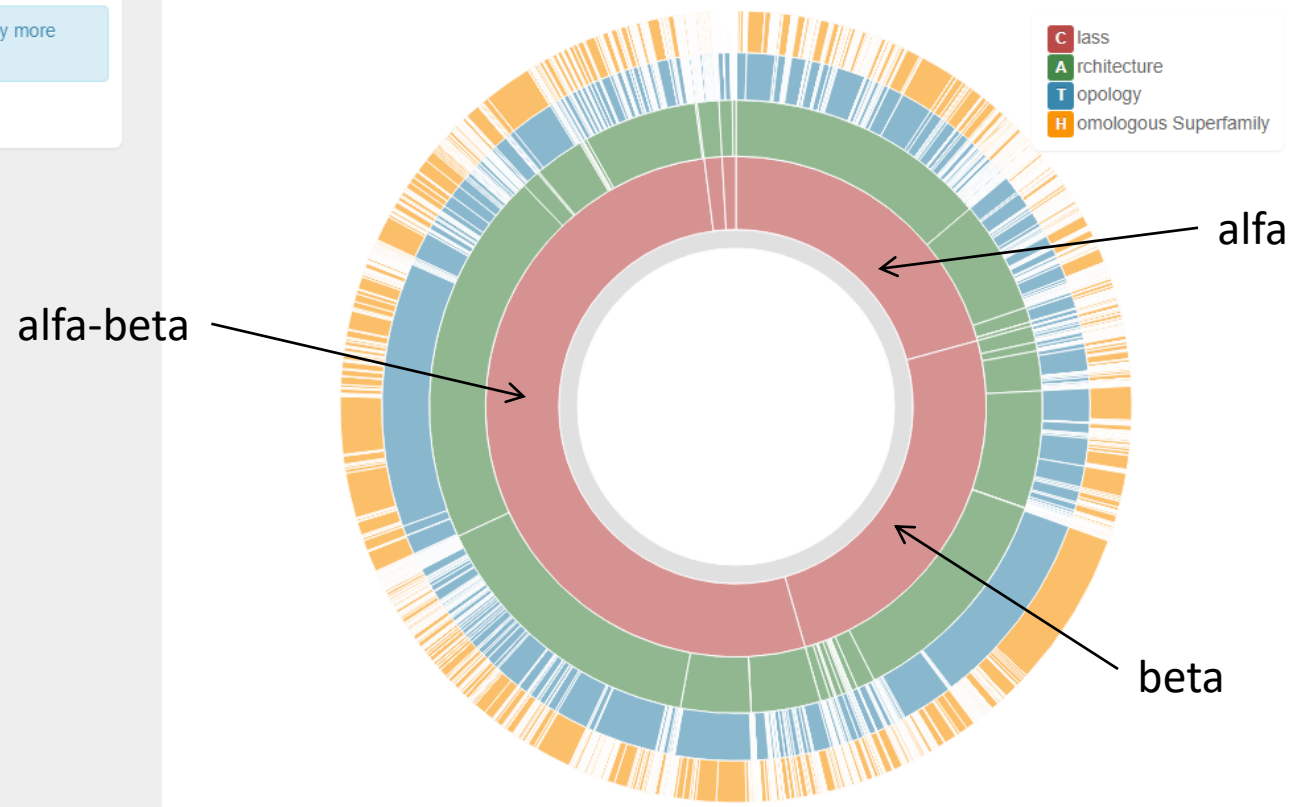
Top of CATH Hierarchy (4 Classes)

▷	C	1	Mainly Alpha	<i>5 Architectures, 404 Folds, 2033 Superfamilies, 103788 Domains</i>
▷	C	2	Mainly Beta	<i>21 Architectures, 244 Folds, 1290 Superfamilies, 124032 Domains</i>
▷	C	3	Alpha Beta	<i>14 Architectures, 634 Folds, 2337 Superfamilies, 262275 Domains</i>
▷	C	4	Few Secondary Structures	<i>1 Architectures, 108 Folds, 181 Superfamilies, 5716 Domains</i>
▷	C	6	Special	<i>2 Architectures, 82 Folds, 790 Superfamilies, 4427 Domains</i>

500238 domínios

Select a CATH node...
Click on a node to display more information in this panel

- C lass
- A rchitecture
- T opology
- H omologous Superfamily



This interactive sunburst plot lets you browse the tree-like hierarchy of CATH.
The innermost layer splits CATH into the 4 main CATH **C** lasses. Successive layers split these **C** lasses into **A**

Class

Classe é o nível mais alto na classificação. Refere-se, de modo geral, ao conteúdo de elementos de estrutura secundária. Há 4 classes:

Mainly Alpha – Predominantemente Alfa

Mainly Beta – Predominantemente Beta

Alpha Beta – Combinação de Alfa e Beta

Few Secondary Structures – Pouca estrutura secundária.

A classificação baseia-se em critérios como:

Composição percentual da estrutura secundária

Percentual de contatos entre elementos de estrutura secundária

Alternância entre elementos de estrutura secundária

Architecture

A arquitetura é descrita pela forma geral do domínio proteico, a qual é determinada pela orientação dos elementos de estrutura secundária, porém sem considerar o modo de conexão entre eles.

Por exemplo: um barril, um sanduíche, um feixe de hélices.

Atualmente há 43 arquiteturas proteicas.

Topology (Fold Group)

A topologia (um grupo de dobramento, *fold*) refere-se à forma geral e ao modo como estão conectados os elementos de estrutura secundária do domínio proteico.

Há casos em que proteínas pertencem à mesma topologia, porém exibem elementos adicionais em torno do domínio principal.

Atualmente há 1472 topologias (*folds*).

Homologous Superfamily

Este nível de classificação agrupa proteínas que compartilham um ancestral comum. São homólogas. São identificadas por comparação de estrutura primária (sequência de aminoácidos) ou sobreposição de estrutura terciária.

Identidade de Sequência $\geq 35\%$, Sobreposição Estrutural $\geq 60\%$

Atualmente há 6631 superfamílias de proteínas homólogas.

BROWSE LINKS

Browse Hierarchy

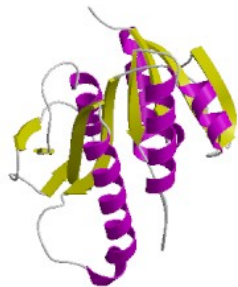
- Highly Diverse Superfamilies
- Superfamily Comparison

Select a CATH node...

C Alpha Beta

3

CATH ID	3
Architectures	14
Topologies	634
Superfamilies	2337
Domains	262275
Example Domain	1hdoA00 [PDB]



Tree



Sunburst

Top of CATH Hierarchy (4 Classes)

- ▶ **C** 1 Mainly Alpha *5 Architectures, 404 Folds, 2033 Superfamilies, 103788 Domains*
- ▶ **C** 2 Mainly Beta *21 Architectures, 244 Folds, 1290 Superfamilies, 124032 Domains*
- ▶ **C** 3 Alpha Beta *14 Architectures, 634 Folds, 2337 Superfamilies, 262275 Domains*
- ▶ **C** 4 Few Secondary Structures *1 Architectures, 108 Folds, 181 Superfamilies, 5716 Domains*
- ▶ **C** 6 Special *2 Architectures, 82 Folds, 790 Superfamilies, 4427 Domains*

BROWSE LINKS

Browse Hierarchy

[Highly Diverse Superfamilies](#)

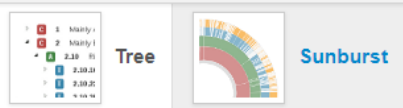
[Superfamily Comparison](#)

Select a CATH node...

C Alpha Beta

3

CATH ID	3
Architectures	14
Topologies	634
Superfamilies	2337
Domains	262275
Example Domain	1hdoA00 [PDB]

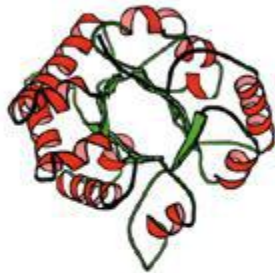


Top of CATH Hierarchy (4 Classes)

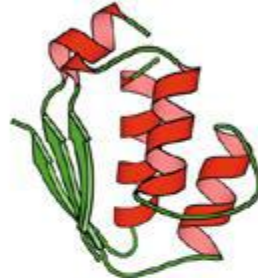
- ▶ **C** 1 Mainly Alpha *5 Architectures, 404 Folds, 2033 Superfamilies, 103788 Domains*
- ▶ **C** 2 Mainly Beta *21 Architectures, 244 Folds, 1290 Superfamilies, 124032 Domains*
- ▶ **C** 3 Alpha Beta *14 Architectures, 634 Folds, 2337 Superfamilies, 262275 Domains*
 - ▶ **A** 3.10 Roll *60 Folds, 227 Superfamilies, 18308 Domains*
 - ▶ **A** 3.15 Super Roll *3 Folds, 8 Superfamilies, 65 Domains*
 - ▶ **A** 3.20 Alpha-Beta Barrel *18 Folds, 60 Superfamilies, 18478 Domains*
 - ▶ **A** 3.30 2-Layer Sandwich *224 Folds, 1125 Superfamilies, 75764 Domains*
 - ▶ **A** 3.40 3-Layer(aba) Sandwich *126 Folds, 552 Superfamilies, 99008 Domains*
 - ▶ **A** 3.50 3-Layer(bba) Sandwich *11 Folds, 25 Superfamilies, 4765 Domains*
 - ▶ **A** 3.55 3-Layer(bab) Sandwich *6 Folds, 13 Superfamilies, 469 Domains*
 - ▶ **A** 3.60 4-Layer Sandwich *16 Folds, 29 Superfamilies, 13069 Domains*
 - ▶ **A** 3.65 Alpha-beta prism *1 Folds, 2 Superfamilies, 494 Domains*
 - ▶ **A** 3.70 Box *1 Folds, 1 Superfamilies, 369 Domains*
 - ▶ **A** 3.75 5-stranded Propeller *1 Folds, 2 Superfamilies, 207 Domains*
 - ▶ **A** 3.80 Alpha-Beta Horseshoe *3 Folds, 5 Superfamilies, 867 Domains*
 - ▶ **A** 3.90 Alpha-Beta Complex *163 Folds, 286 Superfamilies, 29927 Domains*
 - ▶ **A** 3.100 Ribosomal Protein L15; Chain: K; domain 2 *1 Folds, 2 Superfamilies, 485 Domains*
 - ▶ **C** 4 Few Secondary Structures *1 Architectures, 108 Folds, 181 Superfamilies, 5716 Domains*
 - ▶ **C** 6 Special *2 Architectures, 82 Folds, 790 Superfamilies, 4427 Domains*

CATH – a hierarchic classification of protein domain structures

CA Orengo, coautores e JM Thornton
Structure 5, 1093-1109 (1997)



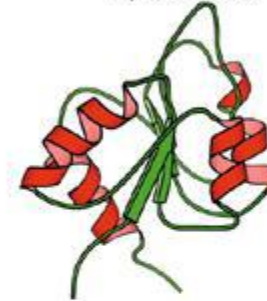
$\alpha\beta$ Barrel (4timA)



$\alpha\beta$ 2-Layer Sandwich (1brsD)



$\alpha\beta$ Roll (1std)



$\alpha\beta$ 3-Layer Sandwich (aba) (1ntr)



$\alpha\beta$ 3-Layer Sandwich (bba) (1pyaB)



$\alpha\beta$ 4-Layer Sandwich (2dnjA)



$\alpha\beta$ Box (1plq)



$\alpha\beta$ Horseshoe (1bnh)



$\alpha\beta$ Complex (1pyp)



$\alpha\beta$ Few SS (1pyaB)

Select a CATH node...

