TITLE

An immunoinformatics approach to study the epitopes contributed by Nsp13 of SARS-CoV-2

AUTHORS

Sushant Kumar[1] and Gajendra Kumar Azad[1#]

AFFILIATIONS

[1]Department of Zoology, Patna University, Patna, Bihar, India, 800005

[#]To whom correspondence should be addressed: Gajendra Kumar Azad, PhD, Assistant Professor, Department of Zoology, Patna University, Patna-800005, Bihar, India.

Email: gkazad@patnauniversity.ac.in

ABSTRACT

The on-going coronavirus disease-19 (COVID-19) pandemic caused by SARS-CoV-2 has infected hundreds of millions of people and killed more than two million people worldwide. Currently, there are no effective drugs available for treating SARS-CoV-2 infections; however, vaccines are now being administered worldwide to control this virus. In this study, we have studied SARS-CoV-2 helicase, Nsp13, which is critical for viral replication. We compared the Nsp13 sequences reported from India with the first reported sequence from Wuhan province, China to identify and characterize the mutations occurring in this protein. To correlate the functional impact of these mutations, we characterised the most prominent B cell and T cell epitopes contributed by Nsp13. Our data revealed twenty-one epitopes, which exhibited high antigenicity, stability and interactions with MHC class-I and class-II molecules. Subsequently, the physiochemical properties of these epitopes were also analysed. Furthermore, several of these Nsp13 epitopes harbour mutations, which were further characterised by secondary structure and per-residue disorderness, stability and dynamicity predictions. Altogether, we report the candidate epitopes of Nsp13 that may help the scientific community to understand the evolution of SARS-CoV-2 variants and their probable implications.

KEYWORDS

Sars-CoV-2; Nsp13; B cell epitopes; T cell epitopes; Mutations; Indian geographical area; immunoinformatics

INTRODUCTION

The coronavirus disease 2019 (COVID-19) disease was first reported from Wuhan province, China in the late 2019 [1–3]. The causative agent of this disease was identified as Severe Acute Respiratory Syndrome-Corona Virus-2 (SARS-CoV-2), which shares considerable similarity with previously known coronavirus, SARS-CoV [4]. Coronaviruses are a group of RNA viruses, which can infect diverse animals, including humans [5]. Earlier reports have shown that at least six different coronaviruses are known to infect humans, including CoV-229E, CoV-OC43, CoV-NL63, CoV-HKU1, SARS-CoV, and MERS-CoV [6]. Four of them do not cause any serious health implication on humans including CoV-229E, CoV-OC43, CoV-NL63, and CoV-HKU1 but serious respiratory issues have been linked to SARS-CoV and MERS-CoV infection [7]. The SARS-CoV-2 is the seventh coronavirus to infect humans and also causes mild to severe respiratory illness in the infected individuals and reported to cause severe symptoms in immune-compromised patients. The SARS-CoV-2 rapidly spread worldwide within few months and become one of the worst pandemic ever reported in human history [8]. Still this virus is spreading and already triggered second and third wave of infections in many countries. As of 18th March 2021, more than 120 million cases of COVID-19 were reported worldwide with more than 2.6 million deaths.

SARS-CoV-2 genome is comprised of a positive sense single stranded RNA of approximately 29kb [9]. Its genome encodes four structural, sixteen non-structural and nine accessory proteins [10]. The non-structural proteins are involved in the maintenance of functional integrity of the virus and also required for infection and virus particle formation. Numerous RNA viruses have been found to encode their own RNA helicases, which are usually indispensable components of the RNA replication complexes [11,12] and recognized as an ideal targets for developing antivirals [13]. Nsp13, an RNA helicase, plays an important role in the folding of its RNA elements or unwinding of double stranded RNAs. Previous studies have shown that SARS-CoV Nsp13 has an NTPase and helicase activity belonging to helicase superfamily-1 [14]. The SARS-CoV-2 Nsp13 helicase shares a 99.8% sequence identity to SARS-CoV (SARS) Nsp13 helicase [15]. Nsp13 helicase is a critical component for viral replication and shares the highest sequence conservation across the CoV family, highlighting their importance for viral viability. RNA dependent RNA polymerase (RdRp) and Nsp13 are required for viral replication, therefore, both of these vital enzyme represents a promising target for anti-CoV drug development [16].

In order to understand the host response towards SARS-CoV-2, we performed in-silico study to characterize Nsp13. We used several bioinformatic tools to predict potential epitopes of Nsp13 and characterised them by studying several immunological parameters. Further, we systematically characterized the mutations in Nsp13 reported from India and their impacts on epitopes are discussed.

MATERIALS AND METHODS

Nsp13 sequence retrieval and multiple sequence alignment (MSA)

The Orf1ab of SARS-CoV-2 contains Nsp1-16 as a single polypeptide which is proteolytically cleaved into individual proteins. The Orf1ab sequences used in this study were downloaded from NCBI-virus database. Till 3rd March 2021, 651 sequences of Orf1ab have been reported from India. We have downloaded all these sequences and their protein accession number are listed in the supplementary table 1. The Nsp13 sequences were retrieved from their respective Orf1ab and used for analysis. The Orf1ab sequence reported from Wuhan, China was used as a reference sequence (Protein accession number: YP_009724389) for MSA studies using Clustal Omega tool [17] as described earlier [18].

B cell epitope prediction

Linear B cell epitopes were predicted using IEDB (Immune Epitope Database and Analysis Resource) [19], an online server tool based on Bepipred linear epitope prediction method (at the threshold value of 0.350). IEDB prediction tool was also used to predict the Nsp13 immunological parameters including the antigenicity, accessibility, flexibility, hydropathicity and beta turn. These standard parameters were estimated by Chou and Fasman beta-turn prediction algorithm, Emini surface accessibility server tool, Karplus and Schulz flexibility prediction tool, Kolaskar and Tongaonkar antigenicity and Parker hydrophilicity prediction algorithms, respectively [20,21]. The DiscoTope 2.0 was used for prediction of Discontinuous B cell epitope using threshold value set at -5.5 and its three dimensional structure were represented by AutoDock software. Antigenicity and allergenicity of B cell epitopes were predicted by Vaxijen 2.0 [22] and AllergenFP v.1.0 [23] servers, respectively.

T cell epitope prediction

On the surface of antigenic presenting cell, T cell epitopes are presented where they are attached to Major Histocompatibility Complex (MHC) molecule. MHC class I and class II molecules were predicted as follows-

MHC class I molecule: IEDB webserver based on NetMHCpanEL 4.1 was used for prediction of MHC class I molecules [24]. For this prediction, we selected HLA reference alleles (a total of 54 alleles) having epitope length of 9 or 10 mers. We finally sorted 9 mers conserved epitopes that show maximum binding interaction at IC50 < 200nm. Antigenicity of all selected epitopes was predicted from Vaxijen 2.0 webserver, whereas allergenicity was predicted by AllergenFP v.1.0.

MHC class II molecules: IEDB recommended 2.22 prediction method was used to predict MHC class II epitopes. For this prediction, we selected the seven standard reference alleles having maximum length of 11 mers. Subsequently, we sorted most conserved 9 mers

epitopes that exhibited maximum binding interaction with other alleles. Antigenicity of all selected epitopes was predicted from Vaxijen 2.0 webserver, whereas allergenicity was predicted by AllergenFP v.1.0.

Physiological profiling of T cell epitopes-

The characterisation of selected MHC class I and MHC class II epitopes were performed by several webservers. The physiological parameters including toxicity, hydrophobicity, hydropathicity, charge PI and molecular weight were calculated by Toxinpred tool [25]. Another webserver tool HLP [26] was used to predict half-life, surface accessibility, flexibility and polarity of the selected epitopes.

Structure modelling, secondary structure analysis and protein disorder prediction

The locations of identified mutations were highlighted in the three-dimensional structure of Nsp13 using UCSF Chimera program [27]. The Nsp13 RCSB-ID: 6ZSL was used for the structural representation. The prediction of secondary structure was performed by CFSSP webserver [28] as described earlier [29]. The input sequence was uploaded on this server, which provides the secondary structure in terms of alpha-helix, beta-sheet and turns. The prediction of per residue intrinsic disorder predisposition contributed by each residue of input Nsp13 polypeptide sequence was performed by PONDR-VSL2 webserver [30]. The value above 0.5 is considered ordered while the value below 0.5 is considered disordered.

Stability of protein structure

The stability of protein structure was predicted by DynaMut webserver [31]. The reported structure of Nsp13 (RCSB-ID: 6ZSL) was used to predict the impact if mutation on the stability of protein. First, the protein structure was uploaded onto the webserver, followed by providing the details of the mutation for analysis as described earlier [32]. This webserver predict the differences in free energy between the wild type and mutant protein.

