Implementation of Chou and Fasman method for secondary structure prediction

Sequence of interest: WHGCITVYWMTV

Say following are the alpha and beta propensities for each amino acid of the peptide:

	W	H	G	C	I	T	V	Y	W	M	T	V
P(H)	1.14	1.24	0.53	0.77	1.0	0.82	1.14	0.61	1.14	1.20	0.82	1.14
P(S)	1.19	0.71	0.81	1.30	1.60	1.20	1.65	1.29	1.19	1.17	1.20	1.65

Predicting helices

Look for all nucleation sites. Scan the entire sequence with a window size of 6.
Possibilities: WHGCIT, HGCITV, GCITVY, CITVYW, ITVYWM, TVYWMT, VYWMTV
The stretch of 6 residues should have atleast 4 residues with P(H) > 1 to be a nucleation site: ITVYWM and VYWMTV

So, let's see VYWMTV
We don't have residues to its right so we will extend only the left side.
Add one more residue, but check the score after including the three nearby residues.
TVYW: sum P(H) = 0.82+1.14+0.61+1.14=3.71 (the score is less than 4 or we can say that the average P(H)<1). So, this stops here. The fragment VYWMTV could be a helix.

Now, let's see ITVYWM.
Firstly, extend to the right:
YWMT: sum P(H) = 0.61+1.14+1.20+0.82=3.77 (the score is less than 4 or we can say that the average P(H)<1). So, this stops here.
Now, extend to the left:
CITV: sum P(H) = 0.77+1.0+0.82+1.14=3.73. No extension. The fragment ITVYWM could be a helix.

Predicting strand

Scan the entire sequence with a window size of 5.
Possibilities: WHGCI, HGCIT, GCITV, CITVY, ITVYW, TVYWM, VYWMT, YWMTV

The stretch of 5 residues should have atleast 3 residues with P(S) > 1 to be a nucleation site. All these candidates qualify to be a nucleation site.

So, let's see WHGCI.
GCIT: P(S) = 0.81+1.3+1.6+1.2=4.91 (score>4 or we can say that the average P(S)>1)
CITV: P(S) = 1.3+1.6+1.2+1.65=5.75
ITVY: P(S)= 1.6+1.2+1.65+1.29=5.74
TVYW: P(S)= 1.2+1.65+1.29+1.19=5.33
VYWM: P(S)= 1.65+1.29+1.19+1.17=5.3
YWMT: P(S)= 1.29+1.19+1.17+1.20=4.85
WMTV: P(S)= 1.19+1.17+1.20+1.65=5.21
So, this entire stretch WHGCITVYWMTV could be a strand.

Similarly, check for all other nucleation sites as well.

Assignment of secondary structure

```
WHGCITVYWMTV
      HHHHHHHH
EEEEEEEEEEEE
```

Conflict resolving for `ITVYWMTV`:

P(H) = 1.0+0.82+1.14+0.61+1.14+1.20+0.82+1.14 =7.87

P(S) = 1.6+1.2+1.65+1.29+1.19+1.17+1.2+1.65=10.95

P(S)>P(H)

Therefore, the final answer is

```
WHGCITVYWMTV
EEEEEEEEEEEE
```