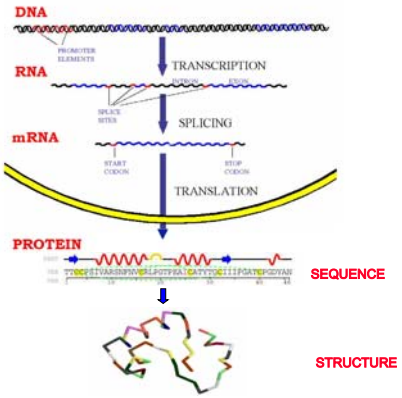
Introduction to Bioinformatics

Iosif Vaisman

Email: ivaisman@gmu.edu

Cell Informatics



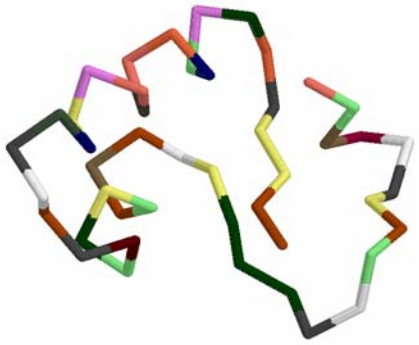
Standard genetic code

```
  AAs = FFLLSSSSYY**CC*WLLLLPPPPHHQQRRRRIIIMTTTTNNKKSSRRVVVVAAAADDEEGGGG
Starts = ---M----------------M----------------M----------------------------
 Base1 = TTTTTTTTTTTTTTTTCCCCCCCCCCCCCCCCAAAAAAAAAAAAAAAAGGGGGGGGGGGGGGGG
 Base2 = TTTTCCCCAAAAGGGGTTTTCCCCAAAAGGGGTTTTCCCCAAAAGGGGTTTTCCCCAAAAGGGG
 Base3 = TCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAGTCAG
```

Frameshift Errors

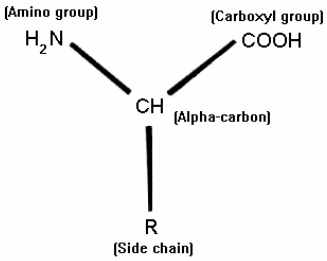
```
ATGAAATTTGGAAACTTCCTTCTCACTTATCAGCCACCTGAGCTATCTCAGACCGAAGTGATGAAGCGATTGGTTAATCT

5'3'Frame1 MKFGNFLLTYQPPELSQTEVMKRLVN
5'3'Frame2 -NLETSFSLISHLSYLRPK--SDWLI
5'3'Frame3 EIWKLPSHLSAT-AISDRSDEAIG-S
3'5'Frame1 RLTNRFITSV-DSSGG--VRRKFPNF
3'5'Frame2 D-PIASSLRSEIAQVADK-EGSFQIS
3'5'Frame3 INQSLHHFGLR-LRWLISEKEVSKFH
```

