

Secondary Structure Prediction Algorithms

Chou - Fasman (1974)

- Assigns Propensities to each classes (α, β , turn) $\langle P_\alpha, P_\beta, P_{turn} \rangle$ for every amino acid residue & its prob. of occurrence in a turn in position $i, i+1, i+2$ or $i+3$ (expressed by $f_i, f_{i+1}, f_{i+2}, f_{i+3}$).

• Table look up:

	P_α	P_β	P_{turn}	f_i	f_{i+1}	f_{i+2}	f_{i+3}
Ala	1.29	0.9	0.78	0.06	0.076	0.035	0.058
⋮							
Gly	0.56	0.92	1.64	0.102	0.085	0.19	0.152
⋮							

Propensity value: $\frac{\text{fraction of no. of occur in all residues}}{\text{\# of Residue in one } \alpha \dots}$
 $\Rightarrow \frac{\text{\# of total } \alpha \text{ in } \alpha \dots}{\text{\# of total } \alpha \text{ in } \alpha \dots}$

Prediction of Helix:

1. Helix Former: (Ala, Cys, Leu, Met, Glu, Gln, His, Lys)

2. Helix Breaker: Total set - Helix Former - {Phe, Asp} \uparrow
 $1.11 > P_\alpha > 1.0$

2. β sheet Former: {Val, Ile, Phe, Tyr, Trp, Thr, Ala}

β -Sheet Breaker: Total set - β -family - {Leu, His}

3. Turn form: {Gly, Ser, Asp, Asn, Pro}

Turn Breaker: Total Set - {Turn form} - {Glu, Tyr, Thr}

Prediction of Helix:

1. choose a cluster of 4 helix formers as residues out of six along the protein seq. initiate a helix. (4 out of 6 residues)
2. extend in both directions until a set of tetra-peptide breakers are reached.
↳ (consecutive breaker in next 4)
3. Any segment of length > 6 , avg. $P_\alpha > 1.06$
 $P_\alpha > P_\beta \Rightarrow \alpha$ -helix.

Prediction of β -strand:

1. 4 out of 5 in a seq. are β -formers initiates the β -strand.
2. extended in both directions until a set of tetra-peptide breakers are reached. ($P_\beta < 1.0$)
3. Any segment with avg. $P_\beta > 1.05$ & $P_\beta > P_\alpha$ is predicted a beta strand.

Prediction of turns

- Prob. of a bend: $P_b = f_i \cdot f_{i+1} \cdot f_{i+2} \cdot f_{i+3}$.
- Tetra peptides with $P_t > 0.75 \times 10^{-4}$
& avg. $P_{turn} > 1.0$, $P_\alpha < P_{turn} < P_\beta > P_\beta$
are predicted as turns.

Overlapping regions:

- Avg. P_α , P_β & P_t are compared for the overlapping region.
- $P_\alpha > P_\beta \Rightarrow \alpha$ -helix, } & P_t is less
else β -sheet } P_t is less

Barrows
 $\alpha \leftarrow P_t \max(P_\alpha, P_\beta, P_t) \Rightarrow$ type structure type

Accuracy:

70-80% reported.
Actually observed: 55%.
Later with more no. of proteins improved to