A Alpha-Beta Barrel	3.20
CATH ID	3.20
Topologies	18
Superfamilies	60
Domains	18478
Example Domain	2vxmA00 [PDB]



- ▶ **C** 1 Mainly Alpha *5 Architectures, 404 Folds, 2033 Superfamilies, 103788 Domains*
- ▶ **C** 2 Mainly Beta *21 Architectures, 244 Folds, 1290 Superfamilies, 124032 Domains*
- ▶ **C** 3 Alpha Beta *14 Architectures, 634 Folds, 2337 Superfamilies, 262275 Domains*
 - ▶ **A** 3.10 Roll *60 Folds, 227 Superfamilies, 18308 Domains*
 - ▶ **A** 3.15 Super Roll *3 Folds, 8 Superfamilies, 65 Domains*
 - ▶ **A** 3.20 Alpha-Beta Barrel *18 Folds, 60 Superfamilies, 18478 Domains*
 - ▶ **T** 3.20.10 D-amino Acid Aminotransferase; Chain A, domain 2 *1 Superfamilies, 282 Domains*
 - ▶ **T** 3.20.14 L-fucose Isomerase; Chain A, domain 3 *1 Superfamilies, 21 Domains*
 - ▶ **T** 3.20.16 Serine Protease, Human Cytomegalovirus Protease; Chain A *1 Superfamilies, 73 Domains*
 - ▶ **T** 3.20.19 Aconitase; domain 4 *1 Superfamilies, 46 Domains*
 - ▶ **T** 3.20.20 TIM Barrel *34 Superfamilies, 17384 Domains*
 - ▶ **T** 3.20.70 Anaerobic Ribonucleotide-triphosphate Reductase Large Chain *1 Superfamilies, 239 Domains*
 - ▶ **T** 3.20.80 Multidrug-efflux Transporter 1 Regulator Bmrr; Chain A *1 Superfamilies, 56 Domains*
 - ▶ **T** 3.20.90 Tubby Protein; Chain A *1 Superfamilies, 7 Domains*
 - ▶ **T** 3.20.100 mRNA Triphosphatase Cet1; Chain A *3 Superfamilies, 38 Domains*
 - ▶ **T** 3.20.110 7-stranded beta/alpha barrel *2 Superfamilies, 74 Domains*
 - ▶ **T** 3.20.120 hfbii hydrophobin *1 Superfamilies, 29 Domains*
 - ▶ **T** 3.20.130 fumarase *1 Superfamilies, 5 Domains*
 - ▶ **T** 3.20.140 nicotinate phosphoribosyltransferase *1 Superfamilies, 22 Domains*
 - ▶ **T** 3.20.160 pseudo Tubby roll *1 Superfamilies, 4 Domains*
 - ▶ **T** 3.20.170 ADP-ribosylation fold *5 Superfamilies, 8 Domains*
 - ▶ **T** 3.20.180 Split barrel-like *2 Superfamilies, 31 Domains*
 - ▶ **T** 3.20.190 N-terminal domain of MutM-like DNA repair proteins *2 Superfamilies, 121 Domains*
 - ▶ **T** 3.20.200 Protein kinase-like fold *1 Superfamilies, 38 Domains*
 - ▶ **A** 3.30 2-Layer Sandwich *224 Folds, 1125 Superfamilies, 75764 Domains*
 - ▶ **A** 3.40 3-Layer(aba) Sandwich *126 Folds, 552 Superfamilies, 99008 Domains*
 - ▶ **A** 3.50 3-Layer(bba) Sandwich *11 Folds, 25 Superfamilies, 4765 Domains*
 - ▶ **A** 3.55 3-Layer(bab) Sandwich *6 Folds, 13 Superfamilies, 469 Domains*
 - ▶ **A** 3.60 4-Layer Sandwich *16 Folds, 29 Superfamilies, 13069 Domains*
 - ▶ **A** 3.65 Alpha-beta prism *1 Folds, 2 Superfamilies, 494 Domains*

Select a CATH node...

T TIM Barrel	
3.20.20	
CATH ID	3.20.20
Superfamilies	34
Domains	17384
Example Domain	2vxmA00 [PDB]



- ▶ **T** 3.20.16 Serine Protease, Human Cytomegalovirus Protease; Chain A 1 Superfamilies, 73 Domains
- ▶ **T** 3.20.19 Aconitase; domain 4 1 Superfamilies, 46 Domains
- ▶ **T** 3.20.20 TIM Barrel 34 Superfamilies, 17384 Domains
 - H** 3.20.20.10 Alanine racemase 311 Domains
 - H** 3.20.20.20 Dihydropteroate synthase-like 249 Domains
 - H** 3.20.20.30 Luciferase-like domain 91 Domains
 - H** 3.20.20.40 1, 4-beta cellobiohydrolase 79 Domains
 - H** 3.20.20.60 Phosphoenolpyruvate-binding domains 708 Domains
 - H** 3.20.20.70 Aldolase class I 6031 Domains
 - H** 3.20.20.80 Glycosidases 4161 Domains
 - H** 3.20.20.100 NADP-dependent oxidoreductase domain 604 Domains
 - H** 3.20.20.105 Queuine tRNA-ribosyltransferase-like 161 Domains
 - H** 3.20.20.110 Ribulose biphosphate carboxylase, large subunit, C-terminal domain 485 Domains
 - H** 3.20.20.120 Enolase-like C-terminal domain 1364 Domains
 - H** 3.20.20.140 Metal-dependent hydrolases 1409 Domains
 - H** 3.20.20.150 Divalent-metal-dependent TIM barrel enzymes 625 Domains
 - H** 3.20.20.190 Phosphatidylinositol (PI) phosphodiesterase 150 Domains
 - H** 3.20.20.210 Not yet named 109 Domains
 - H** 3.20.20.220 Not yet named 71 Domains
 - H** 3.20.20.240 Methylmalonyl-CoA mutase 43 Domains
 - H** 3.20.20.300 Glycoside hydrolase, family 3, N-terminal domain 212 Domains
 - H** 3.20.20.330 Homocysteine-binding-like domain 33 Domains
 - H** 3.20.20.350 Diol/glycerol dehydratase, large subunit 32 Domains
 - H** 3.20.20.360 Malate synthase, domain 3 71 Domains
 - H** 3.20.20.370 Glycoside hydrolase/deacetylase 111 Domains
 - H** 3.20.20.380 Copper homeostasis (CutC) domain 22 Domains
 - H** 3.20.20.390 FMN-linked oxidoreductases 61 Domains
 - H** 3.20.20.410 Protein of unknown function UPF0759 4 Domains
 - H** 3.20.20.440 D-Lysine 5,6-aminomutase alpha subunit 25 Domains
 - H** 3.20.20.450 EAL domain 92 Domains
 - H** 3.20.20.460 Mammalian-like methylase domain 60 Domains

DOMAIN LINKS

- Summary
- Structure
- Sequence
- Neighbourhood

CATH Classification

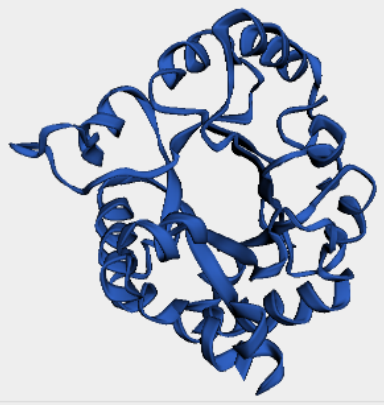
Level	CATH Code	Description
3	3	Alpha Beta
A	3.20	Alpha-Beta Barrel
T	3.20.20	TIM Barrel
H	3.20.20.70	Aldolase class I

PDB Structure

PDB	2VXN
External Links	<ul style="list-style-type: none">PDBSumProteopedia
Method	X-RAY DIFFRACTION
Organism	

Domain Context

View Domain in Chain



2vxn (A)

Primary Citation Atomic Resolution Crystallography of a Complex of Triosephosphate Isomerase with a Reaction-Intermediate Analog: New Insight in the Proton Transfer Reaction Mechanism. Alahuhta, M., Wierenga, R.K. *Proteins*

DOMAIN LINKS

Summary

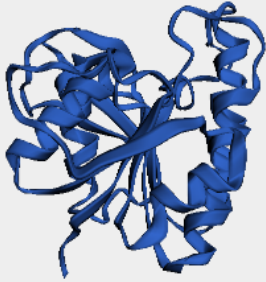
- Structure
- Sequence
- Neighbourhood

CATH Classification

Level	CATH Code	Description
🌀	3	Alpha Beta
🟢	3.40	3-Layer(aba) Sandwich
📘	3.40.50	Rossmann fold
🟡	3.40.50.720	NAD(P)-binding Rossmann-like Domain

Domain Context

View Domain in Chain ▾



1hdo (A)

PDB Structure

PDB 1HDO

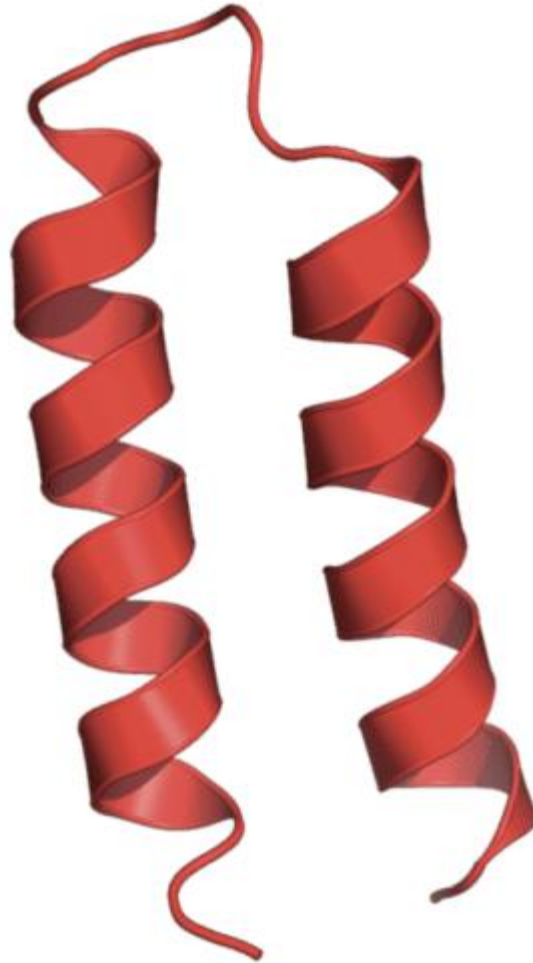
- External Links**
- [PDBSum](#)
 - [Proteopedia](#)

Method X-RAY DIFFRACTION

Organism

Primary Citation **Structure of Human Biliverdin Ix Beta Reductase, an Early Fetal Bilirubin Ix Producing Enzyme**
Pereira, P.J.B., Macedo-Ribeiro, S., Parraga, A., Perez-Luque, R., Cunningham, O., Darcy, K., Mantle, T.J., Coll, M.
Nat. Struct. Biol.