RESULTS

Identification of variations in nsp13 protein among Indian isolates

We compared the Nsp13 sequences reported from India with the first reported sequence from Wuhan province, China using Clustal omega tool. Our analysis revealed twenty-seven mutations have taken place among Indian isolates, which is distributed all over the Nsp13 protein as shown in figure 1A. Next, we observed the location of these mutations in the three-dimensional structure of Nsp13 using UCSC Chimera tool as shown in figure 1B. We also looked at the polarity change and charge alteration due to the mutation in Nsp13, our data revealed that most of mutation does not led to any change (neutral to neutral) however,

at two positions it changed from acidic to neutral (E142V and E168A) while at three locations it altered from basic to neutral (H164Y, R392C and R442Q) as shown in table 1.

Prediction of B cell epitope contributed by Nsp13

IEDB webserver tool was used for predicting the continuous B-cell epitopes of Nsp13 (figure 2A). Our analysis revealed seven best antigenic epitopes that were more than 8 amino acid residue in length (figure 2B). Subsequently, these epitopes were further characterised by analysing various parameters, including vaxijen score, allergenicity, and toxicity (figure 2B). Our prediction data revealed that all of them are non-toxic however, 5 peptides (CNAPGCDVT, CVGSDNVT, VGKPRPPLN, TFEKGDYGDA and GDPAQLPAP) possess alleregenicity properties and two peptides are non-allergen (TQTVDSSQGSEY and STLQGPPGTGKS). Similarly, three of them exhibited antigenic property while four are non-antigenic (figure 2B).

Subsequently, we predicted the B cell epitopes of Nsp13 based on its three dimensional structure using DiscoTop 2.0 webserver tool [33]. Our analysis revealed the eleven discontinuous epitopes of Nsp13 having high score. The location of these epitopes are highlighted on the 3D structure of Nsp13 (figure 2C) and their additional details are shown in figure 2D. Altogether, our data revealed B –cell epitopes contributed by Nsp13.

Analysis of essential parameters of B cell epitopes

Subsequently, all essential parameters of B Cell epitopes including Beta turn, accessibility of surface, flexibility, antigenicity and hydrophilicity were also calculated for Nsp13 (figure 3). Chou and Fasman's beta-turn prediction algorithm (with threshold 0.984) resulted in a minimum score of 0.677 and a maximum score of 1.384, and our selected peptide showed a propensity score of 1.164 (figure 3A). Emini surface accessibility prediction tool (with threshold 1.000) resulted in a minimum score of 0.036 and a maximum score of 4.888, and our selected peptide scored 1.33 in surface accessibility (figure 3B). Karplus and Schulz's flexibility prediction tool (with threshold 0.989) resulted in a minimum score of 0.904 and a maximum score of 1.150, and our selected peptide showed 1.033 in flexibility score (figure 3C). Kolaskar and Tongaonkar antigenicity scale (with threshold 1.052) resulted in a minimum score of 0.893 and a maximum score of 1.284, and our selected peptide showed an antigenicity score of 1.0163 (figure 3D). The parker hydrophilicity prediction algorithms (with threshold 1.325) resulted in a minimum score of -3.714 and a maximum score of 7.000, and our selected peptide showed 1.33 in hydrophilicity score (figure 3E).

Prediction of T cell epitopes contributed by Nsp13 protein

The IEDB webserver tool was used for prediction of Cytotoxic T- Lymphocyte epitopes and its interaction with MHC class I molecules. Our analysis with Nsp13 revealed 8 potential T cell epitopes (table 2). NetMHCpanEL 4.1 MHC-class I binding prediction tools was used to predict the binding of these epitopes with MHC class I molecules with high affinity are listed in table 2. The peptides FAIGLALYY from start (11) to end (19) had highest immunogenicity and affinity to interact with 9 alleles (HLA-B*35:01, HLA-A*26:01, HLA-B*53:01, HLA-A*01:01, HLA-A*30:02, HLA-B*58:01, HLA-B*15:01, HLA-A*68:01, HLA-B*57:01) and also showed allergenicity (table 2). The allergenicity of these epitopes was predicted by Allergen FP tool, which categorises peptides into allergen/non-allergen based on the Tanimoto coefficients.

Similarly, the prediction of Helper T-Lymphocyte epitopes of Nsp13 and its interaction with MHC Class II molecules was predicted by IEDB webserver (based on IEDB recommended 2.22 method). We selected top 6 epitopes that exhibited maximum binding affinity with MHC class II molecules (table 3). Our analysis revealed that only one epitope (HKLVLSVNP) possesses both antigenicity and allergenicity properties and affinity to interact with 5 alleles as shown in table 3 (HLA-DRB1*15:01, HLA-DRB1*03:01, HLA-DRB1*07:01, HLA-DRB3*01:01, HLA-DRB5*01:01). The epitopes EHYVRITGL had highest immunogenicity and affinity to interact with 4 alleles (HLA-DRB5*01:01, HLA-DRB1*15:01, HLA-DRB4*01:01, HLA-DRB1*07:01) (table 3). The allergenicity of these epitopes were also predicted based on Tanimoto coefficients.

Physiochemical profiling of promising T-cell epitopes

To substantiate our data, we examined several vital physiochemical features of the promising T cell epitopes. The half-life of MHC class I and II epitopes were calculated and our data revealed that the maximum half-life was observed for HVISTSHKL (MHC class I) and ETFKLSYGI exhibited for MHC class II molecule (table 4). The toxicity prediction was performed by ToxinPred tool show that all analysed molecules were non-toxic (table 4). We further analysed hydrophobicity, hydropathicity, hydrophilicity, polarity, charge, flexibility, pI, molecular weight and surface accessibility of both MHC class I and II molecules (table 4).

The secondary structure prediction revealed several mutations that might lead to structural change of Nsp13.

To understand the effect of mutation on Nsp13 protein, we performed secondary structure prediction by CFFSP webserver. This prediction tool identifies the variation in secondary structure contributed by the mutant residue. Our data revealed that out of 27 mutations, only 12 show alterations in secondary structure (figure 4). The minor variations were observed at positions E142V, H164Y, G206V, T214I, S259T, R392C, P419S, and R499L (figure 4A-D,

G, I, J and K). S236I and A237T mutation lead to replacement of helical structure into beta-sheet (figure 4E and F). The coiled coil and turn structure were shifted to helix and beta-sheet in P300L and P504L mutants (figure 4H and L).

Protein disorder predictions due to Mutations in Nsp13 protein

Disordered regions (DRs) are defined as entire proteins or regions of proteins that lack a fixed tertiary structure. Alteration of amino acid sequences in polypeptide chain due to the mutation might causes change in protein disorder. Here, we used PONDR-VSL2 webserver to measure the protein disorder of those mutants that showed alteration in secondary structure. Our analysis revealed that ten mutations E142V, H164Y, S166A, E168A, G206V, G206C, T214I, S236I, R392C and P504L (figure- 5A, B, C, D, E, F, G, H, J and K) decreased the protein disorder, while two mutations Y253H and I545M (figure- 5I and L) led to increase in protein disorder. Altogether, our data revealed alteration in protein disorder due to the mutation in Nsp13 of SARS-CoV-2.

Mutations cause alteration in dynamic stability of Nsp13.

We used DynaMut program to predict the effect of mutations on the stability of the protein and calculated the differences in free energy (ΔΔG) between wild type and mutants. The positive ΔΔG corresponds to increase in stability while negative ΔΔG corresponds to decrease in stability. Our analysed data revealed the noticeable increase or decrease in free energy in various mutations as shown in table 5. The maximum positive ΔΔG was observed for H164Y (0.974 kcal/mol) and the maximum negative ΔΔG was obtained for G54S (-1.694kcal/mol) (table 5). Similarly, we also measured the change in vibrational entropy energy (ΔΔSVibENCoM) between the wild type and the mutants. The highest positive ΔΔSVibENCoM was Obtained for P53S (0.504kcal.mol-1.K-1) and negative ΔΔSVibENCoM was obtained for P419S (-0.661 kcal.mol-1.K-1) (table 5). Altogether, the data obtained from ΔΔG and ΔΔSVibENCoM suggests that the mutations identified in this study can influence Nsp13 protein stability and dynamicity.

