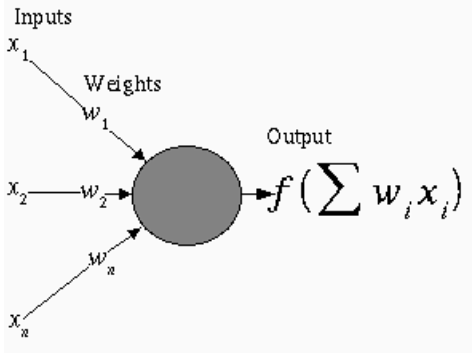
Lecture 13

Basic Neuron Model In A Feedforward Network

- **Inputs x_i** arrive through pre-synaptic connections
- Synaptic efficacy is modeled using real **weights w_i**
- The response of the neuron is a **nonlinear function *f*** of its weighted inputs



Output

- The response function is normally nonlinear
- Samples include
  - Sigmoid

$$f(x)=\frac{1}{1+e^{-\lambda x}}$$

  - Piecewise linear

$$f(x)=\begin{cases} x_i & \text{if } x \geq t \\ 0_i & \text{if } x < t \end{cases}$$

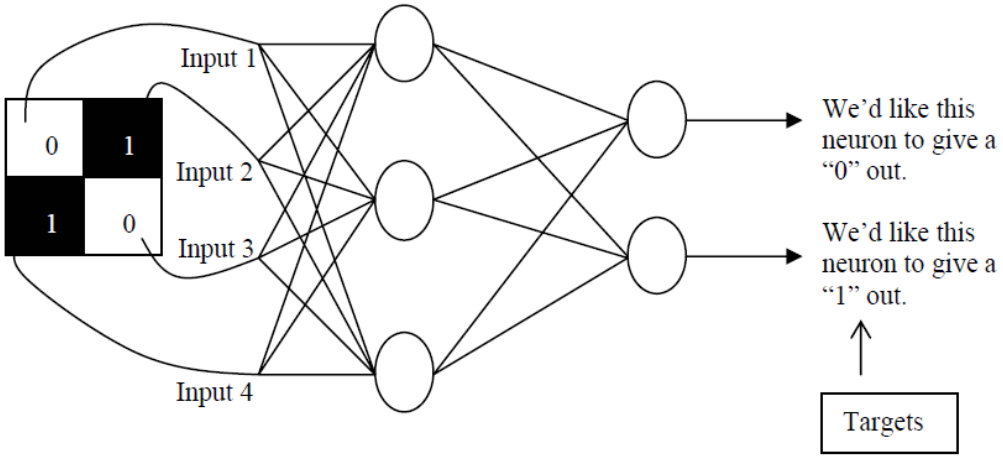
Backpropagation Preparation

- **Training Set**
  A collection of input-output patterns that are used to train the network
- **Testing Set**
  A collection of input-output patterns that are used to assess network performance
- **Learning Rate-η**
  A scalar parameter, analogous to step size in numerical integration, used to set the rate of adjustments

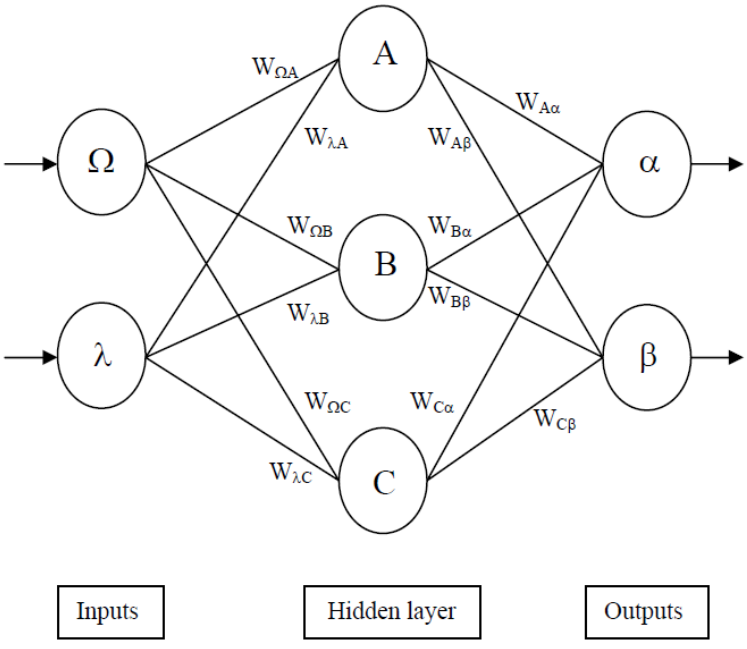
Network Error

- Total-Sum-Squared-Error (TSSE)

- Root-Mean-Squared-Error (RMSE)



A reverse pass of Back Propagation



Pseudo Code

1. Apply the inputs to the network and work out the output – remember this initial output could be anything, as the initial weights were random numbers.

2. Next work out the error for neuron B. The error is *What you want – What you actually get*, in other words:

$$Error_B = Output_B (1-Output_B)(Target_B – Output_B)$$

The "*Output(1-Output)*" term is necessary in the equation because of the Sigmoid Function – if we were only using a threshold neuron it would just be *(Target – Output)*.

3. Change the weight. Let W^+_{AB} be the new (trained) weight and W_{AB} be the initial weight.

$$W^+_{AB} = W_{AB} + (Error_B \times Output_A)$$

Notice that it is the output of the connecting neuron (neuron A) we use (not B). We update all the weights in the output layer in this way.

Pseudo Code

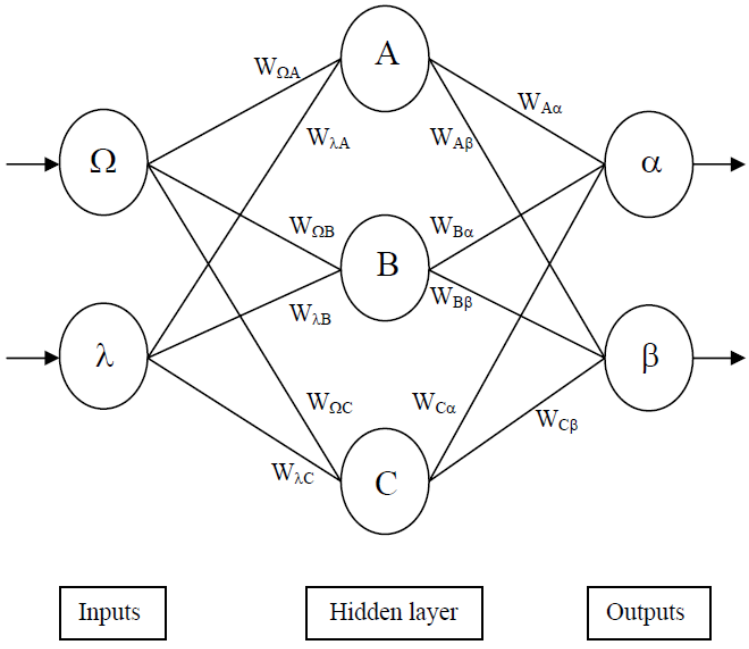
4. Calculate the Errors for the hidden layer neurons. Unlike the output layer we can't calculate these directly (because we don't have a Target), so we *Back Propagate* them from the output layer (hence the name of the algorithm). This is done by taking the Errors from the output neurons and running them back through the weights to get the hidden layer errors. For example if neuron A is connected as shown to B and C then we take the errors from B and C to generate an error for A.

$$Error_A = Output_A\ (1 - Output_A)(Error_B\ W_{AB} + Error_C\ W_{AC})$$

Again, the factor "*Output (1 - Output)*" is present because of the sigmoid squashing function.

5. Having obtained the Error for the hidden layer neurons now proceed as in stage 3 to change the hidden layer weights. By repeating this method we can train a network of any number of layers.