α - α



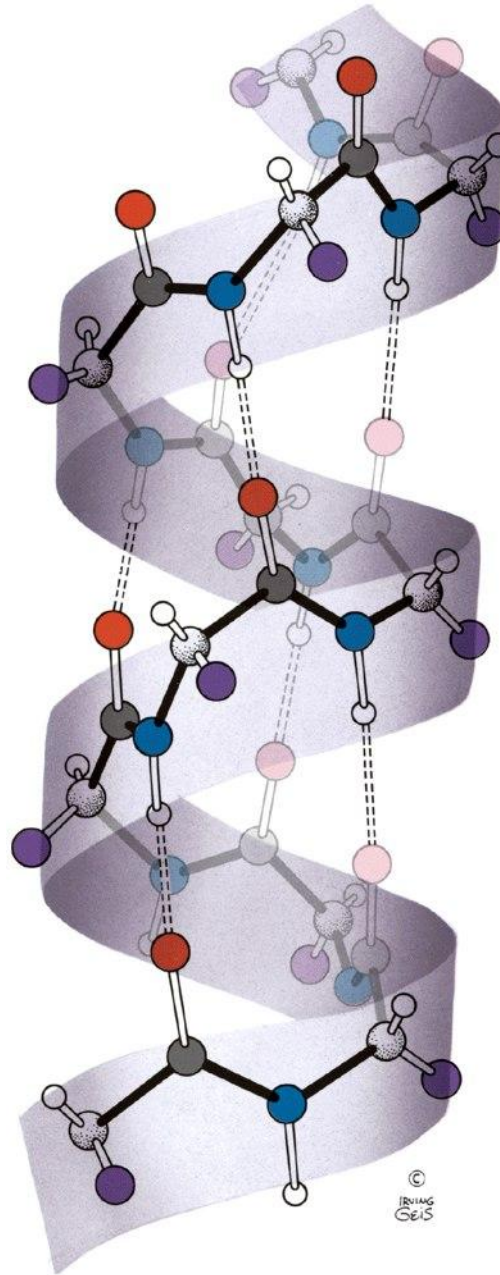


Figure 6-7 Fundamentals of Biochemistry, 2/e

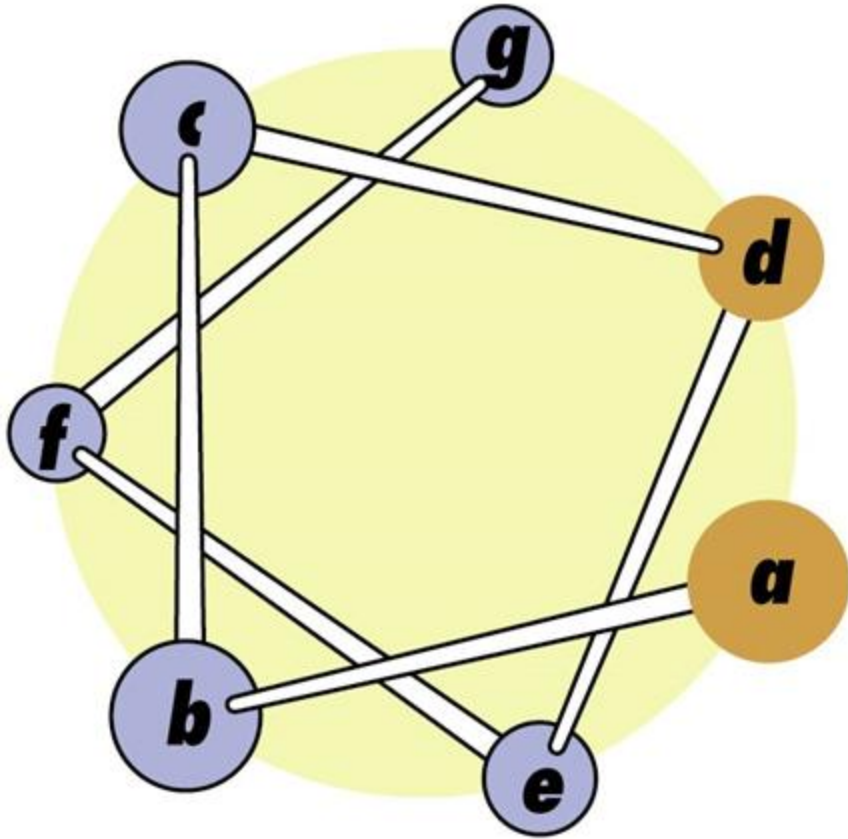


Figure 6-14a Fundamentals of Biochemistry, 2/e
© 2006 John Wiley & Sons

Apolar
Polar

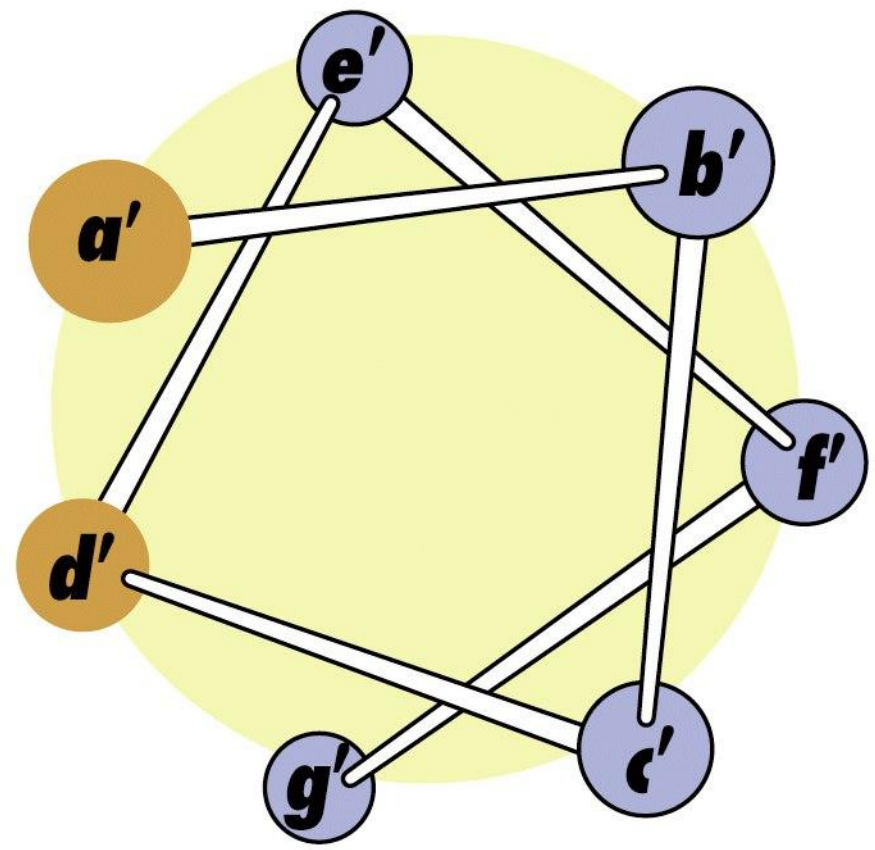
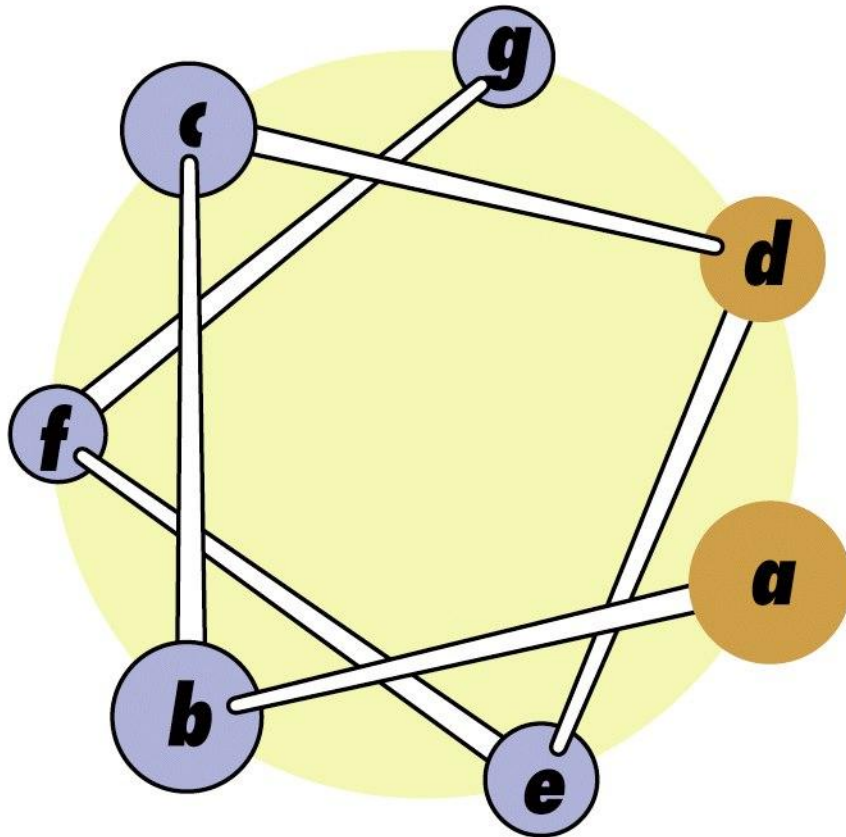
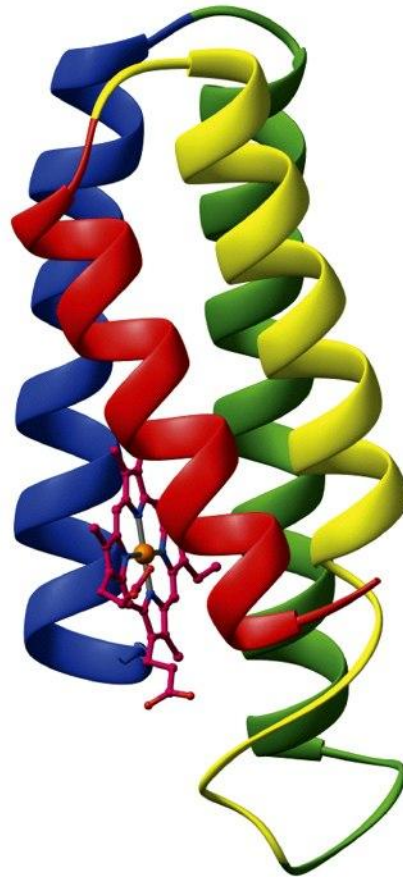


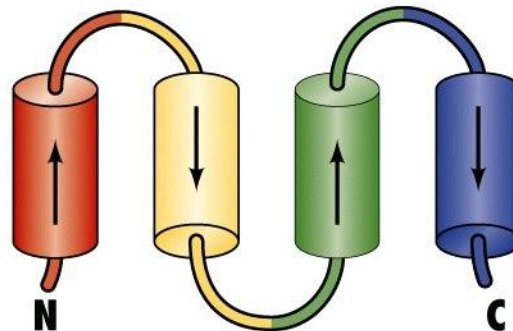
Figure 6-14a Fundamentals of Biochemistry, 2/e
 © 2006 John Wiley & Sons