Protein backbone

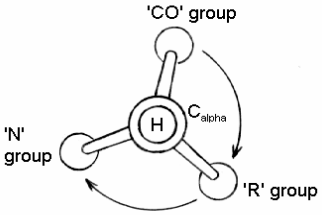


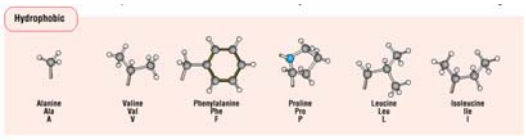
```
NWVLSTAADMQGVVTDGMASGLDKD
```

Amino Acid Residue

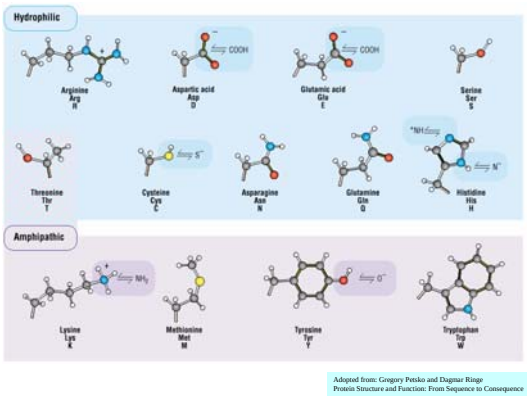


Amino Acid Residue



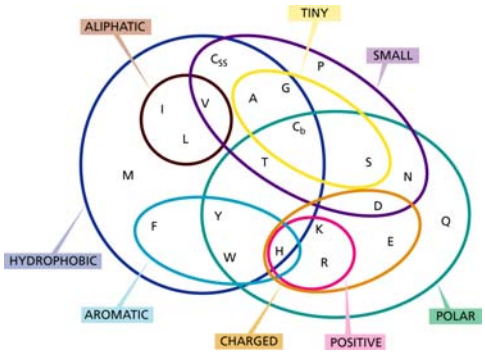


Amino Acids

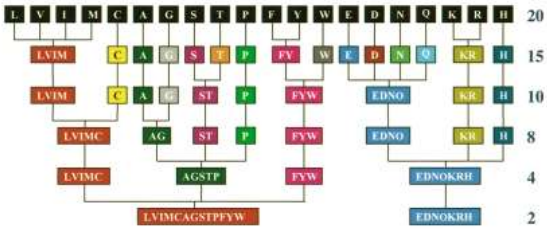


Adopted from: Gregory Petsko and Dagmar Ringe
Protein Structure and Function: From Sequence to Consequence

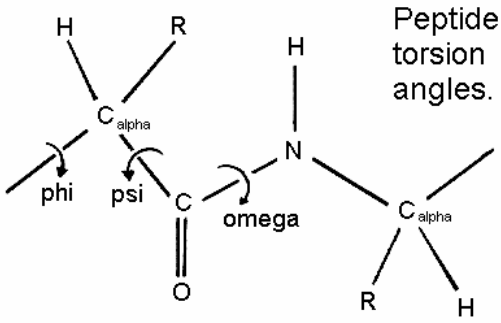
Amino Acid Residue Clustering



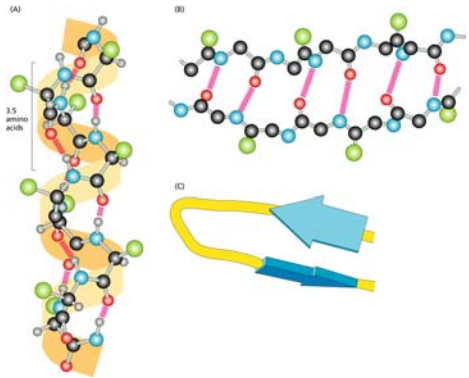
Amino Acid Residue Clustering



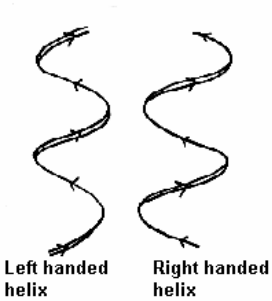
Adopted from: L.R.Murphy et al., 2000



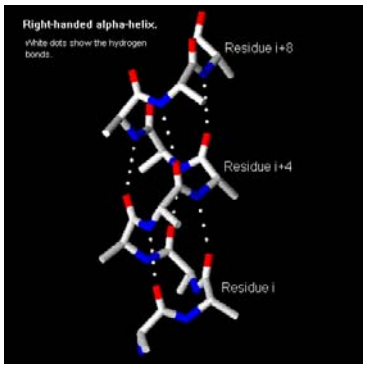
Secondary Structures



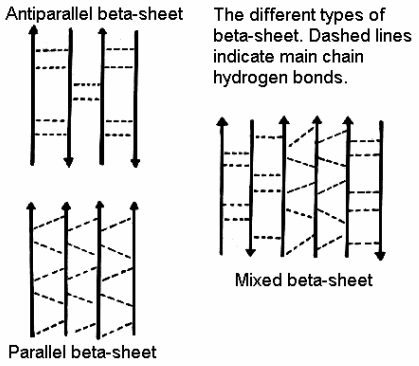
Secondary Structure (Helices)