A reverse pass of Back Propagation



A reverse pass of Back Propagation

1. Calculate errors of output neurons

$$\delta_\alpha = out_\alpha\,(1 - out_\alpha)\,(Target_\alpha - out_\alpha)$$
$$\delta_\beta = out_\beta\,(1 - out_\beta)\,(Target_\beta - out_\beta)$$

2. Change output layer weights

$$W^+_{A\alpha} = W_{A\alpha} + \eta\delta_\alpha\, out_A \qquad W^+_{A\beta} = W_{A\beta} + \eta\delta_\beta\, out_A$$
$$W^+_{B\alpha} = W_{B\alpha} + \eta\delta_\alpha\, out_B \qquad W^+_{B\beta} = W_{B\beta} + \eta\delta_\beta\, out_B$$
$$W^+_{C\alpha} = W_{C\alpha} + \eta\delta_\alpha\, out_C \qquad W^+_{C\beta} = W_{C\beta} + \eta\delta_\beta\, out_C$$

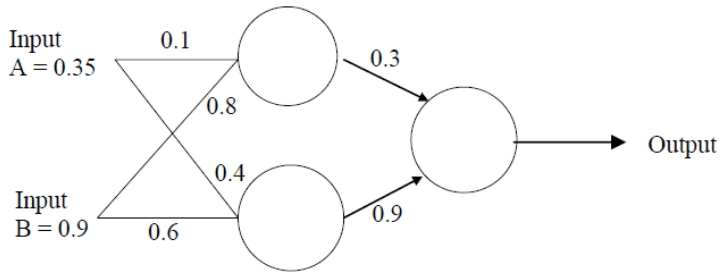
3. Calculate (back-propagate) hidden layer errors

$$\delta_A = out_A\,(1 - out_A)\,(\delta_\alpha W_{A\alpha} + \delta_\beta W_{A\beta})$$
$$\delta_B = out_B\,(1 - out_B)\,(\delta_\alpha W_{B\alpha} + \delta_\beta W_{B\beta})$$
$$\delta_C = out_C\,(1 - out_C)\,(\delta_\alpha W_{C\alpha} + \delta_\beta W_{C\beta})$$

4. Change hidden layer weights

$$W^+_{\lambda A} = W_{\lambda A} + \eta\delta_A\, in_\lambda \qquad W^+_{\Omega A} = W^+_{\Omega A} + \eta\delta_A\, in_\Omega$$
$$W^+_{\lambda B} = W_{\lambda B} + \eta\delta_B\, in_\lambda \qquad W^+_{\Omega B} = W^+_{\Omega B} + \eta\delta_B\, in_\Omega$$
$$W^+_{\lambda C} = W_{\lambda C} + \eta\delta_C\, in_\lambda \qquad W^+_{\Omega C} = W^+_{\Omega C} + \eta\delta_C\, in_\Omega$$

Example



Assume that the neurons have a Sigmoid activation function and

(i) Perform a forward pass on the network.
(ii) Perform a reverse pass (training) once (target = 0.5).
(iii) Perform a further forward pass and comment on the result.

Example

(i)
Input to top neuron = (0.35x0.1)+(0.9x0.8)=0.755. Out = 0.68.
Input to bottom neuron = (0.9x0.6)+(0.35x0.4) = 0.68. Out = 0.6637.
Input to final neuron = (0.3x0.68)+(0.9x0.6637) = 0.80133. Out = 0.69.

(ii)
Output error $\delta=(t-o)(1-o)o = (0.5-0.69)(1-0.69)0.69 = -0.0406$.

New weights for output layer
$w1^+ = w1+(\delta \times input) = 0.3 + (-0.0406x0.68) = 0.272392$.
$w2^+ = w2+(\delta \times input) = 0.9 + (-0.0406x0.6637) = 0.87305$.

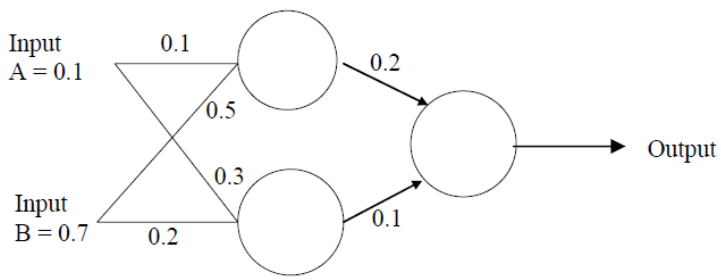
Errors for hidden layers:
$\delta1 = \delta \times w1 = -0.0406 \times 0.272392 \times (1-o)o = -2.406x10^{-3}$
$\delta2 = \delta \times w2 = -0.0406 \times 0.87305 \times (1-o)o = -7.916x10^{-3}$

New hidden layer weights:
$w3^+=0.1 + (-2.406 \times 10^{-3} \times 0.35) = 0.09916$.
$w4^+ = 0.8 + (-2.406 \times 10^{-3} \times 0.9) = 0.7978$.
$w5^+ = 0.4 + (-7.916 \times 10^{-3} \times 0.35) = 0.3972$.
$w6^+ = 0.6 + (-7.916 \times 10^{-3} \times 0.9) = 0.5928$.

(iii)
Old error was -0.19. New error is -0.18205. Therefore error has reduced.

Homework

Try a training pass on the following example. Target = 1, Learning rate = 1:



Lecture 14-15

PSSPred / PSIpred: Prediction of secondary structure from amino acid sequences

Input:
- Amino acid sequence

Generates:
- MTX file by running PSI-BLAST

Output:
- Confidence on H/E/C

PSIPred Flow Diagram

Generation of sequence profiles

Intermediate PSIBLAST profiles as a direct input to a secondary structure prediction method rather than using an explicit multiple sequence alignment

The final position-specific scoring matrix (log odds values) from PSIBLAST (after three iterations) is used as input to the neural network.

This matrix has $20 \times M$ elements, where M is the length of the target sequence, and each element represents the log-likelihood of that particular residue substitution at that position in the template (based on a weighted average of BLOSUM62 matrix scores for the given alignment position).

The profile matrix elements (typically in the range $\pm$ seven) are scaled to the required 0-1 range by using the standard logistic function (where x is the raw profile matrix value):

$$\frac{1}{1+e^{-x}}$$

PSIPred Flow Diagram

Neural Network Architecture

A standard feed-forward back-propagation network architecture with a single hidden layer was used for PSIPRED.

A window of 15 amino acid residues was found to be optimal.

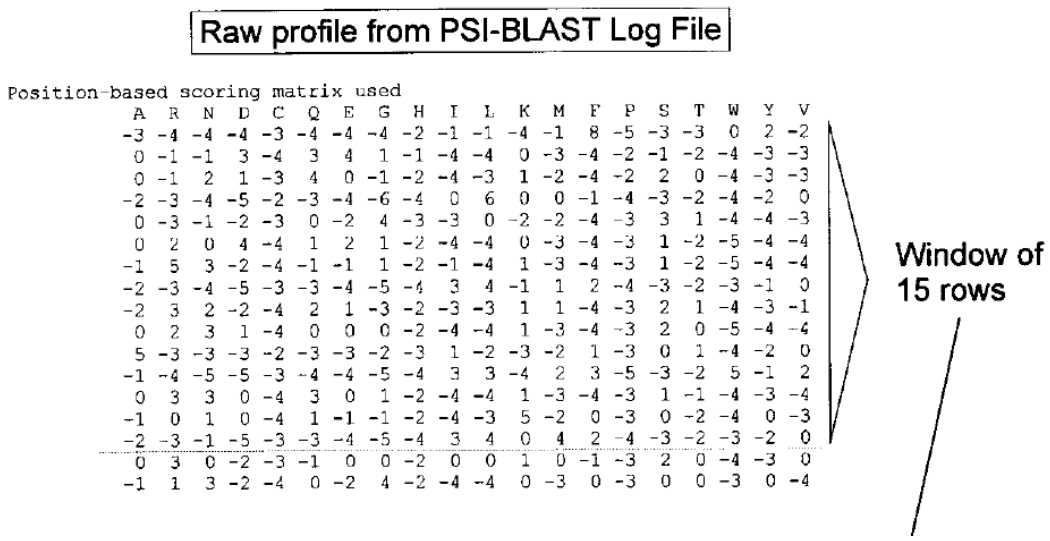
Thus the final input layer comprises 315 input units, divided into 15 groups of 21 units.

A large hidden layer of 75 units was used, with another three units making the output
layer where the units represent the three-states of secondary structure (helix, strand or coil).

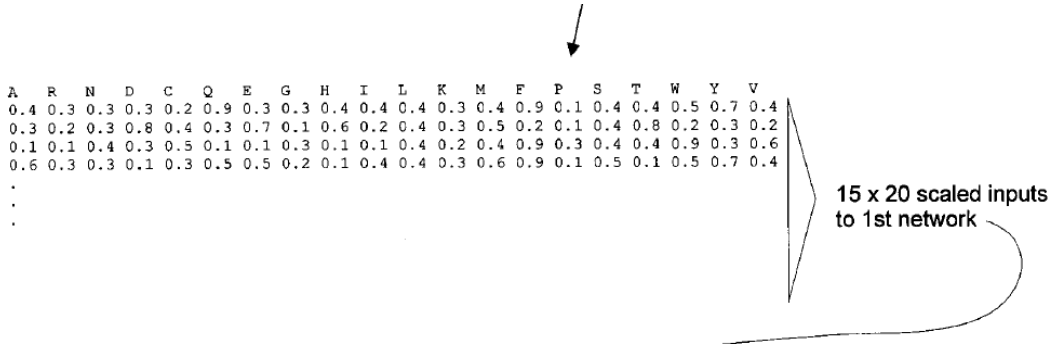
A second network is used to filter successive outputs from the main network.

