

- The half-life of a cytosolic protein is determined to a large extent by its amino-terminal residue

proteasome complex (or 26S protease  
complex)  $\leftarrow$  digests target protein  
 $\nwarrow$   
recycles ubiquitin

$E_1, E_2, E_3$

ATP hydrolysis

in the target protein

it covalently binds to Lys residues

it tags proteins for degradation

small protein (76 aa, 8.5 kDa)

- in eukaryotic cells: ubiquitin system

- in E. coli: ATP-dependent protease Lon

• Protein destruction systems:

days. e.g., half-life for hemoglobin = 110 days

• The half-life of most proteins vary from 30s to many

unwanted proteins and permits the recycling of amino acids

• Protein degradation prevents the buildup of abnormal or

systems in all cells.

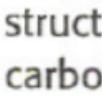
• Protein degradation and destruction is mediated by specialized

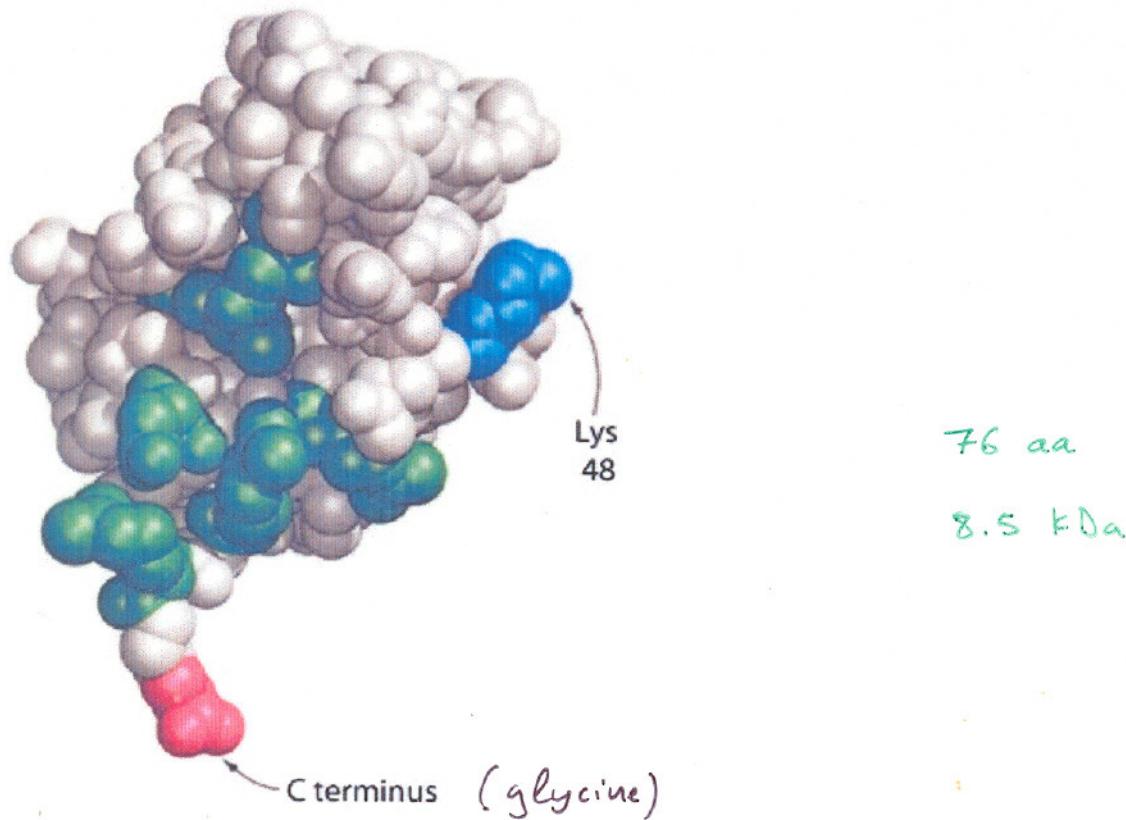
**TABLE 23.1 Dependence of the half-lives of cytosolic yeast proteins on the nature of their amino-terminal residues**

Highly stabilizing residues ( $t_{1/2} > 20$ hours)			
Ala	Cys	Gly	Met
Pro	Ser	Thr	Val
Intrinsically destabilizing residues ( $t_{1/2} = 2$ to 30 minutes)			
Arg	His	Ile	Leu
Lys	Phe	Trp	Tyr
Destabilizing residues after chemical modification ( $t_{1/2} = 3$ to 30 minutes)			
Asn	Asp	Gln	Glu

Source: J. W. Tobias, T. E. Schrader, G. Rocap, and A. Varshavsky. *Science* 254(1991):1374.

