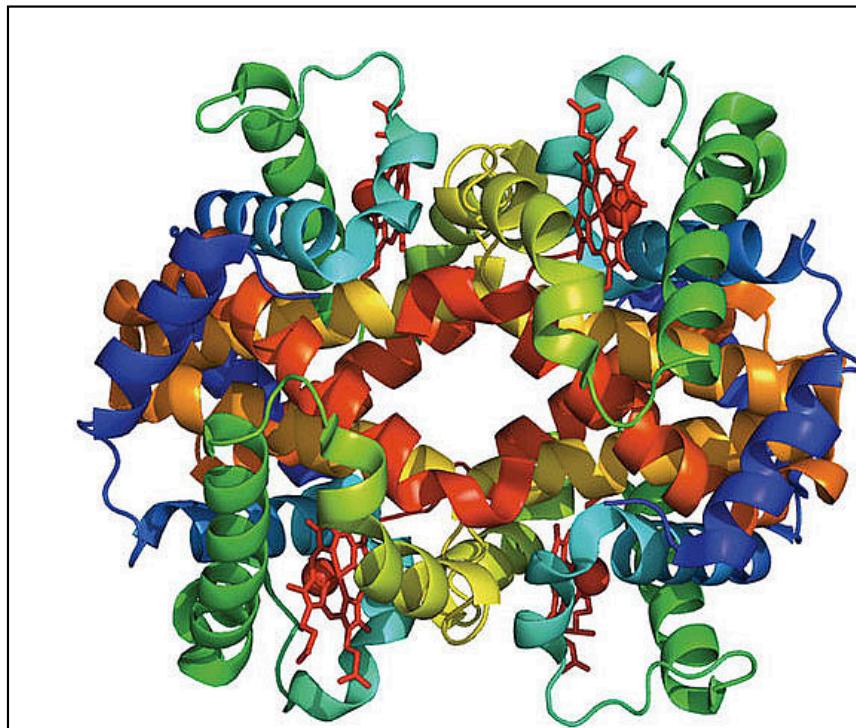


PROTEINS: SEQUENCE-STRUCTURE-FUNCTION



Ozkan Lab : banu.Ozkan@asu.edu

**Welcome to
CBP Summer Workshop**

Classes

- All classes and the lab will be at PSH 530.

Schedule

- Lecture 10:00 AM-11:00 AM
- Intro to Lab 11:00 AM-11:30 AM
- Break 11:30 AM -12:00 PM
- Lunch 12:00 PM - 1:00 PM
- Lab 1:00 PM - 3:30 PM
- For logistics please communicate with Juliet Speas

Instructors for Lectures



Banu Ozkan

Banu.Ozkan@asu.edu

[480/965-2890](tel:4809652890)