Apolar
 Polar

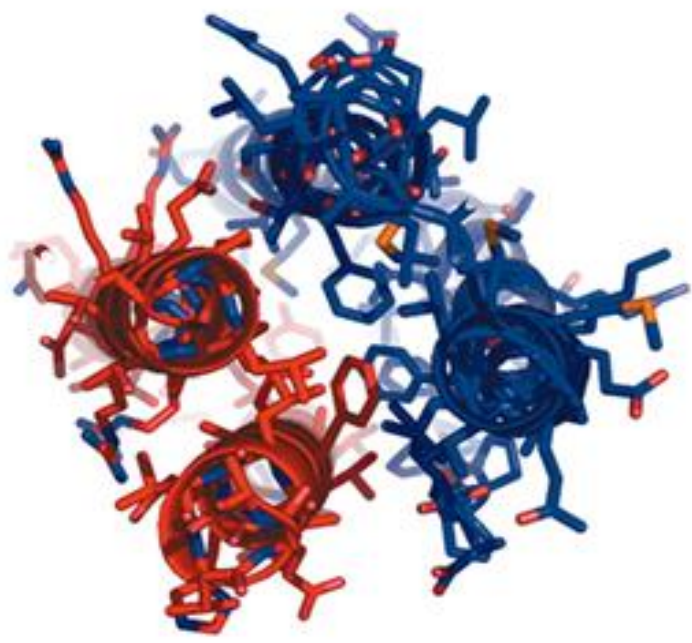
Up and Down Bundle

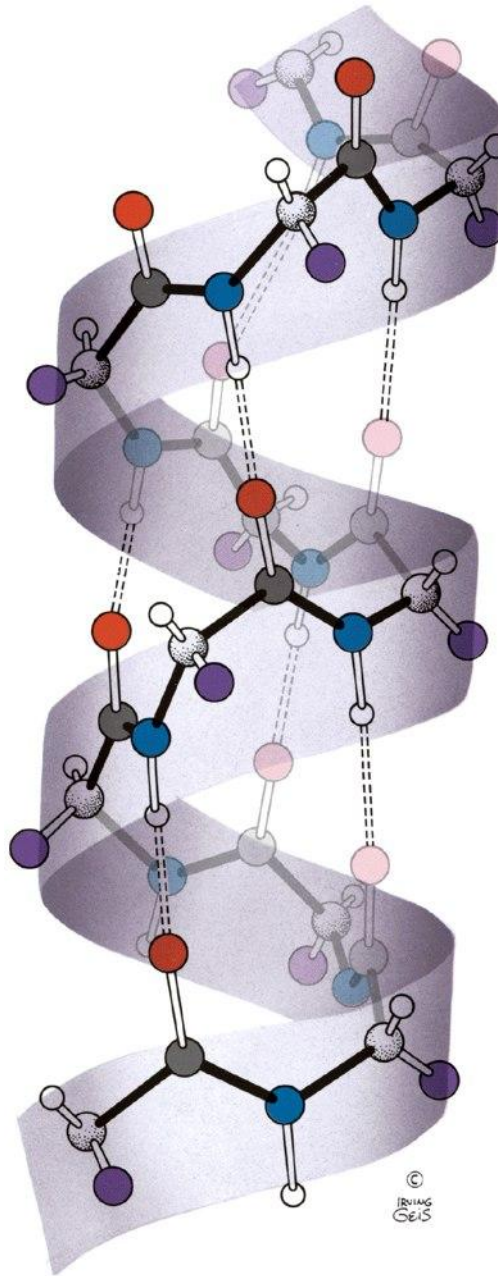


Four Helix Bundle
Feixe de 4 hélices



citocromo b562





©
IRVING
GELFOND

Figure 6-7 Fundamentals of Biochemistry, 2/e

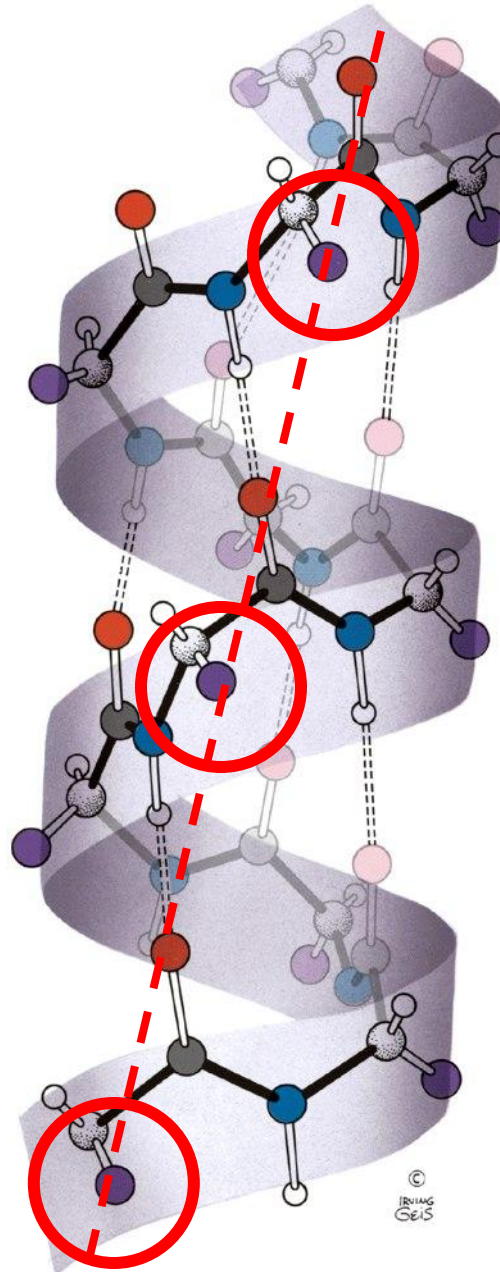


Figure 6-7 Fundamentals of Biochemistry, 2/e

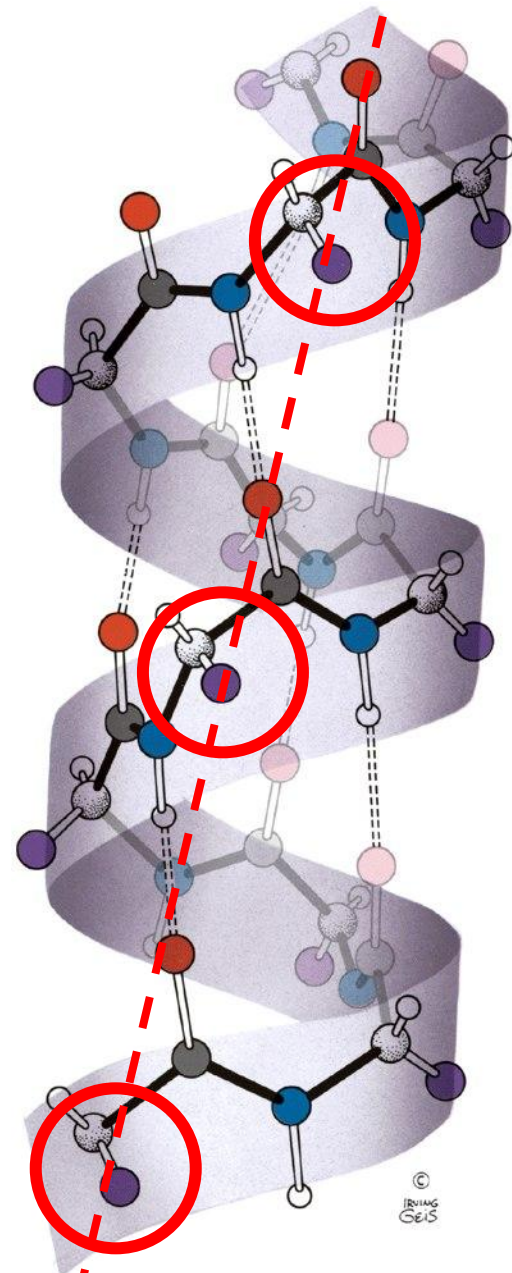
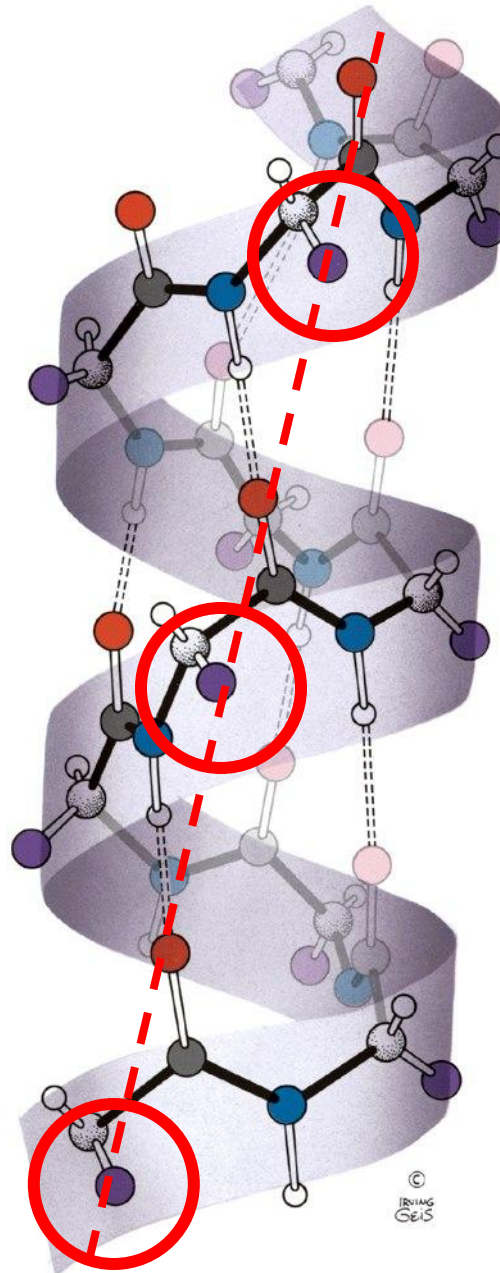
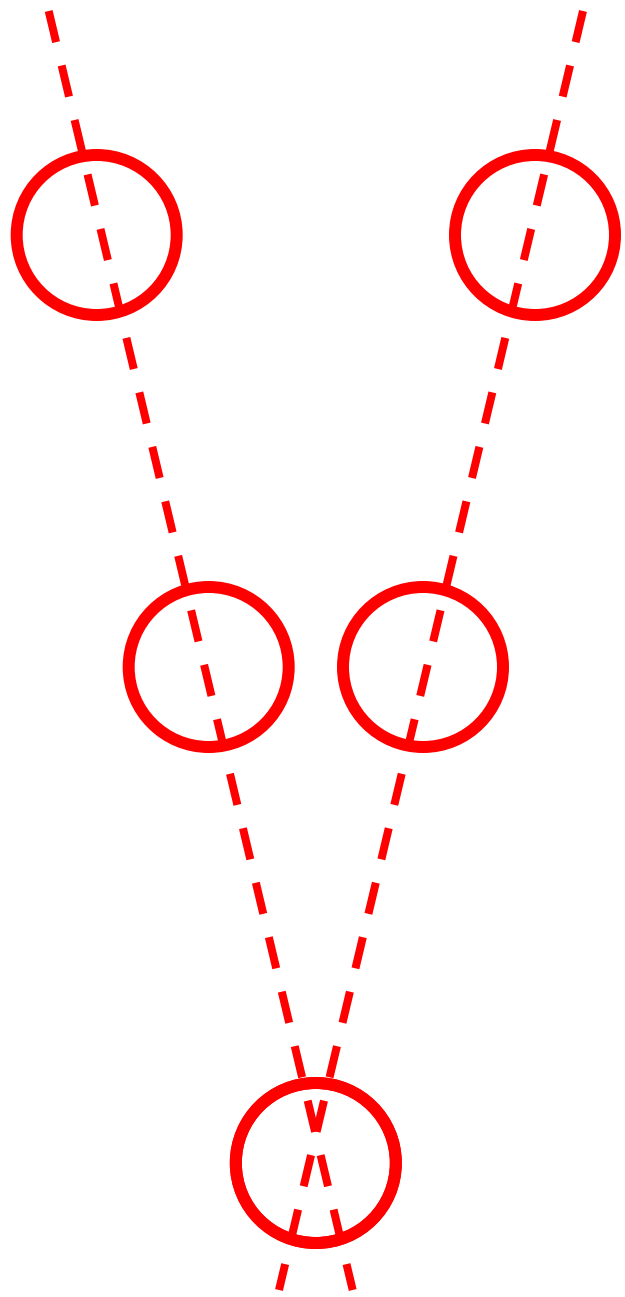
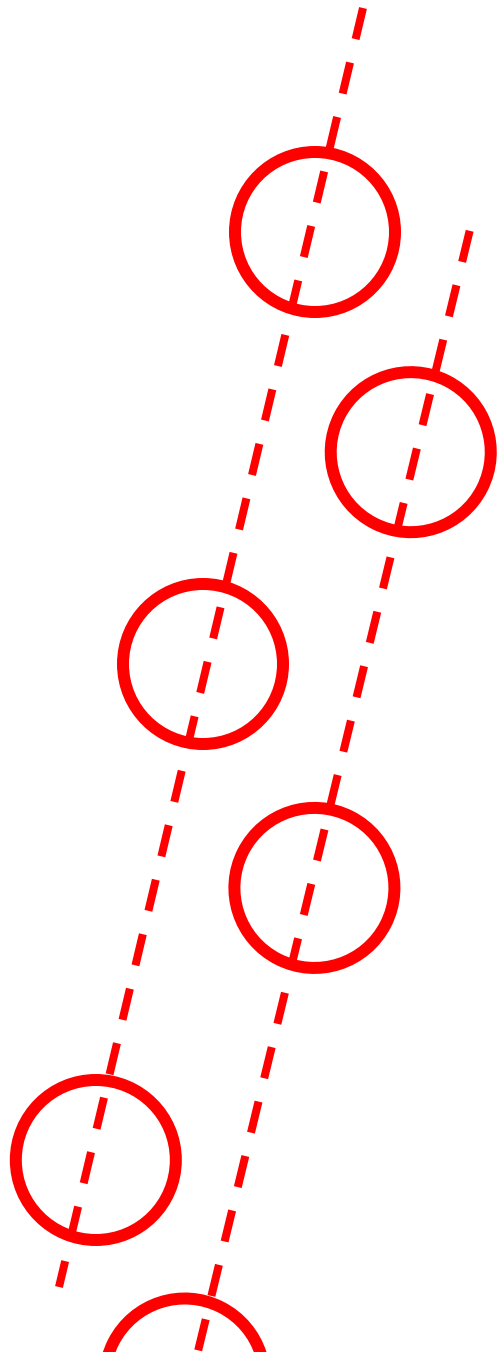
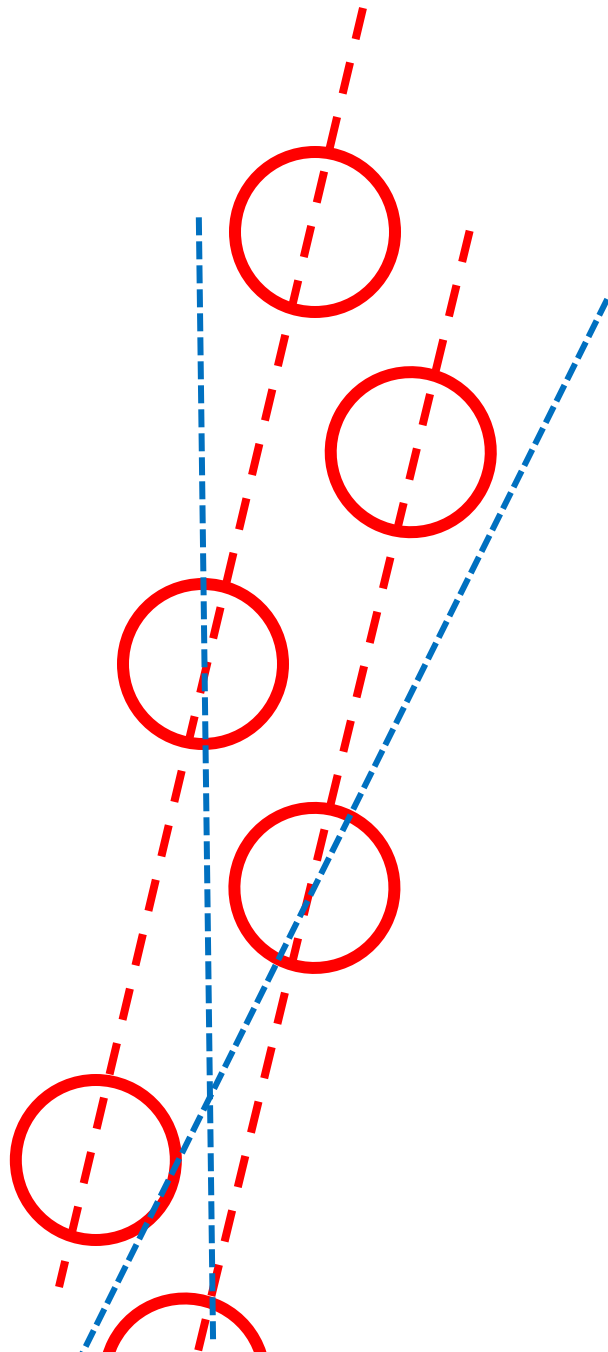
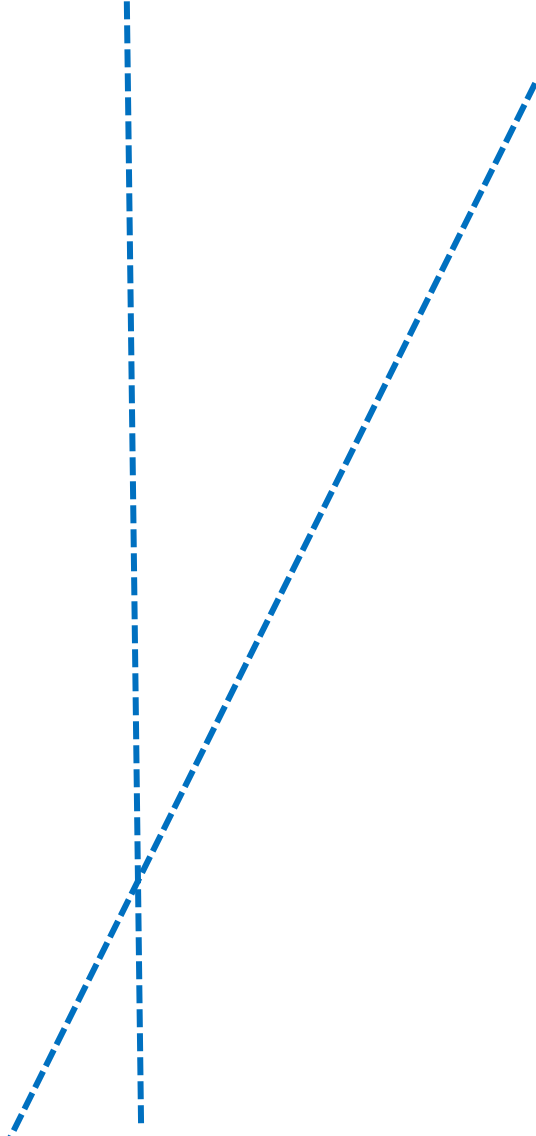


Figure 6-7 Fundamentals of Biochemistry, 2/ Figure 6-7 Fundamentals of Biochemistry, 2/e