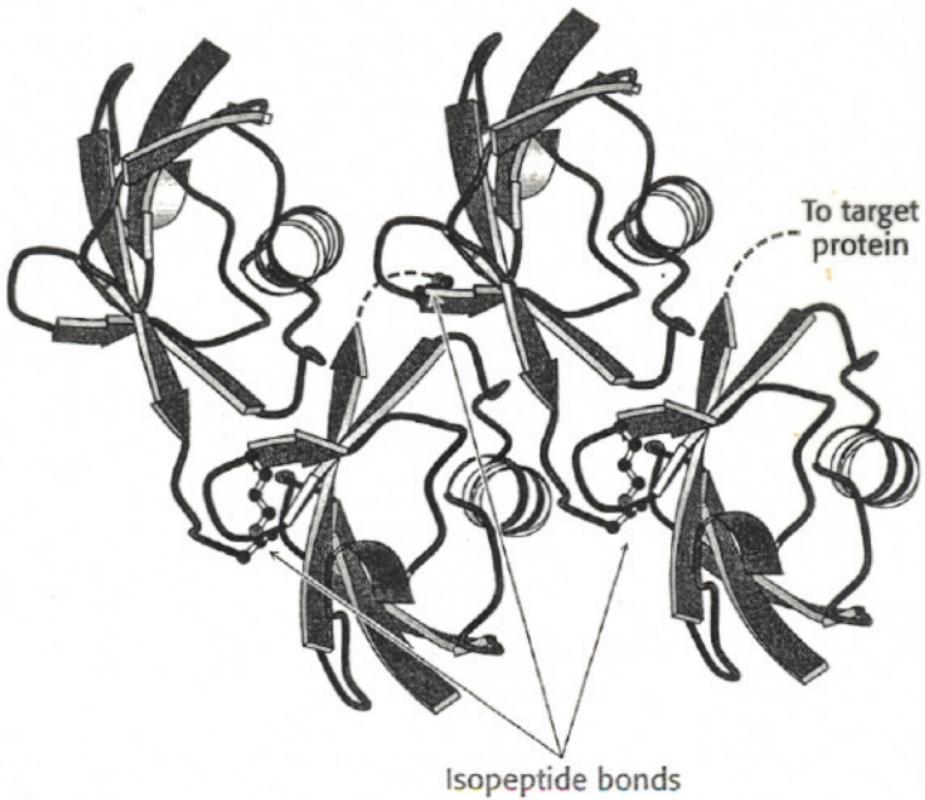


 **FIGURE 23.2 Ubiquitin.** The structure of ubiquitin reveals an extended carboxyl terminus that is activated and linked to other proteins. Lysine residues also are shown, including lysine 48, the major site for linking additional ubiquitin molecules.