Steve Presse

spresse@asu.edu



Dmitry Matyushov

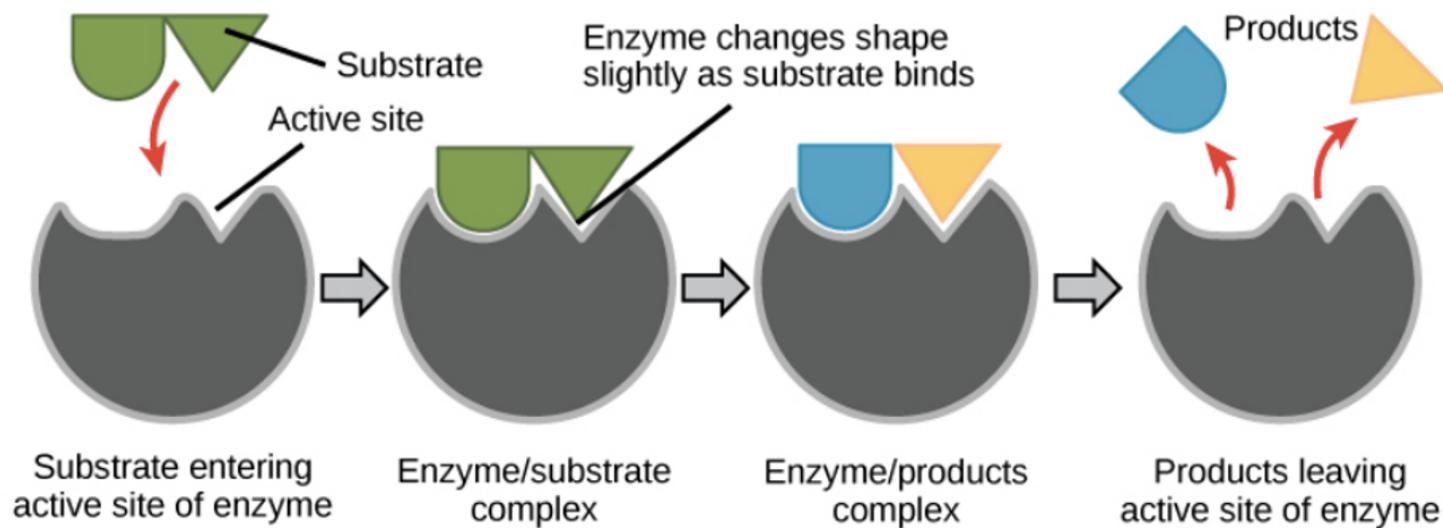
dmitrym@asu.edu

Proteins

Specific Polypeptide sequences : Work horses of a cell

Proteins: Crucial Workhorses of a cell

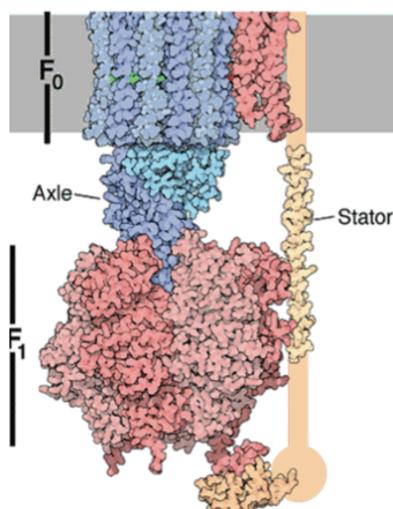
Enzymes: Efficient Catalytic Reaction Centers



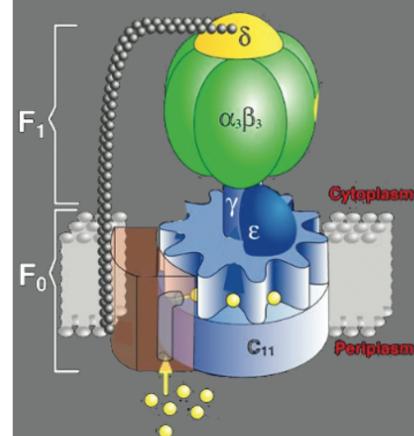
Proteins : workhorses of cell

ATPASE: Energy source in mitochondria

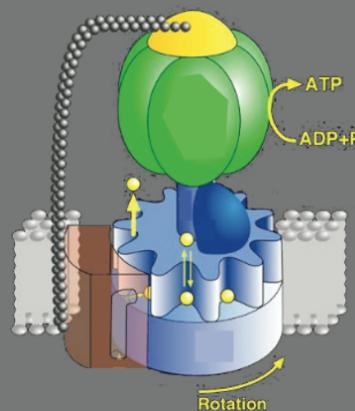
(A) Atomic Structure
of F0F1 ATPase



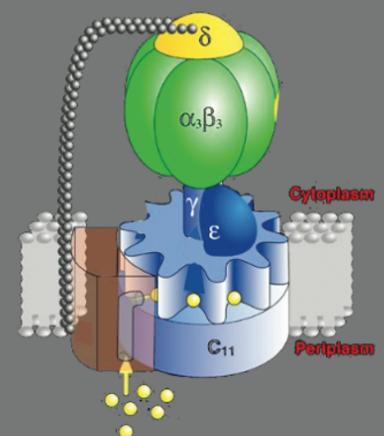
(B) Proton gradient drives three protons to enter F₀



