

Length of the sequence: 1495 amino acids.

Molecular weight is: 169,186.16.

Isoelectric point: 7.31.

4.4 The Secondary Structural Conformations of the Region 4406–5900

SOPMA is an online executable tool for the structural annotation of proteins [2]. The tool predicts the detailed secondary structural confirmation of protein with respect to each amino acid in it. Based on the SOPMA annotations the composition percentage of each kind of secondary structures like helix, coil, extended strand etc. can be calculated (Fig. 4.3).

SOPMA :			
Alpha helix	(Hh)	: 567 is	37.93%
3 ₁₀ helix	(Gg)	: 0 is	0.00%
Pi helix	(Ii)	: 0 is	0.00%
Beta bridge	(Bb)	: 0 is	0.00%
Extended strand	(Ee)	: 315 is	21.07%
Beta turn	(Tt)	: 99 is	6.62%
Bend region	(Ss)	: 0 is	0.00%
Random coil	(Cc)	: 514 is	34.38%
Ambiguous states (?)		: 0 is	0.00%
Other states		: 0 is	0.00%

Fig. 4.3 Partial output of SOPMA tool

As evident, residues forming helix are more than those of sheets, confirming the hydrophilic nature of the peptide.

4.5 3D Structure prediction using Phyre

The phyre [3] result indicates that the selected sequence (region 4406–5900 of ORF1ab polyprotein of NCoV) can be structurally represented by the pdb id 7btfA. The subject sequence being represented by 7btfA shares 100% identity to the submitted query sequence with the confidence score of 100%. The pdb id codes for the structure of SARS Corona Virus 2 RNA-dependent RNA polymerase. This structure can be used for further analysis (Fig. 4.4).