As only three possible inputs are necessary for each amino acid position, this network has an input layer comprising just 60 input units, divided into 15 groups of four.

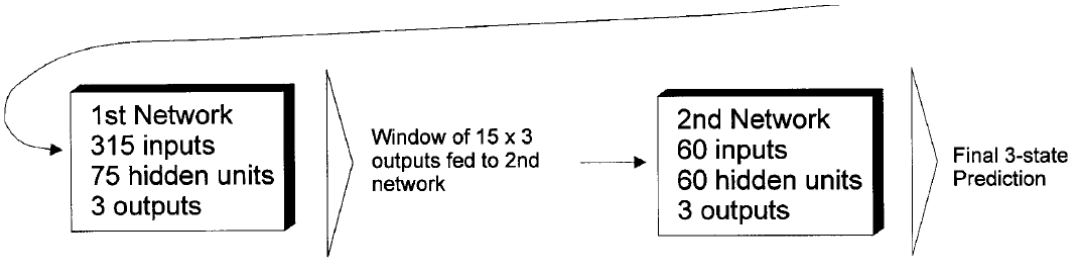
PSIPred Flow Diagram



PSIPred Flow Diagram



PSIPred Flow Diagram



PSIPred

Neural network training

On-line back-propagation training procedure was used to optimize the network weights.

A learning rate of 0.005 was found to be effective.

Cross validation (to avoid overfitting)

Hold-out, K-fold cross validation, leave-one-out cross validation

Testing

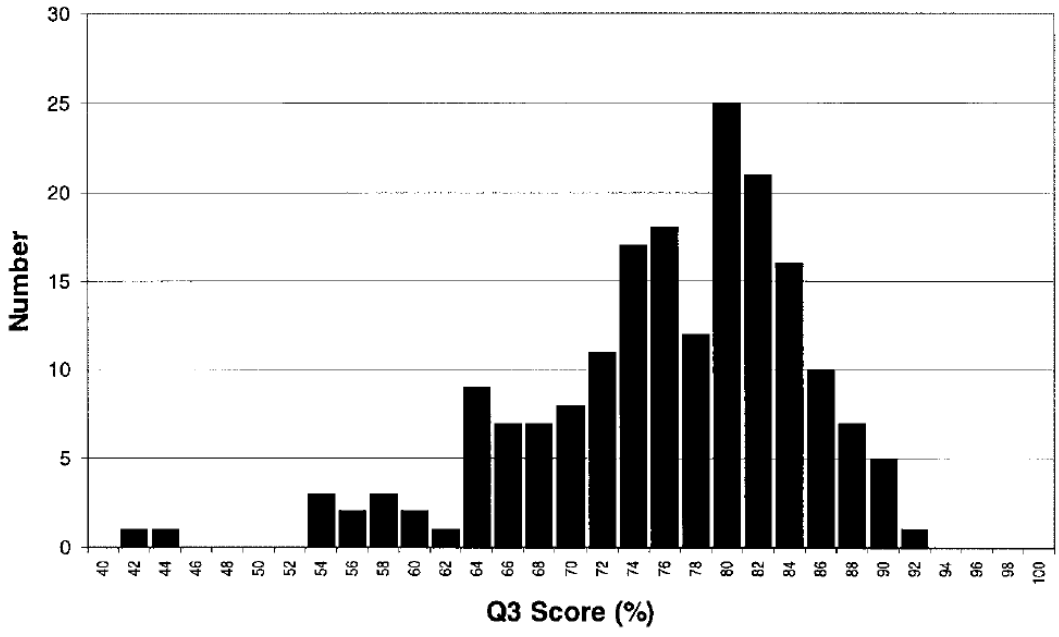
Result

Average Q3 score for 187 test protein chains is 76.0(±7.8)%

Liam J. McGuffin, Kevin Bryson, David T. Jones, The PSIPRED protein structure prediction server , *Bioinformatics*, Volume 16, Issue 4, April 2000, Pages 404–405, https://doi.org/10.1093/bioinformatics/16.4.404

PSIPred Result

http://bioinf.cs.ucl.ac.uk/psipred/



Prediction of secondary structure from single amino acid sequences

Input:

- Amino acid sequence

Generates:

- Amino acid composition score
- Secondary structure propensity score
- BLOSUM62 scoring matrix

Output:

- Confidence on H/E/C

An Evolution-Based Approach to *De Novo* Protein Design and Case Study on *Mycobacterium tuberculosis* Mitra P, Shultis D, Brender JR, Czajka J, Marsh D, et al. (2013) An Evolution-Based Approach to *De Novo* Protein Design and Case Study on *Mycobacterium tuberculosis*. PLOS Computational Biology 9(10): e1003298. https://doi.org/10.1371/journal.pcbi.1003298

Prediction of secondary structure from single amino acid sequences

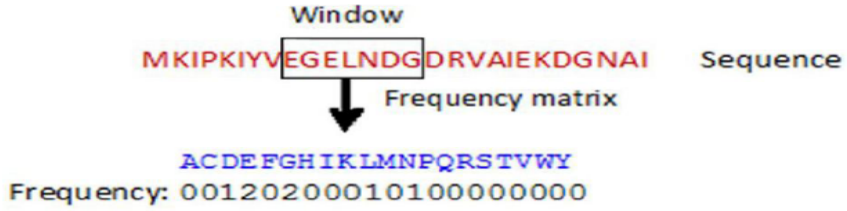


Figure S1. Illustration of amino acid frequency counting within a window size of three residues.

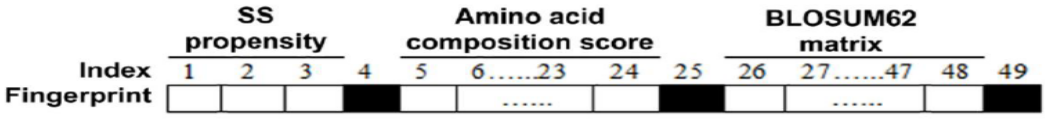
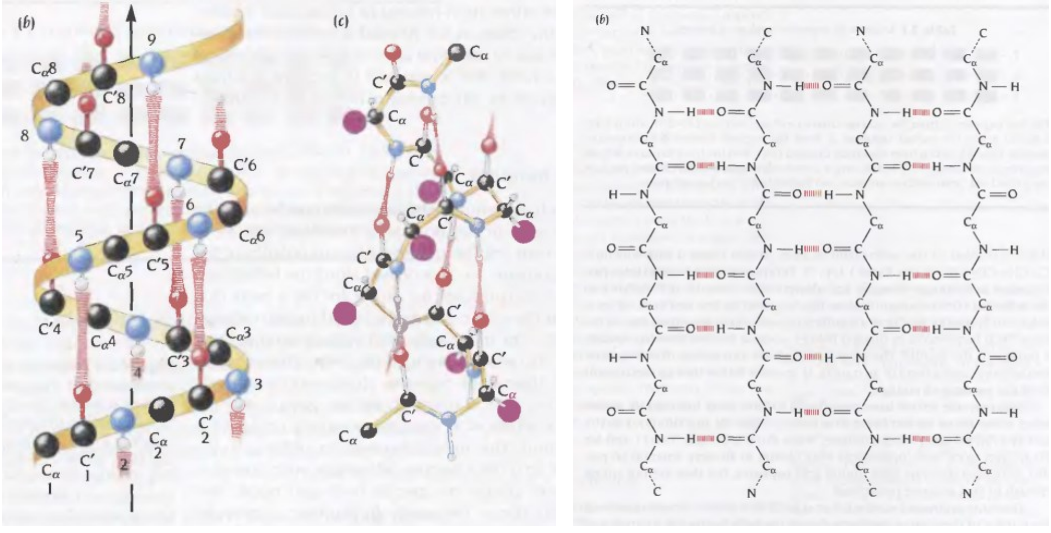


Figure S2. Illustration of fingerprint assignments from neural network SS training where SS propensity score, amino acid composition score, and BLOSUM62 substitution matrix are listed side-by-side with a separation of a noise (black filled cell).

