Supplementary to "Genome-wide Analysis of Indian SARS-CoV-2 Genomes to Identify T-cell and B-cell Epitopes from Conserved Regions based on Immunogenicity and Antigenicity"

Nimisha Ghosh[a,e], Nikhil Sharma[b,e], Indrajit Saha[c,*], Sudipto Saha[d]

[a] *Department of Computer Science and Information Technology, Institute of Technical Education and Research,*
Siksha 'O' Anusandhan (Deemed to be University), Bhubaneswar, Odisha, India
[b] *Department of Electronics and Communication Engineering, Jaypee Institute of Information Technology, Noida, Uttar Pradesh, India*
[c] *Department of Computer Science and Engineering,*
National Institute of Technical Teachers' Training and Research, Kolkata, West Bengal, India
[d] *Division of Bioinformatics Bose Institute, Kolkata, West Bengal, India*
[e] *Equally contributed*

S1. Preliminaries

S1.1. Epitope based Vaccine

Development of vaccine is one of the most remarkable achievements of the modern era [1]. Basic concept of vaccine revolves around evoking the immune response inside host body. This can be achieved with the help of a catalyst in the form of effective micro particles which then helps in developing an adaptive immunity for curing diseases such as small pox, tuberculosis, cancer, AIDS, Alzheimer etc. [2]. With the rapid advancement in technology at molecular level, epitope based vaccine design has become very popular e.g. in case of SARS-CoV-2 [3]. This idea of epitope based vaccine works around the concept of identifying the immunodominant B-cell epitopes conjugated with T-cell epitopes for inducing specific human immune responses [4]. More details about epitope based vaccine can be found in [1, 5].

---

[*] Corresponding author: indrajit@nitttrkol.ac.in

S1.2. T-Cell Epitope Prediction

T-cells have a marked importance in adaptive immunity. T-cells have specific receptors on their surface known as T-cell receptors which are responsible for recognising antigens when they mark their presence on the surface of antigen presenting cells (APCs) [6, 7]. The identification of T-cell epitopes evokes the cellular immune response by bounding to major histocompatibility (MHC) molecules. MHCs are the polymorphic proteins which are present in complex living organisms and are of two different categories: MHC-I and MHC-II interacting binders as CD8+ and CD4+ respectively. These MHC proteins have highly polymorphic alleles of class I human leukocyte antigens (HLA I) and class II human leukocyte antigens (HLA II). As the MHC molecules are recognised by CD8+ and CD4+ T-cells respectively, they give rise to CD8+ and CD4+ T-cell epitopes. Thus, the main aim of T-cell epitope prediction is to identify the peptides in the pathogen that can stimulate CD8+ and CD4+ T-cells. This is known as immunogenicity. Therefore, finding of optimal set of CD8+ and CD4+ T-cell epitopes is an important and challenging task. The online tools for predicting MHC-I and MHC-II T-cell epitopes are given in Table S1.

S1.3. B-Cell Epitope Prediction

B-cells are also important for adaptive immunity [8]. It is important to note that T-cell stimulates B-cells with the help of T-cell receptors (TCR). In this regard, TCR on T helper cells bind to the antigen-complexed MHC II molecules on the surface of B-cell resulting in T-cell activation. This activated T-cell then provides a second activation signal to the B-cell to stimulate. B-cells invoke humoral immunity by recognising solvent-exposed antigens with the help of B-cell receptors. Once activated, B-cells secrete antibodies which evoke humoral immunity [9]. B-cell epitope is the portion of the antigen which binds to the antibodies. This is known as antigenicity. These epitopes recognized by B-cells constitute of exposed solvent regions in the antigen and are usually of different chemical nature. However, most antigens are proteins and are the targets for B-cell epitope prediction. This can be achieved through many online tools as shown in Table S2.

S1.4. Physico-Chemical Properties of Epitope

The measure of therapeutic effects of a vaccine against a disease is highly dependent on the interaction among the different bio-molecules. This in turn initiates the immune response which is influenced by the physico-chemical properties of the chemical substances present on the surface of the molecules. Such physico-chemical properties are charge, polarity, alphaticity, aromaticity, acidicity, basicity, hydrophobicity, hydrophilicity, neutrality, hydroxylicity and sulphur content of the epitopes. Hence, such molecular attributes are important to report and identify physical hazards and ecotoxicity [10]. Therefore, Pfeature server [11] is used to evaluate the physico-chemical properties of the potential T-cell (MHC-I and MHC-II restricted) and B-cell epitopes.

S1.5. Docking of T-cell Epitope

A critical step in the path of epitope based vaccine design is to visualize the identified epitopes with 3D modelling to predict the optimal binding-conformation of molecule ligands to the appropriate binding site to represent a stable molecule. Hence, it is also important to evaluate the predicted epitope pairs using 3D docking during vaccine design [12]. In order to obtain a stable conformation state, a ranking procedure is adapted on the basis of scoring function which depends on the energy of the pose within the binding site. Most of the scoring functions are based on the molecular mechanisms such as solvent effect, free energy, internal rotation, association energy and free energy between ligand and receptor complex [13]. Autodock Vina [14] predicts such protein–peptide docking. AutoDock Vina consists of two main programs for docking of the ligand to a set of grids describing the target protein and for pre-calculating these grids. It also has as an improved local search routine and makes use of multicore/multi-CPU computer setups.

References

[1] M. Skwarczynski, I. Toth, Peptide-based synthetic vaccines, Chemical Science 7 (2015). `doi:10.1039/C5SC03892H`.

[2] F. Zepp, Principles of vaccination, Methods in molecular biology (Clifton, N.J.) 1403 (2016) 57–84. `doi:10.1007/978-1-4939-3387-7_3`.

[3] G. Poland, I. Ovsyannikova, S. Crooke, R. Kennedy, Sars-cov-2 vaccine development: Current status, Mayo Clinic Proceedings 95 (10) (2020) 2172–2188. `doi:https://doi.org/10.1016/j.mayocp.2020.07.021`.

[4] A. Grifoni, D. Weiskopf, S. Ramirez, J. Mateus, J. Dan, C. Moderbacher, S. Rawlings, A. Sutherland, L. Premkumar, R. Jadi, D. Marrama, A. Silva, A. Frazier, A. Carlin, J. Greenbaum, B. Peters, F. Krammer, D. Smith, S. Crotty, A. Sette, Targets of t cell responses to sars-cov-2 coronavirus in humans with covid-19 disease and unexposed individuals, Cell 181 (05 2020). `doi:10.1016/j.cell.2020.05.015`.

[5] S. Parvizpour, M. M. Pourseif, J. R., M. A. Rafi, Y. Omidi, Epitope-based vaccine design: a comprehensive overview of bioinformatics approaches, Drug Discovery Today 25 (6) (2020) 1034 – 1042. `doi:https://doi.org/10.1016/j.drudis.2020.03.006`.

[6] I. Saha, G. Mazzocco, D. Plewczynski, Consensus classification of human leukocyte antigen class ii proteins, Immunogenetics 65 (2013) 97–105. `doi:https://doi.org/10.1007/s00251-012-0665-6`.

[7] D. L. Farber, Form and function for t cells in health and disease, Nature Reviews Immunology 20 (2020) 83–84. doi:https://doi.org/10.1038/s41577-019-0267-8.

[8] D. Tarlinton, B cells still front and centre in immunology, Nature Reviews Immunology 19 (2019) 85–86. doi:https://doi.org/10.1038/s41577-018-0107-2.

[9] J. G. Cyster, C. D. Allen, B cell responses: Cell interaction dynamics and decisions, Cell 177 (3) (2019) 524 – 540. doi:https://doi.org/10.1016/j.cell.2019.03.016.

[10] M. Sied, Filli, A. Abdalla, Ibrahim, A. Abbas, A. Morshed, K. Boakye-Yiadom, S. Kesse, M. Farooq, M. Bazezy, B. Wang, M. B. Wang, The impact of physicochemical characteristics on therapeutic efficacy of anticancer nanomaterials: A review, International Journal of Pharmaceutical Sciences and Drug Research 11 (03 2019). doi:10.25004/IJPSDR.2019.110203.

[11] A. Pande, S. Patiyal, A. Lathwal, C. Arora, D. Kaur, A. Dhall, G. Mishra, H. Kaur, N. Sharma, S. Jain, S. Usmani, P. Agrawal, R. Kumar, V. Kumar, G. Raghava, Computing wide range of protein/peptide features from their sequence and structure, bioRxiv (04 2019). doi:10.1101/599126.

[12] G. S. Krishnan, A. Joshi, V. Kaushik, T cell epitope designing for dengue peptide vaccine using docking and molecular simulation studies, Molecular Simulation 46 (10) (2020) 787–795. doi:10.1080/08927022.2020.1772970.

[13] T. Pantsar, A. Poso, Binding affinity via docking: Fact and fiction, Molecules 23 (2018) 1899. doi:10.3390/molecules23081899.

[14] M. A. Rauf, Ligand docking and binding site analysis with pymol and autodock/vina, International Journal of Basic & Applied Sciences 4 (2015) 168–177. doi:10.14419/ijbas.v4i2.4123.

[15] I. Dimitrov, P. Garnev, D. Flower, I. Doytchinova, Epitop—a proteochemometric tool for mhc class ii binding prediction, Bioinformatics (Oxford, England) 26 (2010) 2066–8. doi:10.1093/bioinformatics/btq324.

[16] C. Lundegaard, K. Lamberth, M. Harndahl, O. Lund, M. Nielsen, Netmhc-3.0: accurate web accessible predictions of human, mouse and monkey mhc class i affinities for peptides of length 8-11, Nucleic Acids Res 36 (2008) W509–W512.

[17] I. Doytchinova, P. Guan, D. Flower, Epijen: A server for multistep t cell epitope prediction, BMC bioinformatics 7 (2006) 131. doi:10.1186/1471-2105-7-131.

[18] P. Dönnes, O. Kohlbacher, Svmhc: a server for prediction of mhc-binding peptides, Nucleic acids research 34 (2006) W194–7. doi:10.1093/nar/gkl284.

[19] M. Bhasin, G. Raghava, Raghava, g.: Prediction of ctl epitopes using qm, svm and ann techniques. vaccine 22(23-24), 3195-3204, Vaccine 22 (2004) 3195–204. doi:10.1016/j.vaccine.2004.02.005.

[20] P. Reche, J. P. Glutting, H. Zhang, E. Reinherz, Enhancement to the rankpep resource for the prediction of peptide binding to mhc molecules using profiles, Immunogenetics 56 (2004) 405–19. doi:10.1007/s00251-004-0709-7.

[21] M. Bhasin, G. Raghava, Prediction of promiscuous and high-affinity mutated mhc binders, Hybridoma and hybridomics 22 (2003) 229–34. doi:10.1089/153685903322328956.

[22] P. Guan, I. Doytchinova, C. Zygouri, D. Flower, Mhcpred: A server for quantitative prediction of peptide-mhc binding, Nucleic acids research 31 (2003) 3621–4. doi:10.1093/nar/gkg510.

[23] H. Singh, G. Raghava, Propred: prediction of hla-dr binding sites, Bioinformatics 17 (2001) 1236–1237. doi:10.1093/bioinformatics/17.12.1236.

[24] K. Parker, M. Bednarek, J. Coligan, Scheme for ranking potential hla-a2 binding peptides based on independent binding of individual peptide side-chains, Journal of immunology (Baltimore, Md. : 1950) 152 (1994) 163–75.

[25] S. Liang, D. Zheng, B. Yao, C. Zhang, EPCES and EPSVR: Prediction of B-Cell Antigenic Epitopes on Protein Surfaces with Conformational Information, Vol. 2131, Methods in

molecular biology (Clifton, N.J.), 2020. doi:10.1007/978-1-0716-0389-5_16.

[26] C. Zhou, Z. Chen, L. Zhang, D. Yan, T. Mao, K. Tang, T. Qiu, Z. Cao, Seppa 3.0-enhanced spatial epitope prediction enabling glycoprotein antigens, Nucleic acids research 47 (05 2019). doi:10.1093/nar/gkz413.

[27] M. Jespersen, B. Peters, M. Nielsen, P. Marcatili, Bepipred-2.0: Improving sequence-based b-cell epitope prediction using conformational epitopes, Nucleic acids research 45 (05 2017). doi:10.1093/nar/gkx346.

[28] W. Chen, P. Sun, Y. Lu, W. Guo, Y. X. Huang, Z. Ma, Mimopro: A more efficient web-based tool for epitope prediction using phage display libraries, BMC bioinformatics 12 (2011) 199. doi:10.1186/1471-2105-12-199.

[29] N. Rubinstein, I. Mayrose, E. Martz, T. Pupko, Epitopia: A web-server for predicting b-cell epitopes, BMC bioinformatics 10 (2009) 287. doi:10.1186/1471-2105-10-287.

[30] M. Sweredoski, P. Baldi, Pepito: Improved discontinuous b-cell epitope prediction using multiple distance thresholds and half sphere exposure, Bioinformatics (Oxford, England) 24 (2008) 1459–60. doi:10.1093/bioinformatics/btn199.

[31] J. Ponomarenko, H. H. Bui, W. Li, N. Fusseder, P. Bourne, A. Sette, B. Peters, Ellipro: A new structure-based tool for the prediction of antibody epitopes, BMC bioinformatics 9 (2009) 514. doi:10.1186/1471-2105-9-514.

[32] S. Saha, G. Raghava, Prediction methods for b-cell epitopes, Methods in molecular biology (Clifton, N.J.) 409 (2007) 387–94. doi:10.1007/978-1-60327-118-9_29.

Tool Name	Link to the tool	Predictive method	Year	References
EpiTOP	http://www.pharmfac.net/EpiTOP	Quantitative Matrix	2010	[15]
NetMHC-3.0	http://www.cbs.dtu.dk/services/NetMHC	ANN based method and artificial neural networks	2008	[16]
EpiJen	http://www.ddg-pharmfac.net/epijen/EpiJen/EpiJen.htm	Multi-step algorithm	2006	[17]
SVMHC	http://www-bs.informatik.uni-tuebingen.de/SVMHC	SVM-based method	2006	[18]
CTLpred	http://www.imtech.res.in/raghava/ctlpred	quantitative matrix (QM), Support Vector Machine (SVM), Artificial Neural Network (ANN)	2004	[19]
RANKPEP	http://bio.dfci.harvard.edu/RANKPEP	PSSM	2004	[20]
MMBPred	http://www.imtech.res.in/raghava/mmbpred	Quantitative matrix	2004	[21]
MHCPred	http://www.ddg-pharmfac.net/mhcpred/MHCPred	Additive method	2003	[22]
ProPred	http://www.imtech.res.in/raghava/propred/	Quantitative matrix	2001	[23]
BIMAS	http://www-bimas.cit.nih.gov/molbio/hla_bind/	Published coefficient tables	1994	[24]

Table S1: List of MHC-I and MHC-II T-cell epitopes prediction servers

Tool Name	Server Link	Type of B-cell Epitope predicted and Method	Year	Reference
EPCES	http://sysbio.unl.edu/EPCES/	Consensus Scoring	2020	[25]
SEPPA 3.0	http://lifecenter.sgst.cn/seppa/index.php	Logistic regression model	2019	[26]
BepiPred	http://www.cbs.dtu.dk/services/BepiPred/	Random Forest Algorithm	2017	[27]
MimoPro	http://informatics.nenu.edu.cn/MimoPro	Mapping a number of mimotopes	2011	[28]
Epitopia	http://epitopia.tau.ac.il/	Naïve Bayes classifier	2009	[29]
BEPro (PEPITO)	http://pepito.proteomics.ics.uci.edu/	A Combination of amino-acid propensity scores and half sphere exposure values at multiple distances	2008	[30]
ElliPro	http://tools.immuneepitope.org/tools/ElliPro/iedb_input	3D Structure Modeling	2008	[31]
ABCPred	http://www.imtech.res.in/raghava/abcpred/	Artificial Neural Network	2006	[32]

Table S2: List of B cell epitope prediction servers

Consensus Conserved Region (CCnR)	Protein Sequence of CCnR	Length of CCnR	BLAST Specificity Score of CCnR	% of BLAST Specificity Score as Query Coverage	Coding Region (CR)	Starting Coordinate of CR	Ending Coordinate of CR	Length of CR	Coded Protein from CR
4012-CACAGAAAACTTGTTACTTTATATTGACATTAATGGCAATCTTCATCCAGATTCTGCCACTCTTGTTAGTGACATTGACATCACTTTCTTAAAGAAAGATGCTCCATATATAGTGGGTGATGTTGTTCAAGAGGGTGTTTTAACTGCTGTGGTTATACCTACTAAAAAGGCTGGTGGCACTACTGAAATGCTAGCGAAAGCTTT-4215	TENLLLYIDINGNLHPDSATLVSDIDITFLKKDAPYIVGDVVQEGVLTAVVIPTKKAGGTTEMLAKA	204	377	100	ORF1ab	266	21555	21290	NSP3
10463-TTAAGGGTTCATTCCTTAATGGTTCATGTGGTAGTGTTGGTTTTAACATAGATTATGACTGTGTCTCTTTTTGTTAC-10539	KGSFLNGSCGSVGFNIDYDCVSFCY	77	143	100	ORF1ab	266	21555	21290	3CL-Proteinase

Consensus Conserved Region (CCnR)	Protein Sequence of CCnR	Length of CCnR	BLAST Specificity Score of CCnR	% of BLAST Specificity Score as Query Coverage	Coding Region (CR)	Starting Coordinate of CR	Ending Coordinate of CR	Length of CR	Coded Protein from CR
13291-TTTTGTGACTTAAAAGGTAAGTATGTACAAATACCTACAACTTGTGCTAATGACCCTGTGGGTTTTACACTTAAAAACACAGTCTGTACCGTCTGCGGTAT-13391	FCDLKGKYVQIPTTCANDPVGFTLKNTVCTVCG	101	187	100	ORF1ab	266	21555	21290	NSP10
5307-TAACACTCCAACAAATAGAGTTGAAGTTTAATCCACCTGCTCTACAAGATGCTTATTACAG-5367	TLQQIELKFNPPALQDAYY	61	113	100	ORF1ab	266	21555	21290	NSP3
9564-ATTCTTACCTGGTGTTTATTCTGTTATTTACTTGTACTTGACATTTTATCTTACTAATGATGTTTCTTTTTTAGCACATATTCAGTGGATGGTT-9657	FLPGVYSVIYLYLTFYLTNDVSFLAHIQWMV	94	174	100	ORF1ab	266	21555	21290	NSP4
5220-ACACTAAAAAGTGGAAATACCCACAAGTTAATGGTTTAACTTCTATTAAATGGGCAGATAACAACTGTT-5288	TKKWKYPQVNGLTSIKWADNNC	69	128	100	ORF1ab	266	21555	21290	NSP3
17812-TTTTGGGACTACCAACTCAAACTGTTGATTCATCACAGGGCTCAGAATATGACTATGTCATATTCACTCAAACCA-17886	FWDYQLKLLIHHRAQNMTMSYSLKP	75	139	100	ORF1ab	266	21555	21290	Helicase
23406-TGTTAACTGCACAGAAGTCCCTGTTGCTATTCATGCAGATCAACTTACTCCTACTTGGCGTGTTTATTCTACAGGTTCTAATGTTTTTCAAACACGTGCAGGCTGTTTAATAGGGGCTGAACATGTCAACAACTCATATGAGTGTGACATACCCATTGGTGCAGGTATATGCGCTAGTTATCAGACTCAGACTAATTCTCCTCGGCGGGCACGTAGTGTAGCTAGT-23631	VNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVAS	226	418	100	Spike	21563	25385	3822	Spike glyco-protein

Consensus Conserved Region (CCnR)	Protein Sequence of CCnR	Length of CCnR	BLAST Specificity Score of CCnR	% of BLAST Specificity Score as Query Coverage	Coding Region (CR)	Starting Coordinate of CR	Ending Coordinate of CR	Length of CR	Coded Protein from CR
9696-TTATATCATTTGTATTTCCACAAAGCATTTCTATTGGTTCTTTAGTAATTACCTAAAGAGACGTGTAGTCTTTAATGGTGTTTCCTTTAGTACTTTTGAAGAAGC-9800	YIICISTKHFYWFFSNYLKRRVVFNGVSFSTFEE	105	195	100	ORF1ab	266	21555	21290	NSP4
29020-AAATCTGCTGCTGAGGCTTCTAAGAAGCCTCGGCAAAAACGTACTGCCACTAAAGCATACAATGTAACACAAGCTTTCGGCAGACGTGGTCCAGAACAAACCCAAGGAAATTTTGGGGACCAGGAACTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTGCACAATTTGCCCCCAGCGCTTCAGCGTTCTT-29219	KSAAEASKKPRQKRTATKAYNVTQAFGRRGPEQTQGNFGDQELIRQGTDYKHWPQIAQFAPSASAF	200	370	100	Nucleocapsid	28274	29534	1260	Nucleocapsid protein
22948-AAACCTTTTGAGAGAGATATTTCAACTGAAATCTATCAGGCCGGTAGCACACCTTGTAATGGTGTTGAAGGTTTTAATTGTTACTTTCCTTTACAATCATATGGTTTCCAACCCACTAATGGTGTTGGTTACCAACC-23084	KPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQ	137	254	100	Spike	21563	25385	3822	Spike glyco-protein
9042-TTCTGGTAAGCCAGTACCATATTGTTATGATACCAATGTACTAGAAGGTTCTGTTGCTTAT-9102	SGKPVPYCYDTNVLEGSVAY	61	113	100	ORF1ab	266	21555	21290	NSP4
16284-TGTGGTGCTTGCATACGTAGACCATTCTTATGTTGTAAATGCTGTTACGACCATGTCATATCAACA-16349	VVLAYVDHSYVVNAVTTMSYQ	66	122	100	ORF1ab	266	21555	21290	Helicase