DISCUSSIONS

In this study, we examined the physiochemical characteristics of the Nsp13, a non-structural protein of SARS-CoV-2. We studied the high rank B cell and T cell epitope candidates based on the immune-informatics tools to identify Nsp13 epitopes that could regulate host immune responses. The immunoinformatics approaches have been used to study the epitopes from several viruses and those information's were used to understand the immune response of viruses [34,35]. We used several physiological parameters including structural protrusion, antigenicity, flexibility, surface accessibility, hydrophilicity of Nsp13 were assessed to predict

potential B-cell epitopes. IEDB resource tool was used to predict B-cell linear epitopes of at least eight amino acid residues. Our analysis revealed that under these conditions, seven linear B-cell epitopes were predicted that are at least 8 amino acids in length and are of non-toxic. The discontinuous epitopes comprised of the residues that might be separated in linear sequence, however, in 3D structure they are in close proximity [36]. Subsequently, the candidate epitopes were further characterised by the tools that can predict various physicochemical properties. Our data revealed that among B-cell epitopes the 'TFEKGDYGDA' peptide exhibited a strong stable immunogenic property as shown by its highest vaxijen score (.9707) and non-toxic. Furthermore, eight MHC class-I and six MHC class-II binding T-cell epitopes assessed as highly antigenic and also predicted to interact with several HLA alleles. Detailed analysis revealed that the best vaxijen score of 1.18 was found for 'FAIGLALYY' peptide for MHC class-I moelcule, and for MHC class-II molecules, 'HKLVLSVNP' has a maximum vaxijen score of 0.50 epitopes.

It has been well established that the replication of coronaviruses are error-prone that lead to creation of highly diverse genotype variants. Our study revealed a considerable alteration in stability and dynamicity due to mutations at various positions of Nsp13 that might alter its function. In-silico analyses were performed to identify and characterise the mutations occurring in Nsp13. Our data demonstrate that seven of the identified mutations reside in those epitopes that include P53, G54, S100, Y205, G206, H245 and V247 positions (figure 2), which can help the SARS-CoV-2 variants to elicit distinct immune response compared to the wild-type SARS-CoV-2. Among these G206V also led to alteration in secondary structure (figure 4) and protein disorder parameters (figure 5). This proposed consequence due to variations in Nsp13 epitopes are substantiated by several findings with similar observations [37–39]. In short, these results could provide some insights in the understanding of Nsp13 epitopes, which is essential for evaluating the immunogenicity and virulence of SARS-CoV-2.

ACKNOWLDEGEMENTS

We would like to acknowledge the Department of Zoology, Patna University, Patna, Bihar (India) for providing infrastructural support for this study.

FIGURE AND TABLE LEGENDS

Figure 1: The diagrammatic representation of Nsp13 of SARS-CoV-2 demonstrating the details and location of mutations identified in this study. The SARS-CoV-2 Nsp13 protein is comprised of 601 amino acid residues. B) Structural representation of three dimensional structure of the Nsp13 (RCSB protein ID-6ZSL) showing the location of mutations (single

amino acid letter code) along with its respective position in polypeptide sequence. The structural data was obtained from UCSF chimera tool.

Figure 2: Prediction of B-cell epitopes of Nsp13. Linear continuous epitopes, A) The Y-axis of the graph corresponds to BepiPred score, while the X-axis depicts the Nsp13 residue positions in the sequence. The yellow area of the graph corresponds to the part of the protein having higher probability to be part of the epitope. B) The top seven peptides of Nsp13 having at least 8 amino acids are shown with vaxijen score. The red font shows the location of mutant residues in the epitope sequence. C) Prediction of discontinuous B-cell epitopes. The position of each predicted epitope on the 3D structural surface of Nsp13 was denoted using Autodock. D) The location and identity of each discontinuous epitopes of Nsp13.

Figure 3: Recognition of B cell epitopes. A) Antigenic determinants of Nsp13 were predicted using Kolaskar and Tongaonkar, B) Hydrophilicity of Nsp13 was predicted by Parker hydrophilicity, C) Surface accessibility analysis shown on Emini surface accessibility scale, D) β variants of structural polyproteins as predicted by Chou and Fasman β metamorphosis prediction, E) Flexibility analysis on Karplus and Schultz flexibility scale.

Figure 4: Effect of identified mutations of Nsp13 on its secondary structure. Panel (i) represents the wild type sequence (Wuhan isolate) and panel (ii) represents mutant sequence (Indian isolate). (A-L) each panel represents the individual mutant as depicted. In each panel the contribution of individual amino acid to the secondary structure of Nsp13 protein are shown (H- helix, C-coiled-Coil, T-turn, E-beta sheet). The data was generated from CFSSP tool. The red box in each panel highlights the location of the wild-type and mutant residue.

Figure 5: The comparisons of the per residue intrinsic disorder predisposition of Nsp13. A disorder threshold is indicated at score = 0.5, residues/regions with the disorder scores >0.5 are considered as disordered. (A-L) Each panel comparatively represents the disorder parameters of wild type and mutant Nsp13 polypeptide sequence.

Table 1: The table show the location and details of mutations identified in Nsp13 of SARS-CoV-2. The frequency of each mutation along with polarity and charge changes are also indicated.

Table 2: The table show the location and details of MHC-Class-I peptides of Nsp13. The MHC-I interaction with top most allele (affinity IC50 value of <200) are mentioned in the table along with Vaxijen score and allergenicity.

Table 3: The table show the location and details of MHC-Class-II peptides of Nsp13. The MHC-II interaction with top most allele (affinity IC50 value of <200) are mentioned in the table along with Vaxijen score and allergenicity.

Table 4: The table show the physiochemical properties of MHC-I and MHC-II peptides of Nsp13.

Table 5: The table show the ΔΔG and ΔΔSvib ENCoM of the mutants present in Nsp13. DynaMut program was used to calculate both parameters.

Supplementary table 1: List of protein accession number used in this study

REFERENCES

[1] C.C. Lai, T.P. Shih, W.C. Ko, H.J. Tang, P.R. Hsueh, Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges, Int. J. Antimicrob. Agents. (2020). https://doi.org/10.1016/j.ijantimicag.2020.105924.

[2] H.A. Rothan, S.N. Byrareddy, The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak, J. Autoimmun. (2020). https://doi.org/10.1016/j.jaut.2020.102433.

[3] F.A. Rabi, M.S. Al Zoubi, A.D. Al-Nasser, G.A. Kasasbeh, D.M. Salameh, Sars-cov-2 and coronavirus disease 2019: What we know so far, Pathogens. (2020). https://doi.org/10.3390/pathogens9030231.

[4] R. Lu, X. Zhao, J. Li, P. Niu, B. Yang, H. Wu, W. Wang, H. Song, B. Huang, N. Zhu, Y. Bi, X. Ma, F. Zhan, L. Wang, T. Hu, H. Zhou, Z. Hu, W. Zhou, L. Zhao, J. Chen, Y. Meng, J. Wang, Y. Lin, J. Yuan, Z. Xie, J. Ma, W.J. Liu, D. Wang, W. Xu, E.C. Holmes, G.F. Gao, G. Wu, W. Chen, W. Shi, W. Tan, Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding, Lancet. (2020). https://doi.org/10.1016/S0140-6736(20)30251-8.

[5] P. Zhou, X. Lou Yang, X.G. Wang, B. Hu, L. Zhang, W. Zhang, H.R. Si, Y. Zhu, B. Li, C.L. Huang, H.D. Chen, J. Chen, Y. Luo, H. Guo, R. Di Jiang, M.Q. Liu, Y. Chen, X.R. Shen, X. Wang, X.S. Zheng, K. Zhao, Q.J. Chen, F. Deng, L.L. Liu, B. Yan, F.X. Zhan, Y.Y. Wang, G.F. Xiao, Z.L. Shi, A pneumonia outbreak associated with a new

coronavirus of probable bat origin, Nature. (2020). https://doi.org/10.1038/s41586-020-2012-7.

[6] J. Cui, F. Li, Z.L. Shi, Origin and evolution of pathogenic coronaviruses, Nat. Rev. Microbiol. (2019). https://doi.org/10.1038/s41579-018-0118-9.

[7] S.R. Weiss, Forty years with coronaviruses, J. Exp. Med. (2020). https://doi.org/10.1084/jem.20200537.

[8] M. Bchetnia, C. Girard, C. Duchaine, C. Laprise, The outbreak of the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2): A review of the current global status, J. Infect. Public Health. (2020). https://doi.org/10.1016/j.jiph.2020.07.011.

[9] F. Wu, S. Zhao, B. Yu, Y.M. Chen, W. Wang, Z.G. Song, Y. Hu, Z.W. Tao, J.H. Tian, Y.Y. Pei, M.L. Yuan, Y.L. Zhang, F.H. Dai, Y. Liu, Q.M. Wang, J.J. Zheng, L. Xu, E.C. Holmes, Y.Z. Zhang, A new coronavirus associated with human respiratory disease in China, Nature. (2020). https://doi.org/10.1038/s41586-020-2008-3.