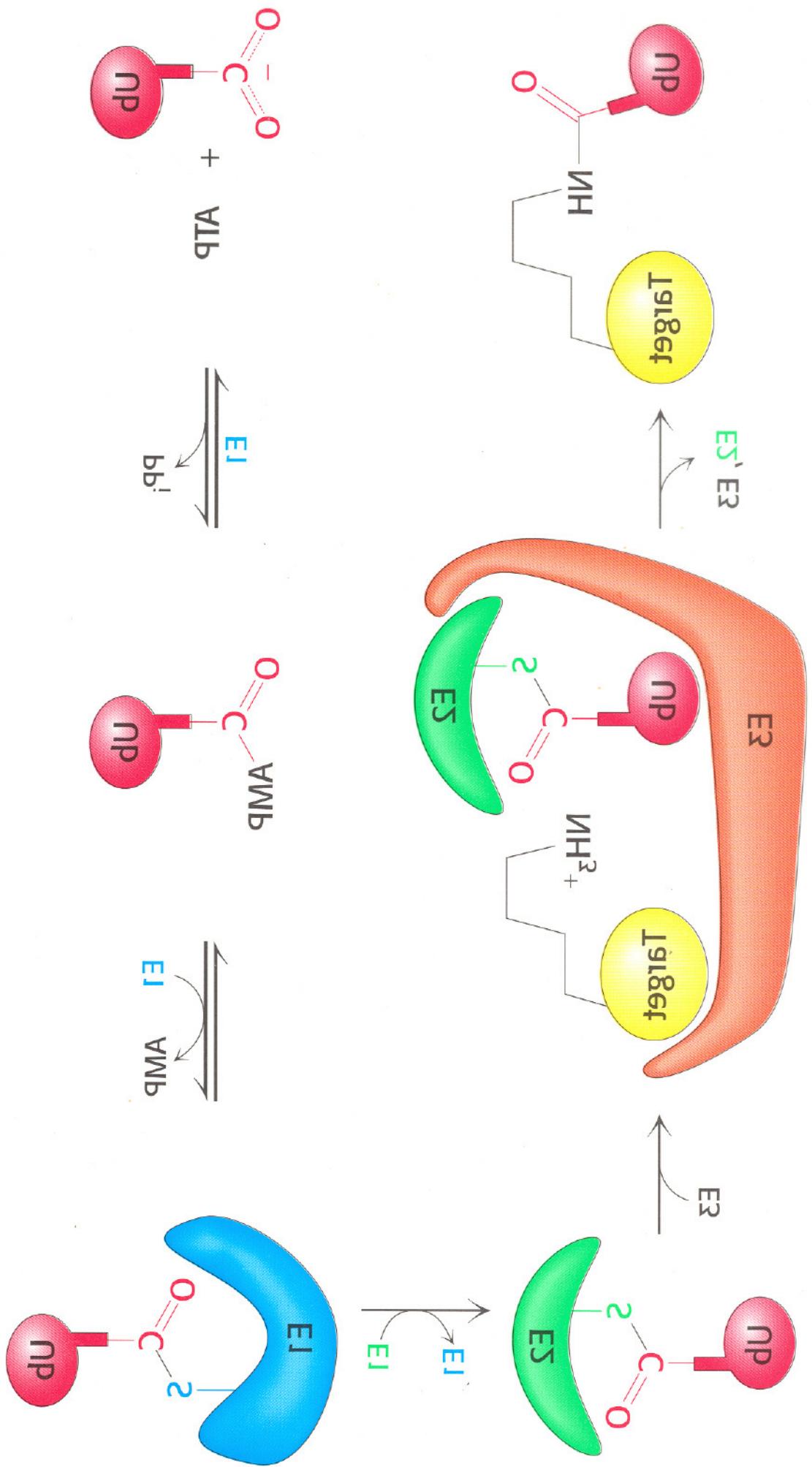


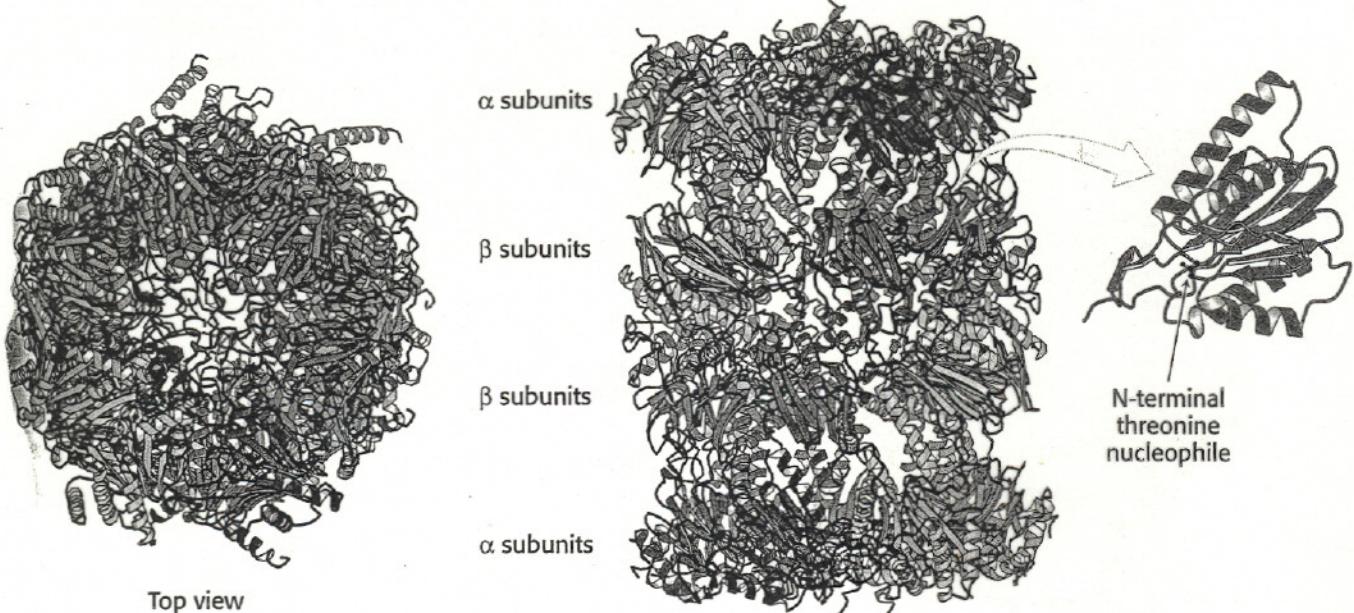
3D structure of ubiquitin

Fig 35-4e

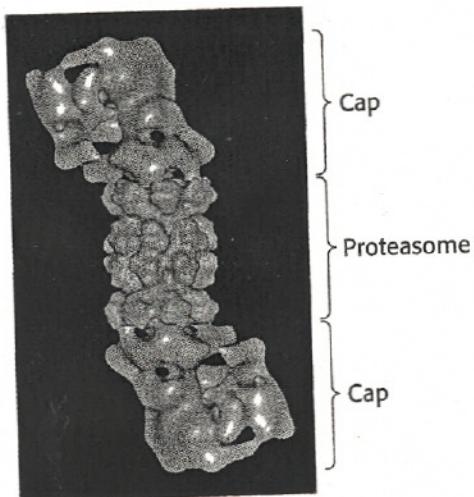


 **FIGURE 23.4 Tetraubiquitin.** Four ubiquitin molecules are linked by isopeptide bonds. This unit is the primary signal for degradation when linked to a target protein.