Rotor rotates 120 degrees,
triggering ADP catalysis

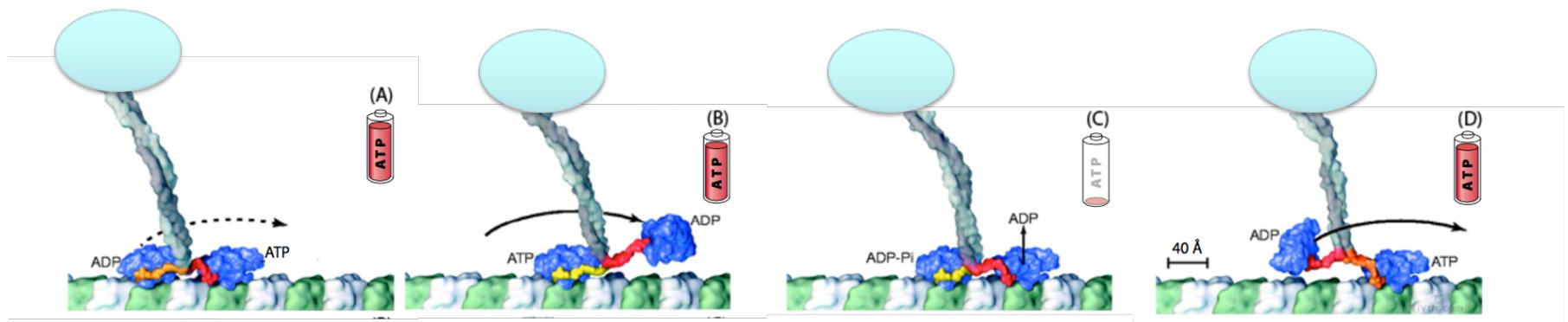


Three more protons enter F₀



Proteins : workhorses of cell

Motor Proteins: Carry Cargo



The blue-prints of proteins:1-D sequence

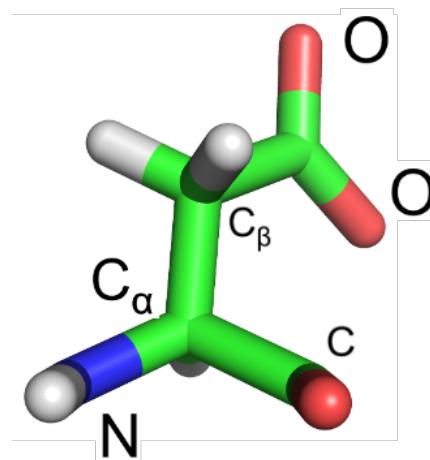
>2GB1:A|PDBID|CHAIN|SEQUENCE

MTYKLILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWTYDDATKTFTVTE

The blue-prints of proteins: 1-D sequence

>2GB1:A|PDBID|CHAIN|SEQUENCE

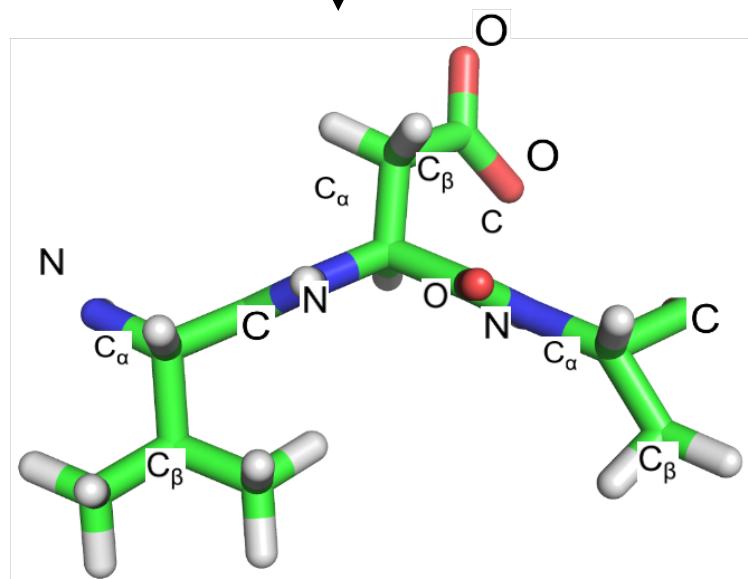
MTYKLILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWTYDDATKTFTVTE