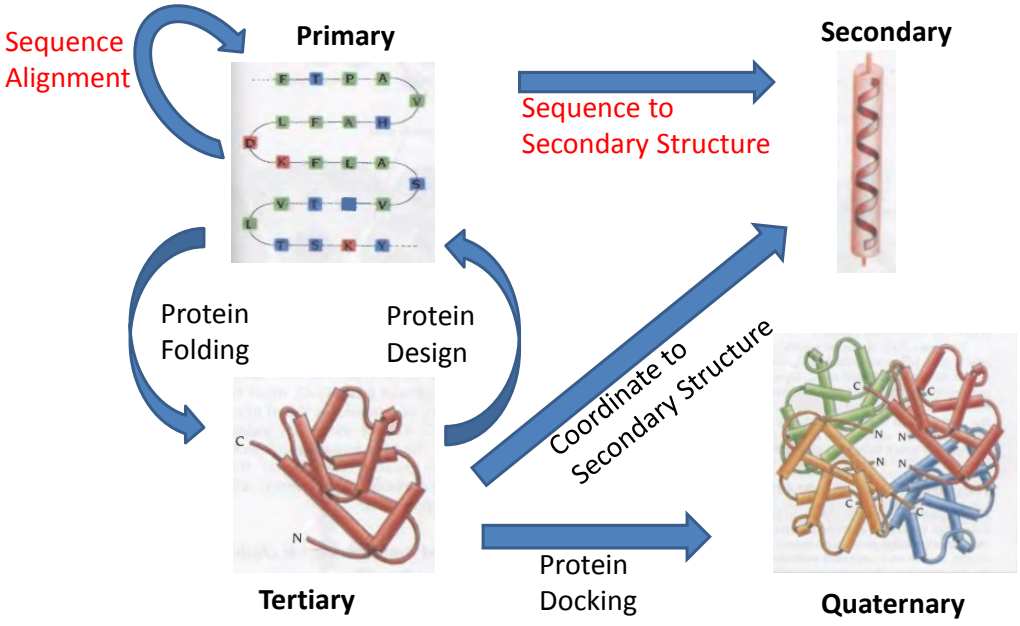
Observation based post-processing



A minimum consecutive amino acids are required for the formation of the helix and sheet.

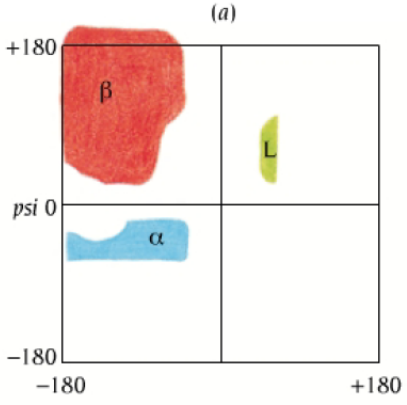
- Improves the accuracy by 3-5%.

Computational Methods in Proteins

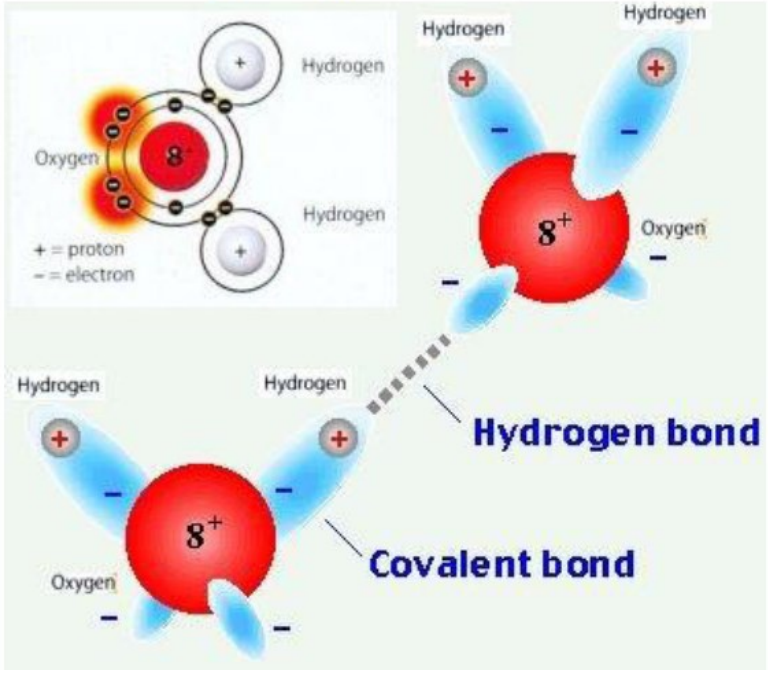


What basic structural information is used ?

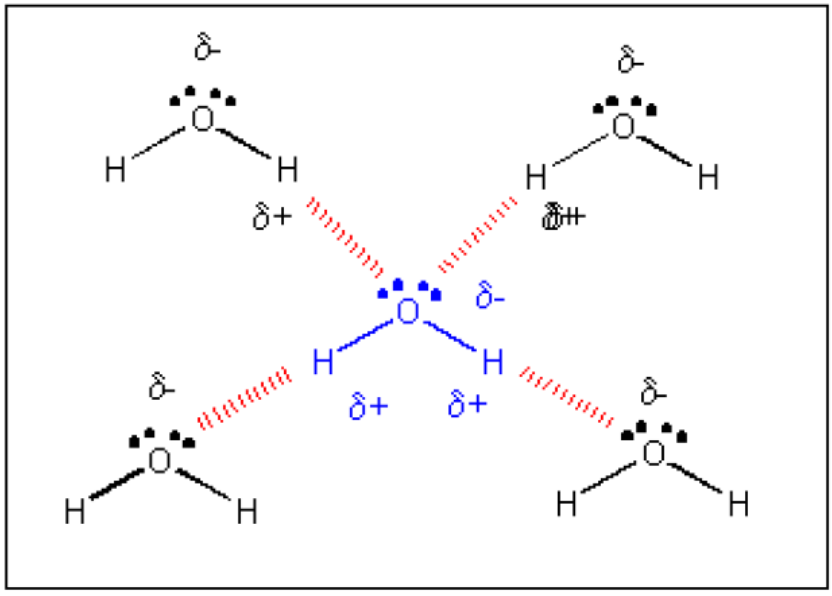
- Hydrogen bond patterns
- Backbone dihedral angles



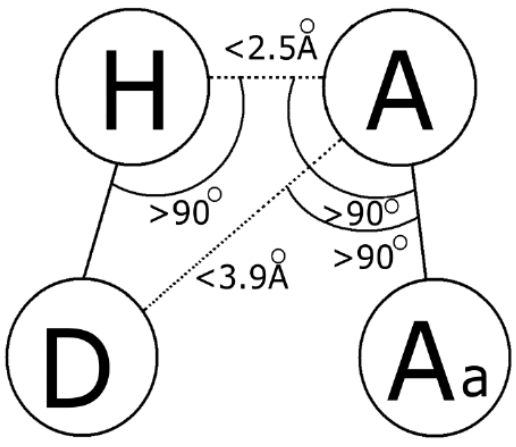
Hydrogen Bond



Hydrogen Bond



Hydrogen Bond



Baker and Hubbard, 1984

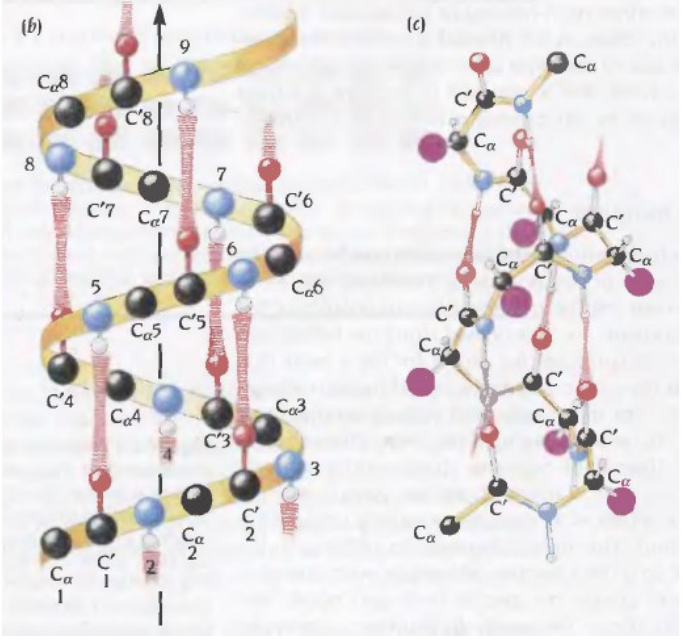
PDB file format

```
ATOM      1  N   ALA A   4      11.751  37.846  29.016  1.00 44.65           N
ATOM      2  CA  ALA A   4      12.501  39.048  28.539  1.00 30.68           C
ATOM      3  C   ALA A   4      13.740  38.628  27.754  1.00 24.74           C
ATOM      4  O   ALA A   4      14.207  37.495  27.890  1.00 25.59           O
ATOM      5  CB  ALA A   4      12.902  39.919  29.730  1.00 16.77           C
ATOM      6  N   TYR A   5      14.235  39.531  26.906  1.00 19.29           N
ATOM      7  CA  TYR A   5      15.552  39.410  26.282  1.00  8.51           C
ATOM      8  C   TYR A   5      16.616  38.913  27.263  1.00  6.11           C
ATOM      9  O   TYR A   5      17.187  37.844  27.068  1.00 17.99           O
ATOM     10  CB  TYR A   5      15.988  40.762  25.702  1.00  2.00           C
ATOM     11  CG  TYR A   5      17.319  40.745  24.982  1.00  2.00           C
ATOM     12  CD1 TYR A   5      17.411  40.331  23.653  1.00 19.29           C
ATOM     13  CD2 TYR A   5      18.476  41.210  25.604  1.00  9.65           C
ATOM     14  CE1 TYR A   5      18.629  40.396  22.953  1.00  2.00           C
ATOM     15  CE2 TYR A   5      19.703  41.271  24.914  1.00  8.78           C
ATOM     16  CZ  TYR A   5      19.763  40.863  23.594  1.00  8.76           C
ATOM     17  OH  TYR A   5      20.971  40.889  22.920  1.00  8.23           O
ATOM     18  N   ILE A   6      16.789  39.630  28.369  1.00 14.56           N
ATOM     19  CA  ILE A   6      17.791  39.281  29.375  1.00 23.27           C
```