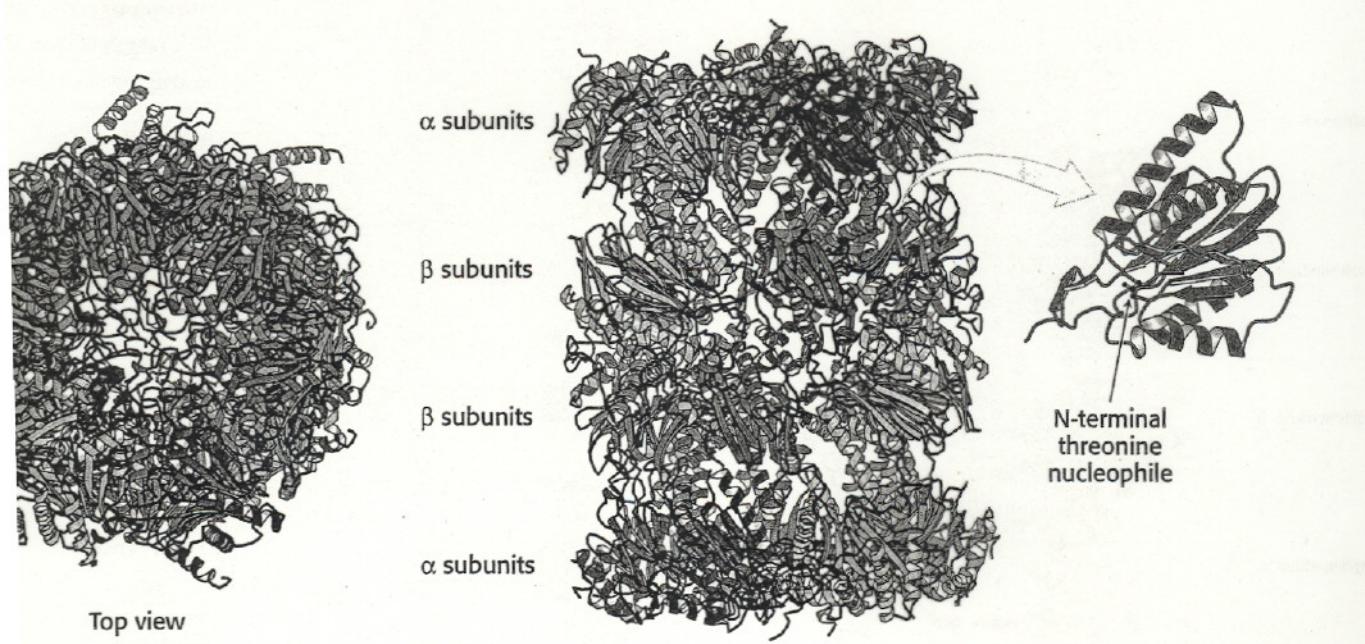




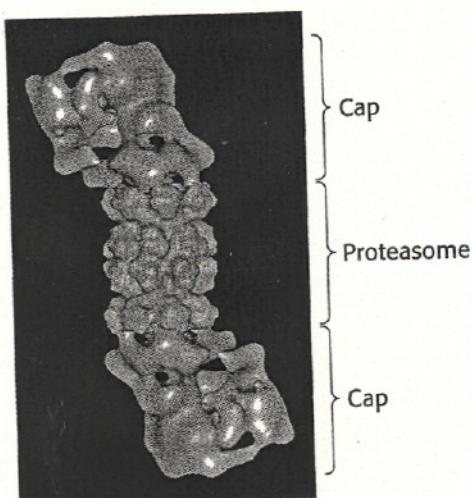
**FIGURE 23.5 20S proteasome.** The 20S proteasome comprises 28 homologous subunits ( $\alpha$ , red;  $\beta$ , blue), arranged in four rings of 7 subunits each. Some of the  $\beta$  subunits (highlighted in yellow) include protease active sites at the amino termini. The top view shows the approximate seven-fold symmetry of the structure.



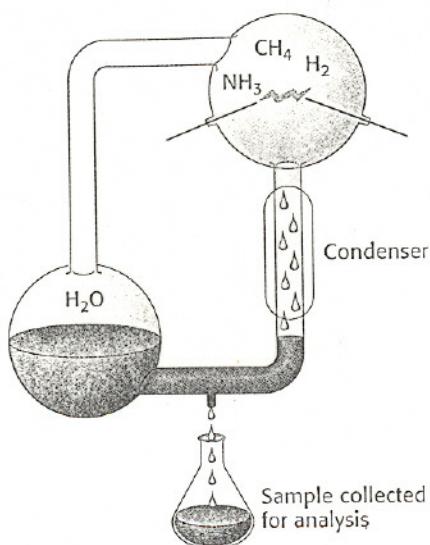
**FIGURE 23.6 26S proteasome.** A 19S cap is attached to each end of the 20S catalytic unit. [From W. Baumeister, J. Walz, F. Zuhl, and E. Seemüller. *Cell* 92(1998):367. Courtesy of Dr. Wolfgang Baumeister.]



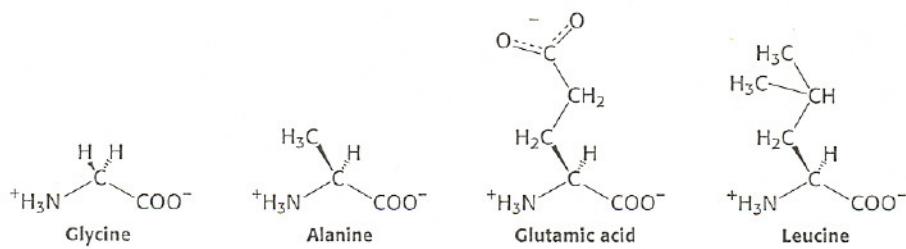