The blue-prints of proteins: 1-D sequence

>2GB1:A|PDBID|CHAIN|SEQUENCE

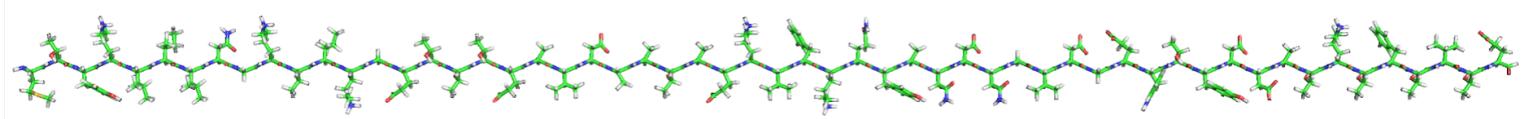
MTYKLILNGKTLKGETTEAVDAATAEKVFKQYANDNGVDGEWTYDDATKTFTVTE



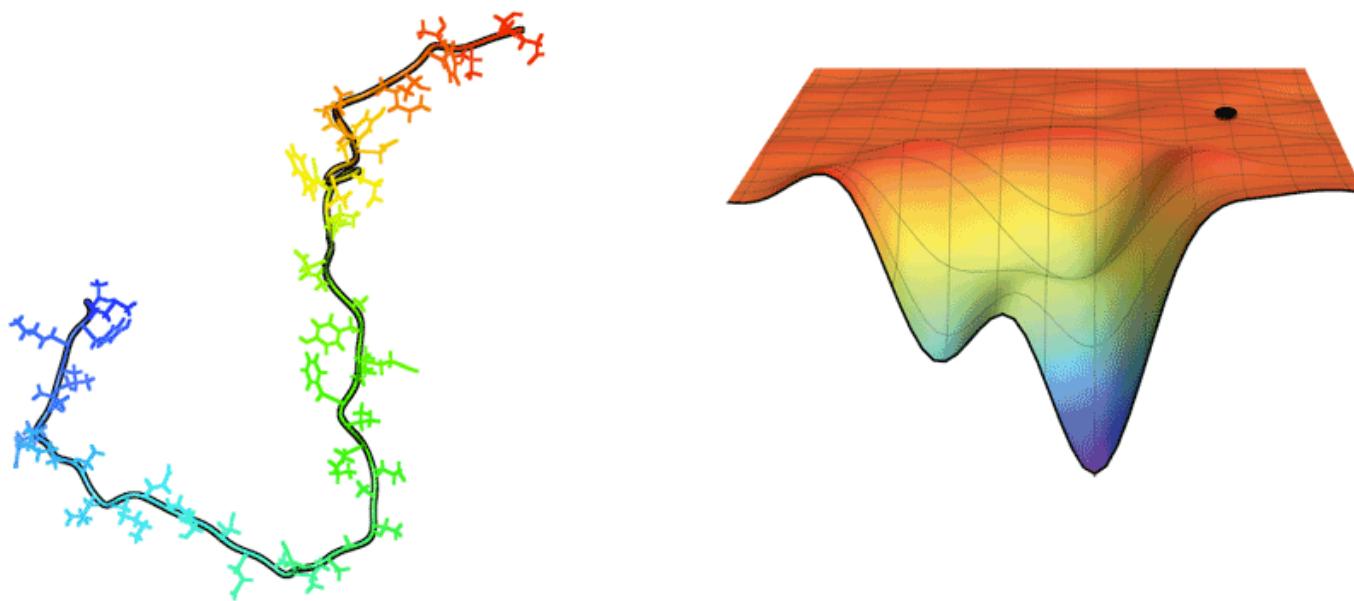
The blue-prints of proteins:1-D sequence

>2GB1:A|PDBID|CHAIN|SEQUENCE

MTYKLILNGKTLKGETTTEAVDAATAEKVFKQYANDNGVDGEWTYDDATKTFTVTE



1-D sequence encodes a unique 3-D structure



Protein structures as free energy minima

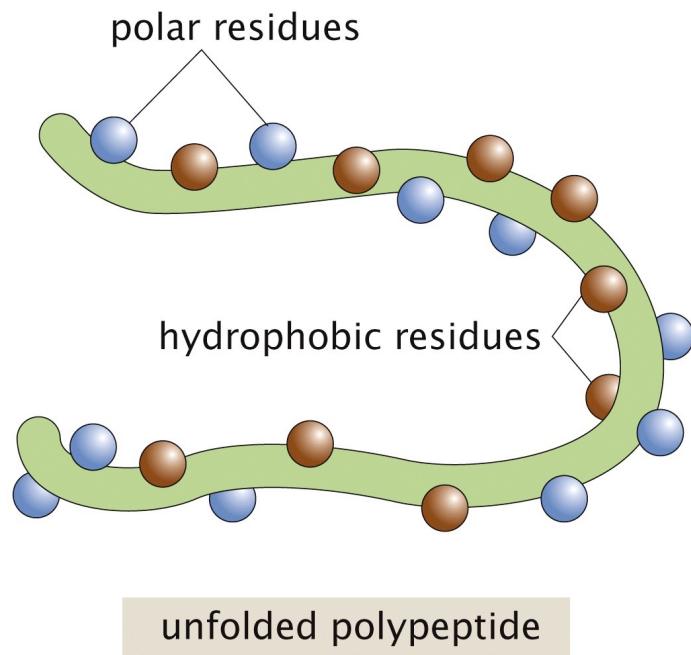


Figure 5.8 (part 1 of 2) Physical Biology of the Cell, 2ed. (© Garland Science 2013)