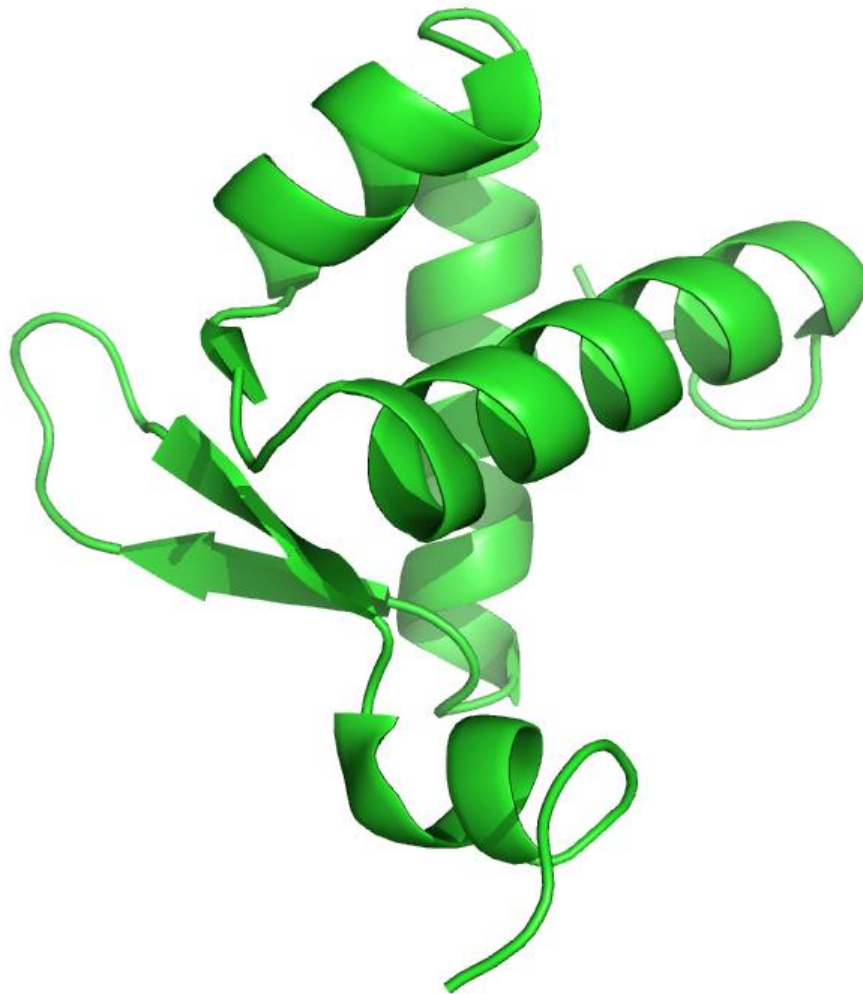








Orthogonal Bundle
Feixe ortogonal



BROWSE LINKS

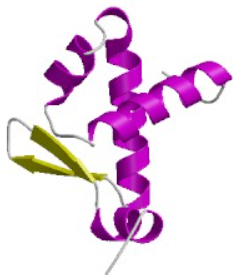
Browse Hierarchy

- Highly Diverse Superfamilies
- Superfamily Comparison

Select a CATH node...

C Mainly Alpha

	1
CATH ID	1
Architectures	5
Topologies	404
Superfamilies	2033
Domains	103788
Example Domain	1xmkA00 [PDB]



Tree



Sunburst

Top of CATH Hierarchy (4 Classes)

- C** **1** Mainly Alpha *5 Architectures, 404 Folds, 2033 Superfamilies, 103788 Domains*
 - A** **1.10** Orthogonal Bundle *290 Folds, 1132 Superfamilies, 69116 Domains*
 - A** **1.20** Up-down Bundle *104 Folds, 788 Superfamilies, 29676 Domains*
 - A** **1.25** Alpha Horseshoe *6 Folds, 103 Superfamilies, 3933 Domains*
 - A** **1.40** Alpha solenoid *2 Folds, 2 Superfamilies, 15 Domains*
 - A** **1.50** Alpha/alpha barrel *2 Folds, 8 Superfamilies, 1048 Domains*
- C** **2** Mainly Beta *21 Architectures, 244 Folds, 1290 Superfamilies, 124032 Domains*
- C** **3** Alpha Beta *14 Architectures, 634 Folds, 2337 Superfamilies, 262275 Domains*
- C** **4** Few Secondary Structures *1 Architectures, 108 Folds, 181 Superfamilies, 5716 Domains*
- C** **6** Special *2 Architectures, 82 Folds, 790 Superfamilies, 4427 Domains*

grampo β

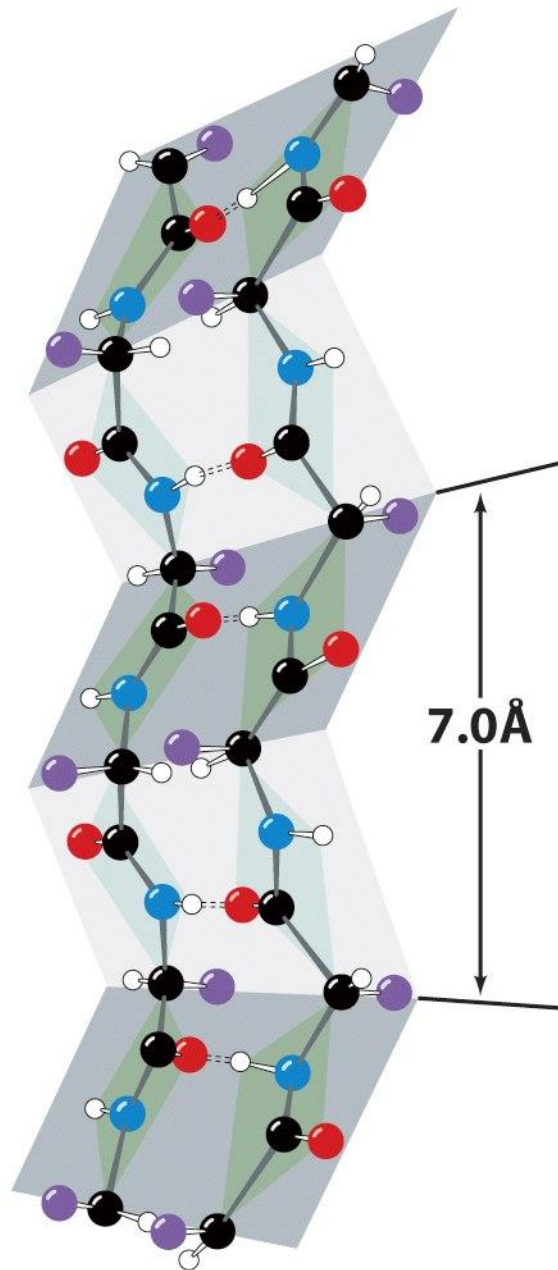
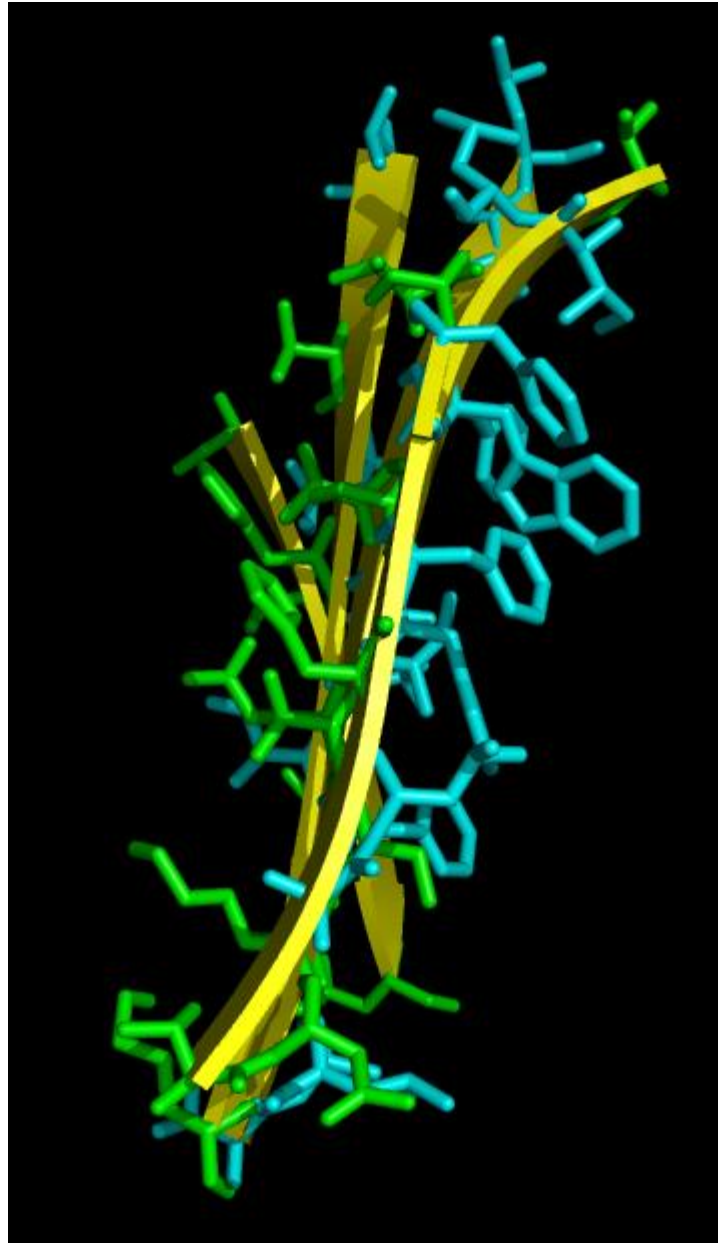


Figure 6-10 Fundamentals of Biochemistry, 2/e

Concanavalina
PDB 1AZD



Cadeia principal:
Cadeia lateral polar
Cadeia lateral apolar

β -barrel

Barril β

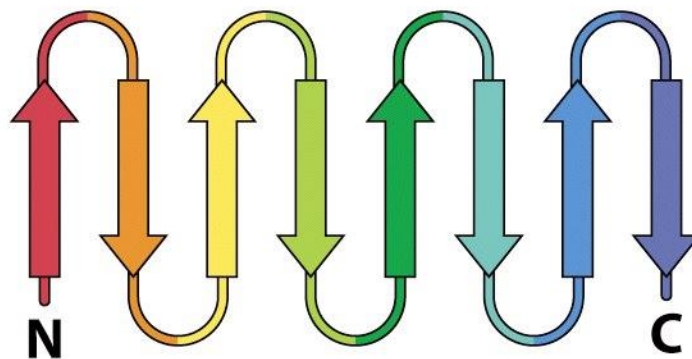
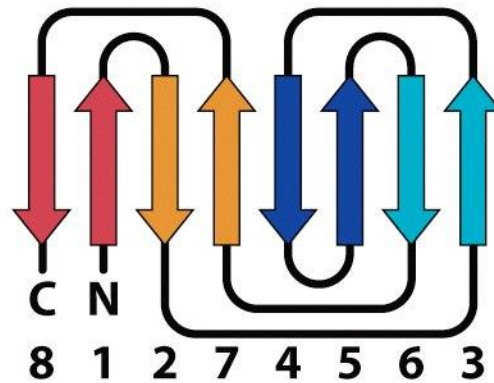
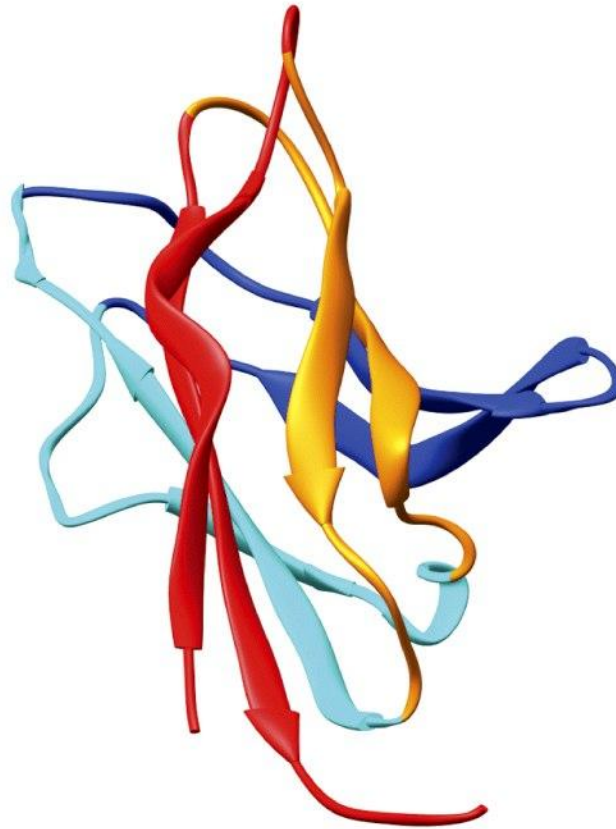


Figure 6-30a Fundamentals of Biochemistry, 2/e

Proteína ligante de retinol
PDB 1RBP

Sandwich

Sanduiche



Greek Key

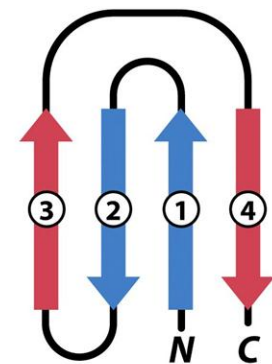


Figure 6-30b Fundamentals of Biochemistry, 2/e

Select a CATH node...