**FIGURE 23.5 20S proteasome.** The 20S proteasome comprises 28 homologous ( $\alpha$ , red;  $\beta$ , blue), arranged in four rings of 7 subunits each. Some of the  $\beta$  subunits (ted in yellow) include protease active sites at the amino termini. The top view shows approximate seven-fold symmetry of the structure.



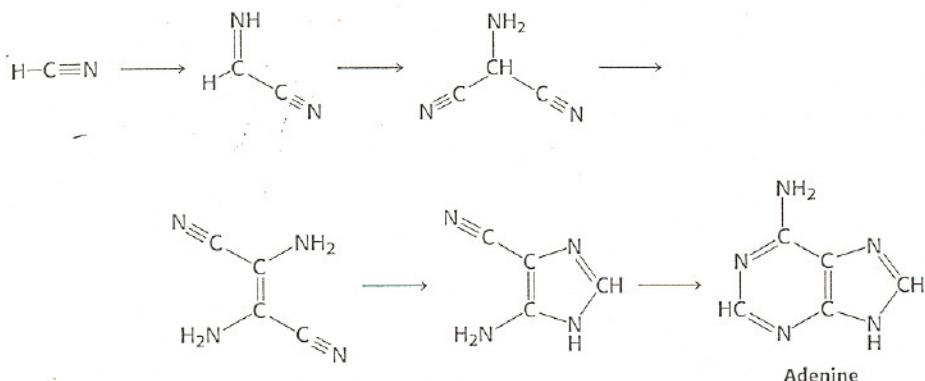
**FIGURE 23.6 26S proteasome.** A 19S cap is attached to each end of the 20S catalytic unit. [From W. Baumeister, J. Walz, F. Zuhl, and E. Seemüller. *Cell* 92(1998):367. Courtesy of Dr. Wolfgang Baumeister.]