Homework

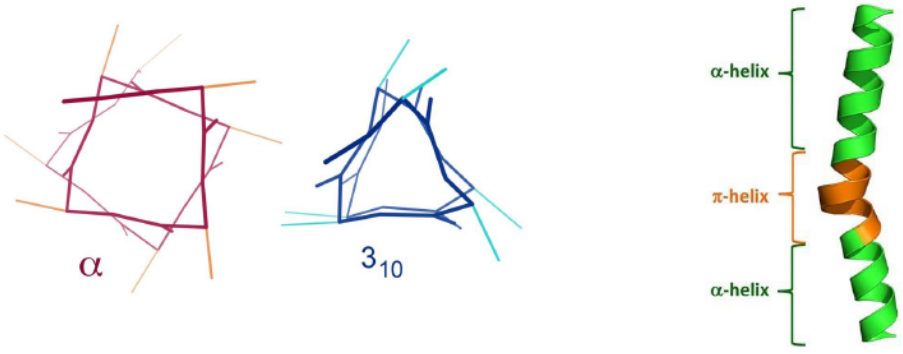
- Write down an algorithm which will check if there is a hydrogen bond between a donor and an acceptor or not.

Formation of helix



Book: Branden and Trooze

Helix in protein structure



Facts:

α-helix (3.6_{13}):

Hydrogen bond between i and i+4 residues

100° turn in the helix (3.6 residues per turn)

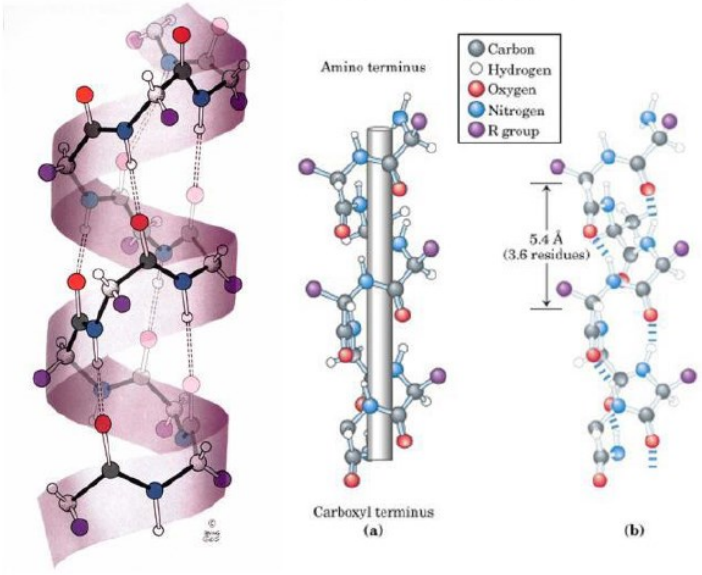
0.15 nm translation along helical axis

3_{10}-helix:

Hydrogen bond between i and i+3 residues

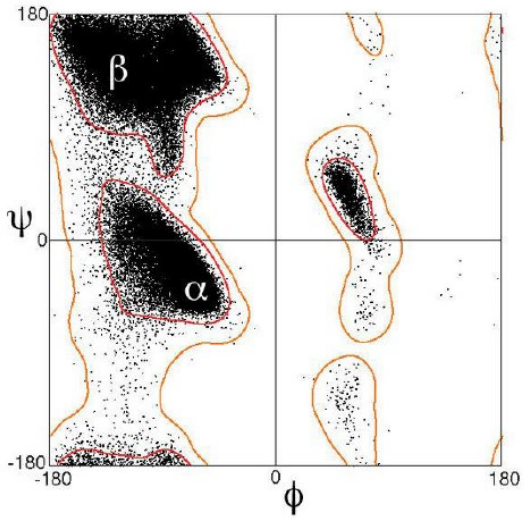
π-helix:

The α-helix

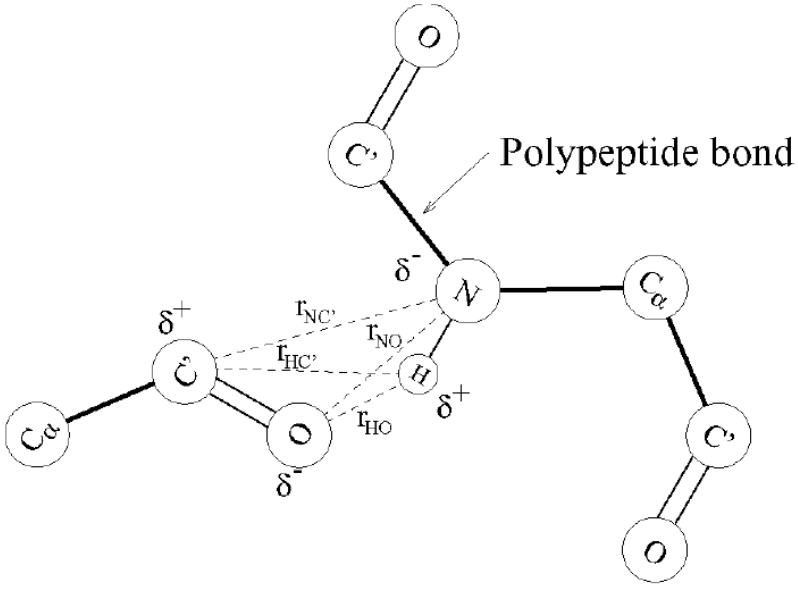


Source: http://xray.bmc.uu.se/Courses/bioinformatik2003/Intro/a-helix.jpg

Ramachandran Plot



$q > 120°$ and $r_{HO} < 2.5$ Å



DSSP Algorithm

•The so-called "Dictionary of Secondary Structure of Proteins" (DSSP) by Kabsch and Sander makes its sheet and helix assignments solely on the basis of backbone-backbone hydrogen bonds.

•The DSSP method defines a hydrogen bond when the bond energy is below -0.5 kcal/mol from a Coulomb approximation of the hydrogen bond energy.

•The structural assignments are defined such that visually appealing and unbroken structures result.

• In case of overlaps, alpha-helix is given first priority.

DSSP Algorithm

•The helix definition does not include the terminal residue having the initial and final hydrogen bonds in the helix.

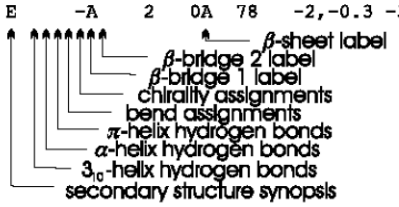
•A minimal size helix is set to have two consecutive hydrogen bonds in the helix, leaving out single helix hydrogen bonds, which are assigned as turns (state 'T').

•beta-sheet residues (state 'E') are defined as either having two hydrogen bonds in the sheet, or being surrounded by two hydrogen bonds in the sheet.

•The minimal sheet consists of two residues at each partner segment.