Consensus Conserved Region (CCnR)	Protein Sequence of CCnR	Length of CCnR	BLAST Specificity Score of CCnR	% of BLAST Specificity Score as Query Coverage	Coding Region (CR)	Starting Coordinate of CR	Ending Coordinate of CR	Length of CR	Coded Protein from CR
6706-ACAAAGTTGTTAGTACAACTACTAACATAGTTACACGGTGTTTAAACCGTGTTTGTACTAATTATATGCCTTATTTCTTTACTTTATTGCTACAATTGTGTACTTTTACTAGAAGTACAAATTCTAGAATTAAA-6839	KVVSTTTNIVTRCLNRVCTNYMPYFFTLLLQLCTFTRSTNSRIK	134	2480	100	ORF1ab	266	21555	21290	NSP3
13041-AGAAGTGCCTGCCAATTCAACTGTATTATCTTTCTGTGCTTTTGCTGTAGATGCTGCTAAAGCTTACAAAGATTATCTAGCTAGTGGGGGACAACCAATCACTAATTGTGTTAAGATGTTGTGTACACACACTGGTACTGGTCAGGCAATAACAGTTACACCGGAAGCCAATATGGATCAAGAATCCTTTGGTGGTGCATCGTGTTGTCTGTACTGCCGTTGCCACATAGATCATCCAAATCCTAAAGG-13289	EVPANSTVLSFCAFAVDAAKAYKDYLASGGQPITNCVKMLCTHTGTGQAITVTPEANMDQESFGGASCCLYCRCHIDHPNPK	249	460	100	ORF1ab	266	21555	21290	NSP10
10094-TTGAGGGTTGTATGGTACAAGTAACTTGTGGTACAACTACACTTAACGGTCTTTGGCTTGATGACGTAGTTTACTGTCCAAGACATGTGATCT-10186	EGCMVQVTCGTTTLNGLWLDDVVYCPRHVI	93	172	100	ORF1ab	266	21555	21290	3CL-Proteinase
584-CTTGTCCCTCATGTGGGCGAAATACCAGTGGCTTACCGCAAGGTTCTTCTTCGTAAGAACGGTAATAAAGGAGC-657	LVPHVGEIPVAYRKVLLRKNGNKG	74	137	100	ORF1ab	266	21555	21290	NSP1
9104-AAAGTTTACGCCCTGACACACGTTATGTGCTCATGGATGGCTCTATTATTCAATTTCCTAACACCTACCTTGAAGGTTCTGTTAGAGTGGTAACAACTTTTGATTCTG-9211	SLRPDTRYVLMDGSIIQFPNTYLEGSVRVVTTFDS	108	200	100	ORF1ab	266	21555	21290	NSP4

Consensus Conserved Region (CCnR)	Protein Sequence of CCnR	Length of CCnR	BLAST Specificity Score of CCnR	% of BLAST Specificity Score as Query Coverage	Coding Region (CR)	Starting Coordinate of CR	Ending Coordinate of CR	Length of CR	Coded Protein from CR
9213-GTACTGTAGGCACGGCACTTGTGAAAGATCAGAAGCTGGTGTTTGTGTATCTACTAGTGGTAGATGGGTACTTAACAATGATTATTACAGATCTTTACCAGGAG-9316	YCRHGTCERSEAGVCVSTSGRWVLNNDYYRSLPG	104	193	100	ORF1ab	266	21555	21290	NSP4
13656-GTTAAGAGACACACTTTCTCTAACTACCAACATGAAGAAACAATTTATAATTTACTTAAGGATTGTCCA-13724	LRDTLSLTTNMKKQFIIYLRIV	69	128	100	ORF1ab	266	21555	21290	RNA-directed RNA polymerase
2371-CTTGAATTTAGGTGAAACATTTGTCACGCACTCAAAGGGATTGTACAGAAAGTGTGTTAAATCCA-2435	LNLGETFVTHSKGLYRKCVKS	65	121	100	ORF1ab	266	21555	21290	NSP2
21661-ACACGTGGTGTTTATTACCCTGACAAAGTTTTCAGATCCTCAGTTTTACATTCAACTCAGGACTTGTT-21728	TRGVYYPDKVFRSSVLHSTQDL	68	126	100	Spike	21563	25385	3822	Spike glycoprotein
2527-GGAAGTTGTCTTGAAAACTGGTGATTTACAACCATTAGAACAACCTACTAGTGAAGCTGTTGAAGC-2592	EVVLKTGDLQPLEQPTSEAVE	66	122	100	ORF1ab	266	21555	21290	NSP2

Table S3: 23 Consensus Conserved Regions (CCnRs) as derived from SARS-CoV-2 with associated details

Consensus Conserved Region (CCnR)	Protein Sequence	Coded Protein	MHC-I restricted T-cell		MHC-II restricted T-cell		B-cell Epitopes		Final Score
			Binding Score	Antigen score	Binding score	Antigen Score	Binding Score	Antigen Score	
4012-CACAGAAAACTTGTTACTTTATATTGACATTAATGGCAATCTTCATCCAGATTCTGCCACTCTTGTTAGTGACATTGACATCACTTTCTTAAAGAAAGATGCTCCATATATAGTGGGTGATGTTGTTCAAGAGGGTGTTTTAACTGCTGTGGTTATACCTACTAAAAAGGCTGGTGGCACTACTGAAATGCTAGCGAAAGCTTT-4215	TENLLLYIDINGNLHPDSATLVSDIDITFLKKDAPYIVGDVVQEGVLTAVVIPTKKAGGTTEMLAKA	NSP3	0.8640	0.7361	0.9804	0.6382	0.8810	1	0.84
10463-TTAAGGGTTCATTCCTTAATGGTTCATGTGGTAGTGTTGGTTTTAACATAGATTATGACTGTGTCTCTTTTTGTTAC-10539	KGSFLNGSCGSVGFNIDYDCVSFCY	3CL-Proteinase	0.6552	0.9049	0.9114	0.7499	0.7143	0.7401	0.77
13291-TTTTGTGACTTAAAAGGTAAGTATGTACAAATACCTACAACTTGTGCTAATGACCCTGTGGGTTTTACACTTAAAAACACAGTCTGTACCGTCTGCGGTAT-13391	FCDLKGKYVQIPTTCANDPVGFTLKNTVCTVCG	NSP10	0.9136	0.7542	0.9818	0.3852	0.9048	0.6813	0.74
5307-TAACACTCCAACAAATAGAGTTGAAGTTTAATCCACCTGCTCTACAAGATGCTTATTACAG-5367	TLQQIELKFNPPALQDAYY	NSP3	0.8106	1	0.9485	0.6714	0.3333	0.8433	0.72
9564-ATTCTTACCTGGTGTTTATTCTGTTATTTACTTGTACTTGACATTTTATCTTACTAATGATGTTTCTTTTTTAGCACATATTCAGTGGATGGTT-9657	FLPGVYSVIYLYLTFYLTNDVSFLAHIQWMV	NSP4	0.9980	0.7866	0.9933	0.3326	0.9762	0.4726	0.70
5220-ACACTAAAAAGTGGAAATACCCACAAGTTAATGGTTTAACTTCTATTAAATGGGCAGATAACAACTGTT-5288	TKKWKYPQVNGLTSIKWADNNC	NSP3	0.8386	0.6718	0.7818	0.4996	0.8095	0.6413	0.70
17812-TTTTGGGACTACCAACTCAAACTGTTGATTCATCACAGGGCTCAGAATATGACTATGTCATATTCACTCAAACCA-17886	FWDYQLKLLIHHRAQNMTMSYSLKP	Helicase	0.8115	0.7035	0.9704	0.4362	0.5476	0.8407	0.69
23406-TGTTAACTGCACAGAAGTCCCTGTTGCTATTCATGCAGATCAACTTACTCCTACTTGGCGTGTTTATTCTACAGGTTCTAATGTTTTTCAAACACGTGCAGGCTGTTTAATAGGGGCTGAACATGTCAACAACTCATATGAGTGTGACATACCCATTGGTGCAGGTATATGCGCTAGTTATCAGACTCAGACTAATTCTCCTCGGCGGGCACGTAGTGTAGCTAGT-23631	VNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVAS	Spike glycoprotein	1	0.4045	0.9559	0.4993	0.9286	0.6167	0.69

Consensus Conserved Region (CCnR)	Protein Sequence	Coded Protein	MHC-I restricted T-cell		MHC-II restricted T-cell		B-cell Epitopes		Final Score
			Binding Score	Antigen score	Binding score	Antigen Score	Binding Score	Antigen Score	
9696-TTATATCATTTGTATTTCCACAAAGCATTTCTATTGGTTCTTTAGTAATTACCTAAAGAGACGTGTAGTCTTTAATGGTGTTTCCTTTAGTACTTTTGAAGAAGC-9800	YIICISTKHFYWFFSNYLKRRVVFNGVSFSTFEE	NSP4	0.9053	0.9879	0.9952	0.5285	0.4524	0.5166	0.69
29020-AAATCTGCTGCTGAGGCTTCTAAGAAGCCTCGGCAAAAACGTACTGCCACTAAAGCATACAATGTAACACAAGCTTTCGGCAGACGTGGTCCAGAACAAACCCAAGGAAATTTTGGGGACCAGGAACTAATCAGACAAGGAACTGATTACAAACATTGGCCGCAAATTGCACAATTTGCCCCCAGCGCTTCAGCGTTCTT-29219	KSAAEASKKPRQKRTATKAYNVTQAFGRRGPEQTQGNFGDQELIRQGTDYKHWPQIAQFAPSASAF	Nucleocapsid protein	0.9950	0.4810	0.9596	0.2835	1	0.8268	0.69
22948-AAACCTTTTGAGAGAGATATTTCAACTGAAATCTATCAGGCCGGTAGCACACCTTGTAATGGTGTTGAAGGTTTTAATTGTTACTTTCCTTTACAATCATATGGTTTCCAACCCACTAATGGTGTTGGTTACCAACC-23084	KPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQ	Spike glycoprotein	0.8334	0.4502	0.9485	0.2781	0.9524	0.7327	0.64
9042-TTCTGGTAAGCCAGTACCATATTGTTATGATACCAATGTACTAGAAGGTTCTGTTGCTTAT-9102	SGKPVPYCYDTNVLEGSVAY	NSP4	0.9744	0.9447	0.8225	0.6486	0.1905	0.6386	0.63
16284-TGTGGTGCTTGCATACGTAGACCATTCTTATGTTGTAAATGCTGTTACGACCATGTCATATCAACA-16349	VVLAYVDHSYVVNAVTTMSYQ	Helicase	0.8803	0.5322	0.9963	0.2949	0.6667	0.5423	0.61
6706-ACAAAGTTGTTAGTACAACTACTAACATAGTTACACGGTGTTTAAACCGTGTTTGTACTAATTATATGCCTTATTTCTTTACTTTATTGCTACAATTGTGTACTTTTACTAGAAGTACAAATTCTAGAATTAAA-6839	KVVSTTTNIVTRCLNRVCTNYMPYFFTLLLQLCTFTRSTNSRIK	NSP3	0.9890	0.6224	1	0.2055	0.7143	0.3847	0.57
13041-AGAAGTGCCTGCCAATTCAACTGTATTATCTTTCTGTGCTTTTGCTGTAGATGCTGCTAAAGCTTACAAAGATTATCTAGCTAGTGGGGGACAACCAATCACTAATTGTGTTAAGATGTTGTGTACACACACTGGTACTGGTCAGGCAATAACAGTTACACCGGAAGCCAATATGGATCAAGAATCCTTTGGTGGTGCATCGTGTTGTCTGTACTGCCGTTGCCACATAGATCATCCAAATCCTAAAGG-13289	EVPANSTVLSFCAFAVDAAKAYKDYLASGGQPITNCVKMLCTHTGTGQAITVTPEANMDQESFGGASCCLYCRCHIDHPNPK	NSP10	0.9983	0.4562	0.9956	0.1602	0.8095	0.5832	0.57

Consensus Conserved Region (CCnR)	Protein Sequence	Coded Protein	MHC-I restricted T-cell		MHC-II restricted T-cell		B-cell Epitopes		Final Score
			Binding Score	Antigen score	Binding score	Antigen Score	Binding Score	Antigen Score	
10094-TTGAGGGTTGTATGGTACAAGTAACTTGTGGTACAACTACACTTAACGGTCTTTGGCTTGATGACGTAGTTTACTGTCCAAGACATGTGATCT-10186	EGCMVQVTCGTTTLNGLWLDDVVYCPRHVI	3CL-Proteinase	0.9124	0.4365	0.9900	0.2052	0.9524	0.4408	0.57
584-CTTGTCCCTCATGTGGGCGAAATACCAGTGGCTTACCGCAAGGTTCTTCTTCGTAAGAACGGTAATAAAGGAGC-657	LVPHVGEIPVAYRKVLLRKNGNKG	NSP1	0.9545	0.5884	0.9596	0.2460	0.7143	0.2965	0.55
9104-AAAGTTTACGCCCTGACACACGTTATGTGCTCATGGATGGCTCTATTATTCAATTTCCTAACACCTACCTTGAAGGTTCTGTTAGAGTGGTAACAACTTTTGATTCTG-9211	SLRPDTRYVLMDGSIIQFPNTYLEGSVRVVTTFDS	NSP4	0.9410	0.3945	0.9633	0.2457	1	0.2862	0.54
9213-GTACTGTAGGCACGGCACTTGTGAAAGATCAGAAGCTGGTGTTTGTGTATCTACTAGTGGTAGATGGGTACTTAACAATGATTATTACAGATCTTTACCAGGAG-9316	YCRHGTCERSEAGVCVSTSGRWVLNNDYYRSLPG	NSP4	0.6165	0.4960	0.9040	0.2563	0.9048	0.3262	0.52
13656-GTTAAGAGACACACTTTCTCTAACTACCAACATGAAGAAACAATTTATAATTTACTTAAGGATTGTCCA-13724	LRDTLSLTTNMKKQFIIYLRIV	RNA-directed RNA polymerase	0.7707	0.6230	0.9485	0.6014	0.1667	0.4223	0.52
2371-CTTGAATTTAGGTGAAACATTTGTCACGCACTCAAAGGGATTGTACAGAAAGTGTGTTAAATCCA-2435	LNLGETFVTHSKGLYRKCVKS	NSP2	0.8932	0.2400	0.8707	0.2500	0.7143	0.3765	0.48
21661-ACACGTGGTGTTTATTACCCTGACAAAGTTTTCAGATCCTCAGTTTTACATTCAACTCAGGACTTGTT-21728	TRGVYYPDKVFRSSVLHSTQDL	Spike glycoprotein	0.7867	0.2226	0.9448	0.3253	0.5476	0.2437	0.44
2527-GGAAGTTGTCTTGAAAACTGGTGATTTACAACCATTAGAACAACCTACTAGTGAAGCTGTTGAAGC-2592	EVVLKTGDLQPLEQPTSEAVE	NSP2	0.6067	0.6804	0.1112	0.6341	0.5714	0.4222	0.44

Table S4: Ranking procedure done on the basis of Geometric Mean of Binding and Antigenic Scores of T-cell and B-cell epitopes from each CCnR

Protein Sequence	Coded Proteins	Type	MHC-I restricted T-cell				MHC-II restricted T-cell				B-cell		
			Epitope	Alleles	Scaled Score of		Epitope	Alleles	Scaled Score of		Epitope	Scaled Score of	
					Immunogenicity	Antigenicity			Immunogenicity	Antigenicity		Immunogenicity	Antigenicity
TENLLLYIDINGNLHPDSATLVSDIDITFLKKDAPYIVGDVVQEGVLTAVVIPTKKAGGTTEMLAKA	NSP3	Immunogenic	FLKKDAPYI	HLA-A*31:01	0.8640	0.3890	ITFLKKDAPYIVGDV	HLA-DRB3*01:01	0.9779	0.5536	TLVSDIDITFLKKDAP	0.8809	0.7313
		Antigenic	TAVVIPTKK	HLA-A*68:01	0.8311	0.7361	IDITFLKKDAPYIVG	HLA-DRB3*01:01	0.9766	0.8882	LHPDSATLVSDIDITF	0.5238	0.7401
KGSFLNGSCGSVGFNIDYDCVSFCY	3CL-Proteinase	Immunogenic	FLNGSCGSV	HLA-A*02:03	0.6552	0.3342	CGSVGFNIDYDCVSF	HLA-DQA1*01:01	0.900375156	1	CGSVGFNIDYDCVSFC	0.7142	1
		Antigenic	GSVGFNIDY	HLA-A*30:02	0.2299	0.9049							
FCDLKGKYVQIPTTCANDPVGFTLKNTVCTVCG	NSP10 glycoprotein	Immunogenic	DLKGKYVQI	HLA-B*08:01	0.9136	0.7542	KGKYVQIPTTCANDP	HLA-DRB1*04:01	0.9818	0.1892	TTCANDPVGFTLKNTV	0.9048	0.6813
		Antigenic					DLKGKYVQIPTTCAN	HLA-DRB1*04:01	0.9815	0.3852			
TLQQIELKFNPPALQDAYY	NSP3	Immunogenic	NPPALQDAY	HLA-B*35:01	0.8106	0.4557	QIELKFNPPALQDAY	HLA-DRB3*02:02	0.9485	0.6409	LQQIELKFNPPALQDA	0.3333	0.8433
		Antigenic	IELKFNPPAL	HLA-B*40:01	0.5146	1	IELKFNPPALQDAYY	HLA-DRB3*02:02	0.9448	0.6714			
FLPGVYSVIYLYLTFYLTNDVSFLAHIQWMV	NSP4	Immunogenic	VSFLAHIQW	HLA-B*57:01	0.9980	0.7866	GVYSVIYLYLTFYLT	HLA-DPA1*01:03/DPB1*02:01	0.9933	0.3326	YSVIYLYLTFYLTNDV	0.9762	0.4726
		Antigenic											
TKKWKYPQVNGLTSIKWADNNC	NSP3	Immunogenic	QVNGLTSIKW	HLA-B*57:01	0.8386	0.6718	PQVNGLTSIKWADNN	HLA-DQA1*01:02/DQB1*06:02	0.7818	0.4667	KYPQVNGLTSIKWADN	0.8095	0.6413
		Antigenic					KYPQVNGLTSIKWAD	HLA-DQA1*01:02/DQB1*06:02	0.7595	0.4996			
FWDYQLKLLIHHRAQNMTMSYSLKP	Helicase	Immunogenic	RAQNMTMSY	HLA-A*30:02	0.8115	0.7035	YQLKLLIHHRAQNMT	HLA-DRB4*01:01	0.9704	0.3655	FWDYQLKLLIHHRAQN	0.5476	0.6379
		Antigenic					DYQLKLLIHHRAQNM	HLA-DRB4*01:02	0.9692	0.4362	IHHRAQNMTMSYSLKP	0.3571	0.8407

Protein Sequence	Coded Proteins	Type	MHC-I restricted T-cell				MHC-II restricted T-cell				B-cell		
			Epitope	Alleles	Scaled Score of		Epitope	Alleles	Scaled Score of		Epitope	Scaled Score of	
					Immunogenicity	Antigenicity			Immunogenicity	Antigenicity		Immunogenicity	Antigenicity
VNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVAS	Spike glycoprotein	Immunogenic	HADQLTPTW	HLA-B*58:01	1	0.4045	DIPIGAGICASYQTQ	HLA-DQA1*05:01/DQB1*03:01	0.9559	0.4993	GCLIGAEHVNNSYECD	0.9286	0.6167
		Antigenic											
YIICISTKHFYWFFSNYLKRRVVFNGVSFSTFEE	NSP4	Immunogenic	ICISTKHFYW	HLA-B*57:01	0.9053	0.9879	KHFYWFFSNYLKRRV	HLA-DPA1*01:03/DPB1*04:01	0.9952	0.4352	ISTKHFYWFFSNYLKR	0.4524	0.5166
		Antigenic					TKHFYWFFSNYLKRR	HLA-DPA1*01:03/DPB1*04:01	0.9937	0.5285			
KSAAEASKKPRQKRTATKAYNVTQAFGRRGPEQTQGNFGDQELIRQGTDYKHWPQIAQFAPSASAF	Nucleocapsid protein	Immunogenic	AQFAPSASAF	HLA-B*15:01	0.9950	0.4810	ATKAYNVTQAFGRR	HLA-DRB5*01:01	0.9596	0.1756	KSAAEASKKPRQKRTA	1	0.4085
		Antigenic					KAYNVTQAFGRRGP	HLA-DRB5*01:01	0.9522	0.2835	GRRGPEQTQGNFGDQE	0.8268	0.8268
KPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQ	Spike glycoprotein	Immunogenic	FERDISTEI	HLA-B*40:01	0.8334	0.0097	VEGFNCYFPLQSYGF	HLA-DQA1*01:01/DQB1*05:01	0.9485	0.1916	GSTPCNGVEGFNCYFP	0.9524	0.2921
		Antigenic	YFPLQSYGF	HLA-A*24:02	0.7813	0.4502	NGVEGFNCYFPLQSY	HLA-DRB3*01:01	0.9263	0.2781	EGFNCYFPLQSYGFQP	0.5238	0.7327
SGKPVPYCYDTNVLEGSVAY	NSP4	Immunogenic	NVLEGSVAY	HLA-B*35:01	0.9744	0.3830	PVPYCYDTNVLEGSV	HLA-DRB1*04:01	0.8225	0.4408	SGKPVPYCYDTNVLEG	0.1904	0.6385
		Antigenic	SGKPVPYCY	HLA-A*30:02	0.6585	0.9447	GKPVPYCYDTNVLEG	HLA-DRB1*04:01	0.7188	0.6486			
VVLAYVDHSYVVNAVTTMSYQ	Helicase protein	Immunogenic	VLAYVDHSY	HLA-B*15:01	0.8803	0.5322	VDHSYVVNAVTTMSY	HLA-DRB3*02:02	0.9958	0.5449	LAYVDHSYVVNAVTTM	0.6667	0.5422
		Antigenic											
KVVSTTTNIVTRCLNRVCTNYMPYFFTLLLQLCTFTRSTNSRIK	NSP3	Immunogenic	NYMPYFFTL	HLA-A*24:02	0.9890	0.6224	CTNYMPYFFTLLLQL	HLA-DPA1*03:01	0.042	0.2206	VCTNYMPYFFTLLLQL	0.7142	0.3847