[10] R.A. Khailany, M. Safdar, M. Ozaslan, Genomic characterization of a novel SARS-CoV-2, Gene Reports. (2020). https://doi.org/10.1016/j.genrep.2020.100682.

[11] R. Jain, J. Coloma, A. García-Sastre, A.K. Aggarwal, Structure of the NS3 helicase from Zika virus, Nat. Struct. Mol. Biol. (2016). https://doi.org/10.1038/nsmb.3258.

[12] T. Shu, T. Gan, P. Bai, X. Wang, Q. Qian, H. Zhou, Q. Cheng, Y. Qiu, L. Yin, J. Zhong, X. Zhou, Ebola virus VP35 has novel NTPase and helicase-like activities, Nucleic Acids Res. (2019). https://doi.org/10.1093/nar/gkz340.

[13] T.-F. Li, M. Hosmillo, H. Schwanke, T. Shu, Z. Wang, L. Yin, S. Curry, I.G. Goodfellow, X. Zhou, Human Norovirus NS3 Has RNA Helicase and Chaperoning Activities, J. Virol. (2017). https://doi.org/10.1128/jvi.01606-17.

[14] K.A. Ivanov, V. Thiel, J.C. Dobbe, Y. van der Meer, E.J. Snijder, J. Ziebuhr, Multiple Enzymatic Activities Associated with Severe Acute Respiratory Syndrome Coronavirus Helicase, J. Virol. (2004). https://doi.org/10.1128/jvi.78.11.5619-5632.2004.

[15] M.A. White, W. Lin, X. Cheng, Discovery of COVID-19 Inhibitors Targeting the SARS-CoV-2 Nsp13 Helicase, J. Phys. Chem. Lett. (2020). https://doi.org/10.1021/acs.jpclett.0c02421.

[16] J. Chen, B. Malone, E. Llewellyn, M. Grasso, P.M.M. Shelton, P.D.B. Olinares, K. Maruthi, E.T. Eng, H. Vatandaslar, B.T. Chait, T.M. Kapoor, S.A. Darst, E.A. Campbell, Structural Basis for Helicase-Polymerase Coupling in the SARS-CoV-2 Replication-Transcription Complex, Cell. (2020). https://doi.org/10.1016/j.cell.2020.07.033.

[17] F. Madeira, Y.M. Park, J. Lee, N. Buso, T. Gur, N. Madhusoodanan, P. Basutkar,

A.R.N. Tivey, S.C. Potter, R.D. Finn, R. Lopez, The EMBL-EBI search and sequence analysis tools APIs in 2019, Nucleic Acids Res. (2019). https://doi.org/10.1093/nar/gkz268.

[18] G.K. Azad, Identification of novel mutations in the methyltransferase complex (Nsp10-Nsp16) of SARS-CoV-2, Biochem. Biophys. Reports. (2020). https://doi.org/10.1016/j.bbrep.2020.100833.

[19] R. Vita, S. Mahajan, J.A. Overton, S.K. Dhanda, S. Martini, J.R. Cantrell, D.K. Wheeler, A. Sette, B. Peters, The Immune Epitope Database (IEDB): 2018 update, Nucleic Acids Res. (2019). https://doi.org/10.1093/nar/gky1006.

[20] M.C. Jespersen, B. Peters, M. Nielsen, P. Marcatili, BepiPred-2.0: Improving sequence-based B-cell epitope prediction using conformational epitopes, Nucleic Acids Res. (2017). https://doi.org/10.1093/nar/gkx346.

[21] P. Koehl, M. Levitt, Structure-based conformational preferences of amino acids, Proc. Natl. Acad. Sci. U. S. A. (1999). https://doi.org/10.1073/pnas.96.22.12524.

[22] I.A. Doytchinova, D.R. Flower, VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines, BMC Bioinformatics. (2007). https://doi.org/10.1186/1471-2105-8-4.

[23] I. Dimitrov, L. Naneva, I. Doytchinova, I. Bangov, AllergenFP: Allergenicity prediction by descriptor fingerprints, Bioinformatics. (2014). https://doi.org/10.1093/bioinformatics/btt619.

[24] B. Reynisson, B. Alvarez, S. Paul, B. Peters, M. Nielsen, NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data, Nucleic Acids Res. (2020). https://doi.org/10.1093/nar/gkaa379.

[25] S. Gupta, P. Kapoor, K. Chaudhary, A. Gautam, R. Kumar, G.P.S. Raghava, In Silico Approach for Predicting Toxicity of Peptides and Proteins, PLoS One. (2013). https://doi.org/10.1371/journal.pone.0073957.

[26] A. Sharma, D. Singla, M. Rashid, G.P.S. Raghava, Designing of peptides with desired half-life in intestine-like environment, BMC Bioinformatics. (2014). https://doi.org/10.1186/1471-2105-15-282.

[27] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera - A visualization system for exploratory research and analysis, J. Comput. Chem. (2004). https://doi.org/10.1002/jcc.20084.

[28] T. Ashok Kumar, CFSSP: Chou and Fasman Secondary Structure Prediction server, Wide Spectr. (2013). https://doi.org/10.5281/zenodo.50733.

[29] G.K. Azad, The molecular assessment of SARS-CoV-2 Nucleocapsid Phosphoprotein variants among Indian isolates, Heliyon. (2021).

https://doi.org/10.1016/j.heliyon.2021.e06167.

[30] Z. Obradovic, K. Peng, S. Vucetic, P. Radivojac, A.K. Dunker, Exploiting heterogeneous sequence properties improves prediction of protein disorder, in: Proteins Struct. Funct. Genet., 2005. https://doi.org/10.1002/prot.20735.

[31] C.H.M. Rodrigues, D.E.V. Pires, D.B. Ascher, DynaMut: Predicting the impact of mutations on protein conformation, flexibility and stability, Nucleic Acids Res. (2018). https://doi.org/10.1093/nar/gky300.

[32] G.K. Azad, Identification and molecular characterization of mutations in nucleocapsid phosphoprotein of SARS-CoV-2, PeerJ. (2021). https://doi.org/10.7717/peerj.10666.

[33] J.V. Kringelum, C. Lundegaard, O. Lund, M. Nielsen, Reliable B Cell Epitope Predictions: Impacts of Method Development and Improved Benchmarking, PLoS Comput. Biol. (2012). https://doi.org/10.1371/journal.pcbi.1002829.

[34] V. Chauhan, T. Rungta, K. Goyal, M.P. Singh, Designing a multi-epitope based vaccine to combat Kaposi Sarcoma utilizing immunoinformatics approach, Sci. Rep. (2019). https://doi.org/10.1038/s41598-019-39299-8.

[35] O. Finco, R. Rappuoli, Designing vaccines for the twenty-first century society, Front. Immunol. (2014). https://doi.org/10.3389/fimmu.2014.00012.

[36] H.R. Ansari, G.P. Raghava, Identification of conformational B-cell Epitopes in an antigen from its primary sequence, Immunome Res. (2010). https://doi.org/10.1186/1745-7580-6-6.

[37] M. Letko, A. Marzi, V. Munster, Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses, Nat. Microbiol. (2020). https://doi.org/10.1038/s41564-020-0688-y.

[38] P. Majumdar, S. Niyogi, ORF3a mutation associated higher mortality rate in SARS-CoV-2 infection, Epidemiol. Infect. (2020). https://doi.org/10.1017/S0950268820002599.

[39] G.K. Azad, P.K. Khan, Variations in Orf3a protein of SARS-CoV-2 alter its structure and function, Biochem. Biophys. Reports. (2021). https://doi.org/10.1016/j.bbrep.2021.100933.