A Sandwich

2.60

CATH ID 2.60

Topologies 44

Superfamilies 536

Domains 60890

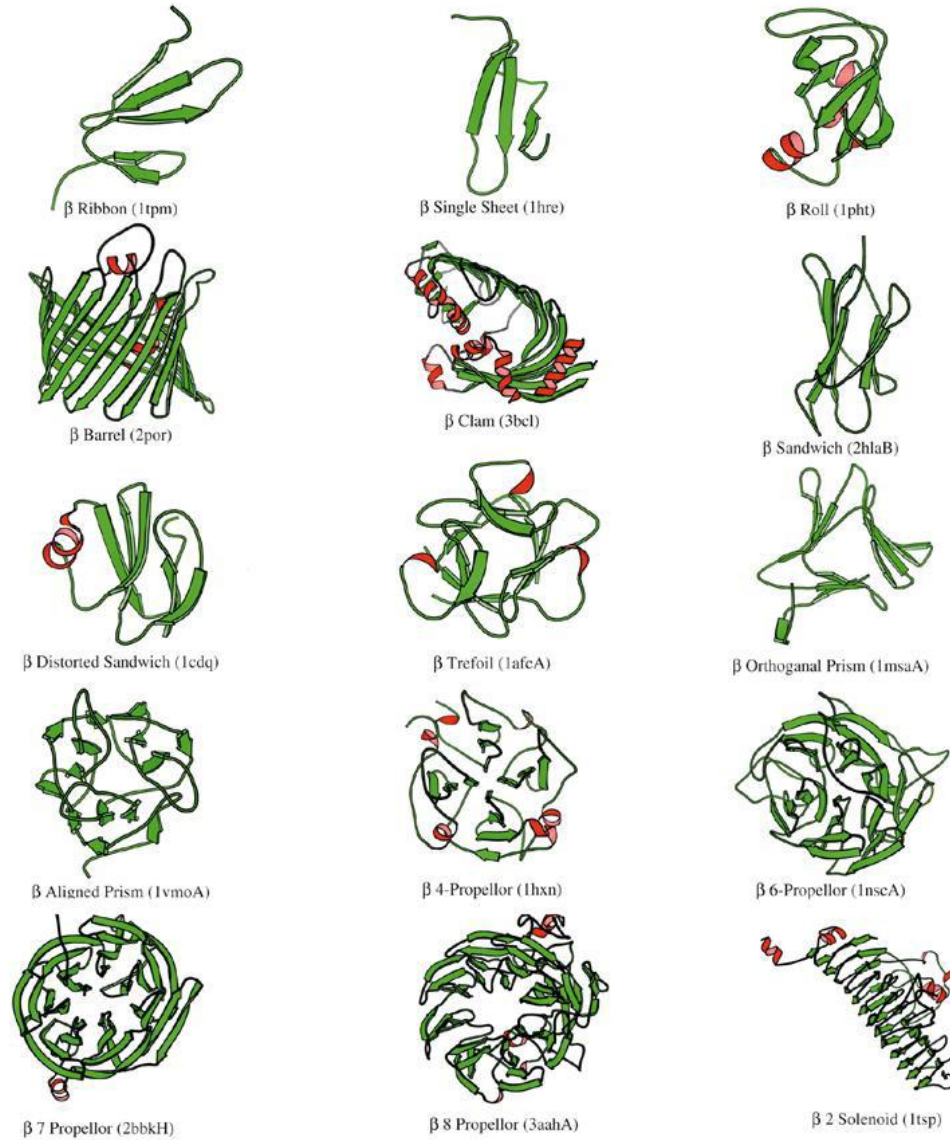
Example Domain [4unuA00 \[PDB\]](#)

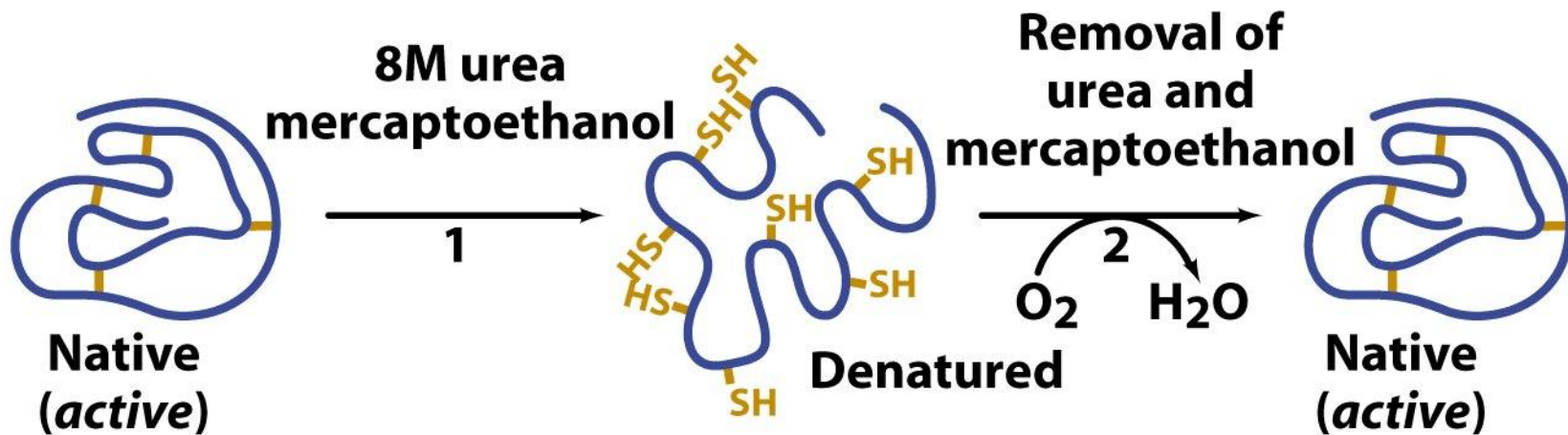
Top of CATH Hierarchy (4 Classes)

▶	C	1	Mainly Alpha	5 Architectures, 404 Folds, 2033 Superfamilies, 103788 Domains
▶	C	2	Mainly Beta	21 Architectures, 244 Folds, 1290 Superfamilies, 124032 Domains
▶	A	2.10	Ribbon	26 Folds, 56 Superfamilies, 4308 Domains
▶	A	2.20	Single Sheet	21 Folds, 103 Superfamilies, 2348 Domains
▶	A	2.30	Roll	40 Folds, 176 Superfamilies, 10458 Domains
▶	A	2.40	Beta Barrel	48 Folds, 278 Superfamilies, 30975 Domains
▶	A	2.50	Clam	2 Folds, 5 Superfamilies, 88 Domains
▶	A	2.60	Sandwich	44 Folds, 536 Superfamilies, 60890 Domains
▶	A	2.70	Distorted Sandwich	18 Folds, 42 Superfamilies, 4847 Domains
▶	A	2.80	Trefoil	2 Folds, 3 Superfamilies, 1557 Domains
▶	A	2.90	Orthogonal Prism	2 Folds, 3 Superfamilies, 159 Domains
▶	A	2.100	Aligned Prism	1 Folds, 5 Superfamilies, 359 Domains
▶	A	2.102	3-layer Sandwich	3 Folds, 3 Superfamilies, 457 Domains
▶	A	2.105	3 Propeller	1 Folds, 1 Superfamilies, 1 Domains
▶	A	2.110	4 Propeller	1 Folds, 1 Superfamilies, 57 Domains
▶	A	2.115	5 Propeller	1 Folds, 2 Superfamilies, 479 Domains
▶	A	2.120	6 Propeller	1 Folds, 7 Superfamilies, 1177 Domains
▶	A	2.130	7 Propeller	1 Folds, 7 Superfamilies, 1641 Domains
▶	A	2.140	8 Propeller	1 Folds, 3 Superfamilies, 428 Domains
▶	A	2.150	2 Solenoid	1 Folds, 2 Superfamilies, 49 Domains
▶	A	2.160	3 Solenoid	3 Folds, 15 Superfamilies, 1166 Domains
▶	A	2.170	Beta Complex	26 Folds, 41 Superfamilies, 2585 Domains
▶	A	2.180	Shell	1 Folds, 1 Superfamilies, 3 Domains
▶	C	3	Alpha Beta	14 Architectures, 634 Folds, 2337 Superfamilies, 262275 Domains
▶	C	4	Few Secondary Structures	1 Architectures, 108 Folds, 181 Superfamilies, 5716 Domains
▶	C	6	Special	2 Architectures, 82 Folds, 790 Superfamilies, 4427 Domains

CATH – a hierarchic classification of protein domain structures

CA Orengo, coautores e JM Thornton
Structure 5, 1093-1109 (1997)





A conformação nativa da cadeia polipeptídica é termodinamicamente estável

Esta conformação depende somente da estrutura primária e as condições do meio

O processo pelo qual a cadeia polipeptídica atinge a conformação nativa, o dobramento, enovelamento, *fold*ing, pode ser estudado *in vitro*

A conformação nativa das proteínas é “levemente” estável

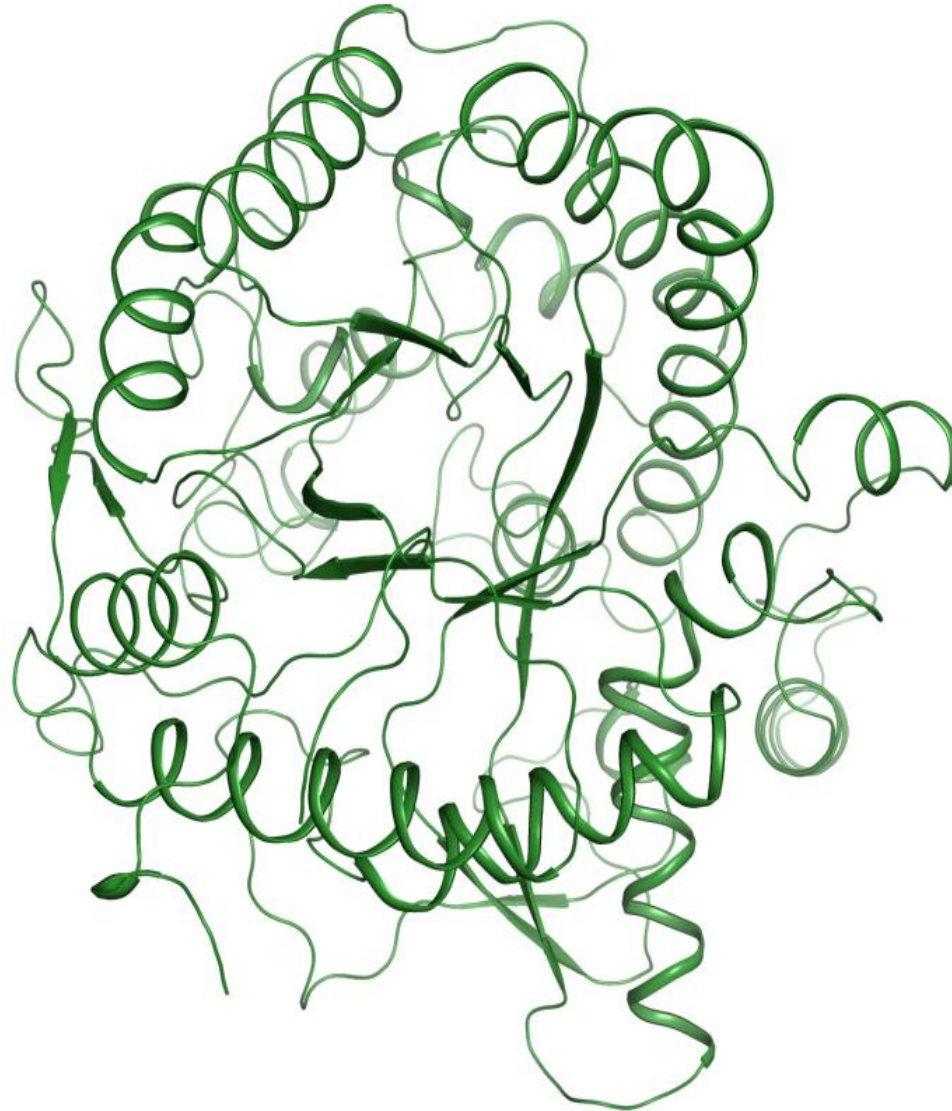
Termodinâmica de enovelamento de proteínas

Proteína	ΔG^0 (kcal/mol)	ΔH^0 (kcal/mol)	$T\Delta S^0$ (kcal/mol)
CI2	-6.62	-32.26	-25.64
EglinC	-8.82	-27.48	-18.66
RNAse T1	-8.96	-67.16	-58.22
Citocromo c	-8.87	-21.27	-12.40
Barnase	-11.69	-73.37	-61.71

Fonte: Lesk, A. *Introduction to protein science*. Oxford University Press, 2ª edição, p. 352 (2010)