Protein Sequence	Coded Proteins	Type	MHC-I restricted T-cell				MHC-II restricted T-cell				B-cell		
			Epitope	Alleles	Scaled Score of		Epitope	Alleles	Scaled Score of		Epitope	Scaled Score of	
					Immunogenicity	Antigenicity			Immunogenicity	Antigenicity		Immunogenicity	Antigenicity
		Antigenic											
EVPANSTVLSFCAFAVDAAKAYKDYLASGGQPITNCVKMLCTHTGTGQAITVTPEANMDQESFGGASCCLYCRCHIDHPNPK	NSP10	Immunogenic	FAVDAAKAY	HLA-B*35:01	0.9983	0.3296	LSFCAFAVDAAKAYK	HLA-DRB3*01:01	0.9949	0.4102	GTGQAITVTPEANMDQ	0.8095	0.3635
		Antigenic	VPANSTVLSF	HLA-B*35:01	0.7694	0.4562					KMLCTHTGTGQAITVT	0.6905	0.5832
EGCMVQVTCGTTTLNGLWLDDVVYCPRHVI	3CL-Proteinase	Immunogenic	GTTTLNGLW	HLA-B*57:01	0.9124	0.4365	TTTLNGLWLDDVVYC	HLA-DQA1*01:01/DQB1*05:01	0.9900	0.0948	QVTCGTTTLNGLWLDD	0.9523	0.4408
		Antigenic					TLNGLWLDDVVYCPR	HLA-DQA1*01:01/DQB1*05:01	0.9878	0.2052			
LVPHVGEIPVAYRKVLLRKNGNKG	NSP1	Immunogenic	HVGEIPVAY	HLA-B*15:01	0.9545	0.4960	VAYRKVLLRKNGNKG	HLA-DRB1*11:01	0.9596	0.1695	PHVGEIPVAYRKVLLR	0.7142	0.2965
		Antigenic	HVGEIPVAYR	HLA-A*68:01	0.9445	0.5884	IPVAYRKVLLRKNGN	HLA-DRB1*11:01	0.9226	0.2460			
SLRPDTRYVLMDGSIIQFPNTYLEGSVRVVTTFDS	NSP4	Immunogenic	RPDTRYVLM	HLA-B*07:02	0.9410	0.3945	LMDGSIIQFPNTYLE	HLA-DRB1*15:01	0.9587	0.2456	GSIIQFPNTYLEGSVR	1	0.1195
		Antigenic									LRPDTRYVLMDGSIIQ	0.3334	0.2861
YCRHGTCERSEAGVCVSTSGRWVLNNDYYRSLPG	NSP4	Immunogenic	VCVSTSGRW	HLA-B*57:01	0.6165	0.4960	TSGRWVLNNDYYRSL	HLA-DRB3*02:02	0.9040	0.2219	YCRHGTCERSEAGVCV	0.9048	0.3262
		Antigenic					STSGRWVLNNDYYRS	HLA-DRB3*02:02	0.8225	0.2563			
LRDTLSLTTNMKKQFIIYLRIV	RNA-directed RNA polymerase	Immunogenic	DTLSLTTNMK	HLA-A*68:01	0.7707	0.5758	TTNMKKQFIIYLRIV	HLA-DPA1*02:01	0.942	0.6014	LNLGETFVTHSKGLYR	0.7142	0.3765
		Antigenic	LSLTTNMKK	HLA-A*11:01	0.5274	0.6230							
LNLGETFVTHSKGLYRKCVKS	NSP2	Immunogenic	VTHSKGLYR	HLA-A*31:01	0.8932	0.2331	ETFVTHSKGLYRKCV	HLA-DRB5*01:01	0.8707	0.0966	LNLGETFVTHSKGLYR	0.7142	0.3765
		Antigenic	VTHSKGLYRK	HLA-A*03:01	0.699	0.24	LGETFVTHSKGLYRK	HLA-DRB5*01:01	0.8522	0.25			
TRGVYYPDKVFRSSVLHSTQDL	Spike glycoprotein	Immunogenic	VYYPDKVFR	HLA-A*31:01	0.7867	0.2018	TRGVYYPDKVFRSSV	HLA-DRB1*03:01	0.9378	0.3252	RGVYYPDKVFRSSVLH	0.5476	0.2437
		Antigenic	GVYYPDKVFR	HLA-A*31:01	0.6607	0.2226							

Protein Sequence	Coded Proteins	Type	MHC-I restricted T-cell				MHC-II restricted T-cell				B-cell		
			Epitope	Alleles	Scaled Score of		Epitope	Alleles	Scaled Score of		Epitope	Scaled Score of	
					Immunogenicity	Antigenicity			Immunogenicity	Antigenicity		Immunogenicity	Antigenicity
EVVLKTGDLQPLEQPTSEAVE	NSP2	Immunogenic	LEQPTSEAV	HLA-B*40:01	0.6067	0.1310	GDLQPLEQPTSEAVE	HLA-DQA1*03:01/DQB1*03:02	0.1112	0.4601	TGDLQPLEQPTSEAVE	0.5714	0.4221
		Antigenic	EVVLKTGDL	HLA-A*26:01	0	0.6804	EVVLKTGDLQPLEQP	HLA-DRB1*08:02	0.0371	0.6341			

Table S5: List of Immunogenic and Antigenic Epitopes for MHC-I, MHC-II restricted T-cell and B-cell Epitopes

Particulars	Link
Details related to 125 CCnRs, 92 MHC-I and MHC-II restricted T-cell epitopes and 61 B-cell epitopes	http://www.nitttrkol.ac.in/indrajit/projects/COVID-EpitopeVaccine-India/downloads/supplementary/Epitope-Details.xlsx
List of proposed epitopes from literature	http://www.nitttrkol.ac.in/indrajit/projects/COVID-EpitopeVaccine-India/downloads/supplementary/AllPaperSummary.xlsx

Table S6: Link of Supplementary excel files

MHC-I restricted T-cell Epitopes	Positively charged	Negatively charged	Polarity	Non Polarity	Aliphaticity	Aromaticity	Acidicity	Basicity	Hydrophobicity	Hydrophilicity	Neutral	Hydroxylic	Sulphur Content
FLKKDAPYI	0.222	0.111	0.111	0.556	0.444	0.222	0.111	0.222	0.556	0.333	0.111	0	0
TAVVIPTKK	0.222	0	0.222	0.556	0.556	0	0	0.222	0.778	0.333	0.222	0.222	0
FLNGSCGSV	0	0	0.333	0.556	0.444	0.111	0	0	0.444	0.111	0.444	0.222	0.111
GSVGFNIDY	0	0.111	0.222	0.556	0.444	0.222	0.111	0	0.333	0.111	0.444	0.111	0
DLKGKYVQI	0.222	0.111	0.222	0.444	0.444	0.111	0.111	0.222	0.333	0.222	0.333	0	0
NPPALQDAY	0	0.111	0.222	0.556	0.556	0.111	0.111	0	0.556	0.333	0.222	0	0
IELKFNPPAL	0.1	0.1	0	0.7	0.6	0.1	0.1	0.1	0.7	0.4	0.1	0	0
VSFLAHIQW	0.111	0	0.222	0.667	0.444	0.222	0	0.111	0.667	0.111	0.222	0.111	0
QVNGLTSIKW	0.1	0	0.3	0.5	0.4	0.1	0	0.1	0.5	0.2	0.4	0.2	0
RAQNMTMSY	0.111	0	0.444	0.333	0.111	0.111	0	0.111	0.444	0.222	0.333	0.222	0.222
HADQLTPTW	0.111	0.111	0.333	0.444	0.333	0.111	0.111	0.111	0.667	0.222	0.444	0.222	0
ICISTKHFYW	0.2	0	0.4	0.4	0.2	0.3	0	0.2	0.6	0.2	0.2	0.2	0.1
AQFAPSASAF	0	0	0.3	0.7	0.5	0.2	0	0	0.7	0.1	0.3	0.2	0
FERDISTEI	0.111	0.333	0.222	0.333	0.222	0.111	0.333	0.111	0.444	0.111	0.556	0.222	0
YFPLQSYGF	0	0	0.444	0.556	0.333	0.444	0	0	0.444	0.111	0.333	0.111	0
NVLEGSVAY	0	0.111	0.222	0.556	0.556	0.111	0.111	0	0.444	0.111	0.333	0.111	0
SGKPVPYCY	0.111	0	0.444	0.444	0.444	0.222	0	0.111	0.444	0.333	0.222	0.111	0.111
VLAYVDHSY	0.111	0.111	0.333	0.444	0.444	0.222	0.111	0.111	0.444	0.111	0.222	0.111	0
NYMPYFFTL	0	0	0.333	0.556	0.222	0.444	0	0	0.667	0.222	0.111	0.111	0.111
FAVDAAKAY	0.111	0.111	0.111	0.667	0.556	0.222	0.111	0.111	0.667	0.111	0.111	0	0
VPANSTVLSF	0	0	0.3	0.6	0.5	0.1	0	0	0.7	0.2	0.3	0.3	0
GTTTLNGLW	0	0	0.333	0.556	0.444	0.111	0	0	0.667	0.111	0.556	0.333	0
HVGEIPVAY	0.111	0.111	0.111	0.667	0.667	0.111	0.111	0.111	0.556	0.222	0.222	0	0
HVGEIPVAYR	0.2	0.1	0.1	0.6	0.6	0.1	0.1	0.2	0.5	0.3	0.2	0	0
RPDTRYVLM	0.222	0.111	0.222	0.444	0.333	0.111	0.111	0.222	0.556	0.333	0.222	0.111	0.111
VCVSTSGRW	0.111	0	0.444	0.444	0.333	0.111	0	0.111	0.556	0.111	0.444	0.333	0.111
DTLSLTTNMK	0.1	0.1	0.4	0.3	0.2	0	0.1	0.1	0.6	0.2	0.5	0.4	0.1
LSLTTNMKK	0.222	0	0.333	0.333	0.222	0	0	0.222	0.556	0.333	0.333	0.333	0.111
VTHSKGLYR	0.333	0	0.333	0.333	0.333	0.111	0	0.333	0.333	0.333	0.333	0.222	0
VTHSKGLYRK	0.4	0	0.3	0.3	0.3	0.1	0	0.4	0.3	0.4	0.3	0.2	0
VYYPDKVFR	0.222	0.111	0.222	0.444	0.333	0.333	0.111	0.222	0.444	0.333	0.111	0	0
GVYYPDKVFR	0.2	0.1	0.2	0.5	0.4	0.3	0.1	0.2	0.4	0.3	0.2	0	0
LEQPTSEAV	0	0.222	0.333	0.444	0.444	0	0.222	0	0.556	0.111	0.556	0.222	0
EVVLKTGDL	0.111	0.222	0.111	0.556	0.556	0	0.222	0.111	0.556	0.111	0.444	0.111	0

Table S7: List of Physico-chemical properties of MHC-I restricted T-cell epitopes

MHC-I restricted T-cell Epitopes	Positively charged	Negatively charged	Polarity	Non Polarity	Aliphaticity	Aromaticity	Acidicity	Basicity	Hydrophobicity	Hydrophilicity	Neutral	Hydroxylic	Sulphur Content
ITFLKKDAPYIVGDV	0.133	0.133	0.133	0.6	0.533	0.133	0.133	0.133	0.6	0.2	0.267	0.067	0
IDITFLKKDAPYIVG	0.133	0.133	0.133	0.6	0.533	0.133	0.133	0.133	0.6	0.2	0.267	0.067	0
CGSVGFNIDYDCVSF	0	0.133	0.333	0.467	0.333	0.2	0.133	0	0.467	0.067	0.4	0.133	0.133
KGKYVQIPTTCANDP	0.133	0.067	0.333	0.4	0.4	0.067	0.067	0.133	0.533	0.333	0.333	0.133	0.067
DLKGKYVQIPTTCAN	0.133	0.067	0.333	0.4	0.4	0.067	0.067	0.133	0.533	0.267	0.333	0.133	0.067
QIELKFNPPALQDAY	0.067	0.133	0.2	0.533	0.467	0.133	0.133	0.067	0.533	0.267	0.267	0	0
IELKFNPPALQDAYY	0.067	0.133	0.2	0.533	0.467	0.2	0.133	0.067	0.533	0.267	0.2	0	0
GVYSVIYLYLTFYLT	0	0	0.467	0.533	0.467	0.333	0	0	0.6	0	0.267	0.2	0
PQVNGLTSIKWADNN	0.067	0.067	0.2	0.467	0.4	0.067	0.067	0.067	0.467	0.333	0.333	0.133	0
KYPQVNGLTSIKWAD	0.133	0.067	0.267	0.467	0.4	0.133	0.067	0.133	0.467	0.267	0.333	0.133	0
YQLKLLIHHRAQNMT	0.267	0	0.267	0.4	0.333	0.067	0	0.267	0.467	0.333	0.2	0.067	0.067
DYQLKLLIHHRAQNM	0.267	0.067	0.2	0.4	0.333	0.067	0.067	0.267	0.4	0.333	0.2	0	0.067
DIPIGAGICASYQTQ	0	0.067	0.4	0.533	0.533	0.067	0.067	0	0.533	0.067	0.467	0.133	0.067
KHFYWFFSNYLKRRV	0.333	0	0.2	0.4	0.133	0.4	0	0.333	0.4	0.4	0.067	0.067	0
TKHFYWFFSNYLKRR	0.333	0	0.267	0.333	0.067	0.4	0	0.333	0.4	0.4	0.133	0.133	0
ATKAYNVTQAFGRR	0.214	0	0.286	0.429	0.357	0.143	0	0.214	0.5	0.286	0.286	0.143	0
KAYNVTQAFGRRGP	0.214	0	0.214	0.5	0.429	0.143	0	0.214	0.429	0.357	0.286	0.071	0
VEGFNCYFPLQSYGF	0	0.067	0.333	0.533	0.333	0.333	0.067	0	0.467	0.133	0.333	0.067	0.067
NGVEGFNCYFPLQSY	0	0.067	0.333	0.467	0.333	0.267	0.067	0	0.4	0.2	0.333	0.067	0.067
PVPYCYDTNVLEGSV	0	0.133	0.333	0.467	0.467	0.133	0.133	0	0.533	0.2	0.333	0.133	0.067
GKPVPYCYDTNVLEG	0.067	0.133	0.267	0.467	0.467	0.133	0.133	0.067	0.467	0.267	0.333	0.067	0.067
VDHSYVVNAVTTMSY	0.067	0.067	0.4	0.4	0.333	0.133	0.067	0.067	0.533	0.133	0.333	0.267	0.067
CTNYMPYFFTLLLQL	0	0	0.4	0.533	0.333	0.267	0	0	0.733	0.133	0.2	0.133	0.133
LSFCAFAVDAAKAYK	0.133	0.067	0.2	0.6	0.467	0.2	0.067	0.133	0.667	0.133	0.133	0.067	0.067
TTTLNGLWLDDVVYC	0	0.133	0.333	0.467	0.4	0.133	0.133	0	0.667	0.067	0.4	0.2	0.067
TLNGLWLDDVVYCPR	0.067	0.133	0.2	0.533	0.467	0.133	0.133	0.067	0.6	0.2	0.267	0.067	0.067
VAYRKVLLRKNGNKG	0.333	0	0.067	0.467	0.467	0.067	0	0.333	0.333	0.467	0.133	0	0
IPVAYRKVLLRKNGN	0.267	0	0.067	0.533	0.533	0.067	0	0.267	0.467	0.467	0.067	0	0
LMDGSIIQFPNTYLE	0	0.133	0.267	0.533	0.4	0.133	0.133	0	0.533	0.133	0.4	0.133	0.067
TSGRWVLNNDYYRSL	0.133	0.067	0.333	0.333	0.267	0.2	0.067	0.133	0.333	0.267	0.333	0.2	0
STSGRWVLNNDYYRS	0.133	0.067	0.4	0.267	0.2	0.2	0.067	0.133	0.267	0.267	0.4	0.267	0
TTNMKKQFIIYLRIV	0.2	0	0.267	0.467	0.333	0.133	0	0.2	0.6	0.267	0.2	0.133	0.067
ETFVTHSKGLYRKCV	0.267	0.067	0.333	0.333	0.267	0.133	0.067	0.267	0.467	0.267	0.333	0.2	0.067
LGETFVTHSKGLYRK	0.267	0.067	0.267	0.4	0.333	0.133	0.067	0.267	0.4	0.267	0.4	0.2	0
TRGVYYYPDKVFRSSV	0.2	0.067	0.333	0.4	0.333	0.2	0.067	0.2	0.4	0.267	0.333	0.2	0
GDLQPLEQPTSEAVE	0	0.267	0.267	0.467	0.467	0	0.267	0	0.467	0.133	0.6	0.133	0
EVVLKTGDLQPLEQP	0.067	0.2	0.2	0.533	0.533	0	0.2	0.067	0.533	0.2	0.467	0.067	0

Table S8: List of Physico-chemical properties of MHC-II restricted T-cell epitopes