DSSP Algorithm

```
PDB:1crn
  #  RESIDUE AA STRUCTURE BP1 BP2  ACC    N-H-->O  O-->H-N  N-H-->O  O-->H-N
  ....
  15   15   V  H >< S+     0   0   99   -4,-1.7   3,-1.3   2,-0.2  -2,-0.2
  16   16   c  H 3<>S+     0   0   18   -4,-2.5   5,-0.8   1,-0.3  -2,-0.2
  17   17   R  H ><5S+     0   0   94   -4,-2.0   3,-1.6   1,-0.2  -1,-0.3
  18   18   L  T <<5S+     0   0  144   -3,-1.3   -1,-0.2 -4,-0.6  -2,-0.2
  19   19   P  T 3 5S-     0   0  107    0, 0.0   -1,-0.3   0, 0.0  -2,-0.1
  20   20   G  T < 5 +     0   0   53   -3,-1.6   -3,-0.2   1,-0.2  -2,-0.1
  21   21   T      < -     0   0   37   -5,-0.8   -1,-0.2   1,-0.1   5,-0.1
  22   22   P    >>  -     0   0   81    0, 0.0   4,-2.2    0, 0.0   3,-0.7
  23   23   E  H 3> S+     0   0   70    1,-0.2   4,-2.5    2,-0.2   5,-0.1
  24   24   A  H 3> S+     0   0   63    1,-0.2   4,-1.7    2,-0.2  -1,-0.2
  25   25   I  H <> S+     0   0   99   -3,-0.7   4,-1.8    2,-0.2  -1,-0.2
  26   26   c  H  X S+     0   0    0   -4,-2.2   4,-1.9    2,-0.2   6,-0.4
  27   27   A  H  X S+     0   0   12   -4,-2.5   4,-2.7   -5,-0.2   5,-0.5
  28   28   T  H  < S+     0   0  120   -4,-1.7   -1,-0.2   1,-0.2  -2,-0.2
  29   29   Y  H  < S+     0   0  176   -4,-1.8   -1,-0.2  -5,-0.2  -2,-0.2
  30   30   T  H  < S-     0   0   24   -4,-1.9   -2,-0.2  -3,-0.2  -3,-0.2
  31   31   G  S  < S+     0   0   35   -4,-2.7   -3,-0.2   1,-0.4  -4,-0.1
  32   32   b        -     0   0    5   -5,-0.5   -1,-0.4  -6,-0.4   2,-0.3
  33   33   I  E     -A     3   0A  51  -30,-2.8  -30,-2.4  -3,-0.1   2,-0.5
  34   34   I  E     -A     2   0A  78   -2,-0.3  -32,-0.2 -32,-0.2   3, 0.0
  ....
```



STRIDE Algorithm

•The secondary STRuctural IDEntification method by Frishman and Argos uses an empirically derived hydrogen bond energy **and phi-psi torsion angle criteria** to assign secondary structure.

•Torsion angles are given alpha-helix and beta-sheet propensities according to how close they are to their regions in Ramachandran plots.

•The parameters are optimized to mirror visual assignments made by crystallographers for a set of proteins.

•By construction, the STRIDE assignments agreed better with the expert assignments than DSSP, at least for the data set used to optimize the free parameters.

STRIDE Algorithm

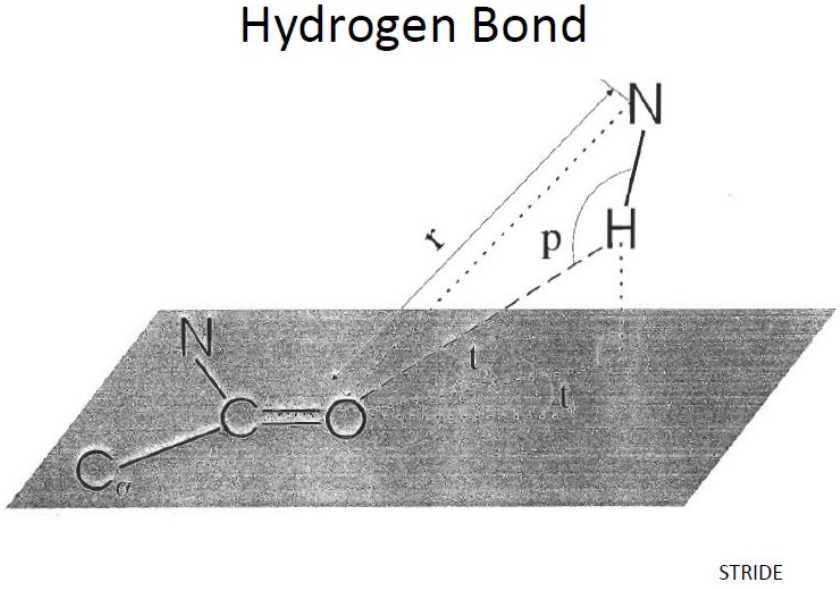
•Like **DSSP, STRIDE** assigns the shortest alpha-helix ('H') if it contains at least two consecutive *i* - *i+4* hydrogen bonds.

•In contrast to DSSP, helices are elongated to comprise one or both edge residues if they have acceptable phi-psi angles, similarly a short helix can be vetoed.

•hydrogen bond patterns may be ignored if the phi-psi angles are unfavorable.

•The sheet category does not distinguish between parallel and anti-parallel sheets. The minimal sheet ('E') is composed of two residues.

•The dihedral angles are incorporated into the final sheet assignment criterion as was done for the alpha-helix.



STRIDE

STRIDE Algorithm

$$E_{hb} = E_r \times E_t \times E_p$$

$$E_r = \frac{C}{r^8} + \frac{D}{r^6}$$

where $C = -3E_m r_m{}^8$ kcal Å^8/mol, $D = -4E_m r_m{}^6$ kcal Å^8/mol

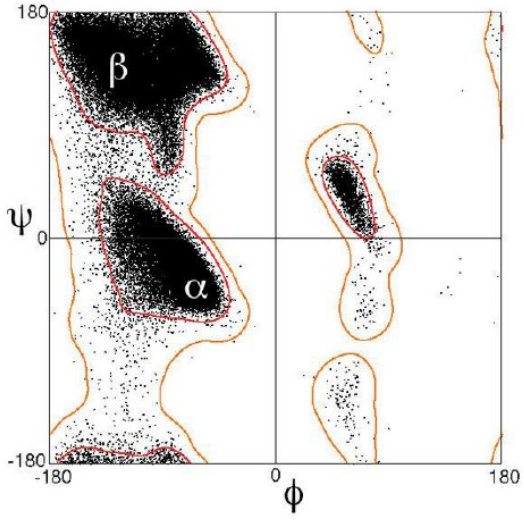
For main chain-main chain hydrogen bonds E_m= -2.8 Kcal/mol and r_m=3.0 Angstrom

$$E_p = \cos^2 p$$

$$E_t = \begin{cases} (0.9 + 0.1 \sin 2t_i)\cos t_o, & 0 < t_i < 90^\circ \quad K_1 = 0.9/\cos^6 110^\circ \\ K_1(K_2 - \cos^2 t_i)^3 \cos t_o, & 90^\circ < t_i < 110^\circ \quad K_2 = \cos^2 110^\circ \\ 0, & t_i > 110^\circ \end{cases}$$

$$E_r = E_m \qquad \text{for } r < r_m$$

Ramachandran Plot

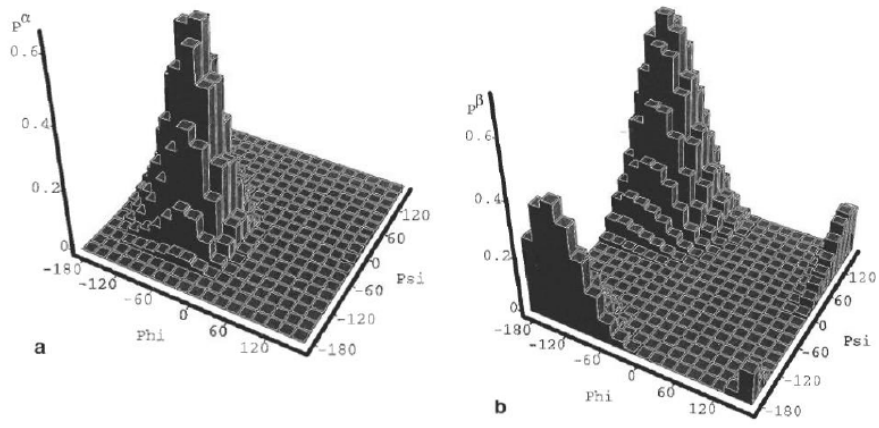


Probabilities for residues in α-helical (a) and β-sheet (b) as calculated in STRIDE