516 Table 1: List of NSP13 mutations identified in this study from India

S. No.	Mutation	Polarity changes	Charge changes	Frequency of mutation among Indian isolates
1	P53S	NP to P	Neutral to Neutral	2
2	G54S	NP to P	Neutral to Neutral	1
3	S100G	P to NP	Neutral to Neutral	5
4	E142V	P to NP	Acidic to Neutral	1
5	H164Y	P to P	Basic(weakly) to Neutral	1
6	S166A	P to NP	Neutral to Neutral	1
7	E168A	P to NP	Acidic to Neutral	1
8	Y205C	P to P	Neutral to Neutral	4
9	G206C	NP to P	Neutral to Neutral	4
10	G206V	NP to NP	Neutral to Neutral	1
11	T214I	P to NP	Neutral to Neutral	1
12	V226L	P to P	Neutral to Neutral	2
13	S236I	P to NP	Neutral to Neutral	3
14	A237T	NP to P	Neutral to Neutral	5
15	H245R	P to P	Basic(weakly) to Basic(strongly)	1
16	V247F	NP to NP	Neutral to Neutral	2
17	Y253H	P to P	Neutral to Basic(weakly)	2
18	S259T	P to P	Neutral to Neutral	1
19	P300L	NP to NP	Neutral to Neutral	2
20	R392C	P to P	Basic(strongly) to Neutral	1
21	P419S	NP to P	Neutral to Neutral	2
22	R442Q	P to P	Basic(strongly) to Neutral	1
23	F499L	NP to NP	Neutral to Neutral	1
24	P504L	NP to NP	Neutral to Neutral	1
25	P504S	NP to P	Neutral to Neutral	1
26	I545M	NP to NP	Neutral to Neutral	1
27	I575L	NP to NP	Neutral to Neutral	1

517

518

519

520

521

522

523

524

525 Table 2: List of MHC-class I molecules

526

Epitopes	Start –End Position in Nsp13 sequence	MHC-I allele with an affinity IC50 value of <200	Vaxijen score	Allergenecity
AQLPAPRTL	403-411	HLA-A*02:06, HLA-B*15:01, HLA-B*40:01, HLA-A*32:01, HLA-A*02:01, HLA-A*02:03, HLA-B*44:03	-0.3560 (probable non-antigen)	probable allergen
AVASKILGL	420-428	HLA-A*02:03, HLA-A*02:06, HLA-A*02:01, HLA-A*68:02, HLA-A*32:01, HLA-B*08:01, HLA-A*30:01	0.4971 (probable antigen)	probable allergen
DVTDVTQLY	56-64	HLA-A*26:01,HLA-B*35:01,HLA-A*68:01,HLA-A*01:01,HLA-A*30:02,HLA-B*53:01	0.9181 (probable antigen)	probable allergen
FAIGLALYY	291-299	HLA-B*35:01,HLA-A*26:01,HLA-B*53:01,HLA-A*01:01,HLA-A*30:02,HLA-B*58:01,HLA-B*15:01,HLA-A*68:01,HLA-B*57:01	1.1890 (probable antigen)	probable allergen
HVISTSHKL	33-41	HLA-A*68:02,HLA-A*02:06,HLA-A*26:01,HLA-A*32:01,HLA-A*02:03,HLA-B*35:01,HLA-A*02:01	0.6850 (probable antigen)	probable non-allergen
KSHKPPISF	73-81	HLA-B*57:01,HLA-B*58:01,HLA-A*32:01,HLA-B*15:01,HLA-A*30:01,HLA-A*30:02,HLA-A*24:02,HLA-A*23:01,HLA-B*07:02	0.9349 (probable antigen)	probable allergen
KVQIGEYTF	192-200	HLA-A*32:01,HLA-B*57:01,HLA-B*58:01,HLA-B*15:01,HLA-A*24:02,HLA-A*23:01	0.8369 (probable antigen)	probable allergen
QLYLGGMSY	62-70	HLA-B*15:01,HLA-A*03:01,HLA-A*30:02,HLA-B*35:01,HLA-A*26:01,HLA-A*32:01,HLA-A*01:01,HLA-A*11:01	0.6598 (probable antigen)	probable allergen

527

528

529

530

531

532

533

534

535

536 Table 3: List of MHC-class II molecules

Epitopes	Start –End Position in Nsp13 sequence	MHC-II allele with an affinity IC50 value of <200	Vaxijen score	Allergenecity
ACSHAAVDA	308-316	HLA-DRB1*07:01 HLA-DRB4*01:01 HLA-DRB1*03:01 HLA-DRB1*15:01	0.1368 (Probable NON-ANTIGEN).	PROBABLE ALLERGEN
HKLVLSVNP	39-47	HLA-DRB1*15:01 HLA-DRB1*03:01 HLA-DRB1*07:01 HLA-DRB3*01:01 HLA-DRB5*01:01	0.5001 (Probable ANTIGEN	PROBABLE ALLERGEN
ETFKLSYGI	143-151	HLA-DRB1*07:01 HLA-DRB1*15:01 HLA-DRB3*01:01 HLA-DRB5*01:01 HLA-DRB1*03:01	0.2915 (Probable NON-ANTIGEN).	PROBABLE ALLERGEN
ERLKLFAAE	128-136	HLA-DRB1*15:01 HLA-DRB1*07:01 HLA-DRB5*01:01 HLA-DRB1*03:01 HLA-DRB3*01:01	-0.0539 (Probable NON-ANTIGEN)	PROBABLE ALLERGEN
EHYVRITGL	244-252	HLA-DRB5*01:01 HLA-DRB1*15:01 HLA-DRB4*01:01 HLA-DRB1*07:01	0.5647 (Probable ANTIGEN	PROBABLE NON-ALLERGEN
QCFKMFYKG	470-478	HLA-DRB1*15:01 HLA-DRB5*01:01 HLA-DRB1*07:01 HLA-DRB3*01:01 HLA-DRB1*03:01	0.4573 (Probable ANTIGEN	PROBABLE NON-ALLERGEN

537

538

539

540

541

542

543

544

545

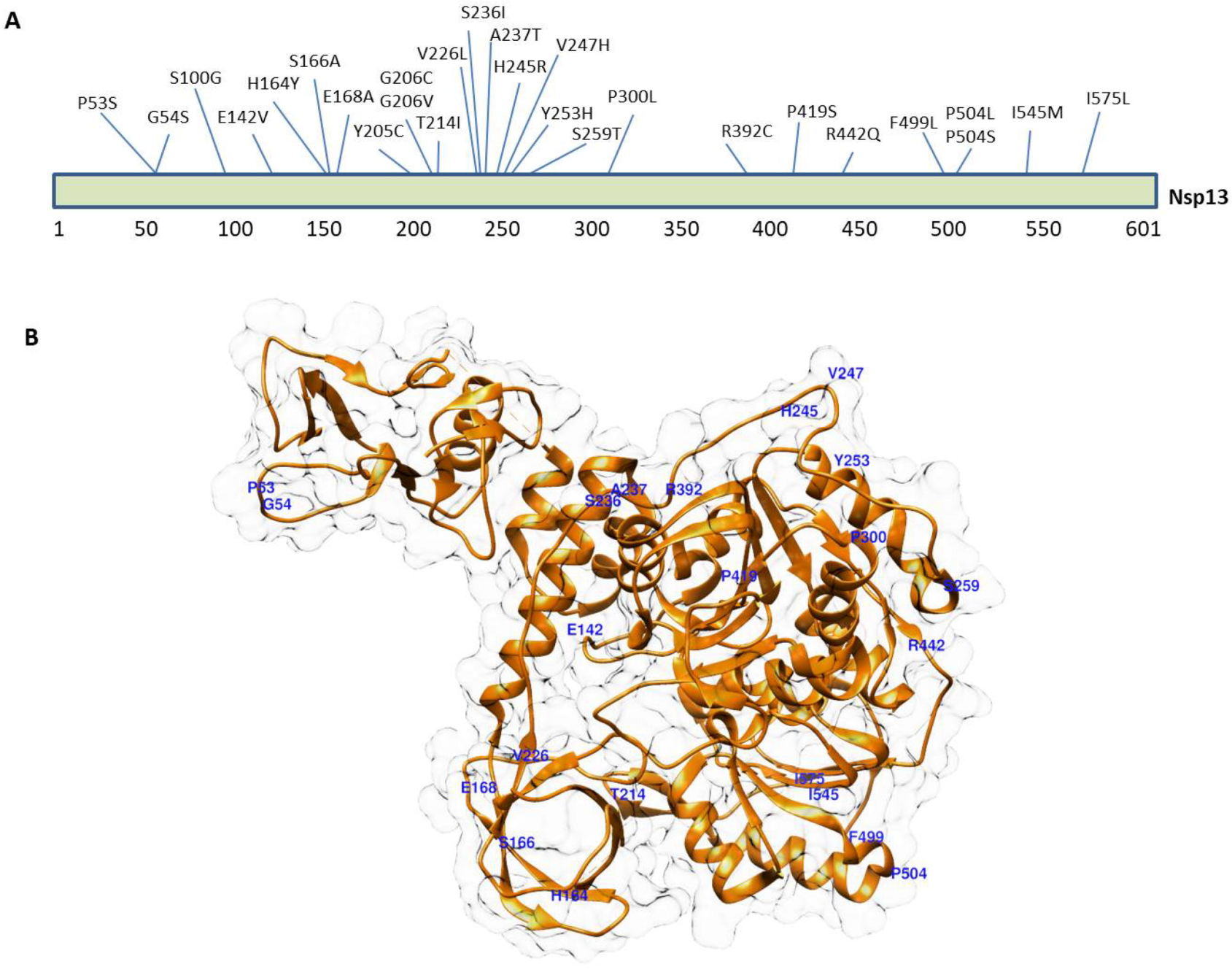
546 Table 4: Physiochemical properties of MHC-I and MHC-II molecules