**FIGURE 2.1** The Urey-Miller experiment. An electric discharge (simulating lightning) passed through an atmosphere of  $\text{CH}_4$ ,  $\text{NH}_3$ ,  $\text{H}_2\text{O}$ , and  $\text{H}_2$  leads to the generation of key organic compounds such as amino acids.



**FIGURE 2.2** Products of prebiotic synthesis. Amino acids produced in the Urey-Miller experiment.



**FIGURE 2.3** Prebiotic synthesis of a nucleic acid component. Adenine can be generated by the condensation of  $\text{HCN}$ .

## Prebiotic Earth :

- atmosphere was highly reduced
- $\text{CH}_4$ ,  $\text{NH}_3$ ,  $\text{H}_2\text{O}$ ,  $\text{H}_2$
- solar radiation and lightning

## Urey - Miller experiment (1950s) :

synthesis of prebiotic molecules (such as aa  
- Gly, Ala, Glu, Leu - , adenine and ribose )

## Basic principles of evolution :

(1) replication or reproduction

(2) variation

(3) competition (evolution by natural selection)

In 1967, Sol Spiegelman showed that evolution can take place at the molecular level

Bacteriophage Q<sub>B</sub> RNA molecules

+

Q<sub>B</sub> replicase

selective pressures (time of replication,  
precursor concentration)

Theory: single-stranded RNA could have become self-replicating

RNA was the first catalyst (ribozyme)

The genetic code established a relationship between RNA and proteins

- Mutations:
- point mutations
  - gene duplication
  - gene deletion

DNA is chemically more stable than RNA

ATP can be generated through the breakdown of organic molecules (degradation of glycine)

Cells were formed by the inclusion of nucleic acids within membranes

Osmosis:

- development of cell wall structures
- energy-dependent ion pumps

Proton gradients across membranes can be used to drive the synthesis of ATP

## CHAPTER 7 :

### Exploring evolution & bioinformatics

Sequence-comparison methods have become a powerful tool in modern biochemistry

Homology : two molecules are said to be homologous if they have been derived from a common ancestor

Homologs

```
graph LR; Homologs --> Paralogs; Homologs --> Orthologs;
```

Paralogs : homologs that are present within one species. Often they differ in their detailed biochemical functions.

Orthologs : homologs that are present within different species and have very similar or identical functions.

Homology is often manifested by significant similarity in nucleotide or amino acid sequence and almost always manifested in 3D structure

(A)



(B)

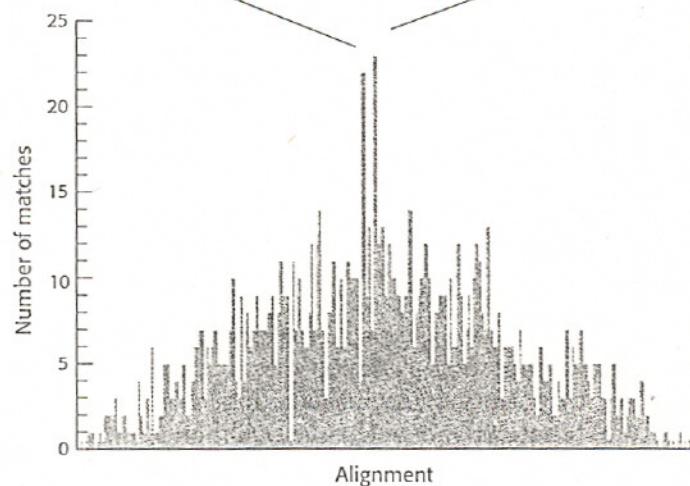
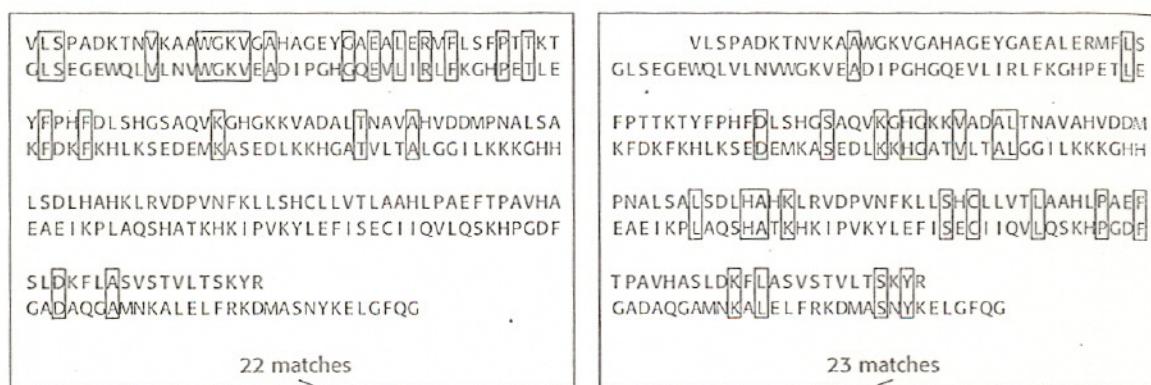