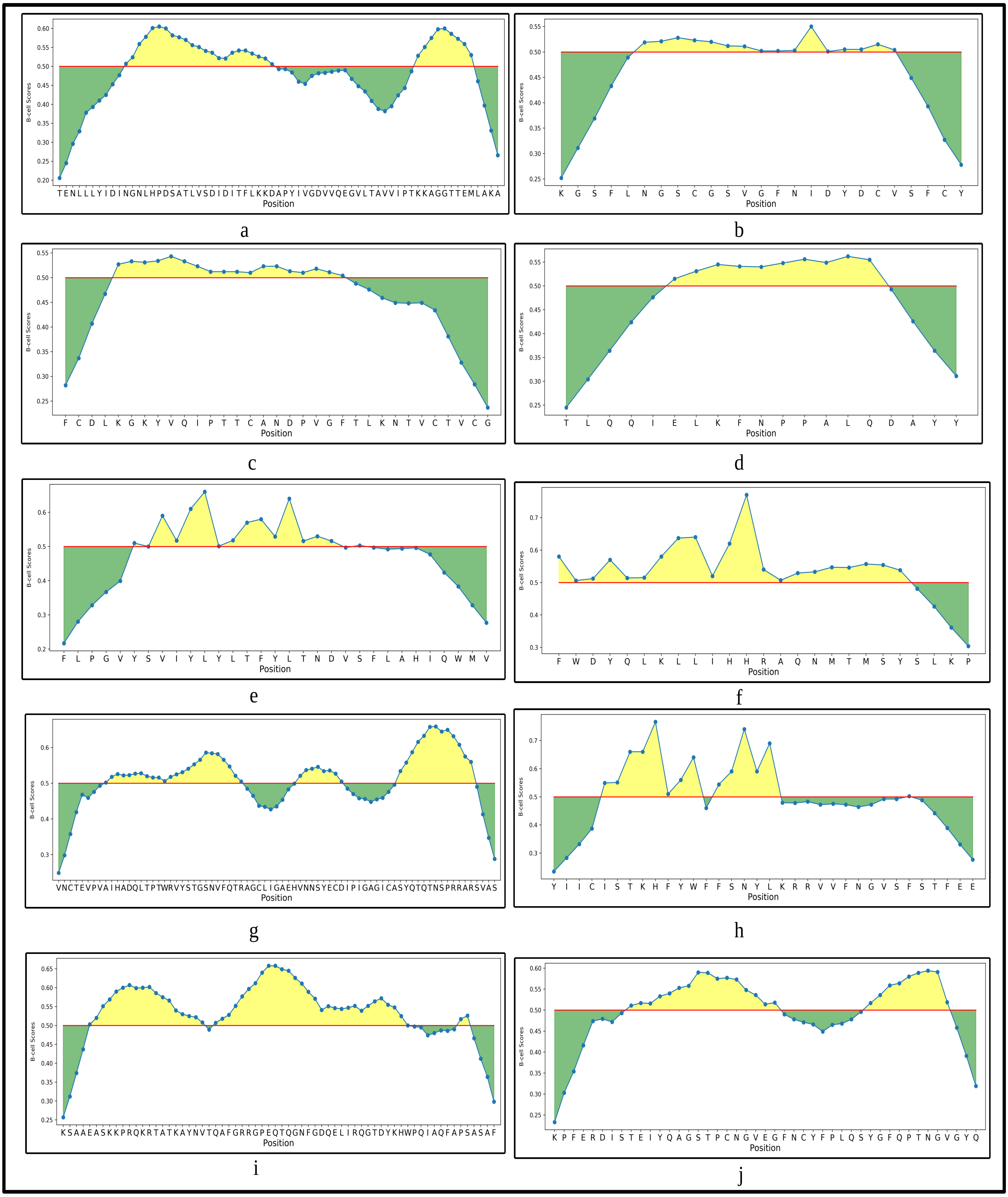
B-cell T-cell Epitopes	Positively charged	Negatively charged	Polarity	Non Polarity	Aliphaticity	Aromaticity	Acidicity	Basicity	Hydrophobicity	Hydrophilicity	Neutral	Hydroxylic	Sulphur Content
TLVSDIDITFLKKDAP	0.125	0.188	0.188	0.5	0.438	0.062	0.188	0.125	0.625	0.188	0.375	0.188	0
LHPDSATLVSDIDITF	0.062	0.188	0.25	0.5	0.438	0.062	0.188	0.062	0.625	0.125	0.438	0.25	0
CGSVGFNIDYDCVSFC	0	0.125	0.375	0.438	0.312	0.188	0.125	0	0.5	0.062	0.375	0.125	0.188
TTCANDPVGFTLKNTV	0.062	0.062	0.312	0.438	0.375	0.062	0.062	0.062	0.688	0.25	0.375	0.25	0.062
LQQIELKFNPPALQDA	0.062	0.125	0.188	0.562	0.5	0.062	0.125	0.062	0.562	0.25	0.312	0	0
YSVIYLYLTFYLTNDV	0	0.062	0.438	0.438	0.375	0.312	0.062	0	0.562	0.062	0.25	0.188	0
KYPQVNGLTSIKWADN	0.125	0.062	0.25	0.438	0.375	0.125	0.062	0.125	0.438	0.312	0.312	0.125	0
FWDYQLKLLIHHRAQN	0.25	0.062	0.188	0.438	0.312	0.188	0.062	0.25	0.438	0.312	0.188	0	0
IHHRAQNMTMSYSLKP	0.25	0	0.312	0.375	0.25	0.062	0	0.25	0.438	0.375	0.25	0.188	0.125
GCLIGAEHVNNSYECD	0.062	0.188	0.25	0.375	0.375	0.062	0.188	0.062	0.375	0.188	0.375	0.062	0.125
ISTKHFYWFFSNYLKR	0.25	0	0.312	0.375	0.125	0.375	0	0.25	0.438	0.312	0.188	0.188	0
KSAAEASKKPRQKRTA	0.375	0.062	0.25	0.312	0.312	0	0.062	0.375	0.375	0.438	0.312	0.188	0
GRRGPEQTQGNFGDQE	0.125	0.188	0.25	0.375	0.312	0.062	0.188	0.125	0.188	0.25	0.688	0.062	0
GSTPCNGVEGFNCYFP	0	0.062	0.312	0.5	0.375	0.188	0.062	0	0.5	0.25	0.375	0.125	0.125
EGFNCYFPLQSYGFQP	0	0.062	0.375	0.5	0.312	0.312	0.062	0	0.438	0.188	0.375	0.062	0.062
SGKPVPYCYDTNVLEG	0.062	0.125	0.312	0.438	0.438	0.125	0.125	0.062	0.438	0.25	0.375	0.125	0.062
LAYVDHSYVVNAVTTM	0.062	0.062	0.312	0.5	0.438	0.125	0.062	0.062	0.625	0.125	0.25	0.188	0.062
VCTNYMPYFFTLLLQL	0	0	0.375	0.562	0.375	0.25	0	0	0.75	0.125	0.188	0.125	0.125
GTGQAITVTPEANMDQ	0	0.125	0.312	0.5	0.438	0	0.125	0	0.562	0.125	0.562	0.188	0.062
KMLCTHTGTGQAITVT	0.125	0	0.438	0.438	0.375	0	0	0.125	0.688	0.125	0.5	0.312	0.125
QVTCGTTTLNGLWLDD	0	0.125	0.375	0.438	0.375	0.062	0.125	0	0.625	0.062	0.562	0.25	0.062
PHVGEIPVAYRKVLLR	0.25	0.062	0.062	0.625	0.625	0.062	0.062	0.25	0.562	0.375	0.125	0	0
GSIIQFPNTYLEGSVR	0.062	0.062	0.312	0.5	0.438	0.125	0.062	0.062	0.438	0.188	0.438	0.188	0
LRPDTRYVLMDGSIIQ	0.125	0.125	0.25	0.5	0.438	0.062	0.125	0.125	0.5	0.188	0.375	0.125	0.062
YCRHGTCERSEAGVCV	0.188	0.125	0.375	0.312	0.312	0.062	0.125	0.188	0.438	0.188	0.375	0.125	0.188
LRDTLSLTTNMKKQFI	0.188	0.062	0.312	0.375	0.25	0.062	0.062	0.188	0.562	0.25	0.375	0.25	0.062
LNLGETFVTHSKGLYR	0.188	0.062	0.25	0.438	0.375	0.125	0.062	0.188	0.438	0.25	0.375	0.188	0
RGVYYPDKVFRSSVLH	0.25	0.062	0.25	0.438	0.375	0.188	0.062	0.25	0.375	0.312	0.25	0.125	0
TGDLQPLEQPTSEAVE	0	0.25	0.312	0.438	0.438	0	0.25	0	0.5	0.125	0.625	0.188	0

Table S9: List of Physico-chemical properties of B-cell epitopes

MHC-I restricted T-cell epitopes	Score from Autodock Vina	Z Score	MHC-II restricted T-cell epitopes	Score from Autodock Vina	Z Score
FLKKDAPYI	-8.2	-9.81	ITFLKKDAPYIVGDV	-9	-5.2
TAVVIPTKK	-8.1	-5.9	IDITFLKKDAPYIVG	-8.8	-5.59
FLNGSCGSV	Not Generated	Not Generated	CGSVGFNIDYDCVSF	Not Generated	Not Generated
GSVGFNIDY	-7.1	-5.4			
DLKGKYVQI	-8.1	-8.81	KGKYVQIPTTCANDP	Not Generated	Not Generated
			DLKGKYVQIPTTCAN	Not Generated	Not Generated
NPPALQDAY	Not Generated	Not Generated	QIELKFNPPALQDAY	Not Generated	Not Generated
IELKFNPPAL	Not Generated	Not Generated	IELKFNPPALQDAYY	Not Generated	Not Generated
VSFLAHIQW	-8.8	-9.26	GVYSVIYLYLTFYLT	-8	-5.02
QVNGLTSIKW	Not Generated	Not Generated	PQVNGLTSIKWADNN	-8.3	-4.01
			KYPQVNGLTSIKWAD	Not Generated	Not Generated
RAQNMTMSY	Not Generated	Not Generated	YQLKLLIHHRAQNMT	Not Generated	Not Generated
			DYQLKLLIHHRAQNM	Not Generated	Not Generated

MHC-I restricted T-cell epitopes	Score from Autodock Vina	Z Score	MHC-II restricted T-cell epitopes	Score from Autodock Vina	Z Score
HADQLTPTW	-8.2	-0.6	DIPIGAGICASYQTQ	Not Generated	Not Generated
ICISTKHFYW	Not Generated	Not Generated	KHFYWFFSNYLKRRV	Not Generated	Not Generated
			TKHFYWFFSNYLKRR	Not Generated	Not Generated
AQFAPSASAF	Not Generated	Not Generated	ATKAYNVTQAFGRR	-8.1	0.64
			KAYNVTQAFGRRGP	-8.8	-5.05
FERDISTEI	Not Generated	Not Generated	VEGFNCYFPLQSYGF	Not Generated	Not Generated
YFPLQSYGF	Not Generated	Not Generated	NGVEGFNCYFPLQSY	Not Generated	Not Generated
NVLEGSVAY	-8	-5.49	PVPYCYDTNVLEGSV	-8.6	-5.4
SGKPVPYCY	-6.9	-5.49	GKPVPYCYDTNVLEG	-9.6	-5.31
VLAYVDHSY	Not Generated	Not Generated	VDHSYVVNAVTTMSY	Not Generated	Not Generated
NYMPYFFTL	-8.3	-0.03	CTNYMPYFFTLLLQL	-8.2	-5.63
FAVDAAKAY	Not Generated	Not Generated	LSFCAFAVDAAKAYK	Not Generated	Not Generated
VPANSTVLSF	Not Generated	Not Generated			
GTTTLNGLW	-8.8	-0.66	TTTLNGLWLDDVVYC	Not Generated	Not Generated
			TLNGLWLDDVVYCPR	Not Generated	Not Generated
HVGEIPVAY	Not Generated	Not Generated	VAYRKVLLRKNGNKG	Not Generated	Not Generated
HVGEIPVAYR	Not Generated	Not Generated	IPVAYRKVLLRKNGN	Not Generated	Not Generated
RPDTRYVLM	-8.1	-5.61	LMDGSIIQFPNTYLE	Not Generated	Not Generated
VCVSTSGRW	Not Generated	Not Generated	TSGRWVLNNDYYRSL	Not Generated	Not Generated
	Not Generated	Not Generated	STSGRWVLNNDYYRS	Not Generated	Not Generated
DTLSLTTNMK	Not Generated	Not Generated	TTNMKKQFIIYLRIV	Not Generated	Not Generated
LSLTTNMKK	Not Generated	Not Generated		Not Generated	Not Generated
VTHSKGLYR	Not Generated	Not Generated	ETFVTHSKGLYRKCV	Not Generated	Not Generated
VTHSKGLYRK	-8.4	-8.84	LGETFVTHSKGLYRK	Not Generated	Not Generated
VYYPDKVFR	Not Generated	Not Generated	TRGVYYYPDKVFRSSV	Not Generated	Not Generated
GVYYPDKVFR	Not Generated	Not Generated		Not Generated	Not Generated
LEQPTSEAV	Not Generated	Not Generated	GDLQPLEQPTSEAVE	Not Generated	Not Generated
EVVLKTGDL	Not Generated	Not Generated	EVVLKTGDLQPLEQP	Not Generated	Not Generated

Table S10: Docking and Z-scores of MHC-I and MHC-II restricted T-cell epitopes for 23 CCnRs


B-cell Scores
TENLLLYIDINGNLHPDSATLVSDIDITFLKKDAPYIVGDVVQEGVLTAVVIPTKKAGGTTEMLAKA
Position
B-cell Scores
K G S F L N G S C G S V G F N I D Y D C V S F C Y
Position
B-cell Scores
F C D L K G K Y V Q I P T T C A N D P V G F T L K N T V C T V C G
Position
B-cell Scores
T L Q Q I E L K F N P P A L Q D A Y Y
Position
B-cell Scores
F L P G V Y S V I Y L Y L T F Y L T N D V S F L A H I Q W M V
Position
B-cell Scores
F W D Y Q L K L L I H H R A Q N M T M S Y S L K P
Position
B-cell Scores
VNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVAS
Position
B-cell Scores
Y I I C I S T K H F Y W F F S N Y L K R R V V F N G V S F S T F E E
Position
B-cell Scores
KSAAEASKKPRQKRTATKAYNVTQAFGRRGPEQTQGNFGDQELIRQGTDYKHWPQIAQFAPSASAF
Position
B-cell Scores
K P F E R D I S T E I Y Q A G S T P C N G V E G F N C Y F P L Q S Y G F Q P T N G V G Y Q
Position

a

b

c

d

e

f

g

h

i

j

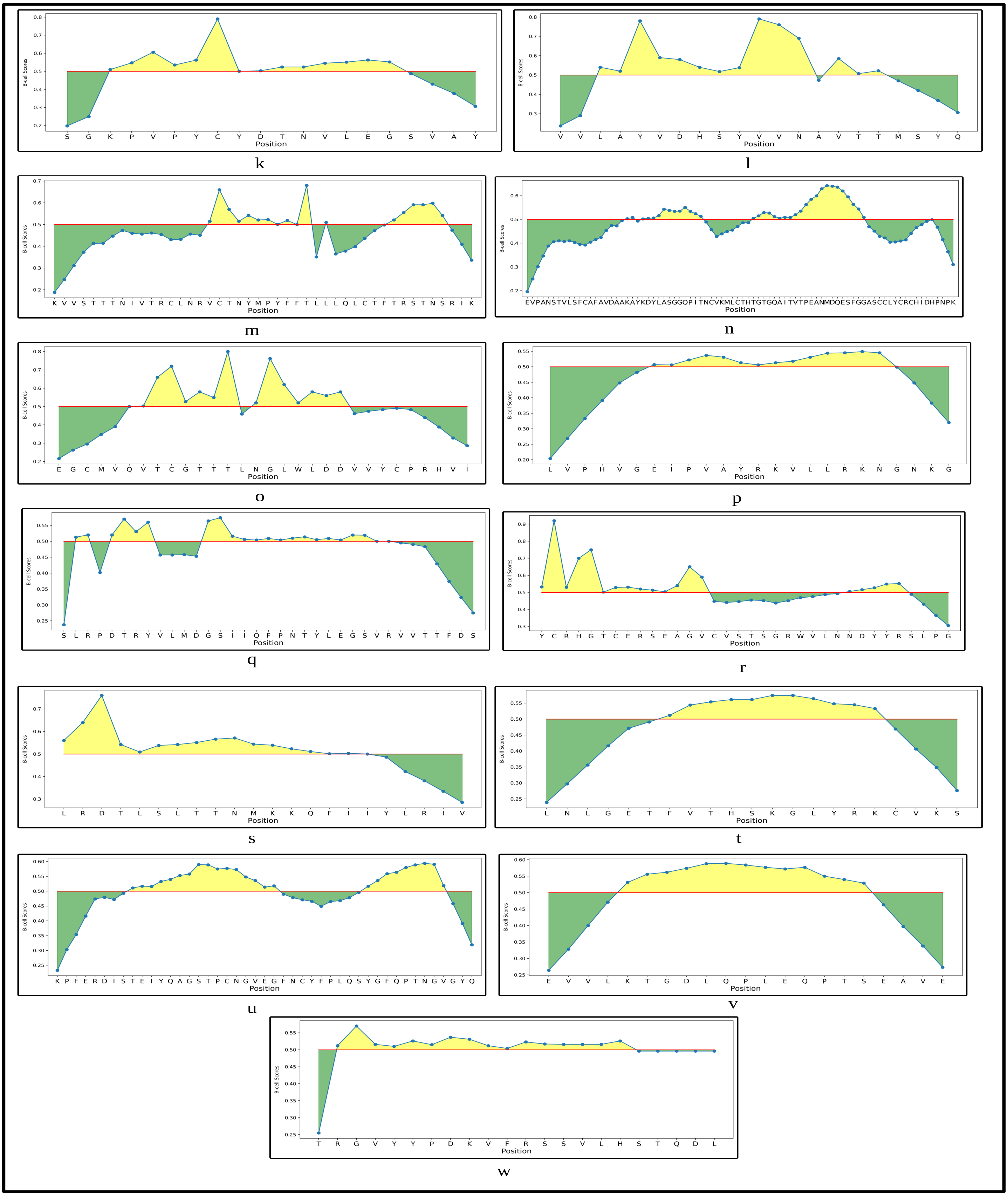
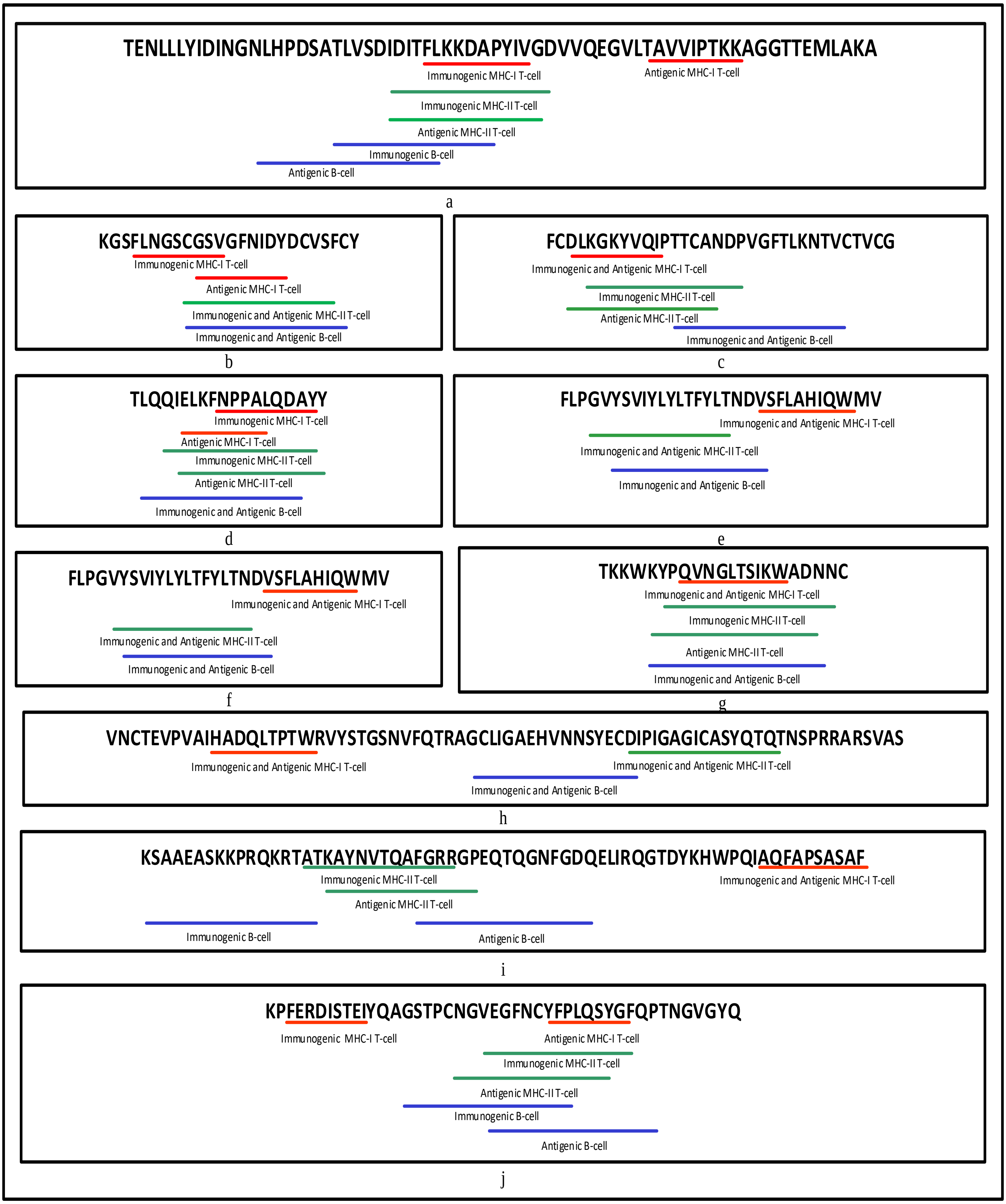


Figure S1: Graphical representation of B-cell epitopes for 23 CCnRS


TENLLLYIDINGNLHPDSATLVSDIDITFLKKDAPYIVGDVVQEGVLTAVVIPTKKAGGTTEMLAKA
Immunogenic MHC-I T-cell
Antigenic MHC-I T-cell
Immunogenic MHC-II T-cell
Antigenic MHC-II T-cell
Immunogenic B-cell
Antigenic B-cell
a
KGSFLNGSCGSVGFNIDYDCVSFCY
Immunogenic MHC-I T-cell
Antigenic MHC-I T-cell
Immunogenic and Antigenic MHC-II T-cell
Immunogenic and Antigenic B-cell
b
FCDLKGKYVQIPTTCANDPVGFTLKNTVCTVCG
Immunogenic and Antigenic MHC-I T-cell
Immunogenic MHC-II T-cell
Antigenic MHC-II T-cell
Immunogenic and Antigenic B-cell
c
TLQQIELKFNPPALQDAYY
Immunogenic MHC-I T-cell
Antigenic MHC-I T-cell
Immunogenic MHC-II T-cell
Antigenic MHC-II T-cell
Immunogenic and Antigenic B-cell
d
FLPGVYSVIYLYLTFYLTNDVSFLAHIQWMV
Immunogenic and Antigenic MHC-I T-cell
Immunogenic and Antigenic MHC-II T-cell
Immunogenic and Antigenic B-cell
e
FLPGVYSVIYLYLTFYLTNDVSFLAHIQWMV
Immunogenic and Antigenic MHC-I T-cell
Immunogenic and Antigenic MHC-II T-cell
Immunogenic and Antigenic B-cell
f
TKKWKYPQVNGLTSIKWADNNC
Immunogenic and Antigenic MHC-I T-cell
Immunogenic MHC-II T-cell
Antigenic MHC-II T-cell
Immunogenic and Antigenic B-cell
g
VNCTEVPVAIHADQLTPTWRVYSTGSNVFQTRAGCLIGAEHVNNSYECDIPIGAGICASYQTQTNSPRRARSVAS
Immunogenic and Antigenic MHC-I T-cell
Immunogenic and Antigenic MHC-II T-cell
Immunogenic and Antigenic B-cell
h
KSAAEASKKPRQKRTATKAYNVTQAFGRRGPEQTQGNFGDQELIRQGTDYKHWPQIAQFAPSASAF
Immunogenic MHC-II T-cell
Immunogenic and Antigenic MHC-I T-cell
Antigenic MHC-II T-cell
Immunogenic B-cell
Antigenic B-cell
i
KPFERDISTEIYQAGSTPCNGVEGFNCYFPLQSYGFQPTNGVGYQ
Immunogenic MHC-I T-cell
Antigenic MHC-I T-cell
Immunogenic MHC-II T-cell
Antigenic MHC-II T-cell
Immunogenic B-cell
Antigenic B-cell
j