Class	S. No	Peptide	Half-life(sec)	Toxicity	Hydrophobicity	Hydropathicity	Hydrophilicity	Charge	pI	Mol Wt	Surface accessibility	flexibility	Polarity
MHC-class I	1	AQLPAPRTL	2.292	NT	-0.13	0.41	-0.20	1.00	10.11	966.27	46.911	3.920	61.25
	2	AVASKILGL	0.792	NT	0.18	1.64	-0.51	1.00	9.11	871.22	36.311	3.810	53.83
	3	DVTDVTQLY	1.221	NT	-0.10	-0.11	-0.19	-2.00	3.57	1053.26	45.567	3.950	90.57
	4	FAIGLALYY	1.107	NT	0.34	1.72	-1.50	0.00	5.87	1030.36	32.667	3.590	4.620
	5	HVISTSHKL	2.582	NT	-0.09	-0.01	-0.32	2.00	9.11	1021.32	45.278	3.780	158.4
	6	KSHKPPISF	1.759	NT	-0.21	-0.94	0.20	2.50	10.02	1040.35	54.667	4.070	157.6
	7	KVQIGEYTF	1.489	NT	-0.06	-0.20	-0.26	0.00	6.35	1084.37	48.511	4.020	106.8
	8	QLYLGGMSY		NT	0.08	0.20	-1.00	0.00	5.87	11.773	39.867	3.930	10.75
MHC-Class II	1	ACSHAAVDA	0.944	NT	0.02	0.71	-0.19	-0.50	5.09	844.01	33.744	3.520	95.580
	2	HKLVLSVNP	1.371	NT	-0.04	0.33	-0.40	1.50	9.11	1006.35	45.544	3.770	108.8
	3	ETFKLSYGI	2.489	NT	-0.01	0.06	-0.28	0.00	6.35	1057.34	45.978	4.010	105.3
	4	ERLKLFAAE	2.445	NT	-0.21	-0.16	0.54	0.00	6.49	1076.38	52.267	3.750	202.5
	5	EHYVRITGL	1.466	NT	-0.11	-0.12	-0.26	0.50	7.10	1087.38	45.822	3.960	157.5
	6	QCFKMFYKG	1.386	NT	-0.13	-0.33	-0.38	2.00	9.25	1151.53	50.489	3.650	107.7

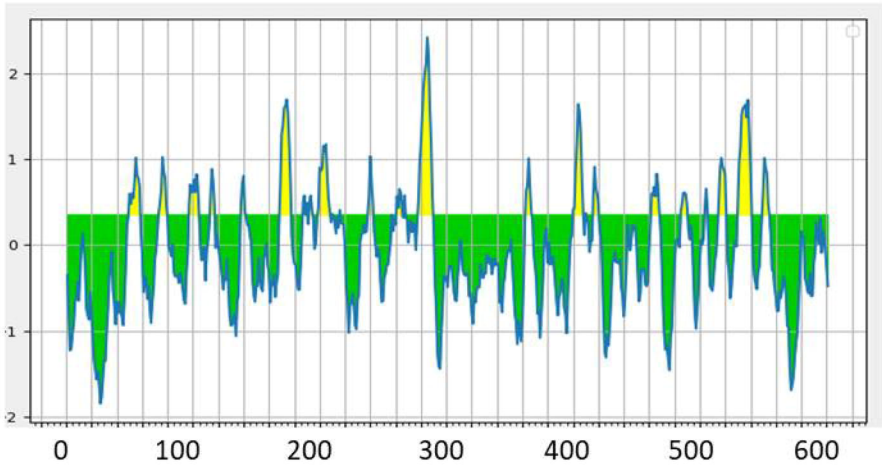
547 Table 5: The table show the ΔΔG and ΔΔSvib ENCoM of the mutants present in Nsp13.

S. No.	Mutation	ΔΔG (kcal/mol)	Effect on Nsp13 stability	$\Delta\Delta S_{Vib}$ ENCoM ($kcal.mol^{-1}.K^{-1}$)	Effect on Nsp13 flexibility
1	P53S	0.024	Stabilizing	0.504	Increase
2	G54S	-1.694	Destabilizing	0.130	Increase
3	E142V	-0.174	Destabilizing	0.260	Increase
4	H164Y	0.975	Stabilizing	-0.375	Decrease
5	S166A	-0.1	Destabilizing	0.145	Increase
6	E168A	-0.192	Destabilizing	0.055	Increase
7	T214I	0.881	Stabilizing	-0.339	Decrease
8	V226L	0.528	Stabilizing	-0.236	Decrease
9	S236I	0.75	Stabilizing	-0.370	Decrease
10	A237T	0.924	Stabilizing	-0.424	Decrease
11	H245R	-0.154	Destabilizing	0.147	Increase
12	V247F	-0.096	Destabilizing	-0.102	Decrease
13	Y253H	0.309	Stabilizing	-0.109	Decrease
14	S259T	0.031	Stabilizing	-0.031	Decrease
15	P300L	0.362	Stabilizing	-0.427	Decrease
16	R392C	-0.347	Destabilizing	0.206	Increase
17	P419S	0.693	Stabilizing	-0.661	Decrease
18	R442Q	-0.18	Destabilizing	0.381	Increase
19	F499L	0.051	Stabilizing	0.444	Increase
20	P504L	0.601	Stabilizing	-0.050	Decrease
21	P504S	-0.122	Destabilizing	-0.052	Decrease
22	I545M	-1.209	Destabilizing	-0.013	Decrease
23	I575L	-0.723	Destabilizing	-0.109	Decrease

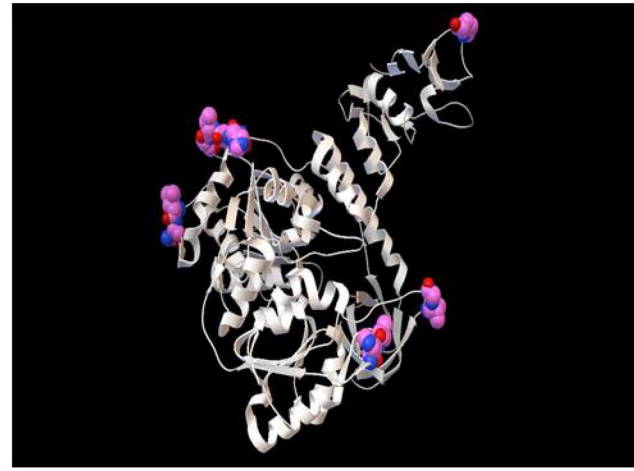
548


A
S236I
A237T
V247H
V226L
H245R
S166A
S100G
G206C
H164Y
E168A
G206V
P53S
P300L
Y253H
T214I
G54S
E142V
Y205C
S259T
R392C
P419S
F499L
P504L
I545M
I575L
R442Q
P504S
Nsp13
1
50
100
150
200
250
300
350
400
450
500
550
601
B
V247
H245
Y253
P53
G54
S236
A237
R392
P300
P419
S259
E142
R442
V226
E168
T214
I575
I545
S166
F499
P504
H164