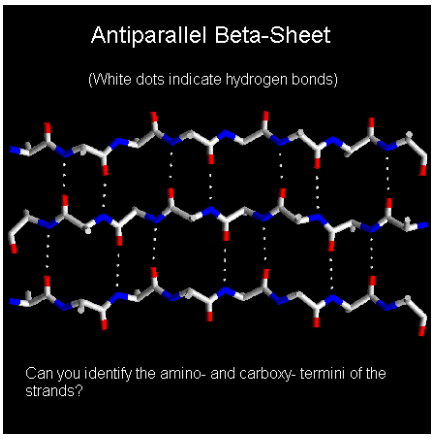


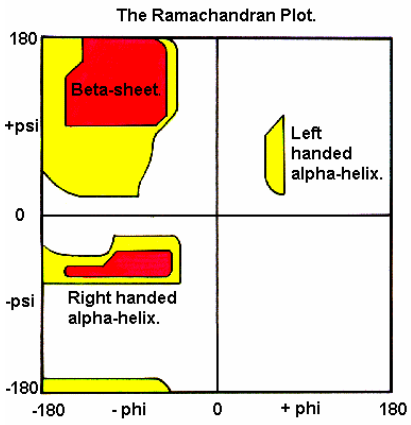
Helix

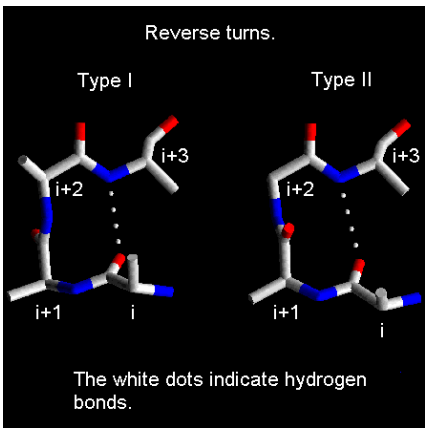


Secondary Structure (Beta-sheets)

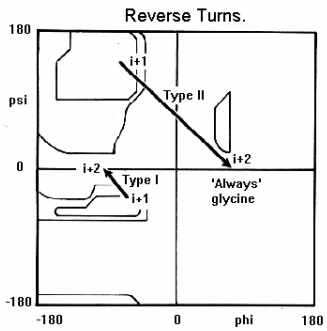


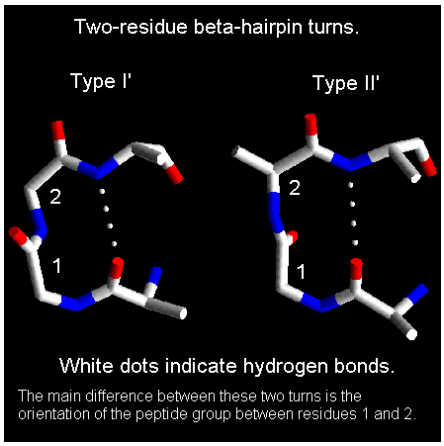




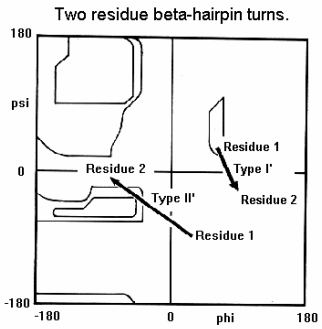


Reverse Turns on a Ramachandran Plot

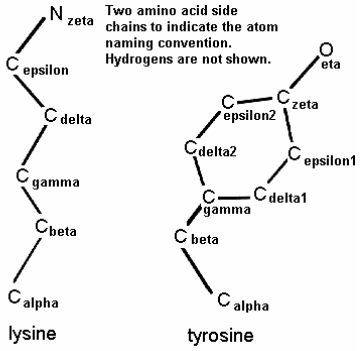




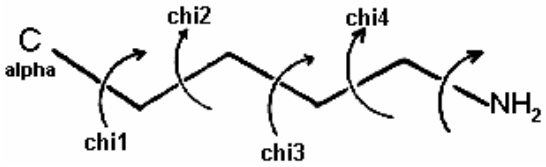
Beta-hairpin Turns on a Ramachandran Plot