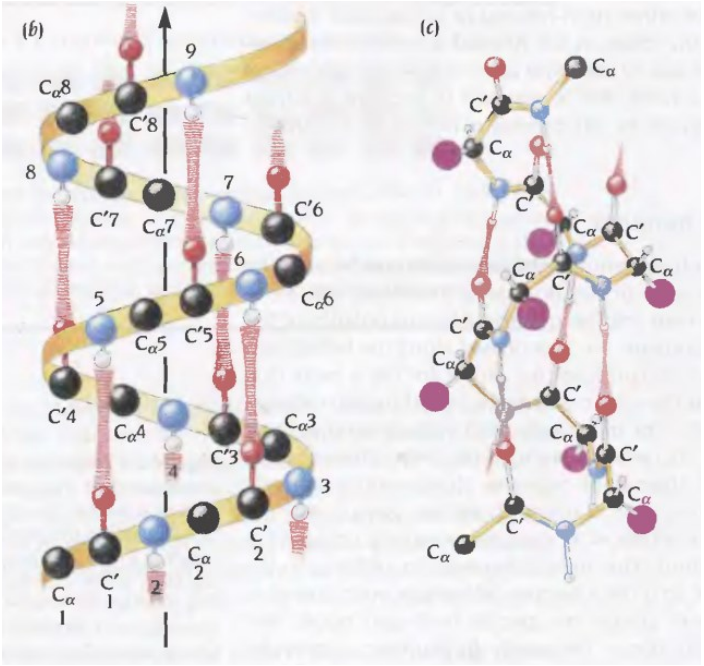
$$P_i^{\alpha} = \begin{cases} \dfrac{N_i^{\alpha}}{N_i^{\text{total}}} & \text{if } -180° < \varphi < 10° \text{ and } -120° < \phi < 45° \\ 0 & \text{otherwise} \end{cases}$$

$$P_i^{\beta} = \begin{cases} \dfrac{N_i^{\beta}}{N_i^{\text{total}}} & \text{if } -180° < \varphi < 0°, -180° < \phi < -120° \text{ or } 45° < \phi < 180° \\ 0 & \text{otherwise} \end{cases}$$

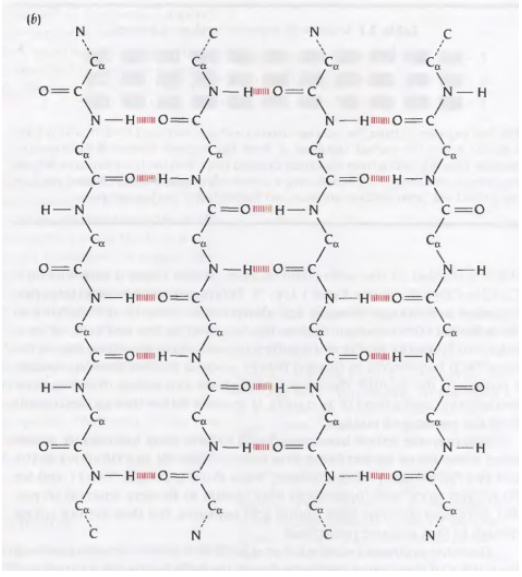
Probabilities for residues in α-helical (a) and β-sheet (b) as calculated in STRIDE

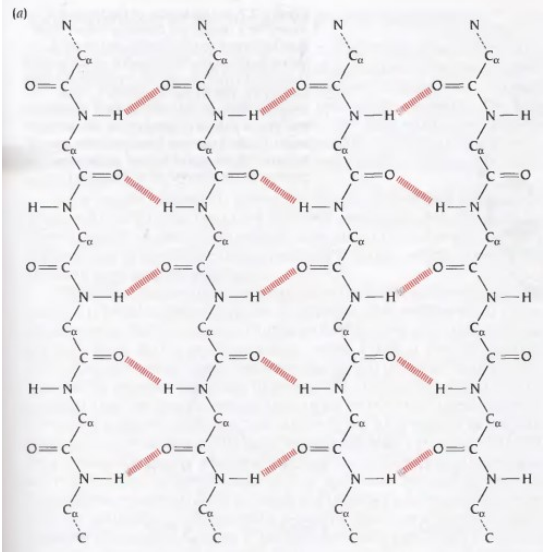


Formation of helix

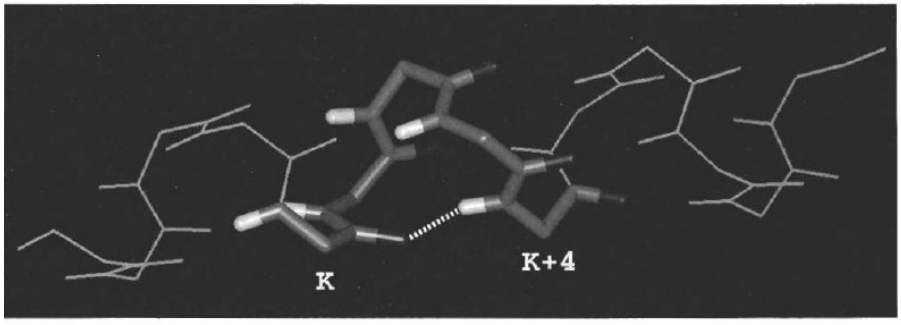


Formation of sheet





Elementary α-helical pattern



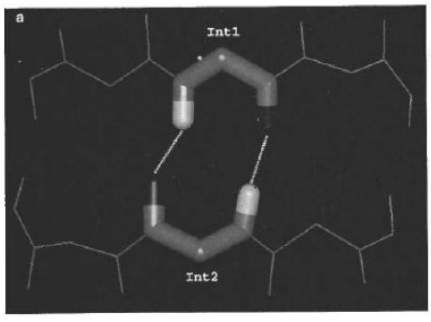
To include k+1, k+2, k+3, k+4

$$E_{\mathrm{hb}}^{k,k+4}\left(1+W_1^{\alpha}+W_2^{\alpha}\cdot\frac{P_k^{\alpha}+P_{k+4}^{\alpha}}{2}\right)<T_1^{\alpha}$$

To include k and k+5

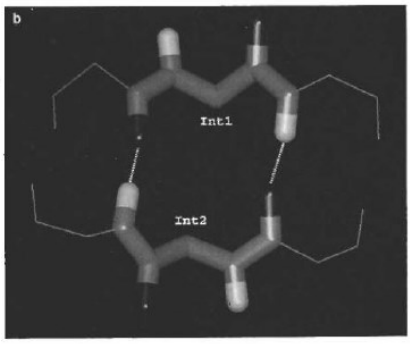
$$P_k{}^{\alpha} < T_2{}^{\alpha} \qquad P_{k+5}{}^{\alpha} < T_3{}^{\alpha}$$

Elementary β-sheet pattern

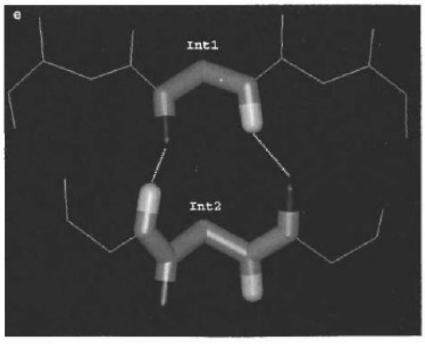


Antiparallel bridge of type I.

Elementary β-sheet pattern

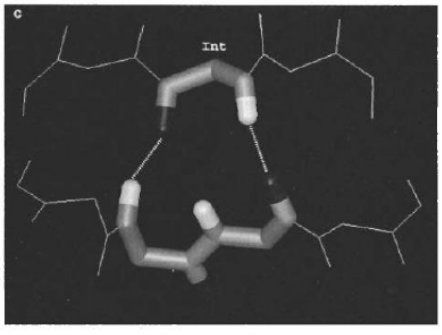


Antiparallel bridge of type II.

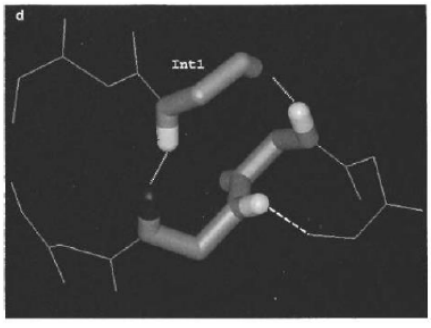


Antiparallel bridge of type IV.

Elementary β-sheet pattern



Antiparallel bridge of type III.



Antiparallel bridge of type III with one additional HBond.