A



C

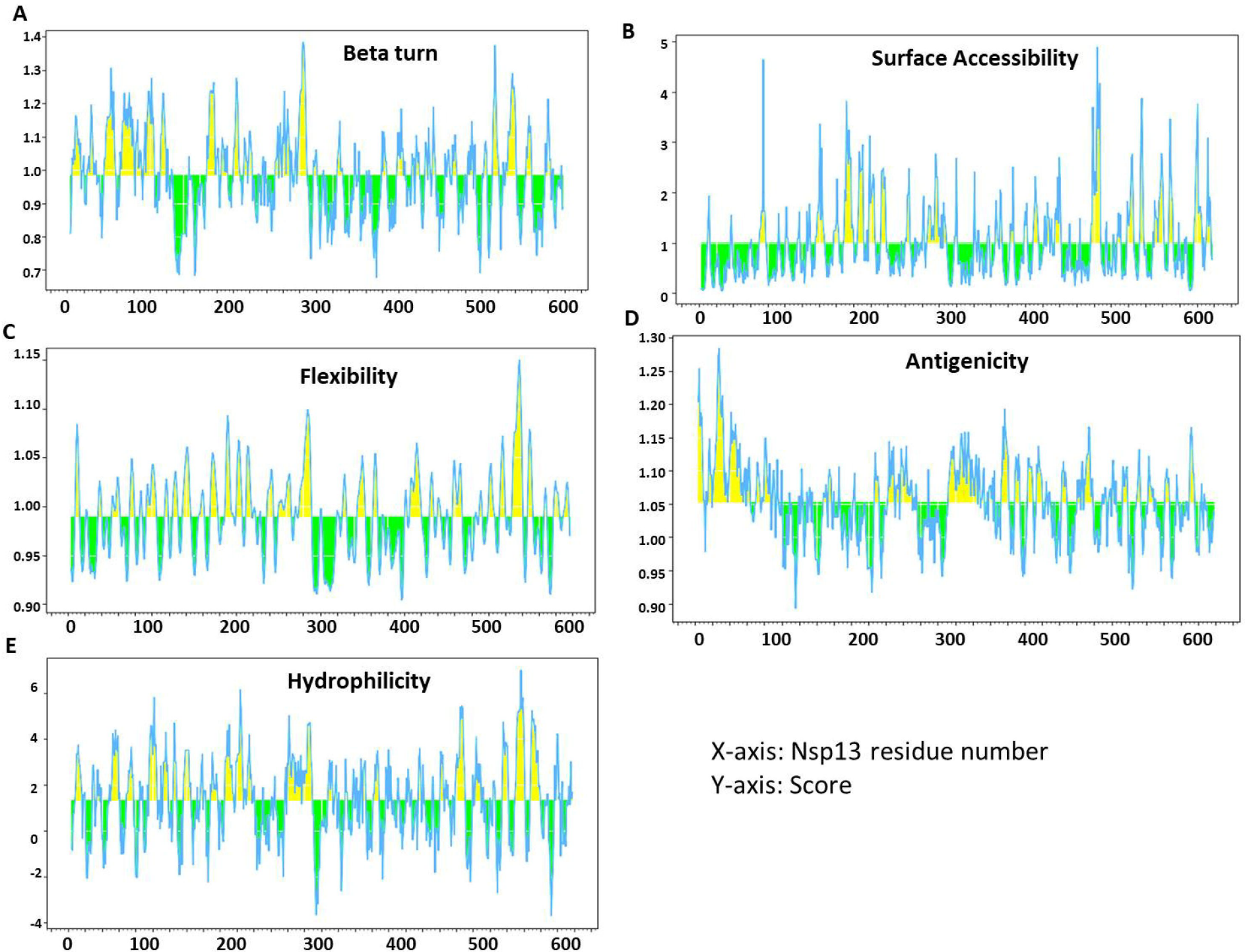


B

Peptide (Start-End)	Peptide Sequence	Vexijen score	Allergenicity	Toxicity
Peptide 1 (50-58)	CNAPGCDVT	-0.3902 (non-antigen)	Probable Allergen	Non Toxin
Peptide 2 (97-104)	CVGSDNVT	0.5270 (antigen)	Probable Allergen	Non Toxin
Peptide 3 (169-177)	VGKPRPPLN	-0.2476 (non-antigen)	Probable Allergen	Non Toxin
Peptide 4 (199-208)	TFEKGDYGDA	0.9707 (antigen)	Probable Allergen	Non Toxin
Peptide 5 (278-289)	STLQGPPGTGKS	0.3085 (non-antigen)	Non Allergen	Non Toxin
Peptide 6 (400-408)	GDPAQLPAP	-0.2856 (non-antigen)	Probable Allergen	Non Toxin
Peptide 7 (530-541)	TQTVDSSQGSEY	0.5969 (antigen)	Non Allergen	Non Toxin