FIGURE 7.5 Comparing the amino acid sequences of hemoglobin  $\alpha$  and myoglobin. (A) A comparison is made by sliding the sequences of the two proteins past one another, one amino acid at a time, and counting the number of amino acid identities between the proteins. (B) The two alignments with the largest number of matches are shown above the graph, which plots the matches as a function of alignment.



Hemoglobin  $\alpha$  VLSPADKTNVKAAGWKVGAHAGEYGAEEALERMFLSFPTTKTYFPHF  
Myoglobin GLSIEGEWQLVLNWMGKVADIPGHQEVLIIRLFKGHPETLEKFDKFKHLKSED

Gap

LSHGSAQVKGHGKKVADALTNAVAHVDDMPNALSAISDLHAKL RVDPVNKKL  
EMKASEDIIKKHGATVLTALGGILKKKGHHEAEIKPLAQSHATKHKIPVKYLEF

LSHCLIVTAAHLPAEFTPASVHASLDKFLASMSVTVLTSKYR  
ISECIQVQLQSKEPGDFGADAQGMNKALELFRKDMASNYKELGFQG

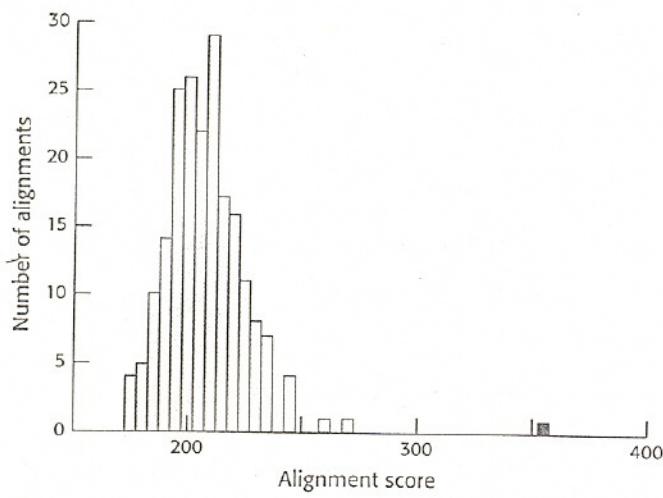
**FIGURE 7.6** Alignment with gap insertion. The alignment of hemoglobin  $\alpha$  and myoglobin after a gap has been inserted into the hemoglobin  $\alpha$  sequence.

THIS IS THE AUTHENTIC SEQUENCE

↓ Shuffling

SNUCSNSEATEEITUHEQIHHTCEI

**FIGURE 7.7** The generation of a shuffled sequence.



**FIGURE 7.8** Statistical comparison of alignment scores. Alignment scores are calculated for many shuffled sequences, and the number of sequences generating a particular score is plotted against the score. The resulting plot is a distribution of alignment scores occurring by chance. The alignment score for hemoglobin  $\alpha$  and myoglobin (shown in red) is substantially greater than any of these scores, strongly suggesting that the sequence similarity is significant.