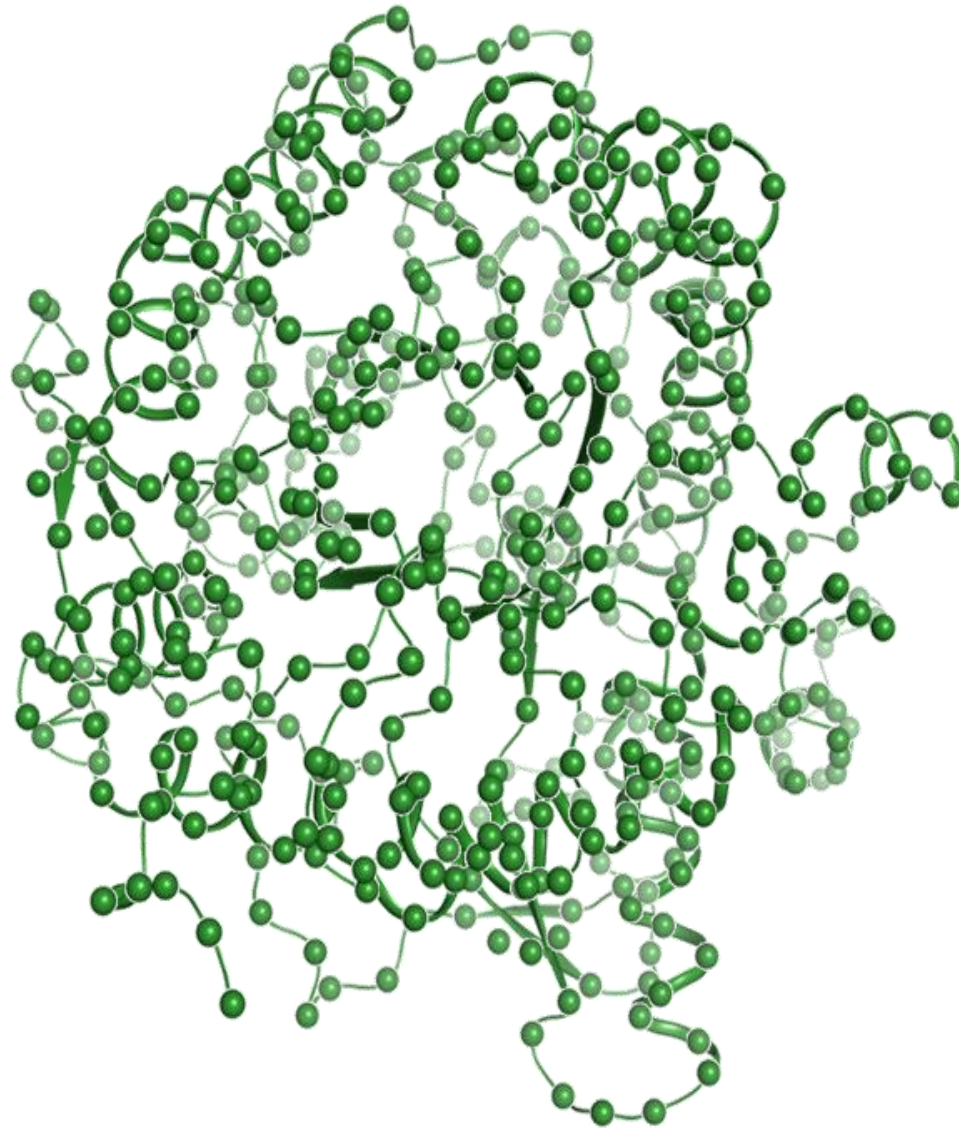
A diferença de energia (ΔG^0) entre a conformação nativa e a desnaturada em condições fisiológicas (25°C, pH 7, 1 atm) está na faixa de 5 a 15 kcal/mol (~ 20 a 60 kJ/mol).

Há milhares de interações não-covalentes na estrutura proteica.



PDB 5CG0

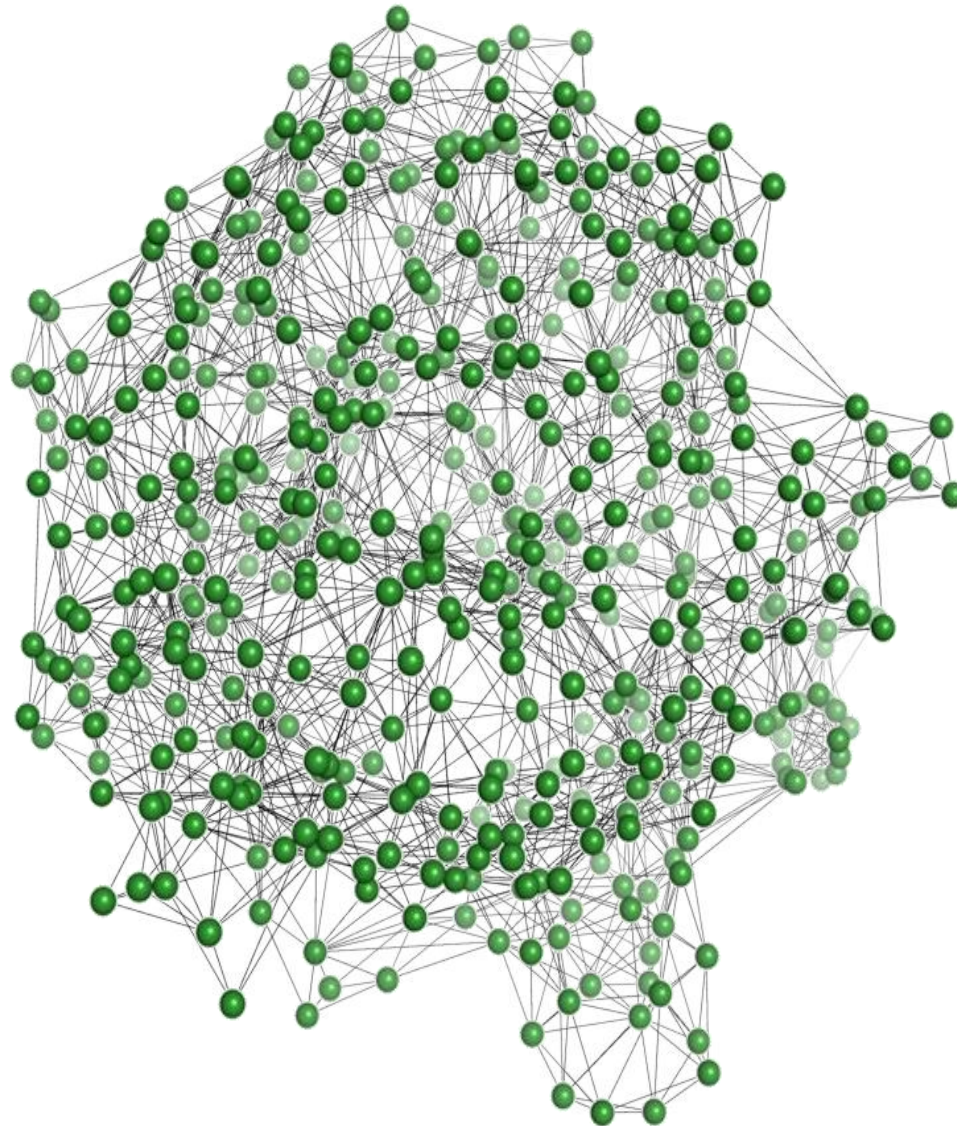
509 aminoácidos



PDB 5CG0

509 aminoácidos

5384 interações não-covalentes aminoácido-aminoácido



PDB 5CG0

Há milhares de interações não-covalentes na estrutura proteica.

Table 1.1 Relationship of Noncovalent Interactions to the Distance Separating the Interacting Molecules, r

Type of Interaction	Distance Relationship
Charge-charge	$1/r$
Charge-dipole	$1/r^2$
Dipole-dipole	$1/r^3$
Charge-induced dipole	$1/r^4$
Dispersion	$1/r^6$
Repulsion	$1/r^{12}$

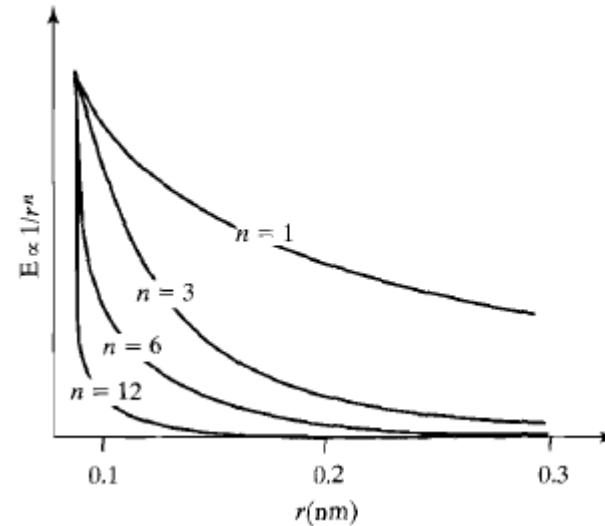
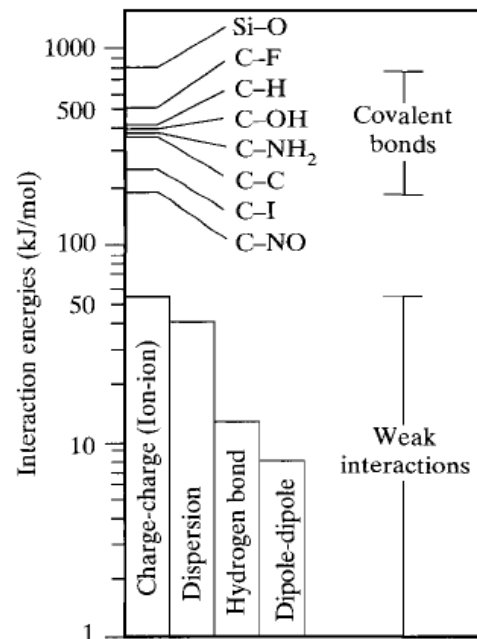


Table 2.1 Noncovalent Interactions Between Molecules

Type of Interaction	Equation*
Ion-ion (Charge-Charge)	$E = \frac{Z_1 Z_2 e^2}{Dr}$
Ion-dipole	$E = \frac{Z_1 e \mu_2 \theta}{Dr^2}$
Dipole-dipole	$E = \frac{\mu_1 \mu_2 \theta^2}{Dr^3} = \frac{3(\mu_1 r \theta^2)(\mu_2 r \theta^2)}{Dr^5}$
Ion-induced dipole	$E = \frac{Z_1 e^2 \alpha_2}{2D^2 r^4}$
Dispersion ^e	$E = \frac{3h\nu_0 \alpha^2}{4r^6}$



Efeito hidrofóbico, resultante da retirada de cadeias laterais apolares do contato com solvente aquoso, é um importante componente do dobramento proteico.

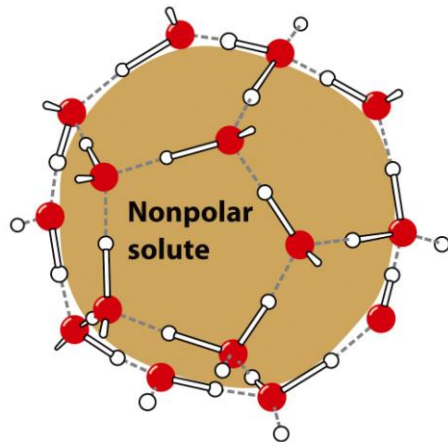


Figure 2-8 Fundamentals of Biochemistry, 2/e
© 2006 John Wiley & Sons

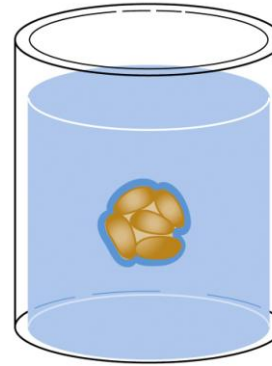


Figure 2-9b Fundamentals of Biochemistry, 2/e
© 2006 John Wiley & Sons

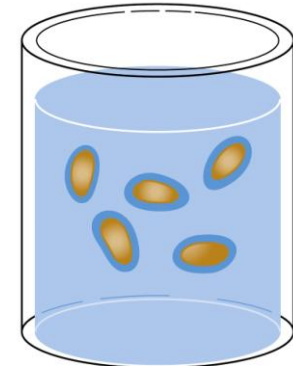


Figure 2-9a Fundamentals of Biochemistry, 2/e
© 2006 John Wiley & Sons

Table 2-2 Thermodynamic Changes for Transferring Hydrocarbons from Water to Nonpolar Solvents at 25°C

Process	ΔH (kJ · mol ⁻¹)	$-T\Delta S$ (kJ · mol ⁻¹)	ΔG (kJ · mol ⁻¹)
CH ₄ in H ₂ O \rightleftharpoons CH ₄ in C ₆ H ₆	11.7	-22.6	-10.9
CH ₄ in H ₂ O \rightleftharpoons CH ₄ in CCl ₄	10.5	-22.6	-12.1
C ₂ H ₆ in H ₂ O \rightleftharpoons C ₂ H ₆ in benzene	9.2	-25.1	-15.9
C ₂ H ₄ in H ₂ O \rightleftharpoons C ₂ H ₄ in benzene	6.7	-18.8	-12.1
C ₃ H ₈ in H ₂ O \rightleftharpoons C ₃ H ₈ in benzene	0.8	-8.8	-8.0
Benzene in H ₂ O \rightleftharpoons liquid benzene ^a	0.0	-17.2	-17.2
Toluene in H ₂ O \rightleftharpoons liquid toluene ^a	0.0	-20.0	-20.0

^aData measured at 18°C.

Source: Kauzmann, W., *Adv. Protein Chem.* **14**, 39 (1959).