Side-Chain Atom Nomenclature



Side-Chain Torsional Angles



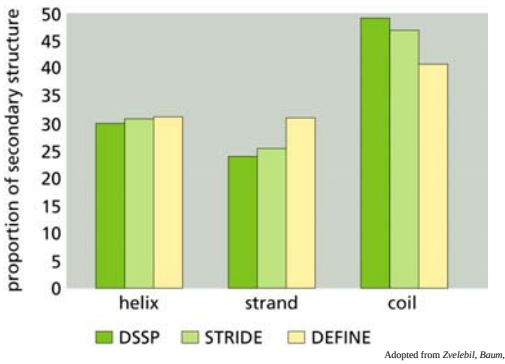
Secondary Structure: Computational Problems

Secondary structure characterization
Secondary structure assignment
Secondary structure prediction
Protein structure classification

Secondary Structure Conformations

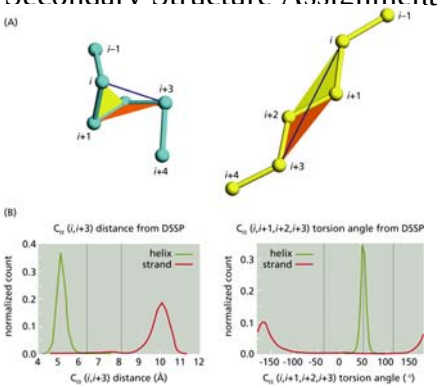
	ϕ	ψ
alpha helix	-57	-47
alpha-L	57	47
3-10 helix	-49	-26
π helix	-57	-80
type II helix	-79	150
β-sheet parallel	-119	113
β-sheet antiparallel	-139	135

Secondary Structure Assignment



Adopted from *Zvelebil, Baum, 2008*

Secondary Structure Assignment



Adopted from *Zvelebil, Baum, 2008*

Secondary Structure Assignment

```
....;....1....;....2....;....3....;....4....;....5....;....6....;....7..
    .-- sequential resnumber, including chain breaks as extra residues
    |    .-- original PDB resname, not nec. sequential, may contain letters
    |    |   .-- amino acid sequence in one letter code
    |    |   |  .-- secondary structure summary based on columns 19-38
    |    |   |  | xxxxxxxxxxxxxxxxxxx recommend columns for secstruc details
    |    |   |  | .-- 3-turns/helix
    |    |   |  | |.-- 4-turns/helix
    |    |   |  | ||.-- 5-turns/helix
    |    |   |  | |||.-- geometrical bend
    |    |   |  | ||||.-- chirality
    |    |   |  | |||||.-- beta bridge label
    |    |   |  | ||||||.-- beta bridge label
    |    |   |  | |||||||   .-- beta bridge partner resnum
    |    |   |  | |||||||   |   .-- beta bridge partner resnum
    |    |   |  | |||||||   |   |.-- beta sheet label
    |    |   |  | |||||||   |   ||   .-- solvent accessibility
    |    |   |  | |||||||   |   ||   |
  #  RESIDUE AA STRUCTURE BP1 BP2  ACC
    |    |   |  | |||||||   |   ||   |
   35   47   I  E     +     0   0    2
   36   48   R  E >  S- K   0  39C  97
   37   49   Q  T 3  S+     0   0   86    (example from 1EST)
   38   50   N  T 3  S+     0   0   34
   39   51   W  E <   -KL  36  98C   6
```