7.2 Materials and Methods

7.2.1 Collection of Sequences

Coronavirus family spike protein sequences were retrieved from National Center for Biotechnology Information (NCBI) protein sequence data base (https://www.ncbi.nlm.nih.gov/protein) with their reference number such as COVID-19 (YP_009724390.1), Bat Coronavirus RatG13 (QHR63300.2), Pangolian coronavirus (QIQ54048.1) and SARS CoV (ACU31032.1).

7.2.2 Phylogenetic Analysis

Mega 6.0 software has been used to construct the phylogenetic tree to establish the relationship between these four coronavirus family. Alignment of the full-length coronavirus spike proteins was performed by MUSCLE with default parameters. The neighbor joining (NJ) tree was computed from the pairwise phylogenetic distance matrix creation [11].

7.2.3 Protein Structure Homology Modeling by ITASSER

I-TASSER (Iterative Threading ASSEmbly Refinement) is a hierarchical approach to protein structure and function prediction. It first identifies structural templates from the PDB by multiple threading approach LOMETS, with full-length atomic models constructed by iterative template-based fragment assembly simulations. To create structural models of the full length COVID-19 spike protein the SARS-CoV spike glycoprotein (5XLR) has been used as a template for modeling [12, 13].

7.2.4 Protein-Protein Docking

The protein–protein complexes from the predicted structure of NCoV and ACE2 human receptor were (PDB id 1R42) downloaded from protein data bank. Further, we use FRODOCK software web based user-friendly protein–protein docking server for interaction between the viral spike protein and ACE2 receptor using molecular docking [14].