Understanding the homology between molecules can reveal

- (i) the evolutionary history of the molecules
- (ii) information about their function

Statistical analysis of sequence alignments can detect homology

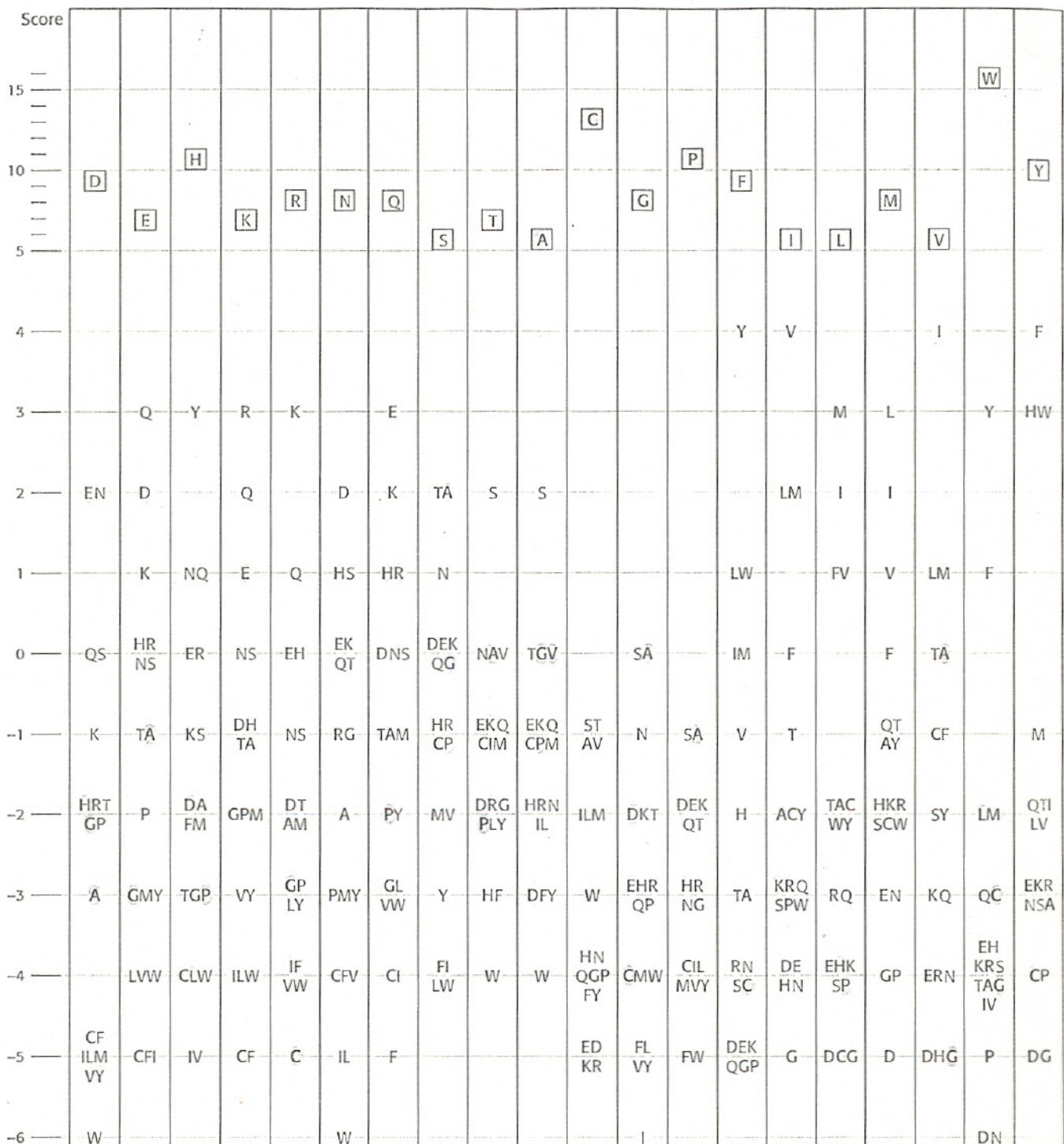
Homology = identical residues

+

conservative substitutions

+

gaps



**FIGURE 7.9** A graphic view of the Blosum-62 substitution matrix. This scoring scheme was derived by examining substitutions that occur within aligned sequence blocks in related proteins. Amino acids are classified into four groups (charged, red; polar, green; large and hydrophobic, blue; other, black). Substitutions that require the change of only a single nucleotide are shaded. To find the score for a substitution of, for instance, a Y for an H, you find the Y in the column having H (boxed) at the top and check the number at the left. In this case, the resulting score is 3.

## Databases:

Databases can be searched to identify homologous sequences

e.g., the first complete sequence of the genome of a free-living organism, the bacterium *Haemophilus influenzae*, was reported in 1995

- 1743 ORFs → potential genes



1007 (58%)  
could be linked to a  
protein of known  
function

347 could be  
linked to a  
protein of unknown  
function

389 did not  
match any sequence  
in the database