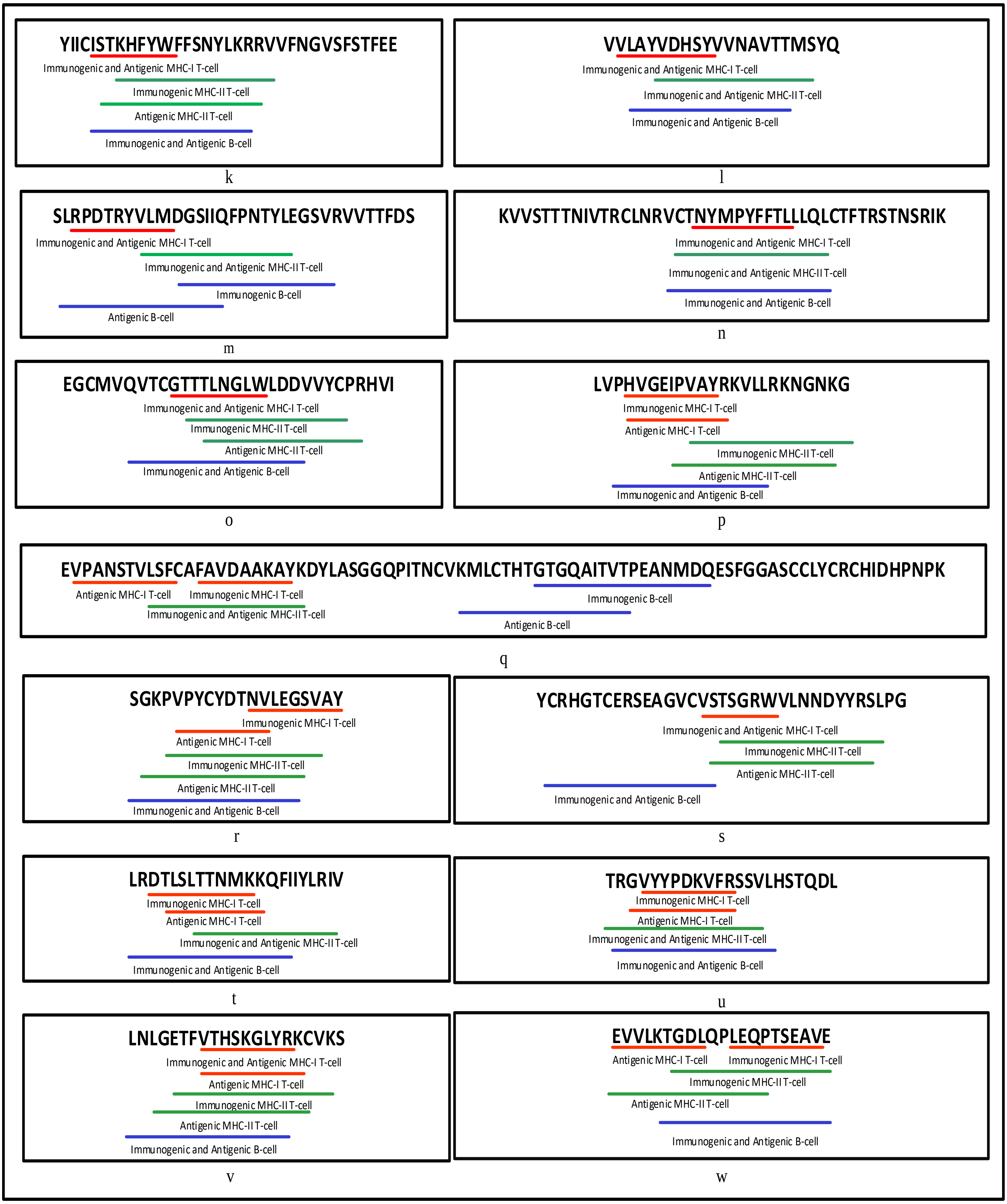
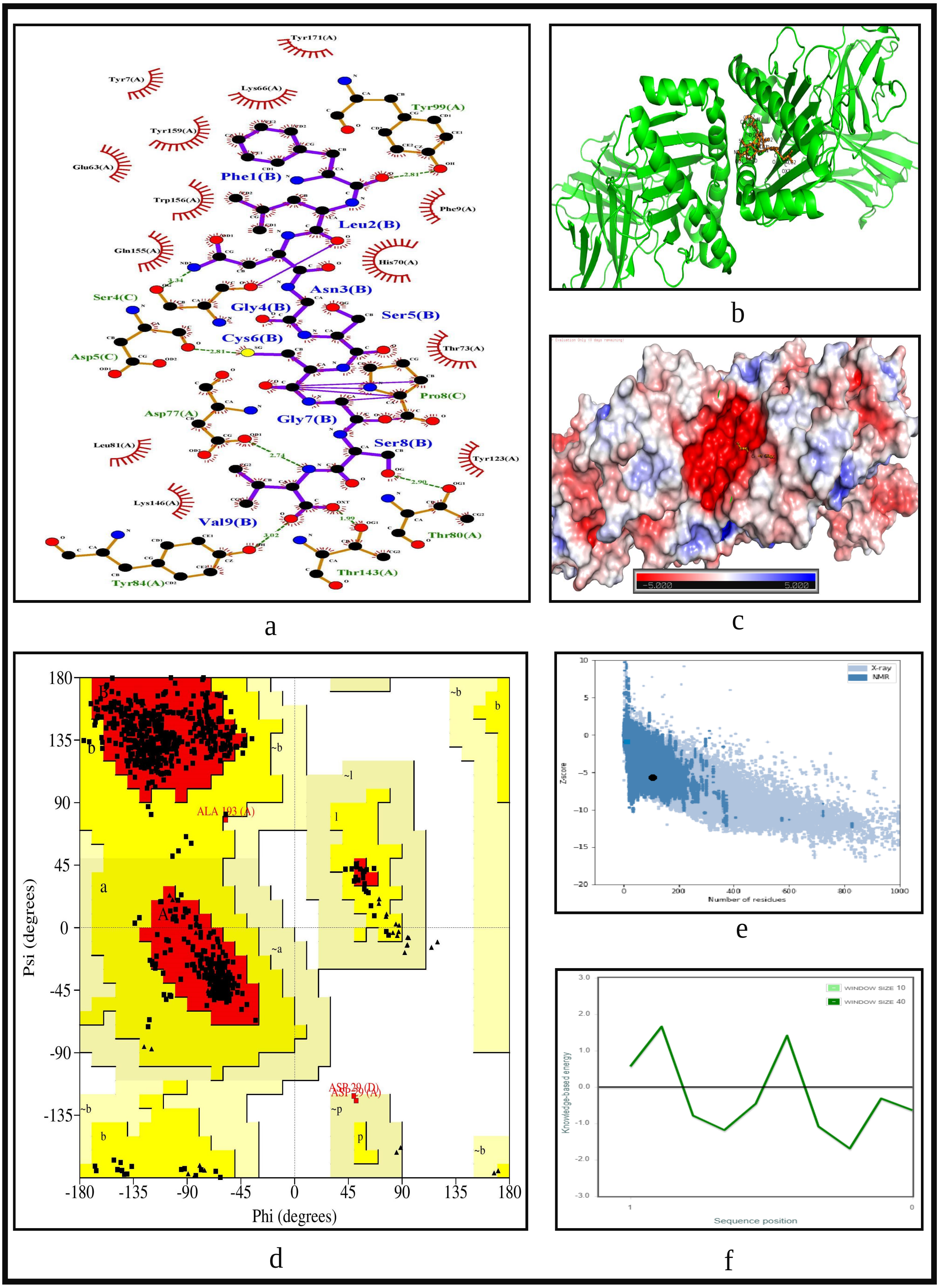
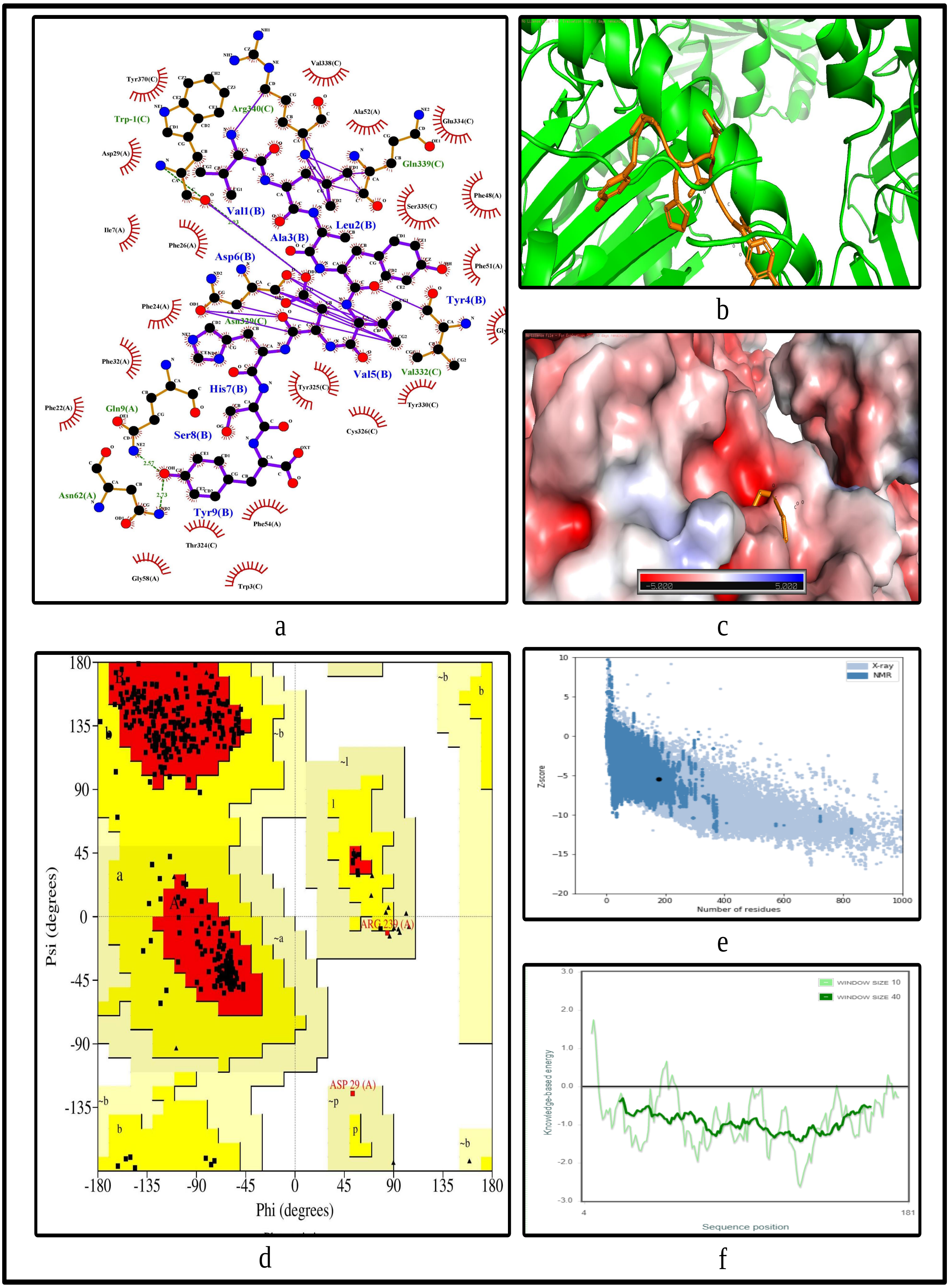


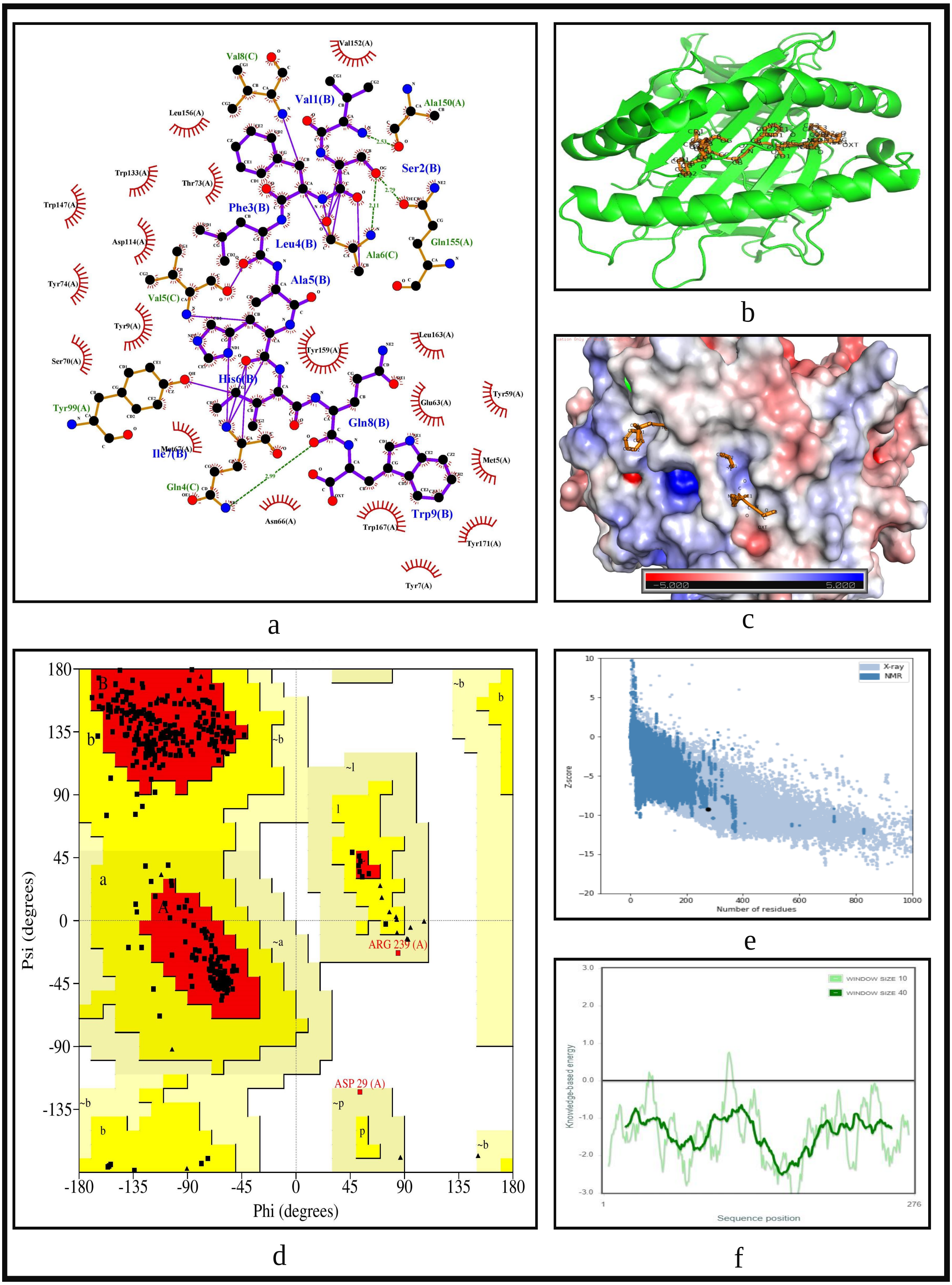
Figure S2: MHC-I, MHC-II restricted T-cell and B-cell epitopes underlined in the protein sequences of the 23 CCnRs


Tyr171(A)
Tyr7(A)
Lys66(A)
Tyr99(A)
Tyr159(A)
Glu63(A)
Phe1(B)
Trp156(A)
Phe9(A)
Leu2(B)
Gln155(A)
His70(A)
3.34
2.81
Ser4(C)
Asn3(B)
Gly4(B)
Ser5(B)
Cys6(B)
Asp5(C)
2.81
Thr73(A)
Pro8(C)
Asp77(A)
Gly7(B)
Ser8(B)
Leu81(A)
2.74
Tyr123(A)
2.90
Lys146(A)
1.99
Val9(B)
3.02
Thr80(A)
Thr143(A)
Tyr84(A)
a
b
-5.000
5.000
c
180
135
90
45
0
-45
-90
-135
-180
Psi (degrees)
Phi (degrees)
B
b
~b
~l
l
ALA 103 (A)
a
A
~a
ASP 29 (D)
ASP 29 (A)
~p
p
d
10
5
0
−5
−10
−15
−20
Z-score
0
200
400
600
800
1000
Number of residues
X-ray
NMR
e
3.0
2.0
1.0
0.0
-1.0
-2.0
-3.0
Knowledge-based energy
WINDOW SIZE 10
WINDOW SIZE 40
1
0
Sequence position
f

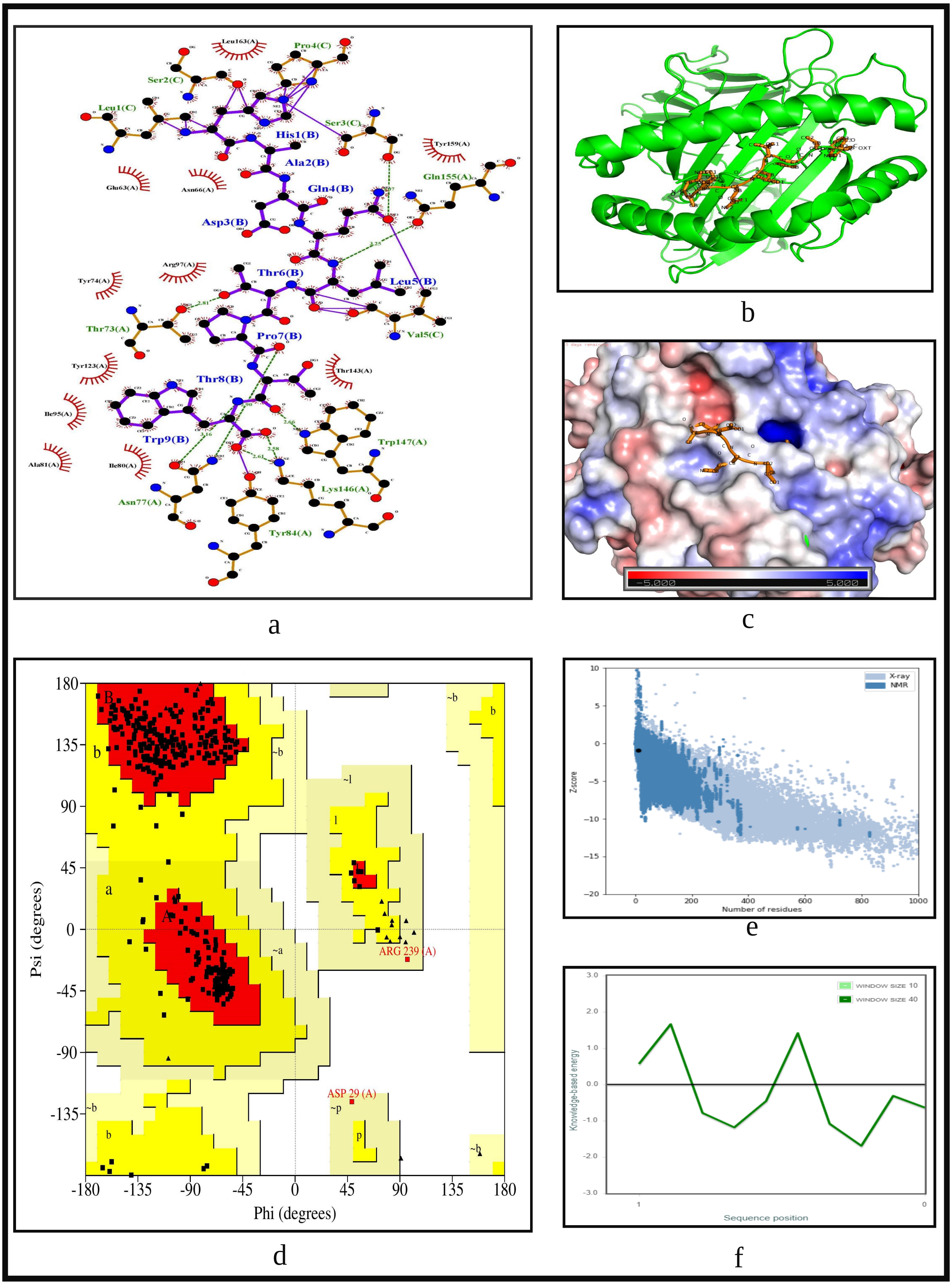
(I)


a
Val1(B)
Leu2(B)
Ala3(B)
Tyr4(B)
Val5(B)
Asp6(B)
His7(B)
Ser8(B)
Tyr9(B)
Trp-1(C)
Arg340(C)
Gln339(C)
Asn329(C)
Val332(C)
Gln9(A)
Asn62(A)
Tyr370(C)
Val338(C)
Ala52(A)
Glu334(C)
Phe48(A)
Ser335(C)
Phe51(A)
Asp29(A)
Ile7(A)
Phe26(A)
Phe24(A)
Phe32(A)
Phe22(A)
Tyr325(C)
Tyr330(C)
Cys326(C)
Phe54(A)
Thr324(C)
Gly58(A)
Trp3(C)
2.57
2.73
b
c
-5.000
5.000
d
Psi (degrees)
Phi (degrees)
ARG 239 (A)
ASP 29 (A)
e
X-ray
NMR
Z-score
Number of residues
f
WINDOW SIZE 10
WINDOW SIZE 40
Knowledge-based energy
Sequence position