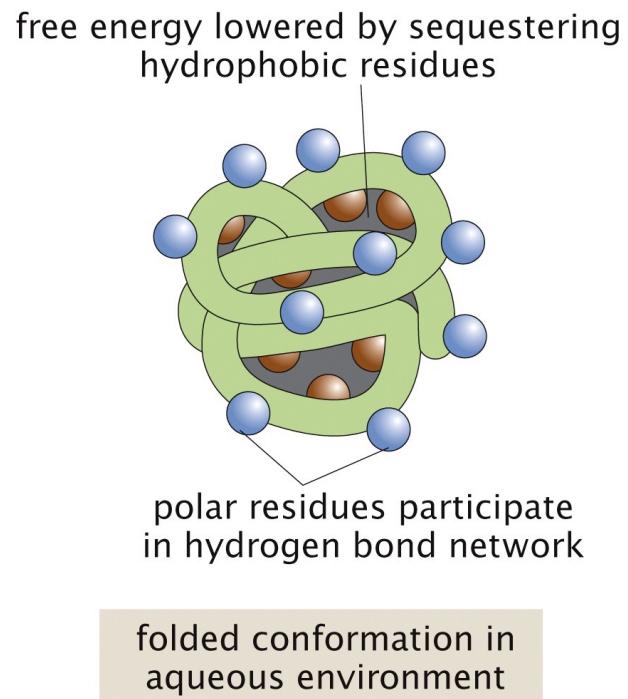
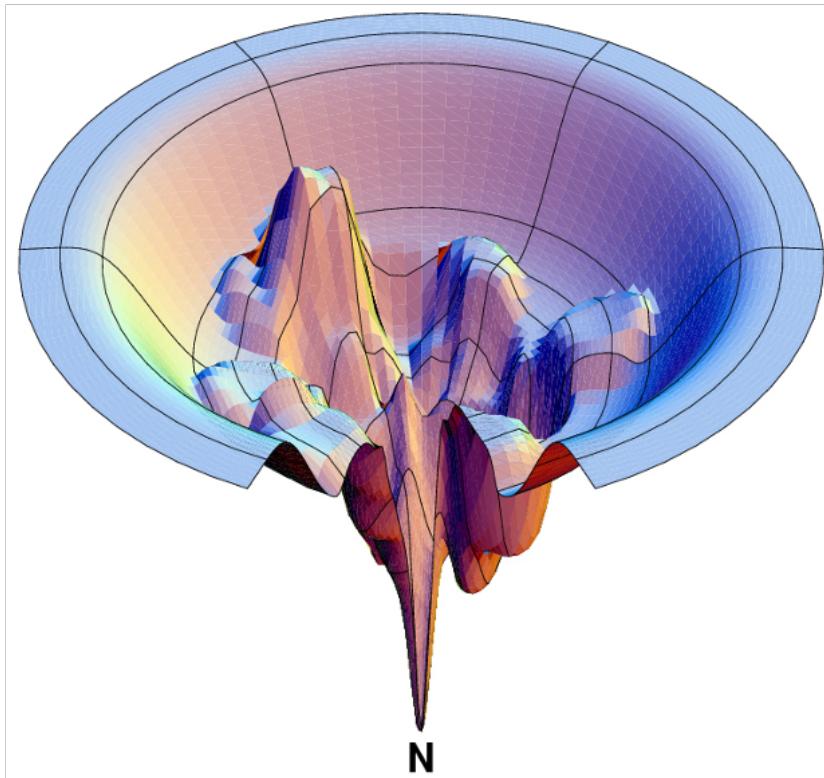


Figure 5.8 (part 2 of 2) Physical Biology of the Cell, 2ed. (© Garland Science 2013)

Levinthal “paradox”: proteins cannot fold by “randomly trying conformations”

Native state is the lowest free energy state



Sequence

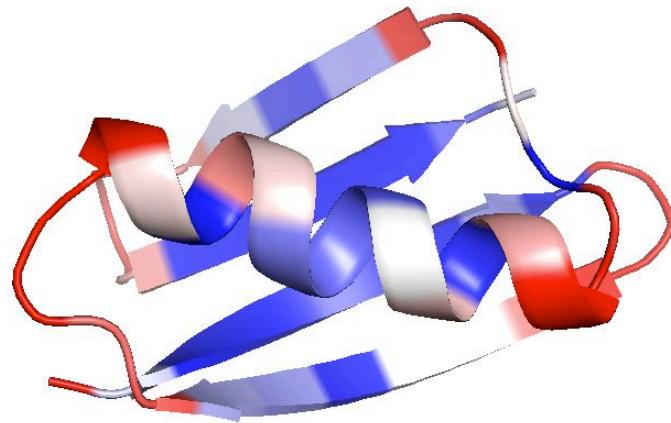
```
>1OKC:A|PDBID|CHAIN|SEQUENCE
SDQALSFLKDFLAGGVAAAISKTAVAPIERVKLLLQVQHASKQISAEKQYKGIIDCVVRIPKEQGFLSFWRGNLANVIRY
FPTQALNFAFKDKYQIIFLGGVDRHKQFWRYFACNLASCGGAAGATSLCFVYPLDFARTRLAADVGKGAAQREFTCLGNCT
TKIFKSDGLRGLYQGFNVSVQGIIYRAAYFGVYDTAKGMLPDPKNVHIIIVSWMIAQTVTAVAGLVSYPPDTVRRRMMQ
SGRGADIMYTGTVDWRKIAKDEGPKAFFKGAWSNVLRGMGCAFVLVLYDEIKKFV
```