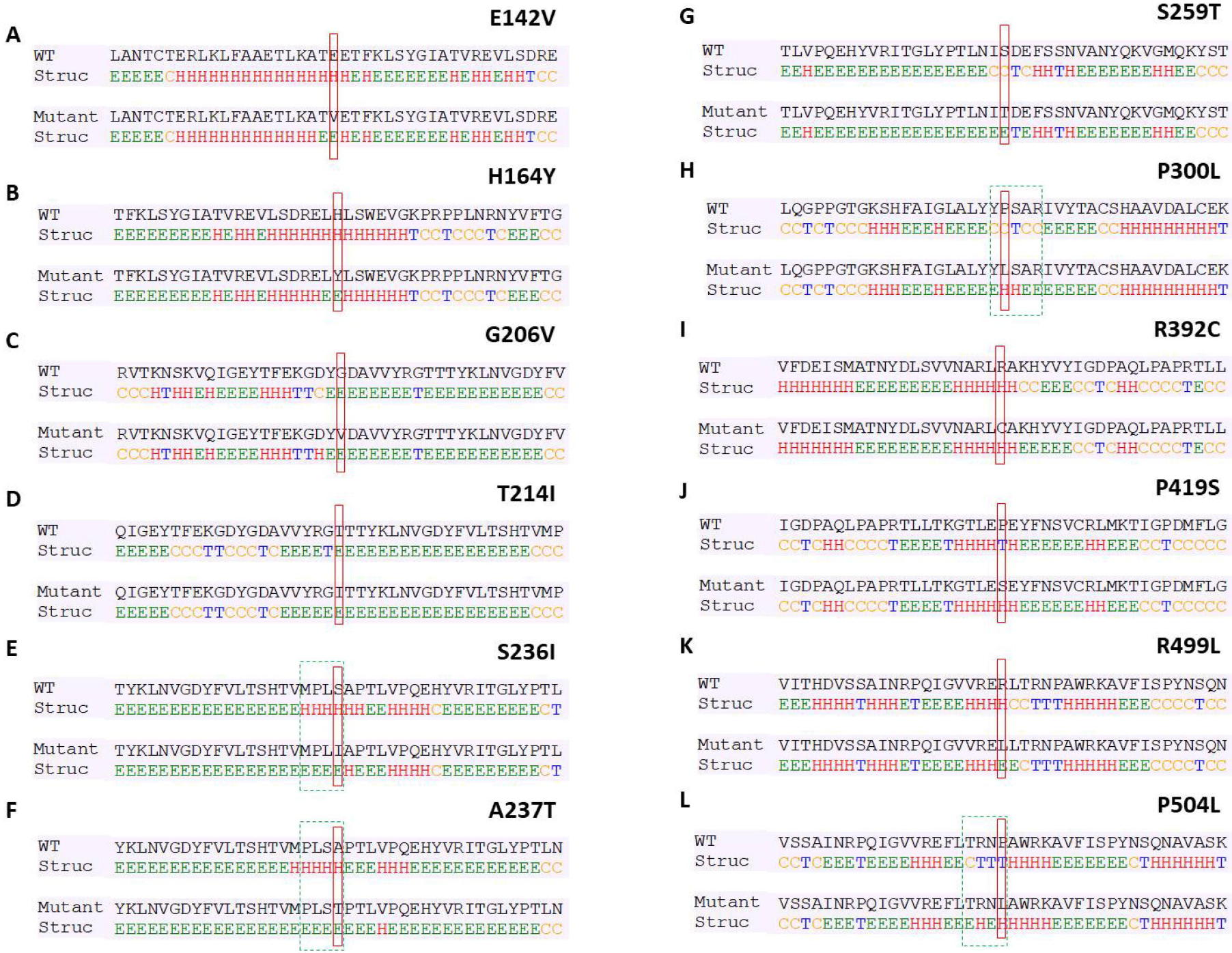
D

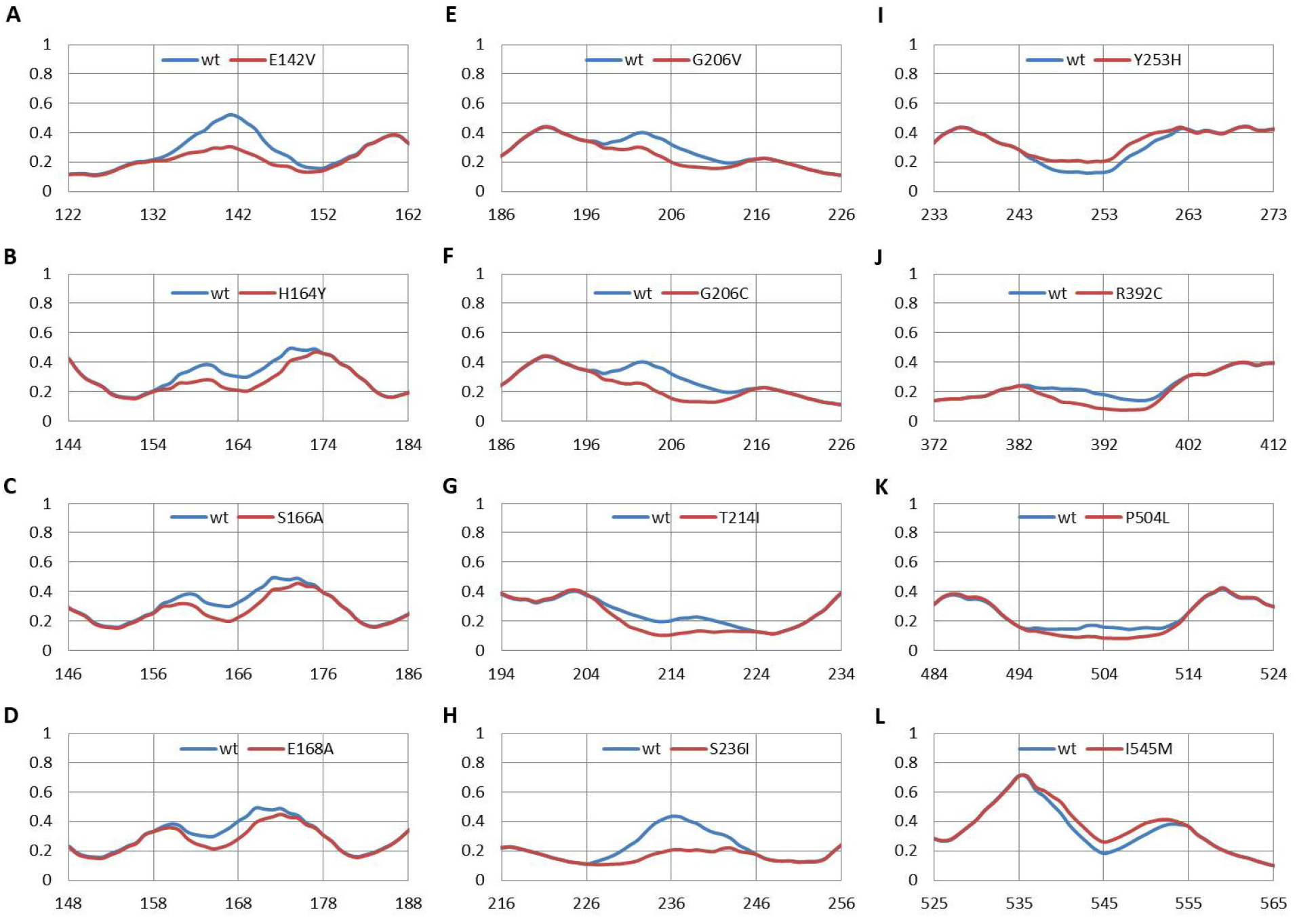
S. No	Nsp13 Residue number	Amino acid
1	78	PRO
2	169	VAL
3	170	GLY
4	245	HIS
5	246	TYR
6	247	VAL
7	248	ARG
8	256	LEU
9	257	ASN
10	482	HIS
11	484	VAL



X-axis: Nsp13 residue number

Y-axis: Score


A
E142V
WT LANTCTERLKLFAAETLKATEETFKLSYGIATVREVLSDRE
Struc EEEEECHHHHHHHHHHHHHHHHEHEEEEEEEEHEHHEHHTCC
Mutant LANTCTERLKLFAAETLKATVETFKLSYGIATVREVLSDRE
Struc EEEEECHHHHHHHHHHHHHHEEHEHEEEEEEEEHEHHEHHTCC
B
H164Y
WT TFKLSYGIATVREVLSDRELHLSWEVGKPRPPLNRNYVFTG
Struc EEEEEEEEEHEHHEHHHHHHHHHHHHHHTCCTCCCTCEEECC
Mutant TFKLSYGIATVREVLSDRELYLSWEVGKPRPPLNRNYVFTG
Struc EEEEEEEEEHEHHEHHHHHHEEHHHHHHTCCTCCCTCEEECC
C
G206V
WT RVTKNSKVQIGEYTFEKGDYGDAVVYRGTTTYKLNVGDYFV
Struc CCCHTHHEHEEEEHHHTTCEEEEEEEETEEEEEEEEEECC
Mutant RVTKNSKVQIGEYTFEKGDYVDAVVYRGTTTYKLNVGDYFV
Struc CCCHTHHEHEEEEHHHTTHEEEEEEEETEEEEEEEEEECC
D
T214I
WT QIGEYTFEKGDYGDAVVYRGITTYKLNVGDYFVLTSHTVMP
Struc EEEEECCCTTCCCTCEEEETEEEEEEEEEEEEEEEECCC
Mutant QIGEYTFEKGDYGDAVVYRGITTYKLNVGDYFVLTSHTVMP
Struc EEEEECCCTTCCCTCEEEEEEEEEEEEEEEEEEEECCC
E
S236I
WT TYKLNVGDYFVLTSHTVMPLSAPTLVPQEHYVRITGLYPTL
Struc EEEEEEEEEEEEEEEEHHHHHHEEHHHHCEEEEEEEECT
Mutant TYKLNVGDYFVLTSHTVMPLIAPTLVPQEHYVRITGLYPTL
Struc EEEEEEEEEEEEEEEEEEEEHEEHHHHCEEEEEEEECT
F
A237T
WT YKLNVGDYFVLTSHTVMPLSAPTLVPQEHYVRITGLYPTLN
Struc EEEEEEEEEEEEEEEEHHHHHEEEHHHEEEEEEEEEEECC
Mutant YKLNVGDYFVLTSHTVMPLSTPTLVPQEHYVRITGLYPTLN
Struc EEEEEEEEEEEEEEEEEEEEEEEHEEEEEEEEEEEEECC
G
S259T
WT TLVPQEHYVRITGLYPTLNISDEFSSNVANYQKVGMQKYST
Struc EEHEEEEEEEEEEEEEEECCTCHHTHEEEEEEHHEECCC
Mutant TLVPQEHYVRITGLYPTLNITDEFSSNVANYQKVGMQKYST
Struc EEHEEEEEEEEEEEEEEEEETEHHTHEEEEEEHHEECCC
H
P300L
WT LQGPPGTGKSHFAIGLALYYPSARIVYTACSHAAVDALCEK
Struc CCTCTCCCHHHEEEHEEEECCTCCEEEEECCHHHHHHHHHT
Mutant LQGPPGTGKSHFAIGLALYYLSARIVYTACSHAAVDALCEK
Struc CCTCTCCCHHHEEEHEEEEEHHEEEEEEECCHHHHHHHHHT
I
R392C
WT VFDEISMATNYDLSVVNARLRAKHYVYIGDPAQLPAPRTLL
Struc HHHHHHHEEEEEEEEEHHHHHHCCEEECCTCHHCCCCTECC
Mutant VFDEISMATNYDLSVVNARLCAKHYVYIGDPAQLPAPRTLL
Struc HHHHHHHEEEEEEEEEHHHHHHEEEEECCTCHHCCCCTECC
J
P419S
WT IGDPAQLPAPRTLLTKGTLEPEYFNSVCRLMKTIGPDMFLG
Struc CCTCHHCCCCTEEEETHHHHTHEEEEEEHHEEECCTCCCCC
Mutant IGDPAQLPAPRTLLTKGTLESEYFNSVCRLMKTIGPDMFLG
Struc CCTCHHCCCCTEEEETHHHHHHEEEEEEHHEEECCTCCCCC
K
R499L
WT VITHDVSSAINRPQIGVVRERLTRNPAWRKAVFISPYNSQN
Struc EEEHHHHTHHHETEEEEHHHHCCTTTHHHHHEEECCCCTCC
Mutant VITHDVSSAINRPQIGVVRELLTRNPAWRKAVFISPYNSQN
Struc EEEHHHHTHHHETEEEEHHHEECTTTHHHHHEEECCCCTCC
L
P504L
WT VSSAINRPQIGVVREFLTRNPAWRKAVFISPYNSQNAVASK
Struc CCTCEEETEEEEHHHEECTTTHHHHEEEEEEECTHHHHHHT
Mutant VSSAINRPQIGVVREFLTRNLAWRKAVFISPYNSQNAVASK
Struc CCTCEEETEEEEHHHEEEHEHHHHHEEEEEEECTHHHHHHT


A
wt
E142V
B
wt
H164Y
C
wt
S166A
D
wt
E168A
E
wt
G206V
F
wt
G206C
G
wt
T214I
H
wt
S236I
I
wt
Y253H
J
wt
R392C
K
wt
P504L
L
wt
I545M
X-axis: Nsp13 residue number, Y-axis: protein disorder score