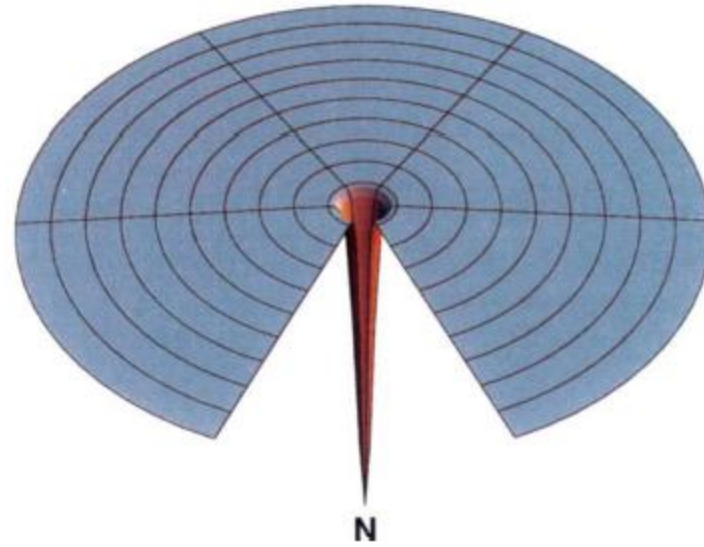
Paradoxo

Proteínas assumem sua conformação nativa em tempos muito curtos (10^{-6} s, 10^{-1} s, 1000 s)

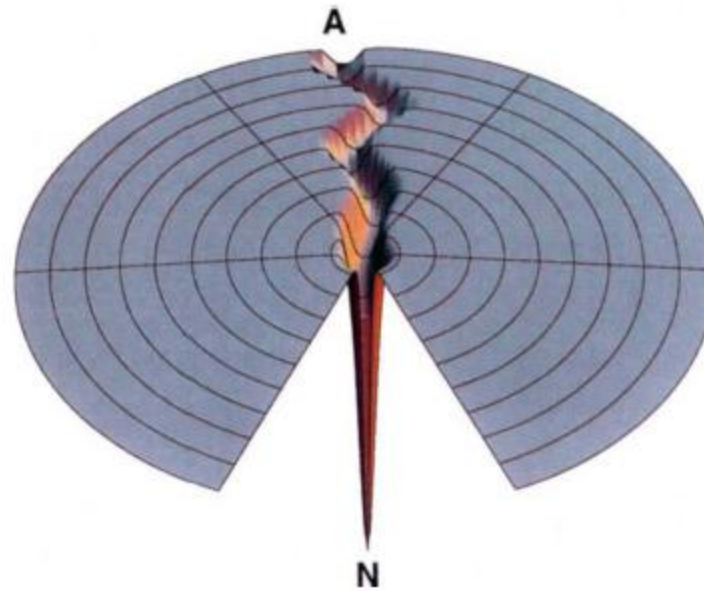
Por outro lado...

Em uma aproximação simplificada, se cada resíduo de aminoácido tiver 3 opções permitidas de conformação, para uma proteína de 100 resíduos haverão 3^{100} conformações possíveis. Se cada uma destas conformações for “visitada” em 10^{-13} s, seriam necessários cerca de 10^{87} s, ou seja, 10^{26} anos para proteína atingir uma conformação particular nesta amostra.

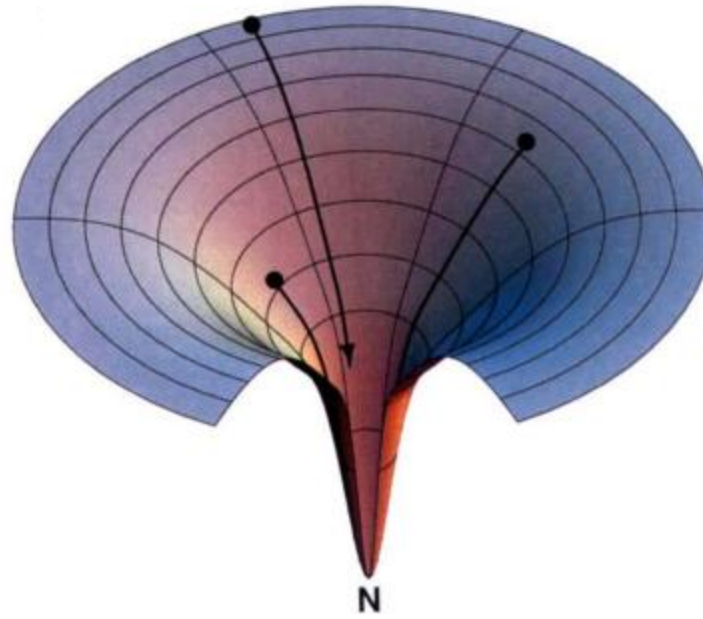
Busca Randômica



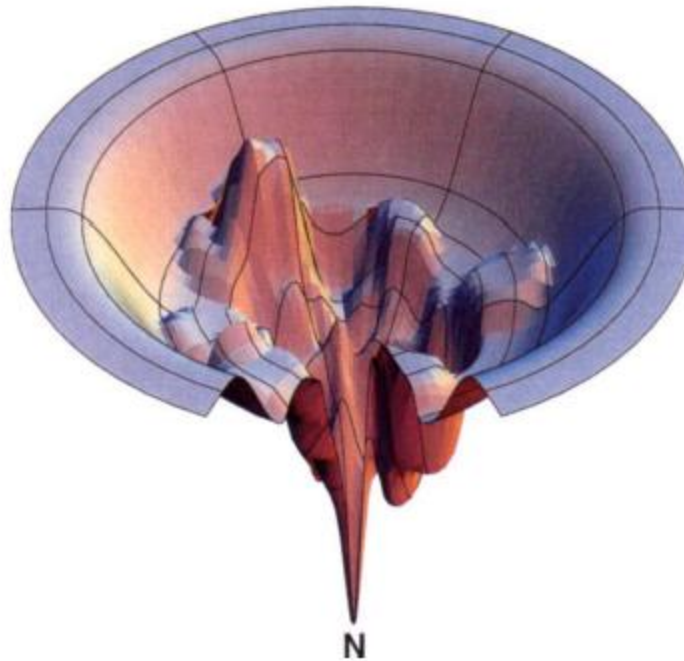
Rota Preferencial (Pathway)



Funil de Enovelamento



Funil "Realista"



Colapso hidrofóbico e formação de hélices

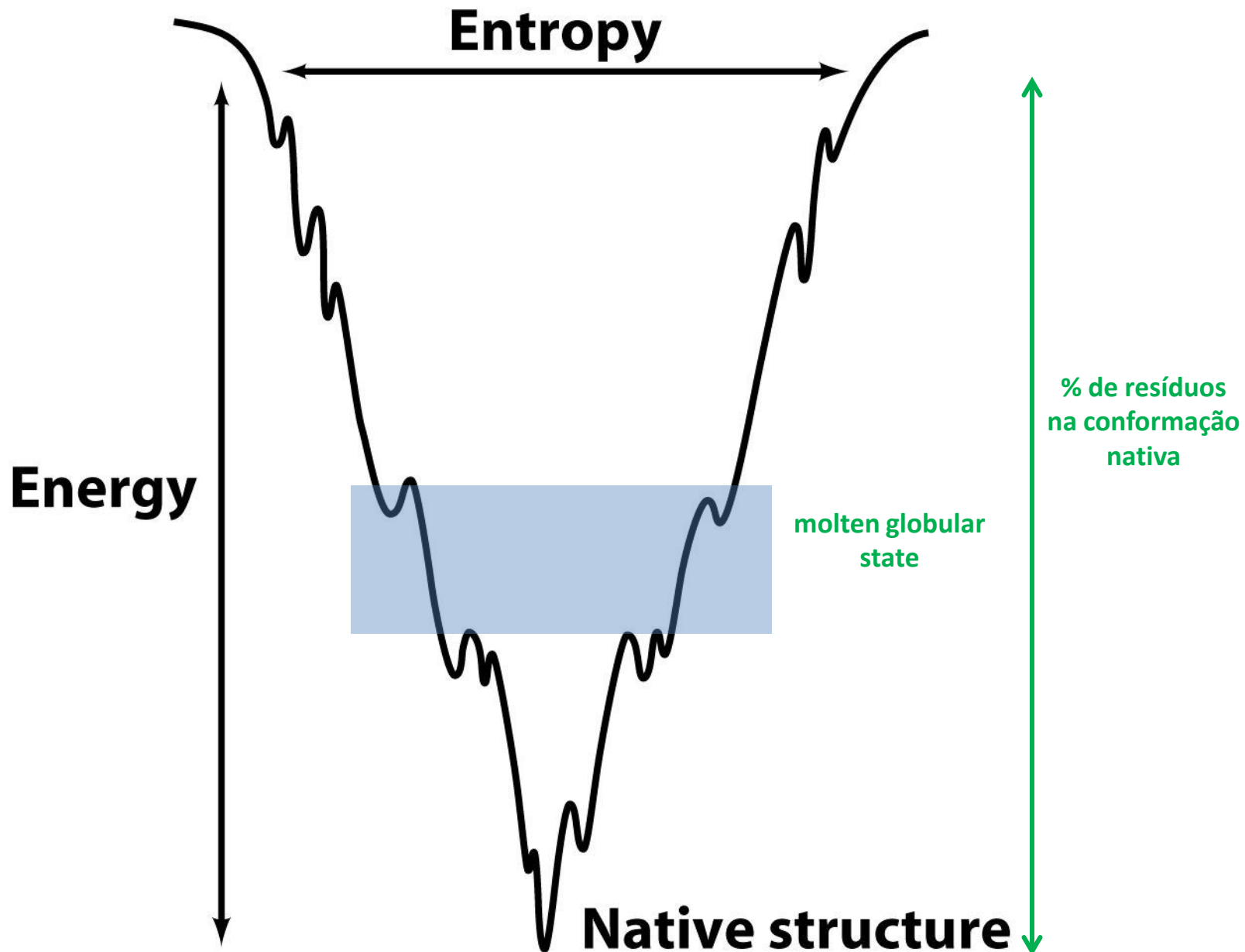
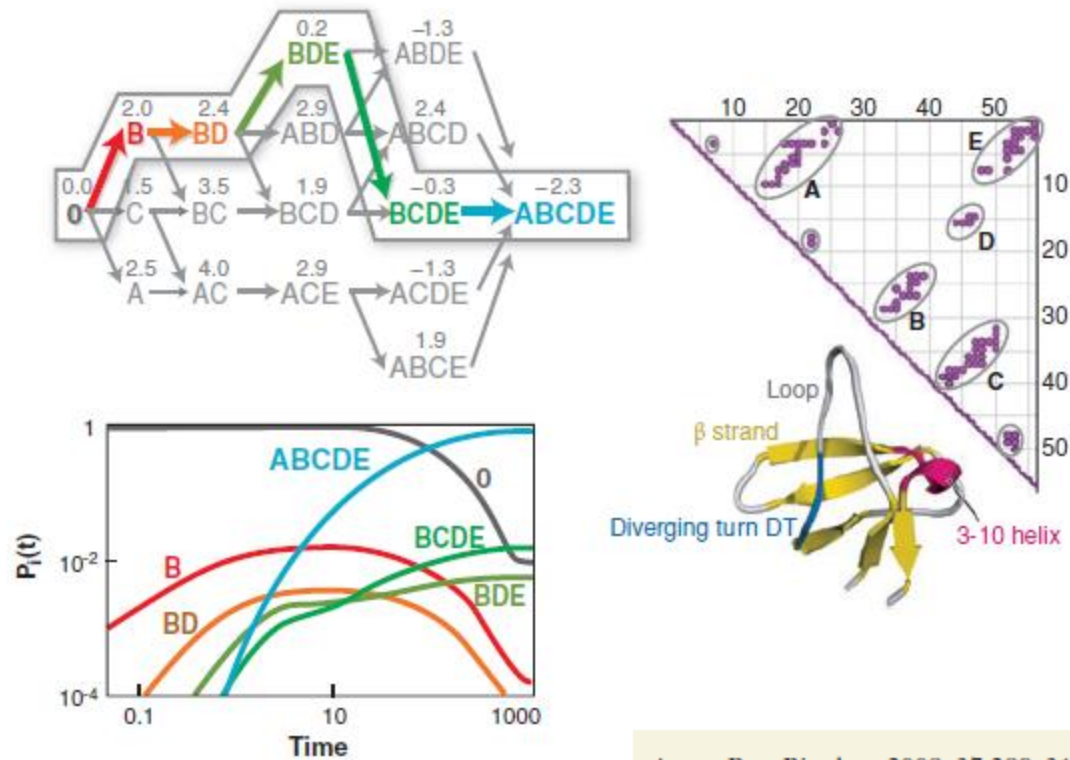


Figure 6-41 Fundamentals of Biochemistry, 2/e
© 2006 John Wiley & Sons

O enovelamento de uma proteína segue rotas paralelas e sequenciais simultaneamente



Annu. Rev. Biophys. 2008. 37:289–316

Proteínas retidas em “acidentes” (vales, cavidades, etc...) da superfície de dobramento podem ser recuperadas por chaperonas, reiniciando seu processo de dobramento.

Enzimas como a *Protein Disulfide Isomerase* (PDI) e *Prolyl Peptide Isomerase* tem atuação equivalente em princípio.