Structure

Sequence

```
>1OKC:A | PDBID | CHAIN | SEQUENCE
SDQALSFLKDFLAGVAAAISKTAVAPIERVKLLLQVQHASKQISAEKQYKGII DCVVRIPKEQGFLSFWRGNLANVIRY
FPTQALNFAFKDKYQIFLGGVDRHKQFWRYFACNLASGGAAGATSLCFVYPLDFARTRLAADVGKGAAQREFTCLGNCI
MVTEDVQDST DDT VVDDDDVQDQST TTTTVA A VEDVIVMVA VMT DDDVVVIT TTYGMLVLT VVQVDDVQVDDVQV
```

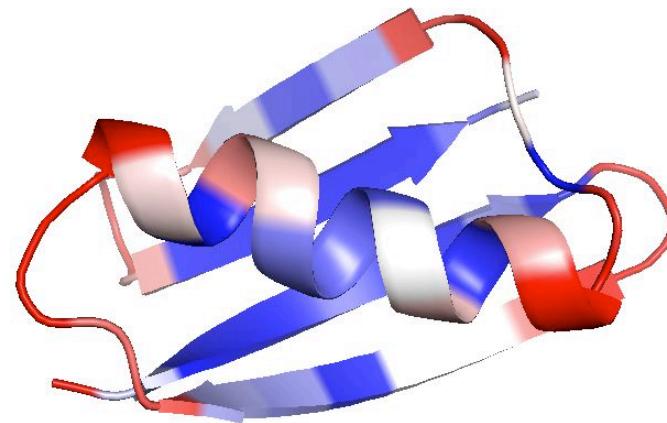
Structure



Sequence

```
>1OKC:A | PDBID | CHAIN | SEQUENCE
SDQALSFLKDFLAGVAAAISKTAVAPIERVKLLLQVQHASKQISAEKQYKGIIDCVVRIPKEQGFLSFWRGNLANVIRY
FPTQALNFAFKDKYKQIFLGGVDRHKQFWRYFACNLASGGAAGATSLCFVYPLDFARTRLAADVGKGAQREFTCLGNCI
TKJ
SGF
```

Structure



Dynamics

Sequence

```
>1OKC:A | PDBID | CHAIN | SEQUENCE  
SDQALSFLKDFLAGVAAAISKTAVAPIERVKLQQVQHASKQISAEKQYKGII DCVVRIPKEQGFLSFWRGNLANVIRY  
FPTQALNFAFKDKYQIIFLGGVDRHKQFWRYFACNLASCGGAAGATSLCFVYPLDFARTRLAADVGKGAQREFTCLGNCT  
TKIFKSDGLRGLYQGFNVSVQGIIIYRAAYFGVYDTAKGMLPDPKNVHIIIVSWMIAQTVTAVAGLVSYPPDTVRMMMQ  
SGRKGADIMYTGTVDWRKIAKDEGPKAFFKGAWSNVLRGMGGAFLVLYDEIKKFV
```