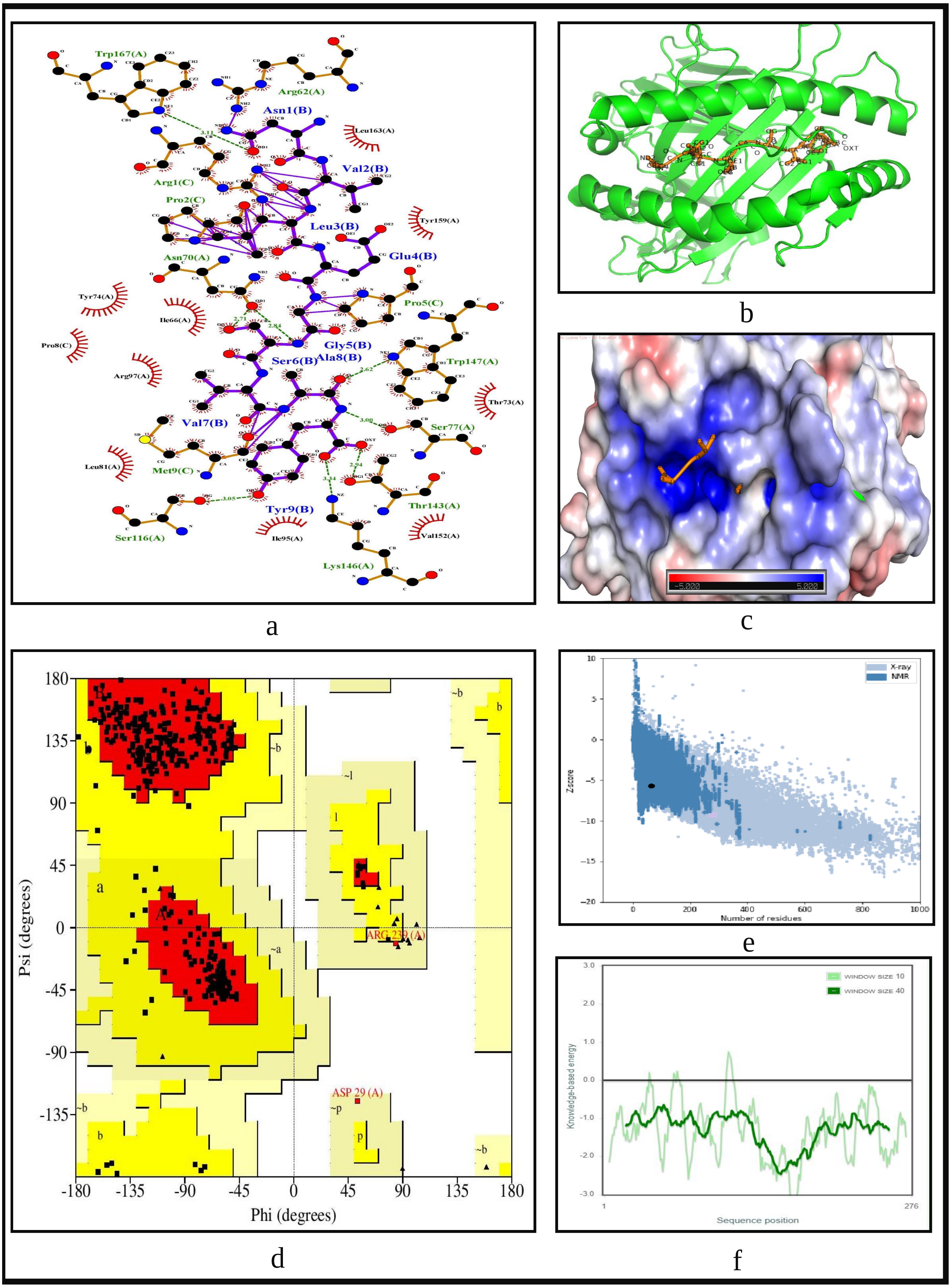
(II)


Val8(C)
Val152(A)
Val1(B)
Ala150(A)
Leu156(A)
Ser2(B)
Trp133(A)
Thr73(A)
Trp147(A)
Phe3(B)
Gln155(A)
Leu4(B)
Asp114(A)
Ala6(C)
Ala5(B)
Tyr74(A)
Val5(C)
Leu163(A)
Tyr9(A)
Tyr159(A)
Ser70(A)
His6(B)
Tyr59(A)
Glu63(A)
Tyr99(A)
Gln8(B)
Ile7(B)
Met67(A)
Met5(A)
Gln4(C)
Asn66(A)
Trp167(A)
Trp9(B)
Tyr171(A)
Tyr7(A)
2.53
2.79
2.11
2.99
a
b
-5.000
5.000
c
180
135
90
45
0
-45
-90
-135
-180
Psi (degrees)
Phi (degrees)
B
b
~b
~l
l
a
A
~a
ARG 239 (A)
ASP 29 (A)
~p
p
d
X-ray
NMR
Z-score
Number of residues
10
5
-5
-10
-15
-20
200
400
600
800
1000
e
WINDOW SIZE 10
WINDOW SIZE 40
Knowledge-based energy
Sequence position
3.0
2.0
1.0
0.0
-1.0
-2.0
-3.0
1
276
f

(III)


Leu163(A)
Pro4(C)
Ser2(C)
Leu1(C)
Ser3(C)
His1(B)
Tyr159(A)
Ala2(B)
Gln155(A)
Glu63(A)
Asn66(A)
Gln4(B)
Asp3(B)
Arg97(A)
Tyr74(A)
Thr6(B)
Leu5(B)
Thr73(A)
Val5(C)
Pro7(B)
Tyr123(A)
Thr143(A)
Thr8(B)
Ile95(A)
Trp9(B)
Trp147(A)
Ala81(A)
Ile80(A)
Lys146(A)
Asn77(A)
Tyr84(A)
2.81
3.16
2.90
2.66
2.58
2.61
3.25
a
b
-5.000
5.000
c
180
135
90
45
0
-45
-90
-135
-180
Psi (degrees)
Phi (degrees)
B
b
~b
~l
l
a
A
~a
ARG 239 (A)
ASP 29 (A)
~p
p
d
10
5
0
-5
-10
-15
-20
200
400
600
800
1000
Z-score
Number of residues
X-ray
NMR
e
3.0
2.0
1.0
0.0
-1.0
-2.0
-3.0
Knowledge-based energy
Sequence position
WINDOW SIZE 10
WINDOW SIZE 40
1
f

(IV)

Trp167(A)
Arg62(A)
Asn1(B)
Leu163(A)
Val2(B)
Arg1(C)
Pro2(C)
Tyr159(A)
Leu3(B)
Glu4(B)
Asn70(A)
Tyr74(A)
Ile66(A)
Pro5(C)
Pro8(C)
Gly5(B)
Ser6(B)
Ala8(B)
Trp147(A)
Arg97(A)
Thr73(A)
Val7(B)
Ser77(A)
Leu81(A)
Met9(C)
Tyr9(B)
Thr143(A)
Ser116(A)
Ile95(A)
Val152(A)
Lys146(A)
3.11
2.71
2.84
2.62
3.00
2.94
3.34
3.05
-5.000
5.000
Psi (degrees)
Phi (degrees)
ARG 239 (A)
ASP 29 (A)
X-ray
NMR
Z-score
Number of residues
WINDOW SIZE 10
WINDOW SIZE 40
Knowledge-based energy
Sequence position
1
276

a

b

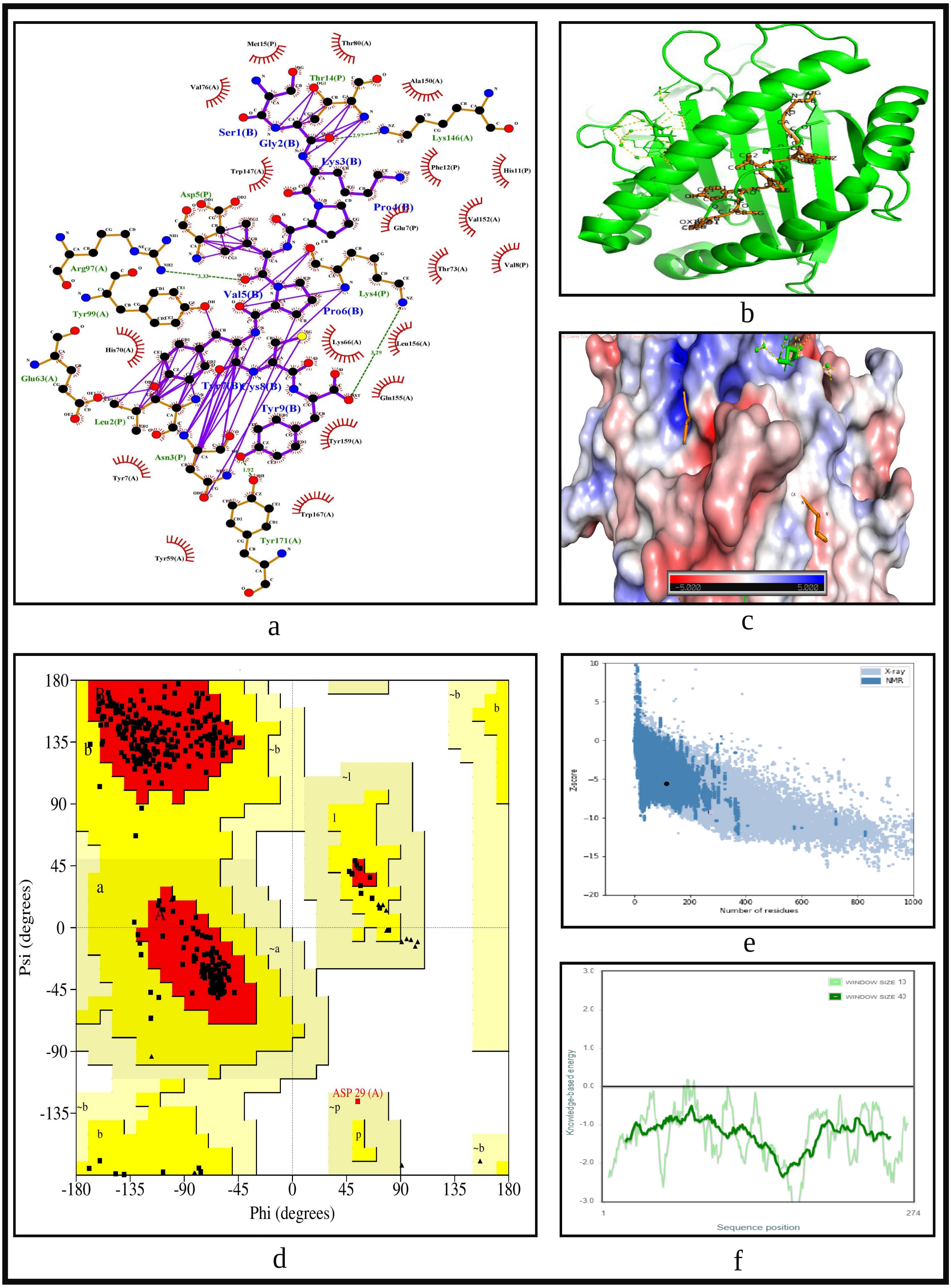
c

d

e

f

(V)


Met15(P)
Thr80(A)
Val76(A)
Thr14(P)
Ala150(A)
Ser1(B)
Gly2(B)
2.97
Lys146(A)
Lys3(B)
Trp147(A)
Phe12(P)
His11(P)
Asp5(P)
Pro4(B)
Glu7(P)
Val152(A)
Arg97(A)
3.33
Thr73(A)
Val8(P)
Val5(B)
Lys4(P)
Tyr99(A)
Pro6(B)
His70(A)
Lys66(A)
Leu156(A)
3.29
Glu63(A)
Tyr7(B)
Cys8(B)
Gln155(A)
Tyr9(B)
Leu2(P)
Tyr159(A)
Asn3(P)
Tyr7(A)
1.92
Trp167(A)
Tyr171(A)
Tyr59(A)
-5.000
5.000
180
135
90
45
0
-45
-90
-135
-180
Psi (degrees)
Phi (degrees)
B
b
~b
a
A
~a
~l
l
~p
p
ASP 29 (A)
X-ray
NMR
Z-score
Number of residues
10
5
-5
-10
-15
-20
200
400
600
800
1000
WINDOW SIZE 10
WINDOW SIZE 40
Knowledge-based energy
Sequence position
3.0
2.0
1.0
0.0
-1.0
-2.0
-3.0
1
274

a

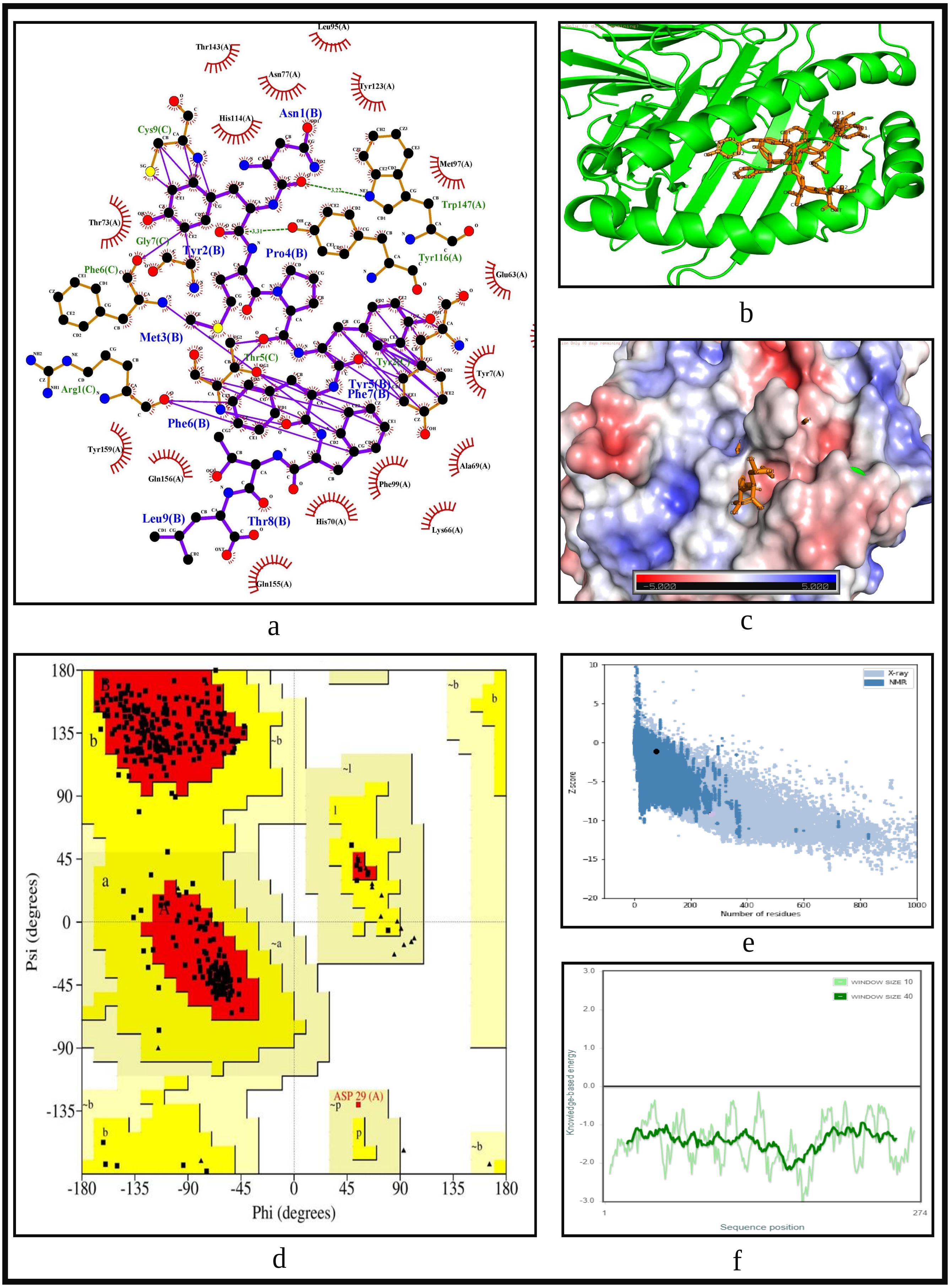
b

c

d

e

f


Leu95(A)
Thr143(A)
Asn77(A)
Tyr123(A)
His114(A)
Asn1(B)
Cys9(C)
Met97(A)
3.22
Trp147(A)
Thr73(A)
3.31
Gly7(C)
Tyr2(B)
Pro4(B)
Tyr116(A)
Phe6(C)
Glu63(A)
Met3(B)
Thr5(C)
Tyr2(C)
Tyr7(A)
Tyr5(B)
Phe7(B)
Arg1(C)
Phe6(B)
Tyr159(A)
Gln156(A)
Ala69(A)
Phe99(A)
Leu9(B)
Thr8(B)
His70(A)
Lys66(A)
Gln155(A)
-5.000
5.000
Psi (degrees)
Phi (degrees)
180
135
90
45
0
-45
-90
-135
-180
ASP 29 (A)
X-ray
NMR
Z-score
Number of residues
10
5
0
-5
-10
-15
-20
200
400
600
800
1000
WINDOW SIZE 10
WINDOW SIZE 40
Knowledge-based energy
Sequence position
3.0
2.0
1.0
0.0
-1.0
-2.0
-3.0
1
274