Recognition of β-sheets

$$\begin{cases} E_{\text{hb1}}(1 + W_1^\beta + W_2^\beta \cdot \text{CONF}_{\text{Antiparallel}}) < T^\beta_{\text{Antiparallel}} \\ E_{\text{hb2}}(1 + W_1^\beta + W_2^\beta \cdot \text{CONF}_{\text{Antiparallel}}) < T^\beta_{\text{Antiparallel}} \end{cases}$$

$$\begin{cases} E_{\text{hb1}}(1 + W_1^\beta + W_2^\beta \cdot \text{CONF}_{\text{Parallel}}) < T^\beta_{\text{Parallel}} \\ E_{\text{hb2}}(1 + W_1^\beta + W_2^\beta \cdot \text{CONF}_{\text{Parallel}}) < T^\beta_{\text{Parallel}} \end{cases}$$

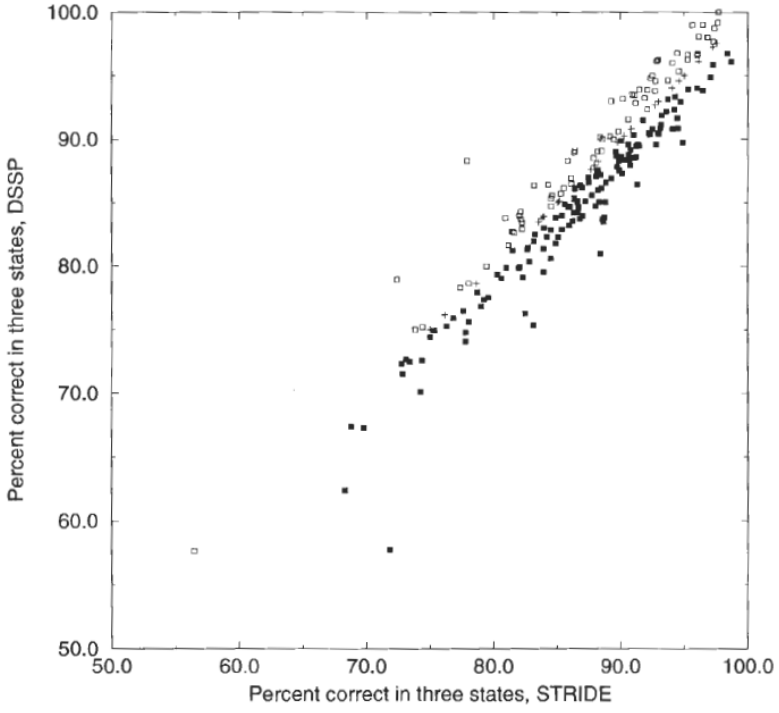
CONF= $\dfrac{(\text{P}^\beta_{\text{Int1}} + \text{P}^\beta_{\text{Int2}})}{2}$

STRIDE Algorithm

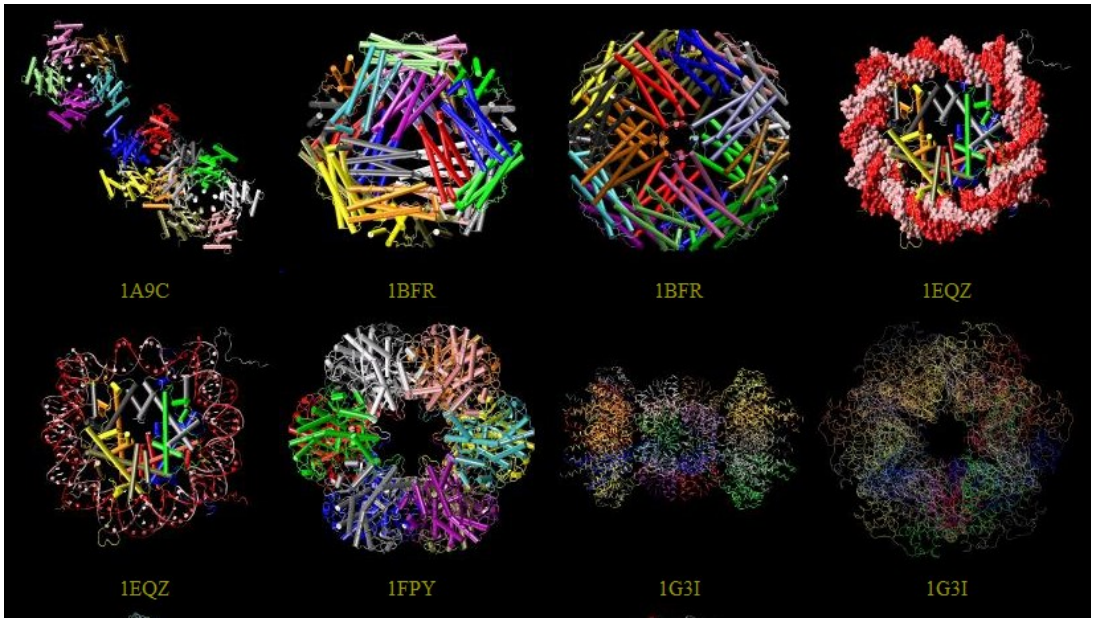
```
PDB:1crn
REM  |---Residue---|    |--Structure--|   |-Phi-|   |-Psi-|  |-Area-|      1CRN
....
ASG  VAL -   15   15    H    AlphaHelix    -69.24    -41.22      93.8      1CRN
ASG  CYS -   16   16    H    AlphaHelix    -56.67    -36.00      18.4      1CRN
ASG  ARG -   17   17    H    AlphaHelix    -77.07    -16.13      94.1      1CRN
ASG  LEU -   18   18    H    AlphaHelix    -53.21    -46.17     143.0      1CRN
ASG  PRO -   19   19    C          Coil    -77.19     -7.60     108.9      1CRN
ASG  GLY -   20   20    C          Coil    106.26      7.31      52.1      1CRN
ASG  THR -   21   21    C          Coil    -52.67    136.34      38.4      1CRN
ASG  PRO -   22   22    C          Coil    -56.98    146.62      81.9      1CRN
ASG  GLU -   23   23    H    AlphaHelix    -56.41    -36.19      68.9      1CRN
ASG  ALA -   24   24    H    AlphaHelix    -63.43    -34.86      61.3      1CRN
ASG  ILE -   25   25    H    AlphaHelix    -74.77    -37.89      98.2      1CRN
ASG  CYS -   26   26    H    AlphaHelix    -64.95    -31.69       0.0      1CRN
ASG  ALA -   27   27    H    AlphaHelix    -62.04    -54.03      11.6      1CRN
ASG  THR -   28   28    H    AlphaHelix    -68.78    -25.49     121.1      1CRN
ASG  TYR -   29   29    H    AlphaHelix    -67.59    -36.30     174.0      1CRN
ASG  THR -   30   30    H    AlphaHelix   -108.96    -18.47      23.4      1CRN
ASG  GLY -   31   31    C          Coil     91.82     -3.07      36.1      1CRN
ASG  CYS -   32   32    C          Coil    -69.52    164.38       4.6      1CRN
ASG  ILE -   33   33    E        Strand   -129.76    157.03      51.0      1CRN
ASG  ILE -   34   34    E        Strand   -111.56    129.59      78.0      1CRN
....
```

Comparison between DSSP and STRIDE

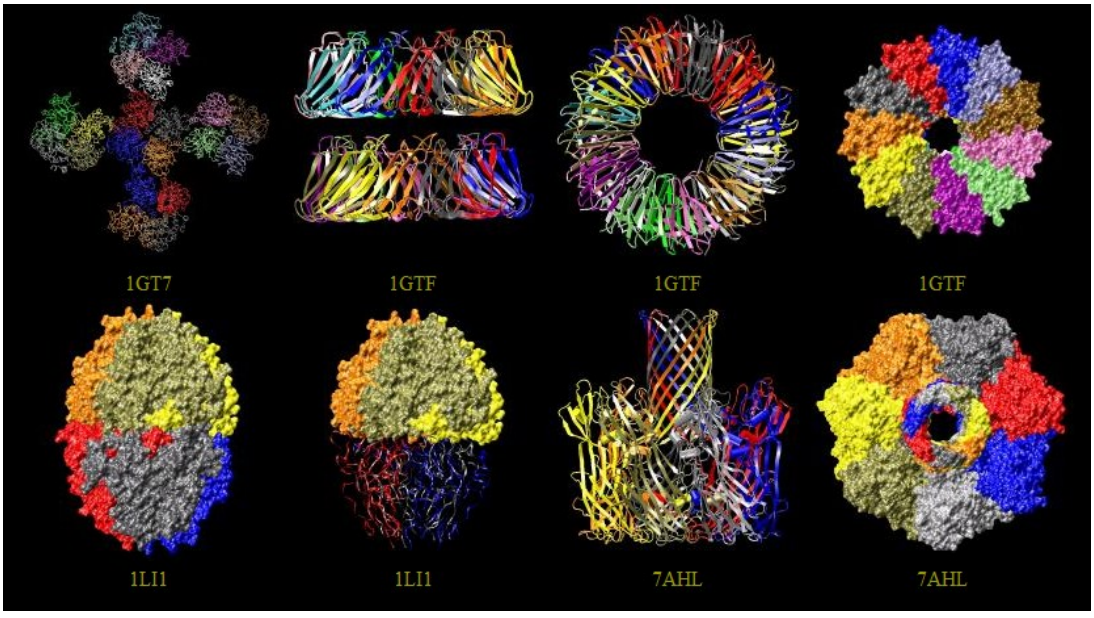
- Size of dataset: 226 protein chains



Visualization of proteins



Visualization of proteins



Bacteriophage

https://medschool.ucsd.edu/som/medicine/divisions/idgph/research/center-innovative-phage-applications-and-therapeutics/patient-care/Pages/default.aspx