Structure

Dynamics

Function

There is structural hierarchy in proteins

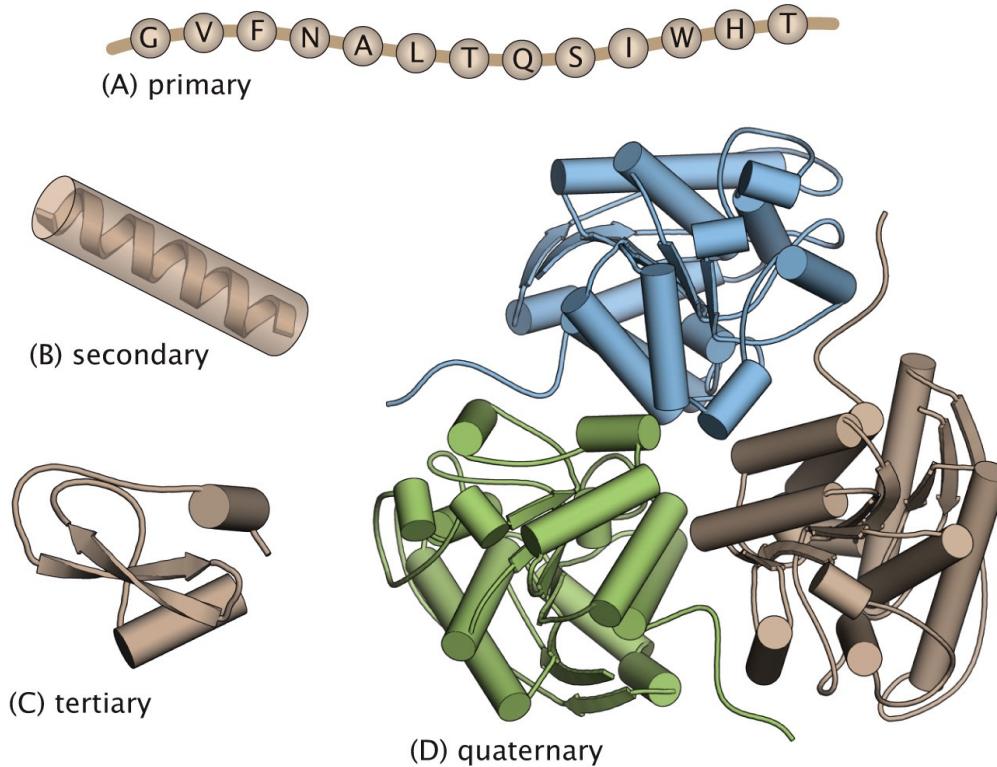


Figure 1.13 Protein Actions (© Garland Science 2017)

Local versus non-local interactions

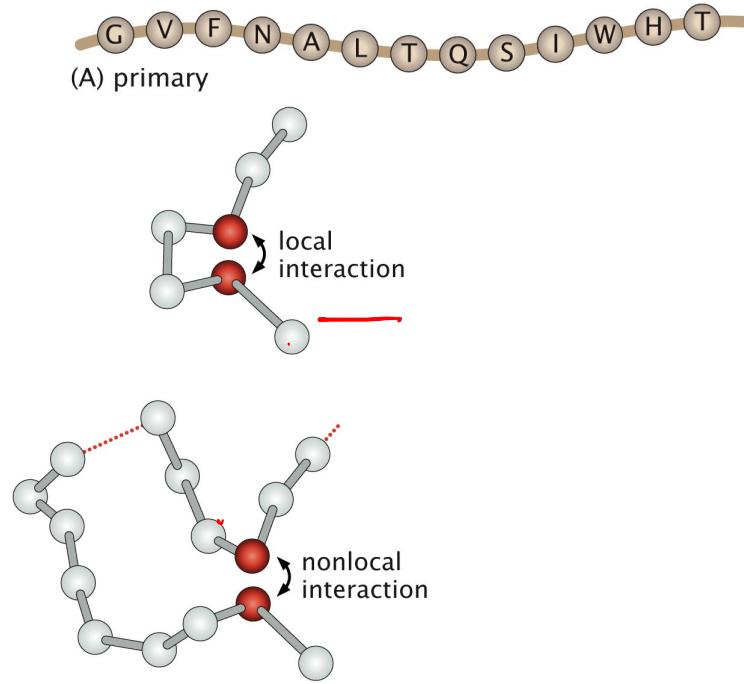
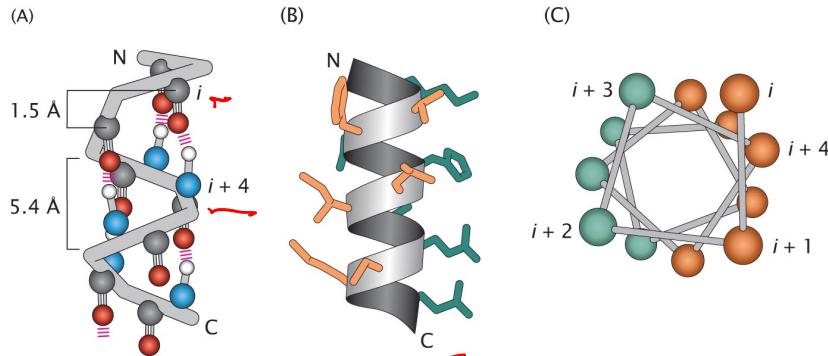


Figure 3.4 Protein Actions (© Garland Science 2017)