a

b

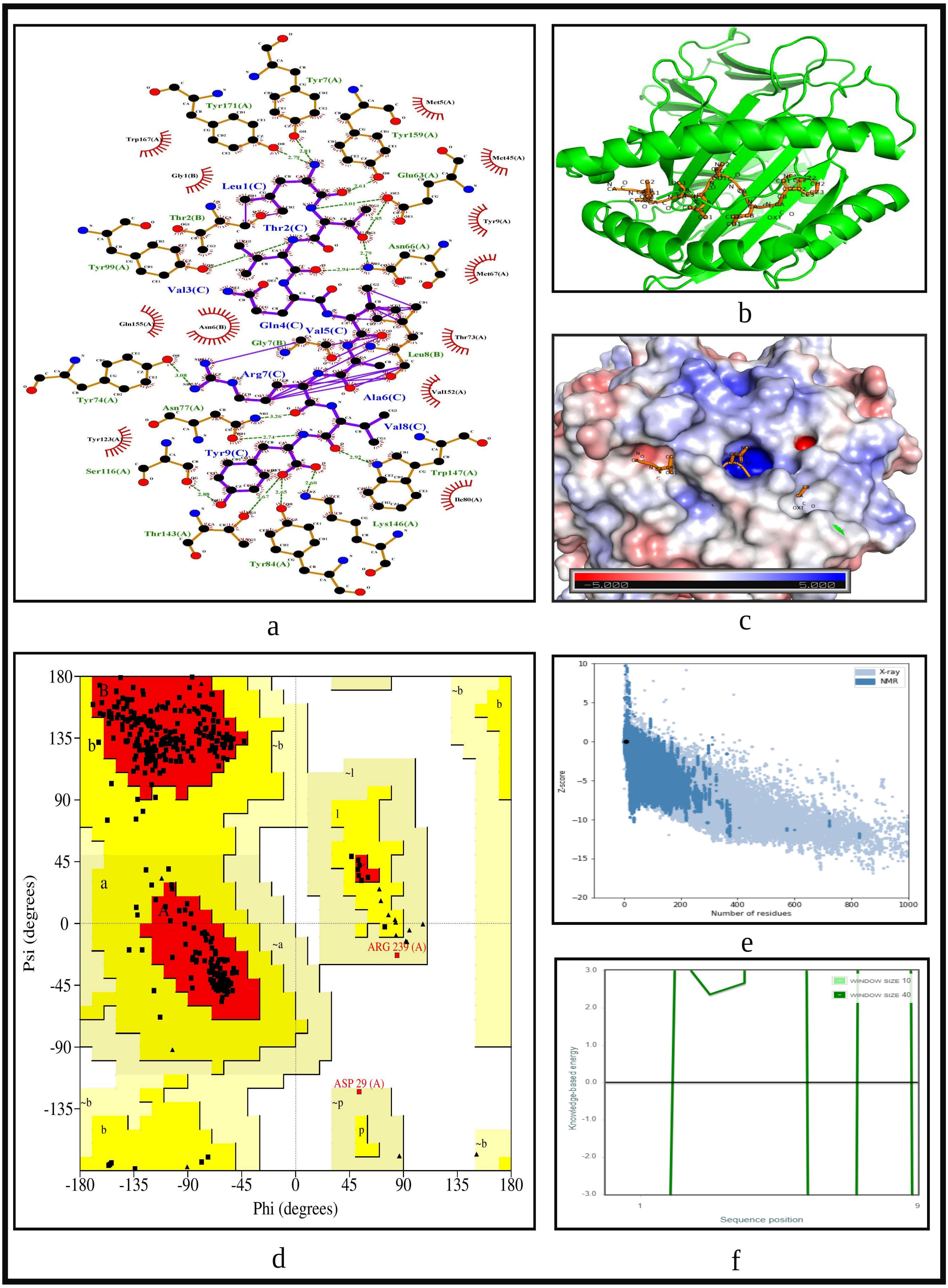
c

d

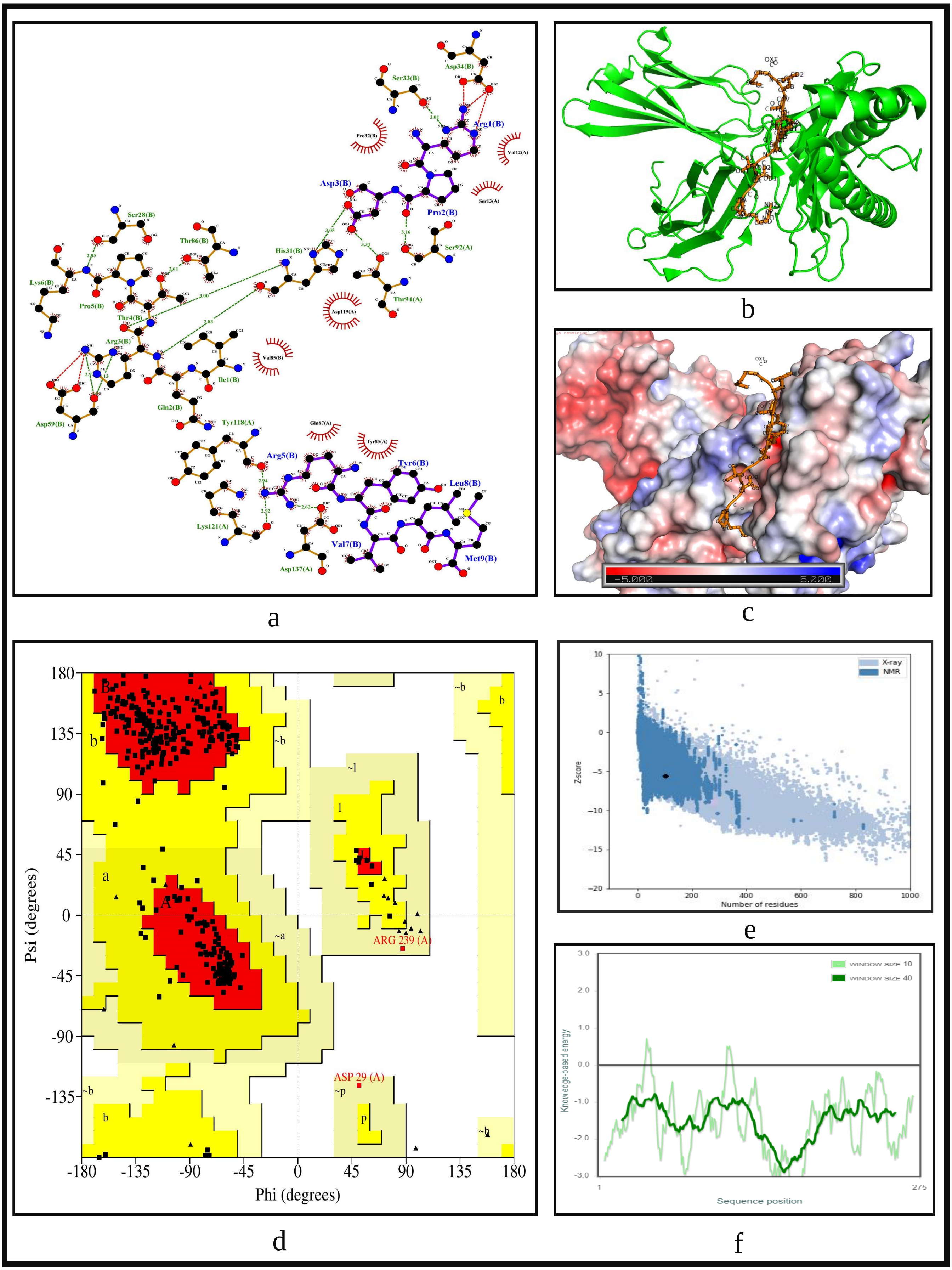
e

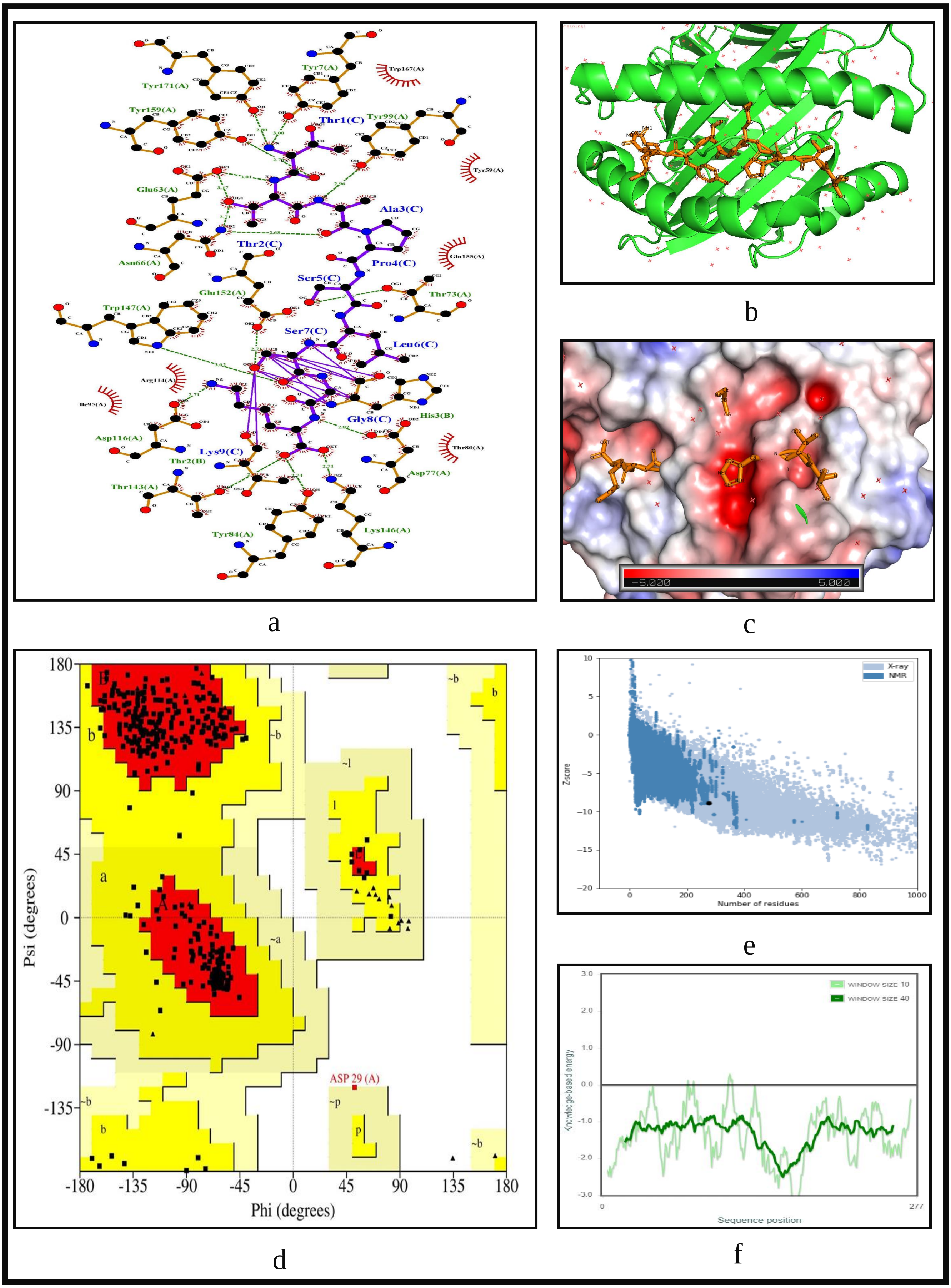
f

(VII)


Tyr7(A)
Tyr171(A)
Met5(A)
Trp167(A)
Tyr159(A)
Met45(A)
Gly1(B)
Leu1(C)
Glu63(A)
2.79
2.81
2.61
3.01
2.85
Thr2(B)
Thr2(C)
Tyr9(A)
Tyr99(A)
Asn66(A)
2.79
2.94
Met67(A)
Val3(C)
Gln155(A)
Asn6(B)
Gln4(C)
Val5(C)
Thr73(A)
Gly7(B)
Leu8(B)
Arg7(C)
3.08
Tyr74(A)
Val152(A)
Ala6(C)
Asn77(A)
3.26
Val8(C)
Tyr123(A)
2.74
Tyr9(C)
2.92
Ser116(A)
Trp147(A)
2.88
2.67
2.65
2.60
Ile80(A)
Lys146(A)
Thr143(A)
Tyr84(A)
a
b
-5.000
5.000
c
180
135
90
45
0
-45
-90
-135
-180
-135
-90
-45
0
45
90
135
180
Psi (degrees)
Phi (degrees)
B
b
~b
A
a
~a
~l
l
ARG 239 (A)
ASP 29 (A)
~p
p
d
10
5
0
−5
−10
−15
−20
0
200
400
600
800
1000
Z-score
Number of residues
X-ray
NMR
e
3.0
2.0
1.0
0.0
-1.0
-2.0
-3.0
WINDOW SIZE 10
WINDOW SIZE 40
Knowledge-based energy
1
9
Sequence position
f

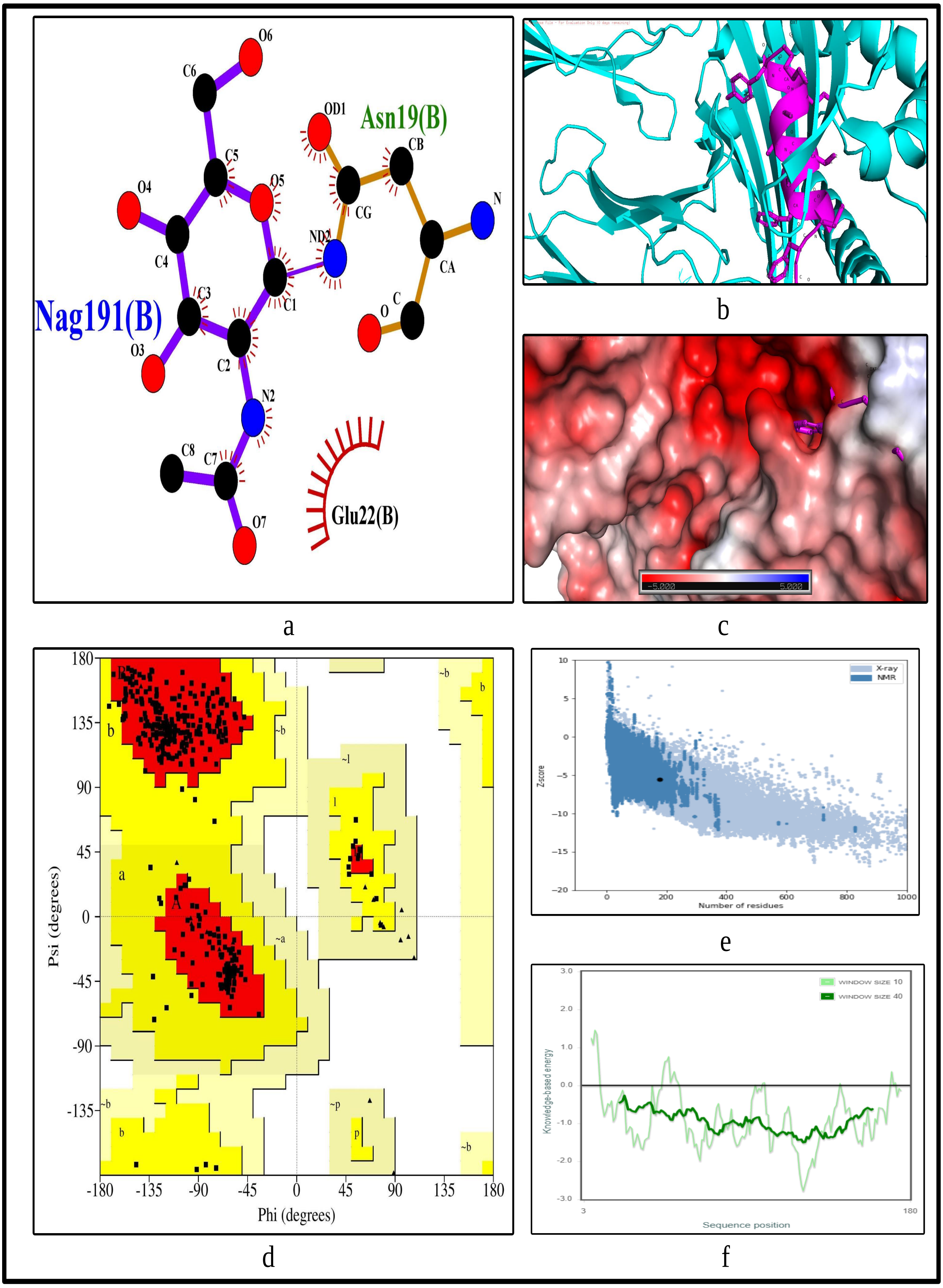
(VIII)


a
Arg1(B)
Pro2(B)
Asp3(B)
Ser33(B)
Asp34(B)
Pro32(B)
Val12(A)
Ser13(A)
Ser92(A)
Thr94(A)
Asp119(A)
His31(B)
Ser28(B)
Thr86(B)
Lys6(B)
Pro5(B)
Thr4(B)
Arg3(B)
Ile1(B)
Gln2(B)
Val85(B)
Asp59(B)
Tyr118(A)
Gln87(A)
Tyr85(A)
Arg5(B)
Tyr6(B)
Leu8(B)
Lys121(A)
Val7(B)
Met9(B)
Asp137(A)
b
c
-5.000
5.000
d
Psi (degrees)
Phi (degrees)
180
135
90
45
0
-45
-90
-135
-180
ARG 239 (A)
ASP 29 (A)
e
X-ray
NMR
Z-score
Number of residues
10
5
0
−5
−10
−15
−20
200
400
600
800
1000
f
WINDOW SIZE 10
WINDOW SIZE 40
Knowledge-based energy
Sequence position
3.0
2.0
1.0
0.0
-1.0
-2.0
-3.0
1
275



(X)

Figure S3: Structure analysis of the obtained MHC-I restricted T-cell epitope for immunogenic and antigenic in all Consensus Conserved Regions of SARS-CoV-2 genome (a) 2D pose representation between the epitope and HLA allele showing the different non-covalent bonds (b) Docking structure of MHC-I restricted T-cell epitope (c) The surface interaction between the allele and epitopes showing the fitting sites in binding grooves (d) Ramachandran plot of the epitope allele structure showing lower energy sites of the residues in different frame (e) Z-score plot and (f) all residue energy for (I) GSVGFNIDY immunogenic peptide of 3CL-Proteinase bound to HLA-A*30:02 (II) DLKGKYVQI peptide of NSP10 bound to HLA-B*08:01 (III) VSFLAHIQW peptide of NS4 bound to HLA-B*57:01 (IV) HADQLTPTW peptide of Spike glycoprotein bound to HLA-B*58:01 (V) NVLEGSVAY peptide of NSP4 bound to HLA-B*35:01 (VI) SGKPVPYCY peptide of NSP4 bound to HLA-A*30:02 (VII) NYMPYFFTL peptide of NSP3 bound to HLA-A*24:02 (VIII) GTTTLNGLW peptide of 3CL-Proteinase bound to HLA-B*57:01 (IX) RPDTRYVLM peptide of NSP4 bound to HLA-B*07:02 (X) VTHSKGLYRK peptide of NSP4 bound to HLA-A*03:01


O6
C6
OD1
Asn19(B)
C5
CB
O4
O5
CG
N
ND2
C4
CA
C3
C1
Nag191(B)
O
C
O3
C2
N2
C8
C7
O7
Glu22(B)
-5.000
5.000
180
135
90
45
0
-45
-90
-135
-180
Psi (degrees)
Phi (degrees)
B
b
~b
~l
l
a
A
~a
~p
p
X-ray
NMR
Z-score
Number of residues
10
5
-5
-10
-15
-20
200
400
600
800
1000
WINDOW SIZE 10
WINDOW SIZE 40
Knowledge-based energy
3.0
2.0
1.0
0.0
-1.0
-2.0
-3.0
3
Sequence position

a

b

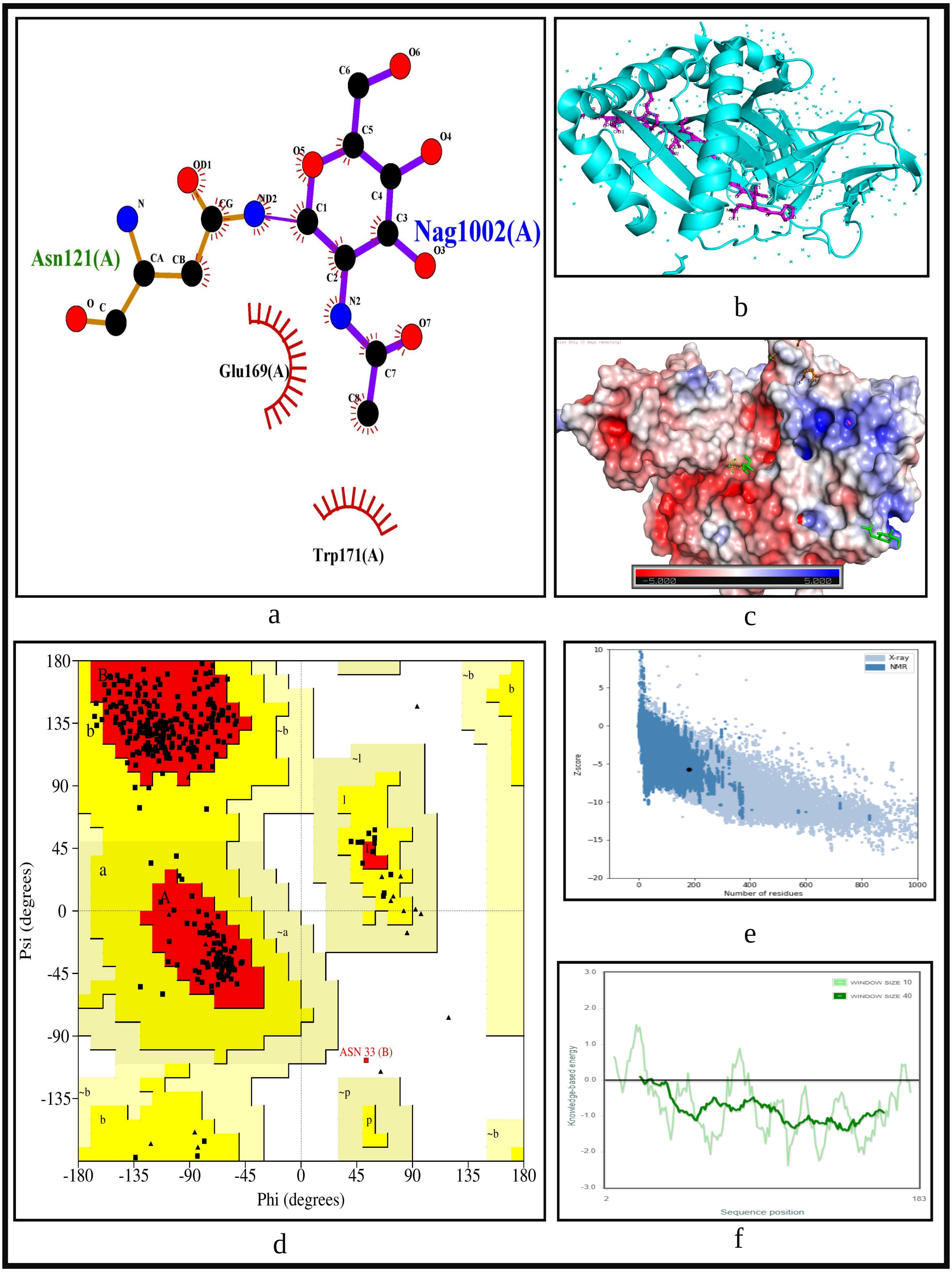
c

d

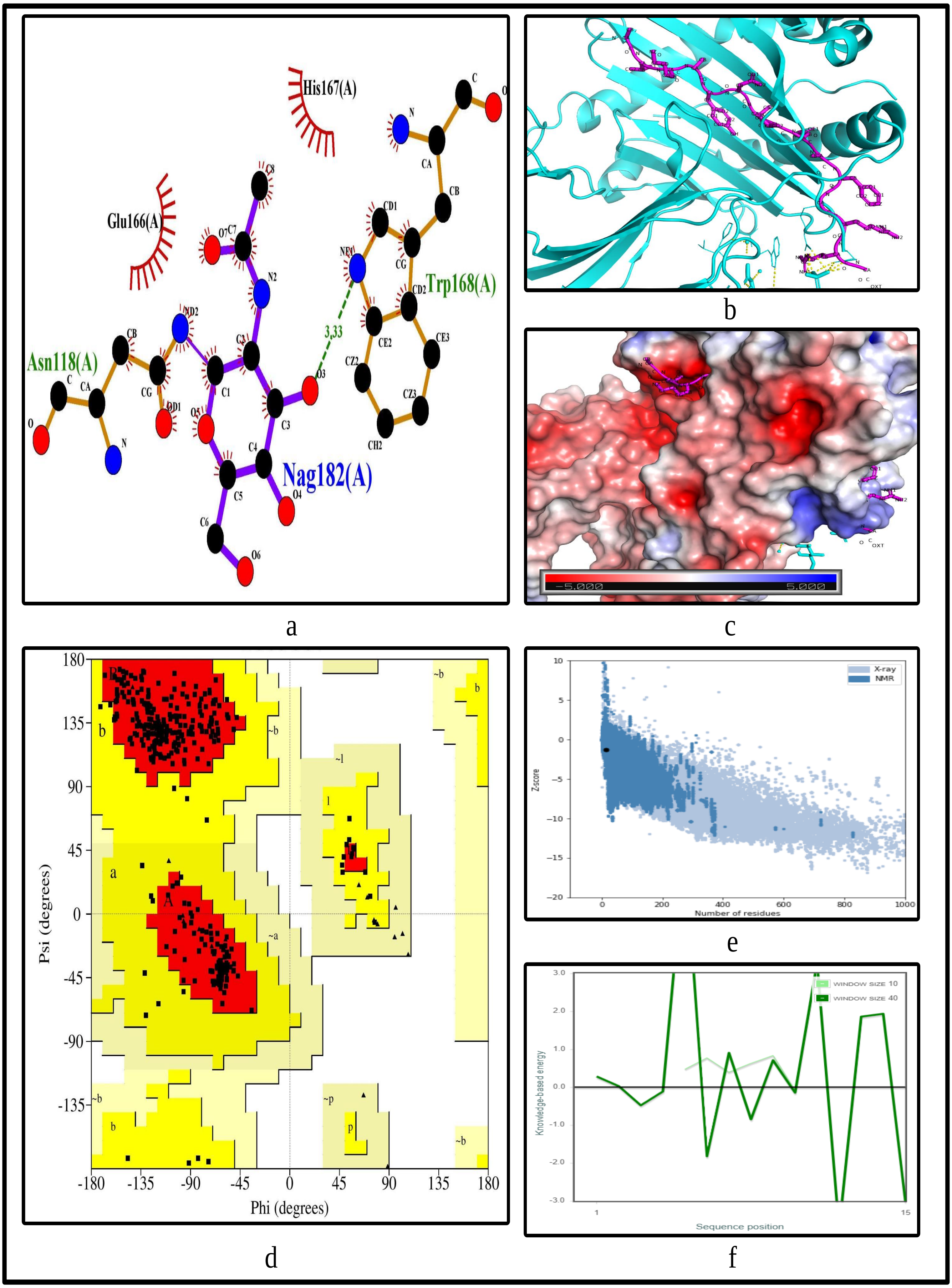
e

f

(I)