The Secondary Structures of Proteins Are Helices and Sheets

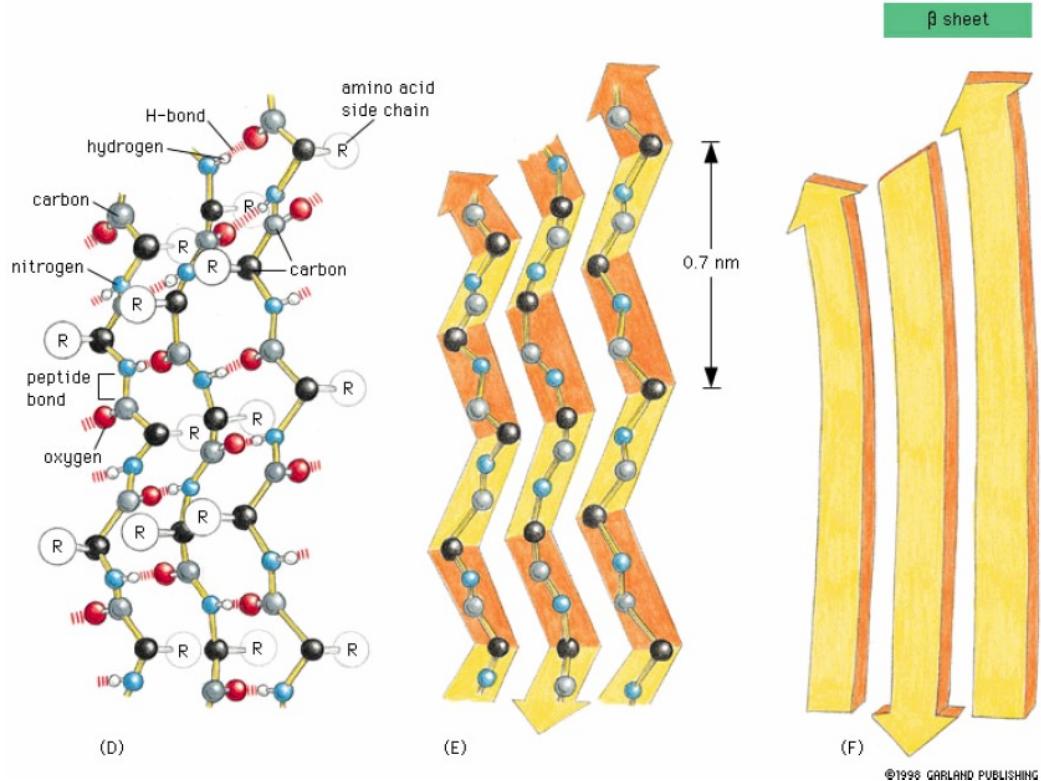
α -helices



The broad occurrence of the α - helical structure arises from two sources.

- (i) it is an energetically accessible conformation due to dihedral angle preferences (see soon) for these (ϕ, ψ) angles
- (ii) the α -helix is stabilized by hydrogen bonds between the carbonyl oxygen of amino acid i and the amide hydrogen of amino acid $i + 4$
- (iii) α -helices can be formed by any of the amino acids (except proline) because the hydrogen bond donors and acceptors are backbone atoms.

Secondary structures: β -sheets



Protein structures can be determined experimentally and computationally

- One can get the spatial coordinates of experimentally determined protein structures from protein data bank
- <https://www.rcsb.org/>

The screenshot shows the homepage of the RCSB PDB (Protein Data Bank) website. The top navigation bar includes links for Deposit, Search, Visualize, Analyze, Download, Learn, About, Documentation, Careers, COVID-19, MyPDB, and Contact us. The main search bar displays "223,166 Structures from the PDB" and "1,068,577 Computed Structure Models (CSMs)". Below the search bar are links for Advanced Search and Browse Annotations, along with social media icons for Facebook, Twitter, YouTube, and LinkedIn. A banner at the top right promotes "Access Computed Structure Models (CSMs) of available model organisms". The left sidebar features a "Welcome" section with links for Deposit, Search, Visualize, Analyze, Download, and Learn. The central content area highlights "Experimentally-determined 3D structures from the Protein Data Bank (PDB) archive" and "Computed Structure Models (CSM) from AlphaFold DB and ModelArchive". It also features a "NEW! Explore NEW Features" section with a lightbulb icon and a "PDB-101 Training Resources" section. A large image of a protein structure labeled "ESCRT-III" is prominently displayed under the "August Molecule of the Month" heading.

Protein Coordinates and Calculating distances