Asn121(A)
Nag1002(A)
Glu169(A)
Trp171(A)
O6
C6
C5
O4
O5
OD1
C4
C3
CG
ND2
C1
N
CA
CB
O3
C2
N2
O
C
O7
C7
C8
a
b
-5.000
5.000
c
180
135
90
45
0
-45
-90
-135
-180
-135
-90
-45
0
45
90
135
180
Psi (degrees)
Phi (degrees)
B
b
A
a
~b
~l
l
~a
~p
p
ASN 33 (B)
d
X-ray
NMR
Z-score
Number of residues
10
5
0
−5
−10
−15
−20
0
200
400
600
800
1000
e
WINDOW SIZE 10
WINDOW SIZE 40
Knowledge-based energy
Sequence position
3.0
2.0
1.0
0.0
-1.0
-2.0
-3.0
2
183
f

(III)


a

b

c

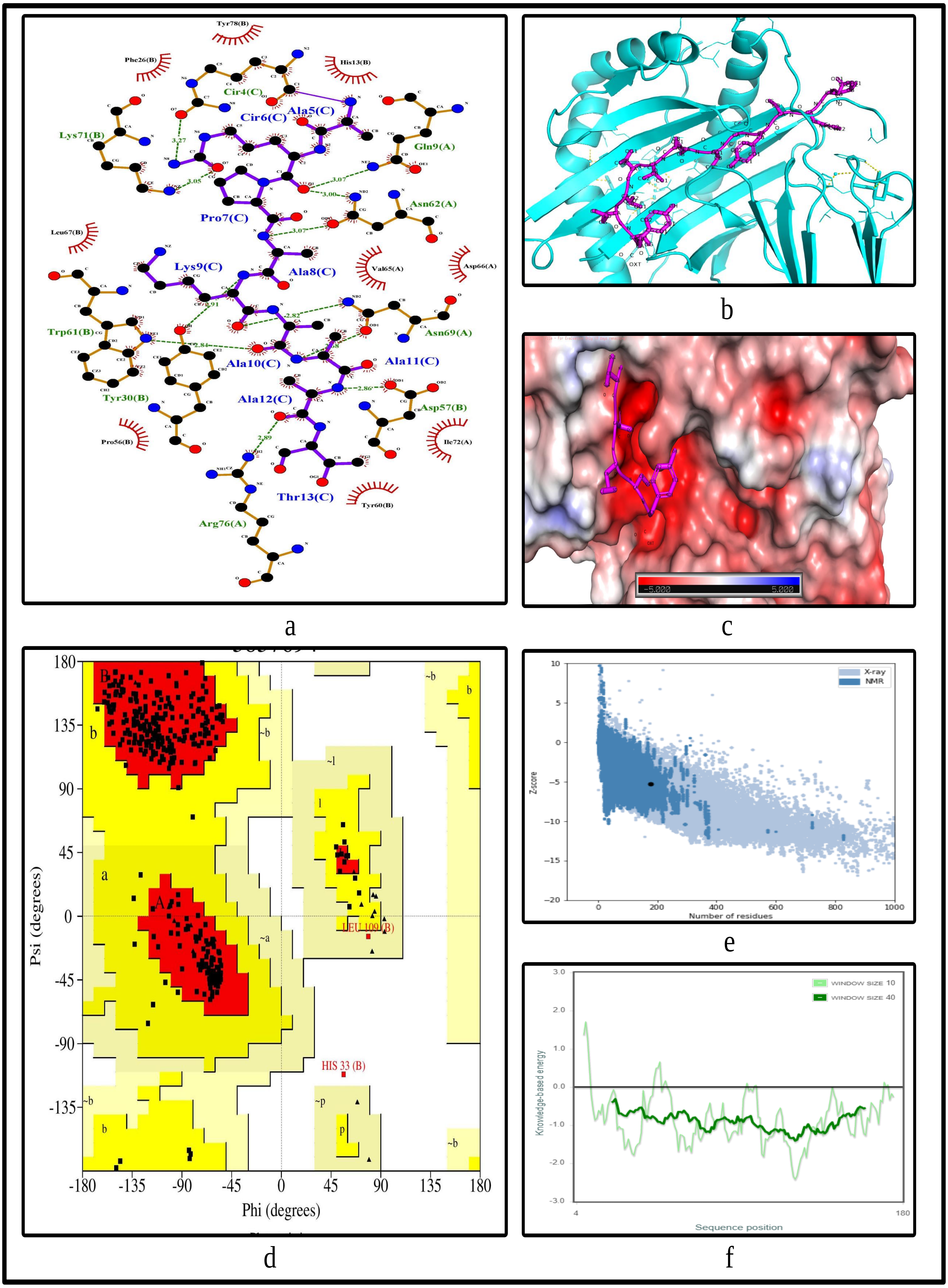
d

e

f

His167(A)
Glu166(A)
Asn118(A)
Trp168(A)
Nag182(A)
3.33
C
O
N
CA
CB
CG
ND2
OD1
C1
C2
C3
C4
C5
C6
O3
O4
O5
O6
O7
C7
C8
N2
NE1
CD1
CD2
CE2
CE3
CZ2
CZ3
CH2
-5.000
5.000
180
135
90
45
0
-45
-90
-135
-180
Psi (degrees)
Phi (degrees)
b
~b
a
A
~a
~l
l
~p
p
X-ray
NMR
Z-score
Number of residues
10
5
-5
-10
-15
-20
200
400
600
800
1000
WINDOW SIZE 10
WINDOW SIZE 40
Knowledge-based energy
Sequence position
3.0
2.0
1.0
0.0
-1.0
-2.0
-3.0
1
15

(II)


Tyr78(B)
Phe26(B)
His13(B)
Cir4(C)
Cir6(C)
Ala5(C)
Lys71(B)
Gln9(A)
Asn62(A)
Pro7(C)
Leu67(B)
Lys9(C)
Ala8(C)
Val65(A)
Asp66(A)
Trp61(B)
Asn69(A)
Ala10(C)
Ala11(C)
Tyr30(B)
Ala12(C)
Asp57(B)
Pro56(B)
Ile72(A)
Thr13(C)
Tyr60(B)
Arg76(A)
3.27
3.05
3.07
3.00
2.91
2.82
2.84
2.86
2.89
OXT
-5.000
5.000
180
135
90
45
0
-45
-90
-135
-180
Psi (degrees)
Phi (degrees)
LEU 109 (B)
HIS 33 (B)
~b
~a
~l
~p
X-ray
NMR
Z-score
Number of residues
10
5
0
-5
-10
-15
-20
200
400
600
800
1000
WINDOW SIZE 10
WINDOW SIZE 40
Knowledge-based energy
Sequence position
3.0
2.0
1.0
0.0
-1.0
-2.0
-3.0
4
180

a

b

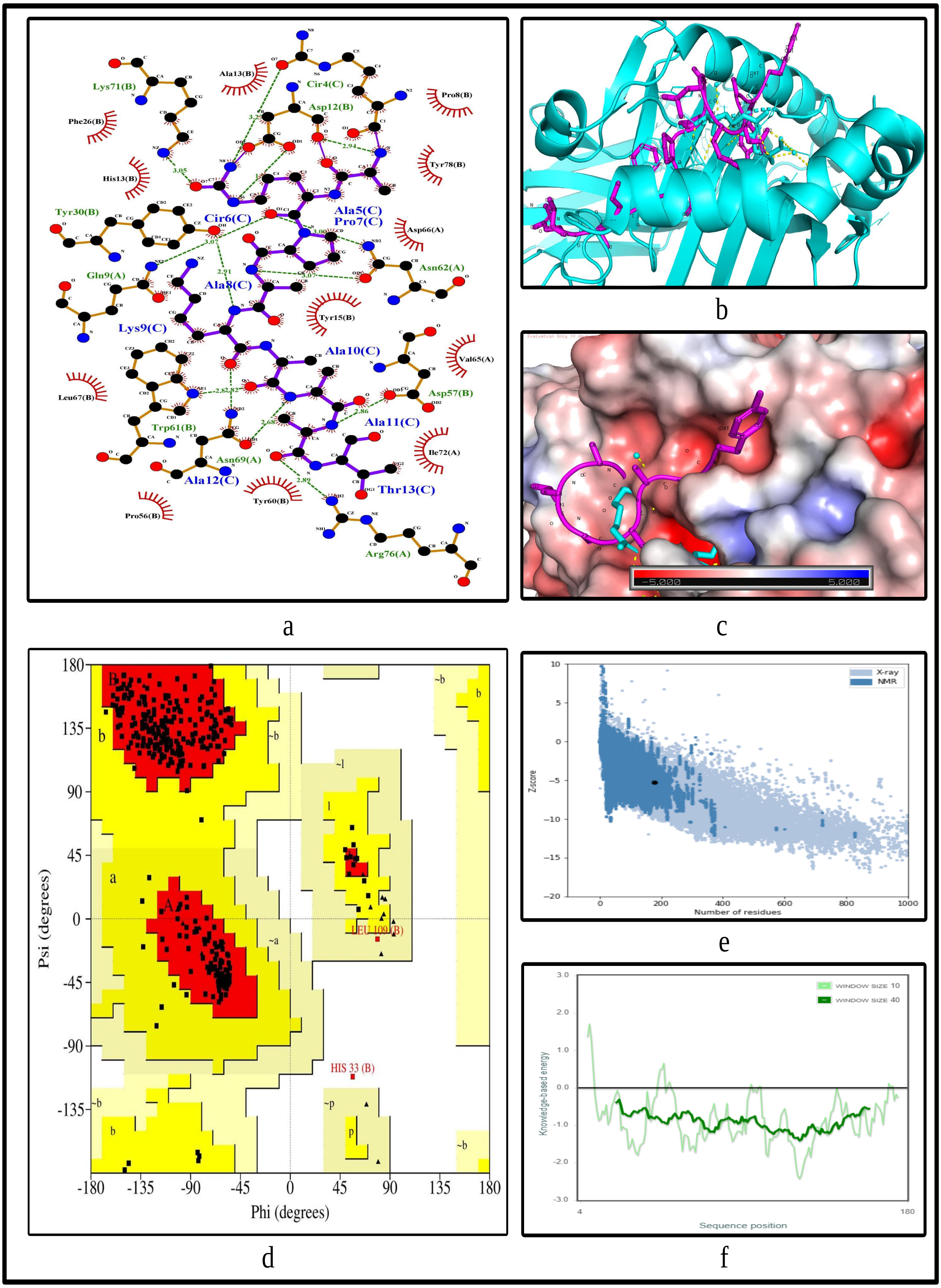
c

d

e

f

(IV)


Lys71(B)
Ala13(B)
Cir4(C)
Pro8(B)
Asp12(B)
Phe26(B)
3.22
2.94
Tyr78(B)
His13(B)
3.05
Tyr30(B)
Cir6(C)
Ala5(C)
Pro7(C)
Asp66(A)
3.07
3.00
Gln9(A)
2.91
Asn62(A)
Ala8(C)
Lys9(C)
Tyr15(B)
Ala10(C)
Val65(A)
Leu67(B)
2.82
2.82
Asp57(B)
2.86
Trp61(B)
2.68
Ala11(C)
Ile72(A)
Asn69(A)
Ala12(C)
2.89
Thr13(C)
Tyr60(B)
Pro56(B)
Arg76(A)
-5.000
5.000
Psi (degrees)
Phi (degrees)
180
135
90
45
0
-45
-90
-135
-180
LEU 109 (B)
HIS 33 (B)
B
b
A
a
~b
~a
~l
l
~p
p
X-ray
NMR
Z-score
Number of residues
10
5
0
−5
−10
−15
−20
200
400
600
800
1000
WINDOW SIZE 10
WINDOW SIZE 40
Knowledge-based energy
Sequence position
3.0
2.0
1.0
0.0
-1.0
-2.0
-3.0
4
180

a

b

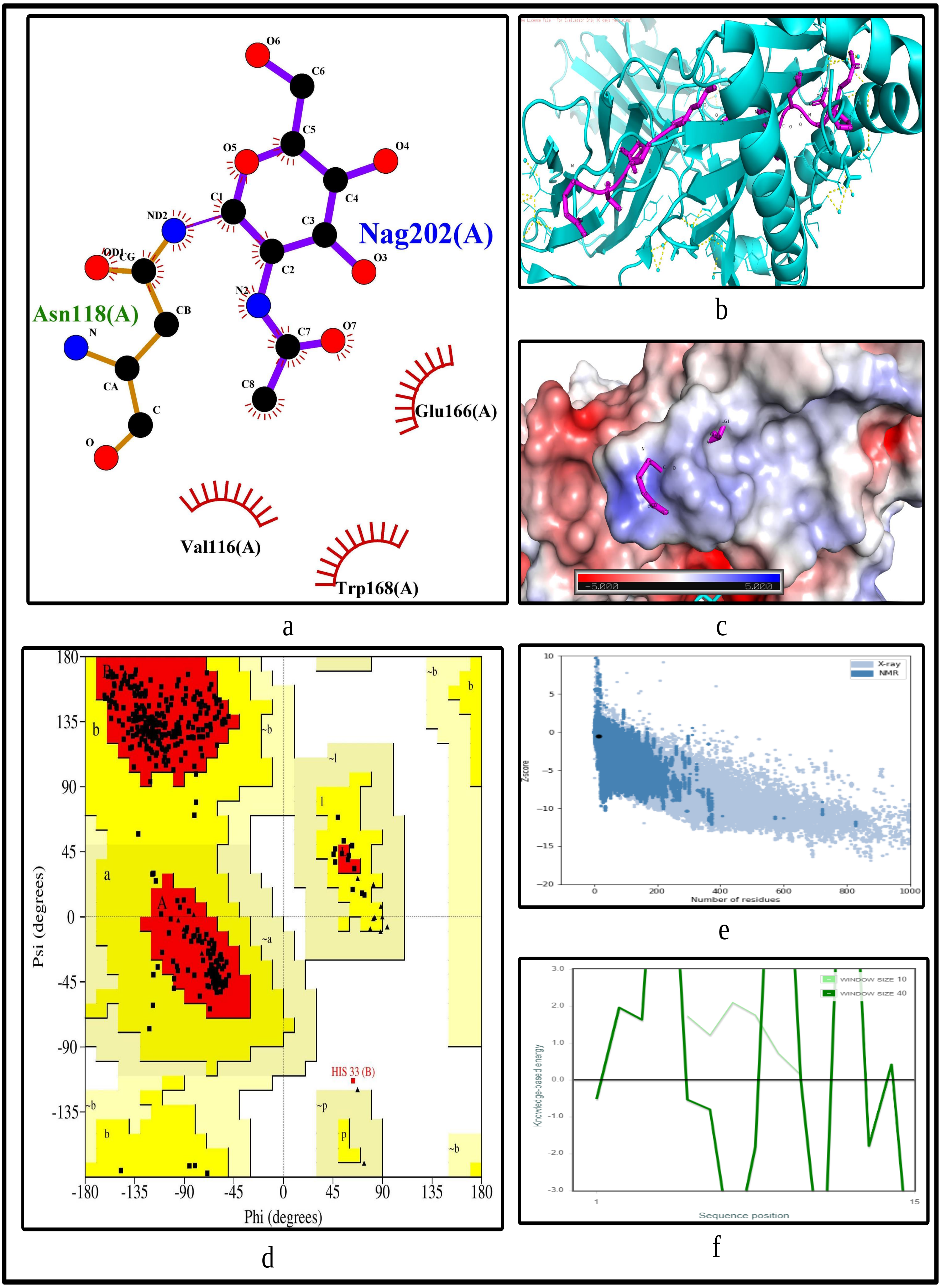
c

d

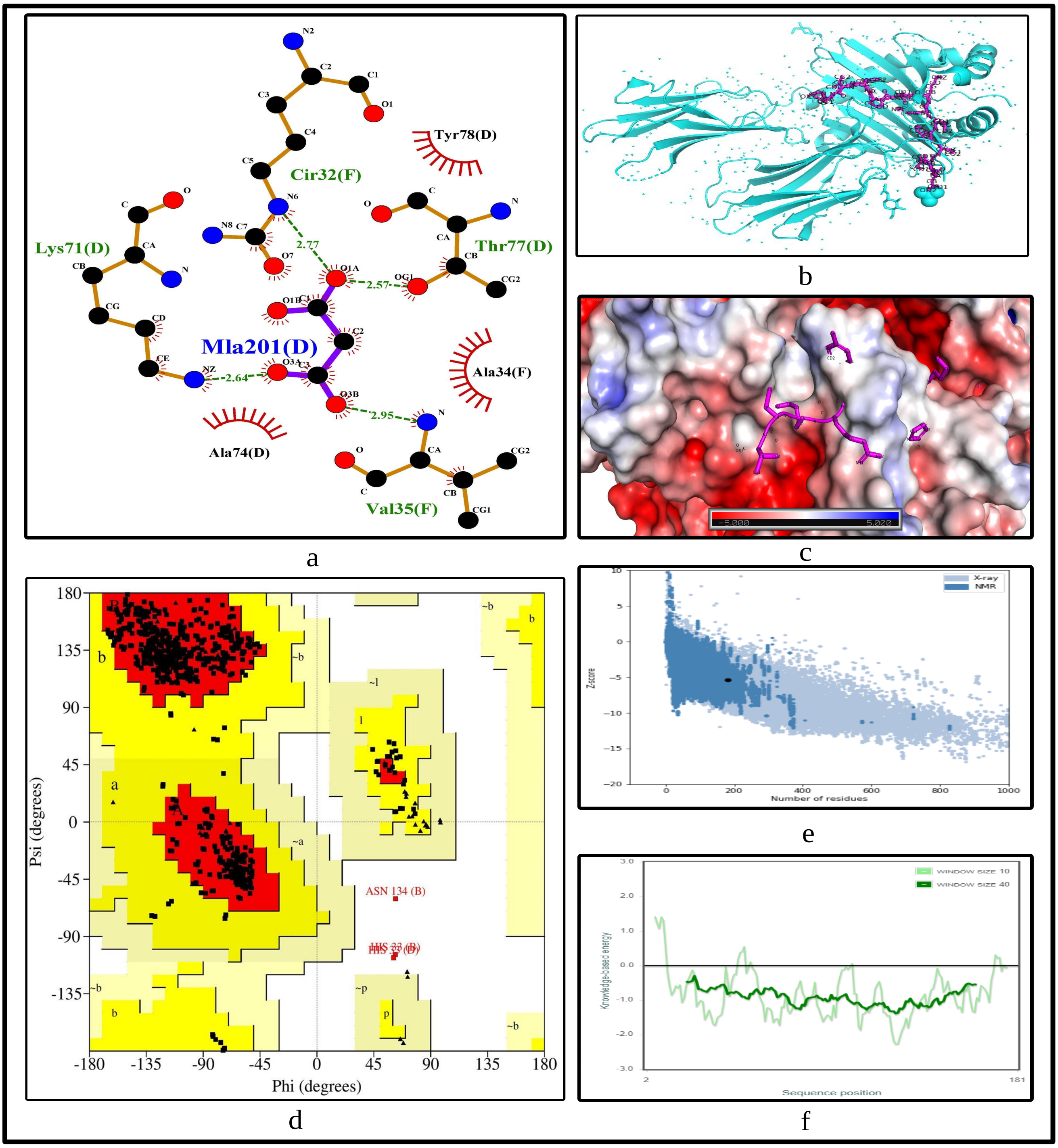
e

f

(V)


O6
C6
C5
O5
O4
C4
C1
ND2
C3
Nag202(A)
OD1
CG
C2
O3
N2
Asn118(A)
CB
N
C7
O7
CA
C8
C
Glu166(A)
O
Val116(A)
Trp168(A)
a
b
-5.000
5.000
c
180
135
90
45
0
-45
-90
-135
Psi (degrees)
-180
-135
-90
-45
0
45
90
135
180
Phi (degrees)
HIS 33 (B)
d
X-ray
NMR
Z-score
Number of residues
e
WINDOW SIZE 10
WINDOW SIZE 40
Knowledge-based energy
Sequence position
f

(VI)



(VII)

Figure S4: SStructure analysis of the obtained MHC-II restricted T-cell epitope for immunogenic and antigenic in all Consensus Conserved Regions of SARS-CoV-2 genome (a) 2D pose representation between the epitope and HLA allele showing the different non-covalent bonds (b) Docking structure of MHC-II restricted T-cell epitope (c) The surface interaction between the allele and epitopes showing the fitting sites in binding grooves (d) Ramachandran plot of the epitope allele structure showing lower energy sites of the residues in different frame (e) Z-score plot and (f) all residue energy for for (I) SVIYLYLTFYLTNDV immunogenic peptide of NSP5 bound to HLA-DPA1*01:03/DPB1*02:01 (II) GVYSVIYLYLTFYLT antigenic peptide of NSP5 bound to HLA-DPA1*01:03/DPB1*02:01 (III) PQVNGLTSIKWADNN immunogenic as well as Antigenic peptide of NSP3 bound to HLA-DQA1*01:02/DQB1*06:02 (IV) IELKFNPPALQDAYY immunogenic as well as Antigenic peptide of NSP3 bound to HLA-DRB3*02:02 (V) PVPYCYDTNVLEGSV immunogenic peptide of NSP4 bound to HLA-DRB1*04:01 (VI) GKPVPYCYDTNVLEG immunogenic as well as anitgenic peptide of NSP10 bound to HLA-DRB1*04:01 (VII) LMDGSIIQFPNTYLE immunogenic as well as Antigenic peptide of NSP4 bound to HLA-DRB1*15:01